University of Novi Sad
DSpace-CRIS Repository https://open.uns.ac.rs

2022-05

Latent, genetic, and molecular genetic structure of the Wisconsin Card Sorting Test

Željka Nikolašević, Vojislava Bugarski Ignjatović, Jasmina Kodžopeljić, Selka Sadiković, Ilija Milovanović, Nataša Vučinić, Mechthild Prinz, Zoran Budimlija, Snežana Smederevac

American Psychological Association

Željka Nikolašević, Vojislava Bugarski Ignjatović, Jasmina Kodžopeljić, Selka Sadiković, Ilija Milovanović, et al. 2022. Latent, genetic, and molecular genetic structure of the Wisconsin Card Sorting Test. Neuropsychology 36(4): 314–329. doi: DOI: 10.1037/neu0000800.
https://open.uns.ac.rs/handle/123456789/32490
Downloaded from DSpace-CRIS - University of Novi Sad

Cite article: Nikolašević, Ž., Bugarski Ignjatović, V., Kodžopeljić, J., Sadiković, S., Milovanović, I., Vučinić, N., ... & Smederevac, S. (2022). Latent, genetic, and molecular genetic structure of the Wisconsin Card Sorting Test. *Neuropsychology*, *36*(4), 314. https://doi.org/10.1037/neu0000800

Running head: LATENT, GENETIC, AND MOLECULAR GENETIC STRUCTURE OF WCST

LATENT, GENETIC, AND MOLECULAR GENETIC STRUCTURE OF THE WISCONSIN CARD SORTING TEST

Željka Nikolašević[a]*, Vojislava Bugarski Ignjatović[a,c], Jasmina Kodžopeljić [b], Selka Sadiković[b], Ilija Milovanović[b], Nataša Vučinić [a], Mechthild Prinz[d] , Zoran Budimlija[e] and Snežana Smederevac[b]

Author Note:

[a] University of Novi Sad, Faculty of Medicine, Department of Psychology, Novi Sad, Serbia, Hajduk Veljkova 3, 21000 Novi Sad, Republic of Serbia.

[b] University of Novi Sad, Faculty of Philosophy, Department of Psychology, Novi Sad, Serbia, Dr Zorana Đinđića 2, 21000 Novi Sad, Republic of Serbia.

[c] Neurology Clinic, Clinical Center of Vojvodina, Novi Sad, Serbia, Hajduk Veljkova 3, 21000 Novi Sad, Republic of Serbia.

[d] City University of New York, John Jay College of Criminal Justice, New York, NY, 524 W 59th St, New York, NY 10019, USA.

[e] NYU School of Medicine, Department of Neurology, New York, NY, 240 E 38th St, New York, NY 10016, USA.

*Corresponding author: Željka Nikolašević, University of Novi Sad, Faculty of Medicine, Hajduk Veljkova 3, 21 000 Novi Sad, Serbia. E-mail: zeljka.nikolasevic@mf.uns.ac.rs, Tel: +381 21 210 2181

Data availability statement:

All data with instructions for their use are available at OSF and can be accessed at OSF link: https://osf.io/de5u2/

Funding

This research was supported by the Science Fund of the Republic of Serbia (#7744418 GENIUS).

LATENT, GENETIC, AND MOLECULAR GENETIC STRUCTURE OF THE WISCONSIN CARD SORTING TEST

Abstract

Objective: The main goal of this study was to explore the latent structure and genetic basis of cognitive processes involved in the Wisconsin Card Sorting Task (WCST) within phenotypic, behavioral genetic, and molecular genetic research paradigms.

Method: The sample used in phenotypic and behavioral genetic analyses comprised 468 twins (154 monozygotic and 80 dizygotic twin pairs), while molecular genetic analyses were performed on 404 twins from the same sample. The zygosity of most twin pairs (96.8%) was determined via DNA analysis of buccal swabs. Trained researchers administered the Wisconsin Card Sorting Test – WCST (Heaton et al., 1993) to the entire sample.

Results: A phenotypic factor analysis of WCST variables suggested a single-factor solution. Overall heritability ranged from 0.19 to 0.23 across different measures of the WCST. The presence of a single general genetic factor, which could be identified from different measures of the WCST, indicated the unity of various WCST indicators and the existence of a common basic ability. Performance on the WCST did not reveal significant differences between the three genotypes on COMT and DRD2. Carriers of the BDNF Met+ genotype exhibited better performance in cognitive functions in comparison to the BDNF Met- genotype.

Conclusions: This study highlighted similarities in the phenotypic and genetic structures of the WCST, suggesting one general factor underlying different cognitive functions. The BDNF Met + genotype showed significant main effects on different WCST measures.

Keywords: behavioral genetics, WCST, COMT, DRD2, BDNF

Key Points

Question: What is the key question this paper addresses?

The main goal of this study was to explore the latent structure and genetic basis of cognitive processes involved in the Wisconsin Card Sorting Task (WCST) within phenotypic, behavioural genetic, and molecular genetic research paradigms.

Findings: What are the primary findings?

In addition to the existence of a single general genetic WCST factor, as suggested by the phenotypic factor analysis and supported by behavioral genetic analyses, the main results of our study include the low heritability of WCST measures, zero evidence for the effects of COMT and DRD2 polymorphisms on WCST measures, and significant main effects of the BDNF Met+ genotype on WCST measures.

Importance: What are the key scientific and practical implications of the findings?

The results point to the existence of a general genetic factor, supporting the thesis on the unity of executive functions measured by the WCST and the existence of a common ability that underlies them, while indicating that the standard WCST performance scores might be insufficient to concurrently assess the distinct cognitive/executive processes required for performing the WCST in healthy adults.

Next Steps: What directions should be explored in future research?

It is paramount for future research directions to include examinations of the etiology of individual differences in executive functions in clinical population samples as well as explorations of the developmental dynamics of genetic and environmental influences on the etiology of individual differences in executive abilities.

1 Introduction

"Executive functions" (EFs) is an umbrella term that encompasses a range of interrelated processes involved in purposeful, goal-oriented behavior (Anderson, 2002; Anderson, et al., 2008; Huges & Graham, 2002). EFs are usually defined as a highly complex, integrated set of cognitive abilities and processes critical to adaptive functioning, including planning, time organization of behavior, goal anticipation, activity monitoring, ability to inhibit or delay responses, evaluation of responses, cognitive flexibility, and selection of problem-solving strategies (Anderson, 2002; Damasio & Anderson, 1993; Zelazo et al.,1997).

However, there are still theoretical disagreements regarding the contents of executive control (Anderson, 2002; Baddeley, 2007; Damasio & Anderson, 1993; Zelazo et al.,1997; Miyake et al., 2000) arising from the use of different methods and strategies as well as the lack of a unique empirical framework for its study. Understanding the core processes tapped by EF tests is further complicated by the multifaceted nature of the available tasks. Additionally, since executive functions are meta-processes, almost all cognitive performances include executive functions to a certain extent, which contributes to task impurity. Although task impurity complicates the study of EF, latent variable analysis has demonstrated a useful approach to addressing this problem, suggesting that EF tests can be modelled very well in careful convergent and discriminant validity research. Following this approach to studying the organization and function of EFs, Miyake and Friedman (2000) established a model of EF supported by a plethora of empirical evidence (Miyake et al., 2000; Diamond, 2013), indicating that EF is a multi-faceted construct consisting of processes related to response inhibition, working memory updating, and cognitive flexibility (or shifting). More precisely, their

theoretical three-factor model of EF (Friedman & Miyake, 2017; Miyake & Friedman, 2012) proposes that individual differences in these EFs can be understood in terms of three different EF components: common EF, shifting-specific, and updating-specific. These three EF abilities are separable (diversity), but are also moderately correlated with one another and thus share a considerable common variance (unity). The unity/diversity pattern has been replicated in several latent-variable studies (Engelhardt et al., 2015; Gustavson et al., 2018, Friedman et al., 2008).

One of the established and most widely used tests for measuring EFs is the Wisconsin Card Sorting Test (WCST), considered to provide quantitative measures of abstraction ability and thinking flexibility. The WCST engages various executive processes, such as strategic planning, organized search, ability to use environmental feedback to shift mental sets, goal-directed behavior, and modulation of impulsive response (Heaton et al.,1993; Demakis, 2003; Eling et. al., 2008).

1.1 Latent structure of the WCST

Although the WCST has been used as a measure of EFs in a large number of clinical and non-clinical studies, only few of them have provided arguments regarding its latent structure and taxonomy of the underlying cognitive processes (e.g., Bowden et al., 1998; Grave et al., 1998; 1999; 2002; 2005; Polgár et al., 2010). However, inconsistencies in the samples, test procedures and factor analysis strategies contributed to the lack of consensus regarding the latent structure of the WCST. For example, a review of 15 factor-analytic studies showed that samples from clinical populations differing in symptomatology (neurological patients, psychiatric patients) were often combined with control groups from the non-clinical population (Polgár et al., 2010), whereas only one of them has been conducted on healthy adolescents (Somsen et al., 2000).

Furthermore, although exploratory factor analysis (EFA) has been predominantly used, confirmatory factor analysis (CFA) has been performed in some studies (Greve et al., 2005). Significant discrepancies in the selection of WCST measures used in the studies also represent a disruptive factor for the replicability of previous findings. Generally, EFA have indicated solutions ranging from 1 to 3 factors, with the first factor explaining 48 - 71% of the variance (Polgár et al., 2010). The WCST measures that most saturate the first factor are: The Number of Categories, Conceptual Level Responses, Perseverative Errors, Perseverative Response, and The Total Number of Correct Responses, while Non-Perseverative Errors and Failure to Maintain Set most saturate factors 2 and 3 (Polgár et al., 2010). The results of the study by Greve et al. (2005), in which CFA was conducted on a large sample comprising 1221 subjects (neurological patients, psychiatric patients, and a non-clinical control sample), support the three-factor solution obtained in many studies with EFA. However, the results of this study also suggest the instability of the latent structure, since only the first factor (general executive functioning) was statistically significant, while the remaining two were less stable, except in subjects who completed all 128 cards. The authors pointed to the fact that early termination of the WCST tasks may lead to loss of information and significantly contributes to the unstable factor structure in research. Additional reason for the instability of the factor structure in previous studies may be relatively small samples, most often ranging from 100 to 200 subjects, which led to a violation of the recommended ratio between the number of variables and sample size in EFA (Greve et al., 2005).

Due to the reduced possibility of replication and unreliability of previous results, limited conclusions can be drawn from these methods, especially since both PCA and EFA research techniques have been used to present the observed data, without formal a priori hypothesis

testing. Namely, analytic decisions regarding EFA (e.g., the number of factors to extract and rotation) can produce misleading findings and result in a faulty foundation for theory or model building. Furthermore, exploratory model fitting is prone to over-fitting due to sample-specific variance (Byrne, 1989). Using orthogonal rotation has generally been the prior rule in factor analytic studies of the WCST, providing possible artifactual results that indicate the independence of its factors. Orthogonal solutions should never be sought unless the less-prejudicial oblique solution shows factor correlations to be zero (e.g., Floyd and Widamin, 1995; Henson and Roberts, 2006). Further complicating the understanding of WCST structure is the disappointingly low reliability of WCST variables in the nonclinical population (Strauss et al. 2006). Inter alia, lower test-retest correlations for executive functioning and memory measures can be explained by non-measurement factors. Namely, Delis, Kramer, Kaplan, and Holdnack (2004) suggest that the complex nature of executive functioning tasks, which involve multiple cognitive processes, may make them more susceptible to performance variability. This was confirmed in a meta-analytic reliability study of neuropsychological measures (Calamia et al., 2013). In this context, specifying the organization of executive functions may necessitate the use of analytical techniques such as latent variable modelling (e.g., confirmatory factor analysis), which can separate measurement error from the measure of individual differences. Latent variable modeling incorporates all of these measurement-related considerations into an analytic strategy, retaining the ability to test relations between important elements in the WCST theoretical model.

1.2 Genetic and Environmental Factors of the WCST

Behavioral genetic studies of the WCST have yielded inconsistent results, since some studies have shown a hereditary basis (Anokhin et al., 2010, Anokhin et al., 2003, Godinez et al., 2012), while others have failed to identify a genetic contribution to WCST indicators (Chou et al., 2009, Kremen et al., 2007). Previous twin studies that have reported a small to moderate contribution to the genetics of variance of individual WCST indicators have mostly been conducted on adolescent samples (Anokhin et al., 2003; Anokhin et al., 2010; Godinez et al., 2012). In a sample of 58 MZ and 25 DZ twin pairs, aged 17–28 years, for four WCST indices, which included the total number of errors, perseverative errors, perseverative responses, and the trials to complete the first category, 37% to 46% of the variance (respectively) was attributed to genetic influence (Anokhin et al., 2003). The rest of the variance of all indicators was explained by non-shared environmental factors. A longitudinal study conducted on a sample of 166 MZ and 201 DZ twin pairs at the ages of 12 and 14 revealed the existence of a gender effect for genetic and environmental influences on WCST performance, with higher heritability of test indicators for female adolescents (Anokhin et al., 2010). This result was particularly significant for the number of perseverative errors, whose heritability tended to increase from 19% at the age of 12 to 49% at the age of 14 in the female sample. Likewise, a study on a sample of 191 MZ and 165 DZ twin pairs with a mean age of 17 years showed a low to moderate heritability for traditionally defined measures such as non-perseverative errors and the number of trials, but also for narrowly defined errors, such as search errors, breaking set errors, and efficient errors (0.10 - 0.42 respectively) (Godinez et al., 2012).

However, in the population of healthy adults, behavioral genetic studies have often shown a complete absence of genetic contributions to achievement on all WCST indicators (Campana et al., 1996; Chou et al., 2010; Kremen et al., 2007; Nicole & Del Miglio 1997;

Taylor, 2007). For example, in a sample of 170 MZ and 190 DZ middle-aged male twin pairs, correlations between twins were low on all standard WCST measures and there were no statistically significant differences between MZ and DZ twins (Kremen et al., 2007). To date, only one study (Godinez et al., 2012) has investigated the etiology of covariance between different WCST indicators and found that covariance was best explained by general / common genetic factors, while differences between indicators were caused by specific genetic and environmental factors.

In general, it is possible that small samples along with the use of different measures and versions of the WCST (computer and classical) contributed to these inconsistencies. It is also possible that the aforementioned characteristics of the WCST, such as its great complexity and multi-determinism, contributed to the not always clear separation of genetic influences.

1.3 The Molecular Genetic Basis of the WCST

Although EFs are polygenically determined, there is still no convincing evidence for all gene candidates that could participate in the regulation of this complex phenotype. We selected two genes from the dopamine system that are associated with cognitive and emotional processing – the catechol-O-methyltransferase (COMT) gene and the dopamine receptor D2 (DRD2) gene – and the brain neurotrophic factor (BDNF) gene to examine possible associations with EFs.

The COMT gene is located on chromosome 22q11. It predominantly shows gene expression in the prefrontal cortex and affects various cognitive functions (Geller et al., 2017). Functional polymorphism of the COMT gene involves the substitution of one amino acid for another (Valine to Methionine) at codon 158 of the COMT gene, which consequently affects the thermal stability and reduces the activity of COMT enzymes in Met carriers (Klaus et al., 2017;

Tunbridge et al., 2019). Previous studies on the association between COMT and executive functions have shown that Met allele carriers tend to perform better in the categories completed score and working memory tests (Bruder et al., 2005), while Val/Met heterozygotes have shown superior cognitive set shifting performance (Khanthiyong et al., 2019). Furthermore, the Val allele has been linked to a higher number of perseverative errors on the WCST (Caldú et al., 2007; Malhotra et al., 2002; Rosa et al., 2004) and a higher number of commission errors on the Continuous Performance Test (Caldú et al., 2007). However, although previous studies have indicated an association between better WCST performance and Met alleles (Barnett et al., 2007; Caldú et al., 2007; Malhotra et al., 2002; Rosa et al., 2004), this result has not been consistently replicated (Geller et al., 2017).

The DRD2 gene is located on chromosome 11q23 and it encodes the DRD2 protein, which regulates the functioning of the D2 receptor for dopamine, which may impact motor output in cognitive tasks (Logue & Gould, 2014; Tunbridge et al., 2019). D2 dopamine receptor function has been linked to the regulation of emotional functioning (Barnes et al., 2011), set shifting (Logue & Gould, 2014), and different aspects of cognition (Klaus et al., 2017). The functional polymorphism of the DRD2 gene implies two types of alleles: A1 and A2 (Barnes et al., 2011). In studies conducted thus far, A1 homozygotes have had difficulties in learning from punishment, error processing, and suppressing irrelevant information (Barnes et al., 2011), accompanied by poorer performance on executive function and working memory tasks (Xu et al., 2007), especially special working memory and planning tasks (Klaus et al., 2017) and attention switching tasks (Gurvich & Rossel, 2015).

The Brain Neurotrophic Factor (BDNF) gene is located on chromosome 11p14.1. It encodes a small protein that plays a key role in synaptic plasticity, regulating the number of synapses, axon growth, and hippocampal neurogenesis (Kautsky et al., 2019). The dominant expression of the BDNF gene is in the prefrontal cortex and hippocampus (e.g., Wishart et al., 2011). The functional polymorphism of the BDNF gene results from the substitution of one amino acid (Val) for another (Met) at codon 66 of the BDNF gene, which alters BDNF gene expression and leads to smaller hippocampal volume (Kautsky et al., 2019). Previous studies on associations between the BDNF gene polymorphism and cognitive functions have provided inconsistent findings (Sanwald et al., 2020). Despite the predominance of findings showing an association between the BDNF Met allele and poorer working memory performance (Gatt et al., 2009) and executive functions (Benzerouk et al., 2013; Wishart et al., 2011), a growing body of evidence suggests that the BDNF Met allele is associated better executive performance (Alfimova et al., 2012; Harris et al., 2006; Hashimoto et al., 2016).

1.4 Current Study

Since previous studies have indicated many controversies regarding the latent structure of the WCST (e.g., Grave et al., 2005; Polgár et al., 2010), the main goal of our study was to explore possible structural aspects of EFs within the WCST by employing three different methodological and conceptual approaches. Specifically, our research focused on examining the latent structure of the WCST and the hereditary and molecular genetic bases of indicators of executive functions. Our assumption was that these different research paradigms would contribute to the elucidation of an important issue regarding the measures of executive functions covered by the WCST. Therefore, our first research question pertained to examining the latent

structure of different WCST indicators using factor analysis, which could contribute to the ongoing debate on whether different measures of the WCST constitute a unique aspect of general executive ability or can be considered as independent cognitive abilities. The second research question pertained to the examination of the relative importance of genetic and environmental influences on individual differences in WCST indices. Moreover, multivariate genetic analyses were expected to determine the extent to which the genes and the environment contribute to the overlap between the indices of the WSCT. In general, this goal could contribute to resolving the dilemma of whether the phenotypic structure reflects the basic genetic architecture. The third goal of the present study was to analyze the possible associations between COMT, DRD2, and BDNF genes and WCST performance. Assessing the molecular genetic associations of these phenotypes could enrich our understanding of how DNA variation gives rise to individual differences in executive abilities. Our design was reinforced by the fact that three different research paradigms applied on the same sample of adult twins of both genders could provide new insights into individual differences in performance on the WCST and their biological underpinnings.

2 Method

2.1. Openness and Transparency

We follow the guidelines given by JARS (Kazak, 2018) and report on how we determined the sample size, the treatment of missing data, and all measures.

All data with instructions for their use are available at OSF and can be accessed at: https://osf.io/de5u2/ The code behind the quantitative behavioral genetic analysis has been made publicly available at the Zenodo and can be accessed at

doi:https://doi.org/10.5281/zenodo.3514218 (Čolović, 2019). This study's design and its analysis were not pre-registered.

2.2 Sample and Procedure

The sample was drawn from the Serbian Twin Registry (for detailed recruitment procedures see Authors, 2019), a national and multidisciplinary twin database. The Serbian Twin Registry sample consists of 1,654 participants, including twin pairs and their family members. For the present study, we relied on data from 308 MZ (154 pairs; 75.3% female) and 160 DZ (80 pairs; 61.3% female) twins reared together. Among dizygotic twin pairs, there were 35 mixed-gender pairs and 45 same-gender pairs. The respondents were 17 to 58 years old and the average age was 24.08 years (SD = 7.02). Detailed information on the sociodemographic characteristics of the sample is presented in Table 1.

< INSERT TABLE 1 ABOUT HERE >>

After excluding some cases due to failed genotyping, the sample consisted of 404 twins for COMT gene (rs4680), DRD2 gene (rs1800497), and BDNF (rs6265) gene polymorphisms in the subsequent molecular genetic analysis.

The zygosity of most twin pairs (96.8%) was determined via DNA analysis of buccal swabs. Buccal swabs were tested using short tandem repeat (STR) megaplex kits – either Investigator 24plex GO! (Qiagen, Valencia, CA, USA) or GlobalFiler (Applied Biosystems, Thermofisher Scientific, Waltham, MA, USA). Both kits detect 21 autosomal STRs. Samples with partial profiles were interpreted if at least 10 loci had results. For a smaller number of twins (3.2%), zygosity was determined using the Twins Physical Resemblance Questionnaire

(Oniszczenko et al.,1993). This questionnaire is a self-report measure containing 31 sets of items (plus 19 demographic questions) referring to the similarities and differences in biological and physical indicators (e.g., eye color, body weight, and body height) between twins within a pair. Zygosity measures of this questionnaire proved to be reliable indicators of zygosity in many studies (e.g., Čolović et al., 2018; Reed et al., 2005).

The research was approved by the Institutional Ethics Committee of the Faculty of Philosophy (#02-374/15), the Committee for Ethics of Clinical Trials at the Faculty of Medicine (#01-39/229/1) at the University of Novi Sad and the data was collected between 2011 and 2018. Twins were recruited from the whole territory of Serbia. A call for participation in the research was published through the media and press. Data collection was carried out in different sites in Novi Sad, Niš, Novi Pazar, Zrenjanin, and Belgrade. The participation of the twins was voluntary and each respondent signed an informed consent for participation. Executive functions were examined by trained researchers and the entire procedure for collecting data is described elsewhere (for detailed recruitment procedures see Smederevac et al., 2019).

2.3 Measures

The Wisconsin Card Sorting Test – WCST (Heaton et al., 1993). The WCST is the most prominent test for the assessment of set shifting, attention, and inhibition. The test assesses the ability to create and change the principles of categorization, using the task of classifying a series of cards according to one of the three classification criteria (color, form, and the number of elements) related to four target cards. The WCST requires the participant to sort a set of cards according to implicit rules and based on the limited corrective feedback provided by the examiner. The participant's responses can be analyzed to produce separate indices of sources of

difficulty on the test. The value of the WCST, therefore, lies in its sensitivity for detecting and characterizing different measures of executive functions: (1) Total Number of Errors: the total number of perseverative and non-perseverative errors during card sorting; (2) Non-Perseverative Errors: the number of errors that are not repetitive during card sorting (3) Perseverative Errors: the number of errors that are repetitive during card sorting, which reflects a tendency towards perseveration; (4) Perseverative Responses: the number of responses that are repetitive, whether or not they are accurate; (5) Categories Completed: the number of series of 10 consecutive correct answers (ranging from 0-6), which reflects the overall success; (6) Conceptual Level Responses: three or more correct answers in a row (individual answers are counted and categories are included), which reflects the insight into the correct pairing principle; (7) Trials to Complete the First Category: the number of attempts to complete the first category (it is counted as 128 if no category is achieved), which reflects the initial conceptual abilities; (8) Failure to Maintain Set: five and more than five – fewer than nine correct answers in a row, which reflects the sorting efficiency. Also included in the analysis was a variable named "Categorizing Efficiency", which in addition to the number of completed categories concerns the number of attempts to single out these categories. That way, participants were awarded additional points for needing fewer attempts to separate the categories (see Cianchetti et al., 2005).

The Genotyping of COMT, DRD, and BDNF Polymorphisms. The genotyping of the COMT gene (rs4680), DRD2 gene (rs1800497) and the BDNF gene (rs6265) was carried out using TaqMan assays (TaqMan SNP, Applied Biosystems®, Warrington, UK), as recommended by the manufacturer. The TaqMan SNP Genotyping Assays uses TaqMan 5′ -nuclease chemistry for amplifying and detecting specific polymorphisms in purified genomic DNA samples and takes advantage of minor groove-binding probes for superior allelic discrimination. The SNP

Genotyping Assays contain a VIC-dye-labelled probe, a FAM-dye-labelled probe with two target-specific primers. PCR was performed using 10 ng of genomic DNA together with 1µl of TaqMan Genotyping assay and 12.5 µl of the genotyping master mix in the final 25µl reaction on a 96-well plate using an ABI Prism 7500 Fast PCR device (Applied Biosystems®, Foster City, California, USA). COMT gene (rs4680), DRD2 gene (rs1800497), and BDNF gene (rs6265) alleles with the specific fluorescence curves were detected and analyzed using the 7500 System SDS program, integrated into the ABI Prism 7500 Fast PCR device.

The COMT gene polymorphism was defined by 3 groups: 124 high-activity homozygotes (Met/Met carriers), 213 intermediate heterozygotes (Met/Val carriers), and 70 low-activity homozygotes (Val/Val carriers). The COMT gene polymorphism was in the Hardy–Weinberg equilibrium (HWE), with no significant differences between the observed and calculated genotype frequencies ($\chi^2 = 4.13$, df = 2, $p > 0.05$).

The DRD2 gene polymorphism was defined by 2 groups, according to the presence of the risk allele: A1 homozygotes and A1/A2 heterozygotes were combined into a single group (A1+, 127 carriers), and the other group was A1– (280 carriers). The DRD2 gene polymorphism was in the HWE ($\chi^2 = 1.42$, df = 1, $p > 0.05$).

All analyses of the BDNF gene polymorphism were also performed at the allele level, according to the presence of the risk Met allele. The BDNF gene was defined by 2 groups: 285 high-activity homozygotes (Met-) formed one group and Val66Met heterozygotes formed the second group together with Met66Met homozygotes (named Met carriers). This group consisted of 122 Met carriers. BDNF was in the HWE and there were no significant differences between the observed and calculated genotype frequencies ($\chi2= 0.20$, df = 1, $p > 0.05$).

2.4 Data Analysis

2.4.1 Preliminary Analysis, Descriptive Analysis, and Twin Intraclass Correlation Analysis

The preliminary analysis involved a partialization of the gender effect, as well as a linear and quadratic partialization of the age effect. The partialization of these effects was conducted by using the standard regression procedures proposed by McGue and Bouchard (1984). Descriptive statistical parameters as well as phenotype and intraclass correlations were calculated in the SPSS v.21 software (IBM Corp, 2012).

2.4.2 Factor Analysis

The latent structure of the WCST indicators was examined via a confirmatory factor analysis (CFA). To provide validation for the obtained latent solutions, the twin sample was split into two cross-validation samples. The Twin 1 subsample comprised all the firstborn twins from the pairs, while the Twin 2 subsample included all the second-born twins from the pairs. The results of the CFA on Twin 1 subsample were then cross-validated in the Twin 2 subsample via a CFA run in the "lavaan" R package (Rosseel, 2012). A robust estimation was used for each model, due to the violation of multivariate normality based on a Mardia's coefficient greater than 3 (Mardia, 1970). Specifically, we performed Satorra–Bentler's scaled Chi-square testing with robust maximum likelihood estimators (MLM). The model fit of the proposed model was assessed based on the following recommendations: (i) the Tucker-Lewis index (TLI) and the comparative fit index (CFI)>0.9; (ii) the standardized root mean square residual (SRMR) and the root mean square error of approximation (RMSEA)<0.08; and (iii) a non-significant chi-square (Hu and Bentler 1999).

2.4.3 Genetic Analysis

Phenotypic similarities between MZ and DZ twins were examined in each WCST dimension by using structural equation modeling (SEM) – employing univariate and multivariate biometric methods. Univariate and multivariate structural equation modeling (SEM) was conducted in the "*lavaan*" R package (Rosseel, 2012). A multivariate genetic analysis was conducted to examine the nature of relationships between WCST factors, by specifying the extent to which they share genetic and environmental influences and the way their influences differ. To specify the form of the observed covariants among WCST factors, multivariate Independent Pathway Models and Common Pathway Models (Rijsdijk & Sham, 2002) were tested in order to estimate: additive (A) and dominance (D) genetic effects, shared (C) and non-shared (E) environmental factors, and specific (s) and common (c) genetic and environmental sources of variance. These models represent different patterns of genetic and environmental influences, which can explain the observed phenotypic correlations between different WCST factors. A comparison of the two model groups and a comparison between the full (ACE, ADE) and reduced (AE, CE) models were carried out by using several fit indicators. Analysis parameters were calculated by using the method of maximum likelihood. Model evaluation was conducted based on the Akaike Information Criterion (AIC: Akaike, 1973), the Bayesian Information Criterion (BIC; Schwarz, 1978), the comparative fit index, the Tucker–Lewis Index (CFI and TLI – optimal values higher than .95, acceptable higher than .90), the root mean square error of approximation (RMSEA - optimal values lower than .05, acceptable lower than .08), and the quotient $\chi 2/df$ (recommended < 2) (Ching–Yun, 2002; Kline, 2010). A series of independent and common pathway models were fitted into multivariate covariance matrices. In accordance with parsimony criteria, we selected the model with the smallest number of parameters and a fit

not significantly worse than the full model. Additionally, Cholesky behavior genetic analyses (e.g, Neale and Cardon, 2013) were calculated to explore genetic and environmental correlations between WCST scores.

2.4.4 Molecular Genetic Analysis

The main effects of gene polymorphisms on executive functions were performed using t-test for independent samples (DRD2 and BDNF allelic gene variants as independent variables) and one-way ANOVA (the COMT gene polymorphism as an independent variable). Cohen's *d* and partial eta squared (η_p^2) were used as measures of effect size. A commonly used interpretation is to refer to effect sizes as small (d = .20; η_p^2 = .01), medium (d = .50; η_p^2= .05), and large (d = .80; η_p^2 = .13), based on the benchmarks suggested by Cohen (1992).

3 Results

3.1 Descriptive Statistics of Phenotypic Characteristics and Twin Intraclass Correlations

Means, standard deviations, and univariate cross-twin (intraclass) correlations for each zygosity group are provided in Table 2. The values of skewness and kurtosis indicated that almost all WCST measures were non-normally distributed. Therefore, all measures were first normalized using the rank-based inverse normal (Rankit) transformation (Solomon & Sawilowsky, 2009). However, after the transformation, the magnitude of both skewness and kurtosis did not fall within the range of −1 to 1 for certain measures, indicating that some distributions did not reach normality. Correlations between MZ twins were consistently higher than correlations between DZ twins for all measures. Almost all correlations between MZ twins were positive, significant, and of low-to-moderate strength, with the exception of Failure to

Maintain Set and Trials to Complete the First Category. Correlations between DZ twins were non-significant for all WCST variables and approximately twice as low as those between MZ twins. This correlation pattern provided evidence that genetic factors were likely to significantly contribute to the covariance between WCST measures.

< INSERT TABLE 2 ABOUT HERE >>

3.2 Confirmatory Factor Analysis (CFA) of the WCST

Total errors were excluded from further analysis since these represented linear combinations of two or more of the obtained variables and provided redundant information. Moreover, the high correlation between perseverative errors and response in the dataset (over 0.99) indicated the redundancy of the use of both scores in the same analysis. Based on the results of earlier studies, we tested different factor solutions (Bowden et al., 1998; Greve et al., 1997, 2005; Polgár et al., 2010). The one-factor model used all six scores as indicators (The Number of Categories, Perseverative Errors, Non-Perseverative Errors, Conceptual Level Responses, Trials to Complete the First Category, and Failure to Maintain Set). The second model was a two-factor model. On Factor 1, the indicators were The Number of Categories, Perseverative Errors, Conceptual Level Responses, and Trials to Complete the First Category, while the second factor consisted of two indicators – Non-Perseverative Errors and Failure to Maintain Set. The CFA results revealed that both models showed indications of a lack of fit to the observed data. Table 3. shows fit indices for the comparative models. Both models yielded a significant Chi-square, suggesting that neither described the data fully. Both models showed an unacceptable fit on the CFI and the TLI (> .95) and an unacceptable fit according to the RMSEA (≤.10). Since the correlation between the factors in the second model was 0.772 and individual factor loadings exceeded the value of 1, the two- factor model was excluded from further testing

with the inclusion of the modification index. Further, we explored the potential for obtaining an improved fit by including correlated residuals between like scores. Correlated residuals may be an important element of any factor analysis that includes 'method' variance, as in cases in which multiple scores from one test are included (Millsap, 2011). Thus, we used the modification indices provided by the software to guide model improvement. The one factor-model was improved by allowing for the following modification indices: 1) correlated residuals between Categories Completed and Conceptual Level Responses, 2) correlated residuals between Conceptual Level Responses and Failure to Maintain Set, and 3) correlated residuals between Perseverative Errors and Non-Perseverative Errors. After these refinements, the model's fit was improved, as evidenced by the following goodness-of-fit statistics: $\chi2(6) = 12.596$, $p = 0.050$, RMSEA= 0.069, SRMR = 0.030, CFI = 0.993, and TLI = 0.983. All loadings were significant (Table 3). To validate the CFA results, a subsequent CFA was tested in the Twin 2 subsample (see Table 4). The single-factor solution demonstrated a poor fit to the data; $\chi2(15) = 102.809$, p= 0.00, RMSEA = 0.211, SRMR = 0.067, CFI =0.885, and TLI = 0.808. We then tested the one-factor model with the modifications suggested on Twin 1. Applied to Twin 2, this solution provided a better fit to the data, but still was not within acceptable limits ($\chi2(6) = 30.178$, p= 0.000, RMSEA = 0.131, SRMR = 0.052, CFI =0.973, TLI = 0.932). All loadings were significant (Table 3).

As an earlier study concluded, the inadequacy of WCST models may be a result of a number of factors including significant overlap among variables; error due to differences in test length; the relatively rare finding of some WCST scores in the normal normative study; non- linear relationship with another indicator (Grave, 2005; Polgár at al., 2010). We decided to test the fourth CFA model, which excluded some parameters such as Categories Completed. The

rationalization for this exclusion was that even after normalization, adequate distribution was not achieved (skewness +1.5). The results of other studies have also shown that the distribution in the normal population is strongly skewed, creating potential problems in statistical evaluation (Obonsavin et al., 1999). Instead of this measure, we included the measures of Categorizing Efficiency, which consider the number of cards used by the subject to complete a maximum of six categories. We also excluded Failure to Maintain Set, since, this variable shows a complex, non-linear association with other WCST scores (Polgár at al., 2010). In the first analysis, Trials to Complete the First Category had very slight loadings (e.g., pattern coefficients greater than or equal to 0.30 were considered salient). We tested the CFA, which included Categorizing Efficiency, Perseverative Error, Non-Perseverative Errors, and Conceptual Level Responses. This solution provided an adequate fit to the data, ($\chi 2(2) = 3.961$, p= 0.139, RMSEA = 0.065, SRMR = 0.006, CFI =0.998, TLI = 0.994). All loadings were significant (Table 3). Additionally, parameter estimates are shown in Fig. 1. The CFA model on Twin 2 exhibited excellent fit – $\chi 2(2) = 0.520$, p = 0.771, RMSEA= 0.000, and SRMR = 0.003, CFI = 1.000, and TLI = 1.000) (Table 4). The model estimated by the CFA analysis on Twin 1 was thus confirmed by the CFA model on Twin 2, in an independent sample.

The results of the CFA, as the factor score, were used as one of the measures in molecular genetic analyses.

<< INSERT TABLE 3 ABOUT HERE >>
<< INSERT TABLE 4 ABOUT HERE >>

<< INSERT FIGURE 1 ABOUT HERE >>

3.3 Genetic analyses

Before conducting a multivariate genetic analysis, we examined univariate models for the used measures (Table A in Supplement). The results of the univariate genetic analysis indicated that the contributions of additive genetic and non-shared environmental effects were substantial for six WCST scores (Number of Categories, Perseverative Errors, Perseverative Response, Non-Perseverative Errors, Conceptual Level Responses, and Categorizing Efficiency). Estimates for non-shared environmental influences ranged between .76 and .81, while the genetic influence was far less dominant, ranging between 18.99 and 24.01 (figure 1). Two WCST measures, Trials to Complete the First Category and Failure to Maintain Set, did not show significant twin correlations, which could be partially due to the distribution properties of these measures in the present sample (a notably skewed distribution; Failure to Maintain Set showed little variability). Therefore, these variables were not used in further genetic analyses.

<< INSERT FIGURE 2 ABOUT HERE >>

Based on the results of the CFA (referring to the arguments provided in the description of the factor analyses), the same WCST scores were used as variables in multivariate genetic analyses. The results of multivariate genetic modeling (Table 5) showed that the most appropriate fit indices were for the AE common pathways model- $\chi^2/df = 1.29$, CFI = .99, TLI = .99, RMSEA = .05, AIC = 16759, BIC = 16863.1. All indices were within acceptable boundaries. The estimation of the parameters of the best-fitting models is given in Table 6.

<< INSERT TABLE 5 ABOUT HERE >>
<< INSERT TABLE 6 ABOUT HERE >>

The results from the common AE pathway model suggested that genetic and environmental factors had similar effect paths for all WCST phenotypes (Table 6). The overall variance of heredity was better explained by common genetic factors in the case of all four variables (20% - 23%). The contribution of the specific genetic factor was only detectable in perseverative errors (3%). In all cases, common environmental effects (56% - 72%) were stronger than specific (5% - 23%) environmental effects. The common environmental effect was the highest for Categorizing Efficiency (72%), while the highest specific environmental effect was observed in the case of Non-Perseverative Errors (23%).

Cholesky behavior genetic analyses were performed to explore the genetic and environmental correlations of WCST scores. Phenotypic correlations between all WCST measures were moderate-to-high and the same was true for genetic and environmental correlations (Supplement table B).

3. 4 Molecular genetic analysis

The mean scores of WCST dimensions and the factor score in the specific allelic configurations of COMT, DRD2, and BDNF are given in Table 7.

<< INSERT TABLE 7 ABOUT HERE >>

<< INSERT TABLE 8 ABOUT HERE >>

In our sample (Table 8), DRD2 and COMT were not significantly related to any specific measures on the WCST or the factor score. However, BDNF-Met66 carriers showed significantly higher scores on Non-Perseverative Errors ($t = 2.095$; $p= 0.037$; $d2= 0.23$), Failure

to Maintain Set (t = 2.233; p= 0.026; d2= 0.25), and Trials to Complete the First Category (t = 2.376; p= 0.018; d2= 0.26) and lower scores on Categorizing Efficiency (t =- 2.158; p= 0.032; d2= 0.23) and the WCST factor score (t =- 2.000; p= 0.047; d2= 0.21) than carriers with other genotype combinations. After applying the false discovery rate (FDR) correction for multiple testing, only Trials to Complete the First Category retained statistical significance.

4 Discussion

Although the WCST is widely used in both clinical and research settings, the factors that contribute to the success of its tasks are still controversial. In the present study, we adopted a comprehensive approach, examining phenotypic, behavioral genetic, and molecular genetic bases of the WCST on the same adult twin sample, with the aim of improving the understanding of the structure of these complex phenotypic features.

In order to examine the latent structure of the WCST, we used latent variable modeling. This allowed us to incorporate all measurement-related considerations into a single analytic strategy while retaining the ability to test relations between important elements in the theoretical model of the WCST. CFA procedures were applied to six measures of the WCST to determine their latent factor structure on one of the twins in each pair (e.g., the Twin 1 subgroup). One- and two-factor solutions as well as the one-factor model with modification and the reduced one-factor model were examined as possible structures. The results of the CFA were then cross-validated in the second subgroup (e.g., Twin 2). Both basic models, i.e, one- and two-factor models, showed indications of a lack of fit to the observed data on both subsamples. The one-factor model with modifications (correlated residuals) provided a better fit to the data on subsample one but was not cross-validated on the second subsample. Relying on conceptual and

584 statistical criteria, we decided to test the dimensionality of a small number of selected measures,
585 including Categorizing Efficiency, Perseverative Responses, Non-Perseverative Errors, and
586 Conceptual Level Responses. The notably high latent factor loadings for all four WCST
587 variables suggested that the single-factor model provided a good representation of these scores.
588 This one-factor model was cross-validated on the Twin 2 subsample, showing an adequate fit.
589 This factor could represent cognitive flexibility as reflected in the ability to identify the sorting
590 rule, efficient problem-solving as well as the ability to switch from an incorrect sorting concept
591 in response to external feedback. Thus, the results of factor analyses support a unidimensional
592 conceptualization of WCST performance and imply that different aspects of the executive
593 function measured by the WCST cannot be clearly distinguished. More specifically, regardless of
594 the number of indicators included in the factor analysis, the one-factor solution proved to be the
595 most plausible. This result is in line with some previous findings (Bowden et al., 1998, Boone et
596 al., 1998; Pineda & Merchan, 2003) that have shown the unidimensionality of WCST measures.

597 The patterns of inheritance obtained in univariate genetic analysis indicate low genetic
598 conditionality of most of the WCST indicators, which at least partially corresponds to the results
599 of related previous studies (Anokhin et al., 2003; Anokhin et al., 2010; Godinez et al., 2012).
600 Inconsistent results regarding genetic foundations of WCST imply the importance of unifying the
601 basic premises of using the test before explicit conclusions. For example, it is difficult to
602 compare results on samples of different ages, as well as results obtained with different test
603 procedures, such as computer-based testing vs. trained researchers (Anokin et al., 2010; Kremen
604 et al., 2007). Our sample mostly included younger adults and the results support the view that
605 during the life course, most environmental influences on cognitive achievement are non-shared
606 (Finkel & Pedersen, 2001; McCartney et al.,1990). These specific (non-shared) environmental

607 factors might reflect the fact that individuals experience considerable life changes during the
608 transition from adolescence to adulthood, including changes in residence, education, social and
609 employment roles, and behaviors such as substance use. Furthermore, some developmental
610 studies (De Luca et al., 2003) have suggested that peak performance occurs sometime between
611 20 and 29 years of age, when the final differentiation of frontal areas takes place and working
612 memory, strategic planning, goal-setting, and problem-solving reach full maturity. This relatively
613 prolonged development of the prefrontal cortex is often discussed in the context of neural
614 changes, since performance improvements coincide with synaptic pruning and myelination
615 throughout the brain (Fuster, 2002; Lebel & Beaulieu, 2011).

616 Despite the complexity of the test, the expectation that WCST represents a
617 multidimensional measure of executive functions does not have strong empirical support, which
618 is further confirmed by the results of multivariate behavioral genetic analyzes. Namely, the
619 results of multivariate genetic modeling show that the AE common pathway model best explains
620 the covariance of four WCST indicators – Categorizing efficiency, Perseverative responses,
621 Non-Perseverative errors and Conceptual Level Responses. More specifically, high phenotypic
622 correlations among the examined measures can be explained by the common/general non-shared
623 environmental factor, while a smaller part of their covariance is a consequence of a common
624 genetic factor (about 20%). Moreover, specific genetic influences did not make a significant
625 contribution. High genetic correlations (range 0.83 – 0.99) additionally indicate that these
626 different measures shared the same genetic influences. This result largely corresponds to the
627 finding of the previous study (Godinez et al., 2012), demonstrating that the etiology of the
628 covariance between different types of errors in WCST can be explained by one general genetic
629 factor and general non-shared environment.

Taken together, the results do not seem to support the classical, clinically justified division into Perseverative and Non-Perseverative Errors, at least when it comes to the non-clinical population, i.e., the population of normal healthy individuals like the ones who participated in this study. In this regard, it seems interesting to examine whether individual differences in achievement in this test (expressed through different WCST scores) explain genetic and environmental influences in the same relationship when it comes to the clinical population (e.g., those with neurological or psychiatric heredity).

Finally, the purpose of the molecular genetics analysis was to examine the association between COMT, DRD2, BDNF, and WCST performance. The association between dopamine genes and executive functions has a solid theoretical and empirical foundation (Bruder et al., 2005; Caldú et al., 2007; Khanthiyong et al., 2019). Therefore, the absence of a relationship between COMT and DRD2 gene polymorphisms and WCST measures in this study cannot be generalized with certainty, since this is precluded by the limitations arising from sample size. Namely, the sample size was the main reason for examining only the main effects of genes on WCST indicators, without analyzing their interactions. The basal level of dopamine has a polygenic basis, with probable significant interaction between different genes. Therefore, it is possible that DRD2 and COMT simply did not cover the extent of variability necessary for a relevant calculation of their effect. In other words, this result should not be viewed as an argument against the association between dopamine genes and executive functions, due to the absence of adequate conditions for observing their effects.

Significant main effects of BDNF genes were found on Non-Perseverative Errors, Trials to Complete the First Category, Failure to Maintain Set, Categorizing Efficiency, and the WCST

factor score. After the correction for multiple testing, only Trials to Complete the First Category still showed significant main effects of BDNF genes. The relatively small sample size may have reduced the statistical power of our analyses. However, in genetics research, overlooking small effects could have substantial consequences on the identification and understanding of the actual determinants of complex psychological phenomena (Götz at al., 2020).

These results have important implications. First, they indicate a common genetic basis of different indicators of the WCST, which further confirms its unitarity. Although Non-Perseverative Errors, Trials to Complete the First Category, Failure to Maintain Set, and Categorizing Efficiency represent different indicators of the WCST, it is possible that they all constitute a unique aspect of executive functions. Second, the results represent a direct replication of previous findings (Alfimova et al., 2012; Harris et al., 2006; Hashimoto et al., 2016), providing additional support for the accumulation of evidence on the role of BDNF in the regulation of cognitive processes. In fact, the association between the BDNF Met allele and better achievement on the WCST is not in line with the assumption that this allele contributes to poorer cognitive efficiency. This result primarily refers to aspects of cognition related to the memory domain (e.g., Egan et al., 2003; Hariri et al., 2003).

Furthermore, after correction for multiple comparisons, this research highlighted a specific relationship between the BDNF polymorphism and the total number of trials to successfully complete the first category. While this measure can be an indicator of executive dysfunction, it also can reflect conceptual ability. Met group carriers have better performance on this measure of initial conception and abstraction ability than Val/Val homozygotes, with significantly fewer attempts to complete the first category. Since the total number of trials to complete the first category reflects the ability to switch from an incorrect sorting concept, which

incorporates set shifting and response inhibition capacity, our finding is consistent with earlier studies that reported that Met carriers had higher average performance in non-hippocampal-based tasks, including the Go/NoGo response inhibition task (Beste et al., 2010) and other executive tasks, such as the Stroop task (Gajewski et al., 2012) and working memory tasks (Alfimova et al., 2012), as well as with better nonverbal reasoning skills (Harris et al., 2006). Although some research does not find any association of the BDNF Val66Met polymorphism with executive function (Mandelman & Grigorenko 2012; Toh et al., 2017), and some find better performance in association with the Val/Val polymorphism (e.g., Altmann et al., 2016; Rybakowski et al., 2003), the Met group in our study mainly included Val-Met heterozygotes, without a representative number of Met homozygotes. Therefore, our result must be considered with great caution. Situational anxiety due to feedback during task solving also may have contaminated the response. An association between anxiety and the BDNF Val allele is a common finding (Lang et al., 2005). Therefore, future studies should clarify how the influences of the BDNF Val66Met polymorphism on cognitive abilities differ depending on the specific neural system and cognitive function, as well as health status, brain development stage, and other potential confounding factors.

Our results diverged from the specificity of the traditional WCST scores posited by authors and indicated an overlap of most of the variance shared between WCST variables. They were inconsistent with the interpretation that these measures represent distinct constructs. Instead, the measures appear to reflect the same dimension of various aspects of WCST performance. The authors' definitions suggest several factors underlying the different types of errors in the WCST, which represent indices of more independent abilities. However, the results of this study seem to support the claim that different WCST scores can be explained by a single

common genetic factor, which would in turn suggest the use of the WCST as a unique measure of a certain general executive function. In other words, the existence of a general genetic factor supports the thesis on the unity of different executive functions measured by the WCST and the existence of a common ability that underlies them.

4.1 Constraints on generality

Limitations arising from our study do not allow for a simple generalization of the results, since the sample size can certainly contribute to the reduced reliability of findings. On average, the sample used in our research encompassed upper-level education participants. Such a sample structure resulted in the reduced variability of executive functions, which somewhat affected the mutual relationships of WCST variables, reducing the magnitude of the correlations of different measures. Furthermore, the study included a greater number of female pairs relative to male pairs and a higher proportion of MZ pairs relative to DZ pair. However, it should be noted that these two disproportions are quite typical in volunteer twin samples (Lykken et al., 1987). Because of this imbalance in sex, we were unable to examine the genetic influences of sex differences. The largest percentage of twins belonged to the young adult category (M = 24.5 and about 80% of the twins in the sample were up to 30 years of age). Considering these limitations, all WCST measures were corrected for sex and age effects by applying McGue and Bouchard's (1984) regression technique, i.e., entering sex and age as predictors and taking each specific WCST measure as a criterion, while retaining the residuals. Since participants in our study were all White, they are unlikely to be representative of the broader population. Moreover, although trained researchers administered the WCST, it is possible that small inconsistencies affected the quality of responses, which could not be controlled as a method factor due to the sample size.

Furthermore, in the era of GWAS, the examination of the associations of single nucleotide polymorphisms yields findings of limited representativeness. Despite all the limitations, our study provides robust arguments from three research paradigms for the conclusion that the WCST encompasses general executive processes in healthy adults.

References

Alfimova, M. V., Korovaitseva, G. I., Lezheiko, T. V., & Golimbet, V. E. (2012). Effect of BDNF Val66Met polymorphism on normal variability of executive functions. B*ulletin of experimental biology and medicine,* 152(5), 606–609. https://doi.org/10.1007/s10517-012-1587-x

Altmann, V., Schumacher-Schuh, A. F., Rieck, M., Callegari-Jacques, S. M., Rieder, C. R., & Hutz, M. H. (2016). Val66Met BDNF polymorphism is associated with Parkinson's disease cognitive impairment. *Neuroscience letters, 615*, 88-91. https://doi.org/10.1016/j.neulet.2016.01.030

Anderson, P. (2002). Assessment and development of executive function during childhood. *Child Neuropsychology*, *8*, 71-82.

Anderson, V., Jacobs, R., & Anderson, P.J. (2008). *Executive functions nad the frontal lobes: A lifespan perspective.* Taylor & Francis Group, New York, London

Anokhin, A. P., Golosheykin, S., Grant, J. D., & Heath, A. C. (2010). Developmental and genetic influences on prefrontal function in adolescents: a longitudinal twin study of WCST performance. *Neuroscience letters*, *472*(2), 119–122. https://doi.org/10.1016/j.neulet.2010.01.067

Anokhin, A. P., Heath, A. C., & Ralano, A. (2003). Genetic influences on frontal brain function: WCST performance in twins. *Neuroreport, 14*(15), 1975–1978. https://doi.org/10.1097/00001756-200310270-00019

Baddeley, A. D. (2007). Working memory, thought and action. Oxford: Oxford University Press.

Barnes, J. J., Dean, A. J., Nandam, L. S., O'Connell, R. G., & Bellgrove, M. A. (2011). The molecular genetics of executive function: role of monoamine system genes. *Biological psychiatry, 69*(12), e127–e143. https://doi.org/10.1016/j.biopsych.2010.12.040

Barnett, J. H., Jones, P. B., Robbins, T. W., & Müller, U. (2007). Effects of the catechol-O-methyltransferase Val 158 Met polymorphism on executive function: a meta-analysis of the Wisconsin Card Sort Test in schizophrenia and healthy controls. *Molecular psychiatry, 12*(5), 502 –509. https://doi.org/10.1038/sj.mp.4001973

Benzerouk, F., Gierski, F., Gorwood, P., Ramoz, N., Stefaniak, N., Hübsch, B., ... & Limosin, F. (2013). Brain-derived neurotrophic factor (BDNF) Val66Met polymorphism and its implication in executive functions in adult offspring of alcohol-dependent probands. *Alcohol, 47*(4), 271–274. https://doi.org/10.1016/j.alcohol.2013.03.001bib/3.2.119.

Beste, C., Baune, B. T., Domschke, K., Falkenstein, M., & Konrad, C. (2010). Paradoxical association of the brain-derived-neurotrophic-factor val66met genotype with response inhibition. *Neuroscience, 166*(1), 178-184. https://doi.org/10.1016/j.neuroscience.2009.12.022

Boone, K. B., Pontón, M. O., Gorsuch, R. L., González, J. J., & Miller, B. L. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology,* 13(7), 585–595. https://doi.org/10.1016/S0887-6177(97)00074-7

764 Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C.
765 M. (1998). The reliability and internal validity of the Wisconsin Card Sorting Test.
766 *Neuropsychological Rehabilitation*, 8(3), 243-254. https://doi.org/10.1080/713755573
767 Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., Long, C. M.
768 (1998). The reliability and internal validity of the Wisconsin Card Sorting Test.
769 *Neuropsychological Rehabilit*ation, *8*, 243-254.
770 Bruder, G. E., Keilp, J. G., Xu, H., Shikhman, M., Schori, E., Gorman, J. M., & Gilliam, T. C.
771 (2005). Catechol-O-methyltransferase (COMT) genotypes and working memory:
772 associations with differing cognitive operations. *Biological psychiatry*, *58*(11), 901–907.
773 https://doi.org/10.1016/j.biopsych.2005.05.010
774 Byrne, B.M. (1989). Aprimer of LISREL. New York: Springer-Verlag
775 Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological
776 measures: meta-analyses of test-retest correlations. *The Clinical neuropsychologist*,
777 *27*(7), 1077–1105. https://doi.org/10.1080/13854046.2013.809795
778 Caldú, X., Vendrell, P., Bartrés-Faz, D., Clemente, I., Bargalló, N., Jurado, M. A., Serra-
779 Grabulosa, J. M., & Junqué, C. (2007). Impact of the COMT Val108/158 Met and DAT
780 genotypes on prefrontal function in healthy subjects. *NeuroImage*, *37*(4), 1437–1444.
781 https://doi.org/10.1016/j.neuroimage.2007.06.021
782 Campana, A., Macciardi, F., Gambini, O., & Scarone, S. (1996). The Wisconsin Card Sorting
783 Test (WCST) performance in normal subjects: a twin study. *Neuropsychobiology*, *34*(1),
784 14–17. https://doi.org/10.1159/000119284
785 Ching–Yun, Y. (2002). Evaluating Cut off Criteria of Model Fit Indices for Latent Var-iable
786 Models with Binary and Continuous Outcomes. (Doctoral thesis, Univer-sity of

787 California, Los Angeles, USA). Retrieved from:
788 http://ww.statmodel2.com/download/Yudissertation.pdf.
789 Chou, L. N., Kuo, P. H., Lin, C. C., & Chen, W. J. (2010). Genetic and environmental influences
790 on the Wisconsin Card Sorting Test performance in healthy adolescents: a twin/sibling
791 study. *Behavior genetics*, *40*(1), 22–30. https://doi.org/10.1007/s10519-009-9299-3
792 Cianchetti, C., Corona, S., Foscoliano, M., Contu, D., & Sannio-Fancello, G. (2007). Modified
793 Wisconsin Card Sorting Test (MCST, MWCST): normative data in children 4-13 years
794 old, according to classical and new types of scoring. *The Clinical neuropsychologist*,
795 21(3), 456–478. https://doi.org/10.1080/13854040600629766
796 Cohen, J. (1992). Statistical Power Analysis for the Behavioral Sciences, 4nd ed., New Jersey:
797 Lawrence Erlbaum Associates, Inc.
798 Čolović, P. (2019). Petarcolovic/R_CqBG 1.0.0-Beta. Zenodo, 2019. Zenodo,
799 doi:https://doi.org/10.5281/zenodo.3514218.
800 Čolović, P., Branovački, B., & Bosić, D. Z. (2018). Validation of zygosity assessment by a self-
801 report questionnaire in a sample of adult Serbian twins. *Primenjena psihologija*, 11(4),
802 419-432.
803 Damasio, A. R., & Anderson, S.W. (1993). The frontal lobes. In K. M. Heilman & E. Valenstein
804 (Eds.), *Clinical neuropsychology* (3rd ed., pp.409-460.). New York: Oxford Univesity
805 Press.
806 De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., &
807 Pantelis, C. (2003). Normative data from the CANTAB. I: development of executive
808 function over the lifespan. Journal of clinical and experimental neuropsychology, 25(2),
809 242–254. https://doi.org/10.1076/jcen.25.2.242.13639

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: an update. Journal of the International Neuropsychological Society, 10(2), 301-303. https://doi.org/10.1017/S1355617704102191

Demakis, G. J. (2003). A meta-analytic review of sensitivity of the Wisconsin Card Sorting Test to frontal and lateralized frontal brain damage. *Neuropsychology, 17*, 255-264.

Diamond A. (2013). Executive functions. Annual review of psychology, 64, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dobson, P. (2000). An investigation into the relationship between neuroticism, extraversion and cognitive test performance in selection. International Journal of Selection and Assessment, 8(3), 99–109. https://doi.org/10.1111/1468-2389.00140

Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., Zaitsev, E., Gold, B., Goldman, D., Dean, M., Lu, B., & Weinberger, D. R. (2003). The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, 112(2), 257–269. https://doi.org/10.1016/s0092-8674(03)00035-7

Eling, P., Derckx, K., & Meas, R. (2008). On the historical and conceptual background of the Wisconsin Card Sorting Test, *Brain and Cognition*, *67*, 247-253.

Engelhardt, L. E., Briley, D. A., Mann, F. D., Harden, K. P., & Tucker-Drob, E. M. (2015). Genes Unite Executive Functions in Childhood. *Psychological science*, *26*(8), 1151–1163. https://doi.org/10.1177/0956797615577209

Erickson, K. I., Kim, J. S., Suever, B. L., Voss, M. W., Francis, B. M., & Kramer, A. F. (2008). Genetic contributions to age-related decline in executive function: a 10-year longitudinal

study of COMT and BDNF polymorphisms. *Frontiers in human neuroscience*, 2, 11. https://doi.org/10.3389/neuro.09.011.2008

Finkel, D., & Pedersen, N. L. (2004). Processing speed and longitudinal trajectories of change for cognitive abilities: The Swedish Adoption/Twin Study of Aging. *Aging, Neuropsychology, and Cognition*, 11(2-3), 325–345. https://doi.org/10.1080/13825580490511152

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. Psychological Assessment, 7(3), 286-299. https://doi.org/10.1037/1040-3590.7.3.286

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. Cortex, 86, 186-204.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. Journal of experimental psychology. General, 137(2), 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Fuster, J. M. (2002). Prefrontal cortex in temporal organization of action. The Handbook of Brain Theory and Neural Networks, 2nd edn. MIT Press, Cambridge, 905-910.

Gajewski, P. D., Hengstler, J. G., Golka, K., Falkenstein, M., & Beste, C. (2012). The Met-genotype of the BDNF Val66Met polymorphism is associated with reduced Stroop interference in elderly. *Neuropsychologia*, *50*(14), 3554-3563. https://doi.org/10.1016/j.neuropsychologia.2012.09.042

Gatt, J. M., Nemeroff, C. B., Dobson-Stone, C., Paul, R. H., Bryant, R. A., Schofield, P. R., ... & Williams, L. M. (2009). Interactions between BDNF Val66Met polymorphism and early

life stress predict brain and arousal pathways to syndromal depression and anxiety. *Molecular psychiatry, 14*(7), 681–695. https://doi.org/10.1038/mp.2008.143

Geller, S., Wilhelm, O., Wacker, J., Hamm, A., & Hildebrandt, A. (2017). Associations of the COMT Val158Met polymorphism with working memory and intelligence–A review and meta-analysis. *Intelligence, 65*, 75–92. https://doi.org/10.1016/j.intell.2017.09.002

Godinez, D. A., Friedman, N. P., Rhee, S. H., Miyake, A., & Hewitt, J. K. (2012). Phenotypic and genetic analyses of the Wisconsin Card Sort. *Behavior genetics, 42*(2), 209–220. https://doi.org/10.1007/s10519-011-9502-1

Götz, F. M., Gosling, S. D., & Rentfrow, J. (2021, January 10). Small effects: The indispensable foundation for a cumulative psychological science. https://doi.org/10.31234/osf.io/hzrxf

Grave, K. W., Bianchini, K. J., Hartley, S. M., Adams, D. (1999). The Wisconsin Card Sorting Test in stroke rehabilitation: Factor structure and relationship to outcome. *Archives of Clinical Neuropsychology, 14*, 497-509.

Grave, K. W., Ingram, F., Bianchini, K. J., (1998). Latent structure of the Wisconsin Card Sorting Test in clinical sample. *Archives of Clinical Neuropsychology, 13*, 597-609.

Grave, K. W., Love, J. M., Sherwin, E., Mathias, C. W., Ramzinski, P., Levy, J. (2002). Wisconsin Card Sorting Test in chronic severe traumatic brain injury: Factor structure and performance subgroups. *Brain Injury, 16*, 29-40.

Grave, K. W., Stickle, T. R., Love, J. M., Bianchini, K. J., Stanford, M. S. (2005). Latent structure of the Wisconsin Card Sorting Test: A confirmatory factor analytic study. *Archives of Clinical Neuropsychology, 20*, 355-364.

Gurvich, C., & Rossell, S. L. (2015). Dopamine and cognitive control: sex-by-genotype interactions influence the capacity to switch attention. *Behavioural brain research*, *281*, 96–101. https://doi.org/10.1016/j.bbr.2014.11.

Gustavson, D. E., Panizzon, M. S., Elman, J. A., Franz, C. E., Reynolds, C. A., Jacobson, K. C., Friedman, N. P., Xian, H., Toomey, R., Lyons, M. J., & Kremen, W. S. (2018). Stability of genetic and environmental influences on executive functions in midlife. Psychology and aging, 33(2), 219–231. https://doi.org/10.1037/pag0000230

Hariri, A. R., Goldberg, T. E., Mattay, V. S., Kolachana, B. S., Callicott, J. H., Egan, M. F., & Weinberger, D. R. (2003). Brain-derived neurotrophic factor val66met polymorphism affects human memory-related hippocampal activity and predicts memory performance. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 23(17), 6690–6694. https://doi.org/10.1523/JNEUROSCI.23-17-06690.2003

Harris, S. E., Fox, H., Wright, A. F., Hayward, C., Starr, J. M., Whalley, L. J., & Deary, I. J. (2006). The brain-derived neurotrophic factor Val66Met polymorphism is associated with age-related change in reasoning skills. *Molecular psychiatry*, 11(5), 505–513. https://doi.org/10.1038/sj.mp.4001799

Harrisberger, F., Smieskova, R., Schmidt, A., Lenz, C., Walter, A., Wittfeld, K., Grabe, H. J., Lang, U. E., Fusar-Poli, P., & Borgwardt, S. (2015). BDNF Val66Met polymorphism and hippocampal volume in neuropsychiatric disorders: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 55, 107–118. https://doi.org/10.1016/j.neubiorev.2015.04.017

Hashimoto, T., Fukui, K., Takeuchi, H., Yokota, S., Kikuchi, Y., Tomita, H., Taki, Y., & Kawashima, R. (2016). Effects of the BDNF Val66Met Polymorphism on Gray Matter

Volume in Typically Developing Children and Adolescents. *Cerebral cortex*, 26(4), 1795–1803. https://doi.org/10.1093/cercor/bhw020

Heaton, R. K. (1981). Wisconsin card sorting test manual. Psychological assessment resources.

Heaton, R. K., Chelune, G. J., Tallei, J. L., Key, G. G. & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and Expanded*. Odessa, FL: Psychological Assessment Resources.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are no WEIRD. *Nature*, *466*(7302), 29. https://doi.org/10.1038/466029a

Henson, R. K., & Roberts, J. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. Educational and Psychological Measurement, 66(3), 393-416. https://doi.org/10.1177/0013164405282485

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit I ndexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118.

Huges, C. & Graham, A. (2002). Measuring executive functions in childhood: Problems and solutions? *Child and Adolescent Mental Health*, *7*, 131-142.

IBM SPSS statistics for windows, version 21.0. Armonk, NY: IBM Corp.

Kautzky, A., James, G. M., Philippe, C., Baldinger-Melich, P., Kraus, C., Kranz, G. S., ... & Wadsak, W. (2019). Epistasis of HTR1A and BDNF risk genes alters cortical 5-HT1A receptor binding: PET results link genotype to molecular phenotype in depression. *Translational psychiatry*, *9*(1), 1–10. https://doi.org/10.1038/s41398-018-0308-2

Kazak, A. E. (2018). Editorial: Journal article reporting standards. *American Psychologist*, *73*(1), 1–2. https://doi.org/10.1037/amp0000263

923 Khanthiyong, B., Thanoi, S., Reynolds, G. P., & Nudmamud-Thanoi, S. (2019). Association
924 study of the functional Catechol-O-Methyltransferase (COMT) Val158Met
925 polymorphism on executive cognitive function in a Thai sample. *International journal*
926 *of medical sciences, 16*(11), 1461. https://doi.org/10.7150/ijms.35789
927 Klaus, K., Butler, K., Durrant, S. J., Ali, M., Inglehearn, C. F., Hodgson, T. L., ... & Pennington,
928 K. (2017). The effect of COMT Val158Met and DRD 2 C957T polymorphisms on
929 executive function and the impact of early life stress. *Brain and behavior*, *7*(5), e00695.
930 https://doi.org/10.1002/brb3.695
931 Kline, R. B. (2010). Principles and practice of structural equation modeling. New York, NY:
932 Guilford
933 Lang, U. E., Hellweg, R., Kalus, P., Bajbouj, M., Lenzen, K. P., Sander, T., Kunz, D., &
934 Gallinat, J. (2005). Association of a functional BDNF polymorphism and anxiety-related
935 personality traits. Psychopharmacology, 180(1), 95–99. https://doi.org/10.1007/s00213-
936 004-2137-7
937 Lebel, C., & Beaulieu, C. (2011). Longitudinal development of human brain wiring continues
938 from childhood into adulthood. *The Journal of neuroscience: the official journal of the*
939 *Society for Neuroscience*, *31*(30), 10937–10947.
940 https://doi.org/10.1523/JNEUROSCI.5302-10.2011
941 Logue, S. F., & Gould, T. J. (2014). The neural and genetic basis of executive function:
942 attention, cognitive flexibility, and response inhibition. *Pharmacology, biochemistry, and*
943 *behavior*, *123*, 45–54. https://doi.org/10.1016/j.pbb.2013.08.007

Lykken, D. T., McGue, M., & Tellegen, A. (1987). Recruitment bias in twin research: the rule of two-thirds reconsidered. Behavior genetics, 17(4), 343–362. https://doi.org/10.1007/BF01068136

Malhotra, A. K., Kestler, L. J., Mazzanti, C., Bates, J. A., Goldberg, T., & Goldman, D. (2002). A functional polymorphism in the COMT gene and performance on a test of prefrontal cognition. *The American journal of psychiatry*, *159*(4), 652–654. https://doi.org/10.1176/appi.ajp.159.4.652

Mandelman, S. D., & Grigorenko, E. L. (2012). BDNF Val66Met and cognition: all, none, or some? A meta-analysis of the genetic association. *Genes, Brain and Behavior*, *11*(2), 127-136. https://doi.org/10.1111/j.1601-183X.2011.00738.x

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. Biometrika, 57(3), 519-530.

McCartney, K., Harris, M. J., & Bernieri, F. (1990). Growing up and growing apart: A developmental meta-analysis of twin studies. *Psychological Bulletin*, 107(2), 226–237. https://doi.org/10.1037/0033-2909.107.2.226

McGue, M., & Bouchard, T. J. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics*, 14(4), 325–343. https://doi.org/10.1007/BF01080045

Millsap, R. E. (2012). Statistical approaches to measurement invariance. Routledge.

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: four general conclusions. Current Directions in Psychological Science, 21, 8–14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex

"Frontal Lobe" tasks: a latent variable analysis. *Cognitive psychology*, 41(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive Psychology, 41, 49–100.

Neale, M. C. C. L., & Cardon, L. R. (2013). *Methodology for genetic studies of twins and families* (Vol. 67). Springer Science & Business Media.

Nicole, S., & Del Miglio, C. (1997). Abstraction skillfulness in monozygotic and dizygotic twin pairs. *Acta geneticae medicae et gemellologiae*, *46*(1), 57–67. https://doi.org/10.1017/s0001566000000751

Obonsawin, M. C., Crawford, J. R., Page, J., Chalmers, P., Low, G., & Marsh, P. (1999). Performance on the Modified Card Sorting Test by normal, healthy individuals: relationship to general intellectual ability and demographic variables. The British journal of clinical psychology, 38(1), 27–41. https://doi.org/10.1348/014466599162647

O'Connor B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. Be*havior research methods, instruments, & computers: a journal of the Psychonomic Society*, Inc, 32(3), 396–402. https://doi.org/10.3758/bf03200807

Oniszczenko, W., Angleitner, A., Strelau, J., & Angert, T. (1993). The questionnaire of twins' physical resemblance. Unpublished questionnaire, University of Warsaw, Warsaw, Poland, and University of Bielefeld, Bielefeld, Germany.

Pineda, D. A., & Merchan, V. (2003). Executive function in young Colombian adults. International *Journal of Neuroscience*, 113(3), 397–410. https://doi.org/10.1080/00207450390162164

Polgár, P., Réthelyi, J. M., Bálint, S., Komlósi, S., Czobor, P., Bitter, I. (2010). Executive function in deficit schizophrenia: What do the dimensions of the Wisconsin Card Sorting Test tell us? *Schizophrenia Research, 122*, 85-93.

R Core Team (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing https://www.R-project.org/.

Reed, T., Plassman, B. L., Tanner, C. M., Dick, D. M., Rinehart, S. A., & Nichols, W. C. (2005). Verification of self-report of zygosity determined via DNA testing in a subset of the NAS-NRC twin registry 40 years later. *Twin research and human genetics*: the official journal of the International Society for Twin Studies, 8(4), 362–367. https://doi.org/10.1375/1832427054936763

Rijsdijk, F. V., & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in bioinformatics*, *3*(2), 119–133. https://doi.org/10.1093/bib/3.2.119

Rosa, A., Peralta, V., Cuesta, M. J., Zarzuela, A., Serrano, F., Martínez-Larrea, A., & Fañanás, L. (2004). New evidence of association between COMT gene and prefrontal neurocognitive function in healthy individuals from sibling pairs discordant for psychosis. *American Journal of Psychiatry, 161*(6), 1110–1112. https://doi.org/10.1176/appi.ajp.161.6.1110

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. Journal of Statistical Software, 48, 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Rybakowski, J. K., Borkowska, A., Czerski, P. M., Skibińska, M., & Hauser, J. (2003). Polymorphism of the brain-derived neurotrophic factor gene and performance on a cognitive prefrontal test in bipolar patients. *Bipolar disorders*, *5*(6), 468-472. https://doi.org/10.1046/j.1399-5618.2003.00071.x

Sanwald, S., Montag, C., & Kiefer, M. (2020). Depressive Emotionality moderates the influence of the BDNF Val66Met polymorphism on executive functions and on unconscious semantic priming. *Journal of Molecular Neuroscience, 70*(5), 699–712. https://doi.org/10.1007/s12031-020-01479-x

Smederevac, S., Mitrović, D., Sadiković, S., Milovanović, I., Branovački, B., Dinić, B. M., ... Milutinović, A. (2019). Serbian twin registry. *Twin Research and Human Genetics,* 22(6), 660–666. https://doi.org/10.1017/thg.2019.114.

Solomon, S. R., & Sawilowsky, S. S. (2009). Impact of rank–based normalizing transformations on the accuracy of test scores. *Journal of Modern Applied Statistical Methods*, 8, 448–462. https://doi.org/10.22237/jmasm/1257034080.

Somsen, R. J., Van der Molen, M. W., Jennings, J. R., van Beek, B. (2000). Wisconsin Card Sorting in adolescents: Analysis of performance, response times and heart rate. *Acta Psychologica, 104,* 227-257.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). Oxford University Press.

Suchy (2009). Executive functioning: Overview, assessment, and reasearch issues for non-neuropsychologists. *Annals of Behavioral Medicine, 37,* 106-116.

Taylor J. (2007). Heritability of Wisconsin Card Sorting Test (WCST) and Stroop Color-Word Test performance in normal individuals: implications for the search for endophenotypes.

Twin research and human genetics: the official journal of the International Society for Twin Studies, 10(6), 829–834. https://doi.org/10.1375/twin.10.6.829

Toh, Y. L., Ng, T., Tan, M., Tan, A., & Chan, A. (2018). Impact of brain-derived neurotrophic factor genetic polymorphism on cognition: A systematic review. *Brain and behavior*, *8*(7), e01009. https://doi.org/10.1002/brb3.1009

Tunbridge, E. M., Narajos, M., Harrison, C. H., Beresford, C., Cipriani, A., & Harrison, P. J. (2019). Which dopamine polymorphisms are functional? systematic review and meta-analysis of COMT, DAT, DBH, DDC, DRD1–5, MAOA, MAOB, TH, VMAT1, and VMAT2. *Biological psychiatry, 86*(8), 608–620. https://doi.org/10.1016/j.biopsych.2019.05.014

Wishart, H. A., Roth, R. M., Saykin, A. J., Rhodes, C. H., Tsongalis, G. J., Pattin, K. A., ... & McAllister, T. W. (2011). COMT Val158Met genotype and individual differences in executive function in healthy adults. *Journal of the International Neuropsychological Society, 17*(1), 174–180. https://doi.org/10.1017/s1355617710001402

Xu, H., Kellendonk, C. B., Simpson, E. H., Keilp, J. G., Bruder, G. E., Polan, H. J., ... & Gilliam, T. C. (2007). DRD2 C957T polymorphism interacts with the COMT Val158Met polymorphism in human working memory ability. *Schizophrenia research, 90*(1-3), 104–107. https://doi.org/10.1016/j.schres.2006.10.001

Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*, 198-226.

Table 1

Sociodemographic characteristics of study subjects

Age (years)	Range (17-58) 24.08 ± 7.02
Gender	
male	28.63%
female	71.37%
Race	
White/Caucasian	100.0%
Education level	
Primary education (8 years in total)	0.86%
Secondary (11-12 years in total)	27.16%
Higher School and University (16-17 years in total)	21.12%
Student	47.84%
Other	3.02%
Employment status	
Unemployed	34.87%
Employed	26.97%
Retired	0.88%
Student	32.24%
Other	5.04%
Material status	
Very bad	0.64%
Bad	4.29%
Average	33.05%
Good	49.36%
Very good	12.66%

Table 2

Descriptive Statistics and Twin Intraclass Correlation Coefficients for the Used Measures With 95% Confidence Intervals

WCST measure	*M (SD)*	MZ	DZ
Total Number of Errors	22.575 (17.229)	0.229** (0.073; 0.373)	-0.042 (-0.262; 0.180)
Non-Perseverative Errors	10.601 (9.891)	0.279** (0.126; 0.419)	-0.169 (-0.377; 0.055)
Perseverative Errors	12.215 (9.197)	0.213** (0.056; .359)	0.041 (-0.182; 0.259)
Perseverative Responses	13.676 (11.405)	0.207** (0.050; 0.353)	0.056 (-0.167; 0.273)
Categories Completed	5.485 (1.268)	0.296** (0.144; 0.434)	-0.116 (-0.330; 0.109)
Conceptual Level Responses	71.930 (17.191)	0.273** (0.120; 0.413)	-0.143 (-0.354; 0.081)
Trials to Complete the First Category	14.911 (10.618)	0.101 (-0.058; 0.255)	0.140 (-0.082; 0.349)
Failure to Maintain Set	0.526 (0.969)	0.143 (-0.016; 0.294)	-0.145 (-0.356; 0.079)
Categorizing Efficiency	68.441 (26.919)	0.292** (0.141; 0.431)	-0.159 (-0.369; 0.064)

Note. M – mean, *SD* – standard deviation; MZ – monozygotic twins, DZ – dizygotic twins; * $p < .05$. ** $p < .01$.

Table 3

Factor Loadings From the One-Factor Result of the CFA on the Measures of the WCST

Model 3. One-factor with correlated residuals	**Twin1**	**Twin2**	Model 4. Reduced one-factor model	**Twin1**	**Twin2**
Number of WCST Categories	0.675	0.805	Categorizing Efficiency	0.971	0.973
Perseverative Errors	-0.918	-0.837	Perseverative Errors	-0.892	-0.884
Non-Perseverative Errors	-0.909	-0.820	Non-Perseverative Errors	-0.895	-0.856
Failure To Maintain Set	-0.455	-0.595	Conceptual Level Responses	0.959	0.953
Trials to Complete the First Category	-0.313	-0.294			
Conceptual Level Responses	0.930	0.988			

Table 4

Fit indices for CFA models Twin 1 and Twin 2.

	Model	$\chi^2(df)$	*p-level*	AIC	BIC	CFI	TLI	SRMR	RMSEA (95% CI)
Twin 1	Model 1.	102.355 (9)	0.000	14007.928	140049.391	0.893	0.822	0.059	0.211 (0.177-0.246)
	Model 2.	86.180 (8)	0.000	13987.631	14032.550	0.914	0.838	0.064	0.204 (0.168- 0.243)
	Model 3.	11.397 (6)	0.077	13909.838	13961.668	0.994	0.985	0.030	0.062 (0.000-0.114)
	Model 4.	**3.961 (2)**	**0.136**	**8321.871**	**8349.514**	**0.998**	**0.995**	**0.006**	**0.065 (0.000-0.156)**
Twin 2	Model 1.	102.809 (9)	0.000	14057.995	14099.459	0.886	0.811	0.067	0.211 (0.179-0.245)
	Model 2.	99.531(8)	0.000	14051.808	14096.727	0.893	0.799	0.073	0.221 (0.186- 0.258)
	Model 3.	30.178 (6)	0.000	13972.265	14024.094	0.973	0.932	0.052	0.131 (0.089-0.177)
	Model 4.	**0.520 (2)**	**0.771**	**8427.141**	**8454.784**	**1.000**	**1.005**	**0.003**	**0.000 (0.000-0.085)**

Note-model 1. One-factor model; Model 2. Two- factor model; Model 3. One-factor model with correlated residuals; Model 4. Reduced one-factor model

Table 5
Fit Indices for Multivariate Models

	Model	χ^{2}/df	AIC	BIC	CFI	TLI	RMSEA (95% CI)
Independent pathways model	ACE	72.526 (48)	16777	16915.274	0.989	0.987	0.066 (0.031 – 0.096)
	ADE	70.255 (48)	16775	16913.003	0.990	0.988	0.063 (0.026 – 0.093)
	AE	74.193 (56)	16763	16873.299	0.992	0.992	0.053 (0.000 – 0.083)
	CE	78.296 (56)	16767	16877.402	0.990	0.990	0.058 (0.021 – 0.087)
Common pathways model	ACE	74.904 (53)	16769	16890.375	0.990	0.990	0.059 (0.022 – 0.089)
	ADE	73.103 (53)	16768	16888.575	0.991	0.991	0.057 (0.016 – 0.087)
	AE	**74.904 (58)**	**16759**	**16863.099**	**0.993**	**0.993**	**0.050 (0.000 – 0.080)**
	CE	78.714 (58)	16763	16866.909	0.991	0.991	0.055 (0.015 – 0.084)

Notes. A – additive genetic variance. D – non-additive genetic variance. C – shared environmental variance. E – non-shared environmental variance and measurement error.

Table 6

Specific and Common Genetic and Environmental Contributions for the Common AE Multivariate Model With 95% Confidence Intervals

	Measures	Ac	As	h2	Ec	Es	e2
AE Common pathways model	Categorizing Efficiency	0.232 (0.070 - 0.375)	0.001 (0.000 - 0.002)	0.233	0.712 (0.572 - 0.869)	0.055 (0.081 - 0.034)	0.767
	Perseverative Errors	0.174 (0.310 - 0.028)	0.025 (0.000 - 0.052)	0.199	0.616 (0.790- .464)	0.185 (0.228 - 0.144)	0.801
	Non-Perseverative Errors	0.211 (0.386 - 0.081)	0.000 (0.000 - 0.000)	0.211	0.557 (0.669 - 0.381)	0.232 (0.193- 0.294)	0.789
	Conceptual Level Responses	0.217 (0.063- 0.362)	0.000 (0.024 - 0.000)	0.217	0.697 (0.549 - 0.845)	0.086 (0.114 - 0.066)	0.782

Note. Ac – common genetic variance. As – specific genetic variance. h2 – hereditary variance. Ec – common environmental variance. Es – specific environmental variance. e2 – environmental variance.

Table 7

Mean Scores of WCST Dimensions in the Seven Allelic Configurations of Interest

	Categories Completed	Perseverative Errors	Perseverative Responses	Non-Perseverative Errors	Failure to Maintain Set	Trials to Complete the First Category	Total Errors	Conceptual Level Responses	Categorizing Efficiency	Factor score on the WCST
COMT										
Met+	8.081 (34.181)	8.900 (42.893)	9.006 (43.417)	10.757 (42.555)	5.619 (40.710)	2.298 (42.850)	11.097 (44.937)	-0.623 (14.983)	-1.412 (25.088)	-1.587 (23.506)
Met+/Met–	4.275 (34.279)	0.218 (45.254)	0.229 (45.145)	-0.369 (45.065)	1.174 (38.275)	0.805 (43.573)	0.001 (48.118)	0.986 (16.974)	1.480 (25.399)	1.641 (24.577)
Met–	-2.720 (35.589)	4.596 (47.048)	4.410 (46.655)	2.934 (49.296)	6.364 (40.980)	13.186 (46.769)	1.505 (51.819)	-1.572 (19.169)	-2.582 (30.419)	-2.112 (28.963)
DRD2										
A1+	2.616 (34.188)	4.791 (44.088)	4.653 (44.219)	4.318 (44.617)	2.819 (38.464)	-0.623 (41.559)	4.656 (47.433)	-0.831 (17.665)	-1.537 (26.885)	-1.270 (25.813)
A1–	4.963 (34.793)	3.121 (45.354)	3.192 (45.341)	3.301 (45.586)	3.718 (39.998)	5.236 (45.082)	3.219 (48.298)	0.448 (16.395)	0.535 (25.950)	0.576 (24.751)
BDNF										
Met66-	2.586 (35.372)	5.975 (45.748)	5.717 (45.693)	6.707 (45.873)	6.308 (40.836)	6.814 (44.120)	5.910 (48.860)	-0.732 (17.141)	-1.866 (27.059)	-1.568 (25.740)
Met66+	8.033 (32.499)	-1.751 (42.622)	-1.137 (42.967)	-3.521 (43.058)	-3.197 (35.430)	-4.463 (43.027)	-1.515 (45.638)	1.853 (15.871)	3.943 (23.819)	3.623 (23.142)

Table 8

Allelic Effects of COMT, DRD2, and BDNF Polymorphisms on WCST Dimensions

	Categories Completed	Perseverative Errors	Perseverative Responses	Non-Perseverative Errors	Failure to Maintain Set	Trials to Complete the First Category	Total Errors	Conceptual Level Responses	Categorizing Efficiency	Factor score on the WCST
COMT	2.196 (2)	1.479 (2)	1.503 (2)	2.384 (2)	0.726 (2)	2.143 (2)	2.184 (2)	0.751 (2)	0.849 (2)	0.947 (2)
DRD2	-0.632 (1)	0.346 (1)	0.303 (1)	0.209 (1)	-0.212 (1)	-1.239 (1)	.0279 (1)	-0.708 (1)	-0.735 (1)	-0.685 (1)
BDNF	-1.505 (1)	1.590 (1)	1.409 (1)	2.095 (1)*	2.361 (1)*	**2.376 (1)***	1.430 (1)	-1.422 (1)	-2.158(1)*	-2.000 (1)*

Note. Values in the cells are values of **t** and **F** tests with df (degrees of freedom) shown in brackets; ** *p< .01, * p<.05.* Correlations that remained significant after the False Discovery Rate (FDR) correction are bolded (for 30 correlation coefficients, with the default critical p-value set on .05, the corrected critical value was p = .01798).

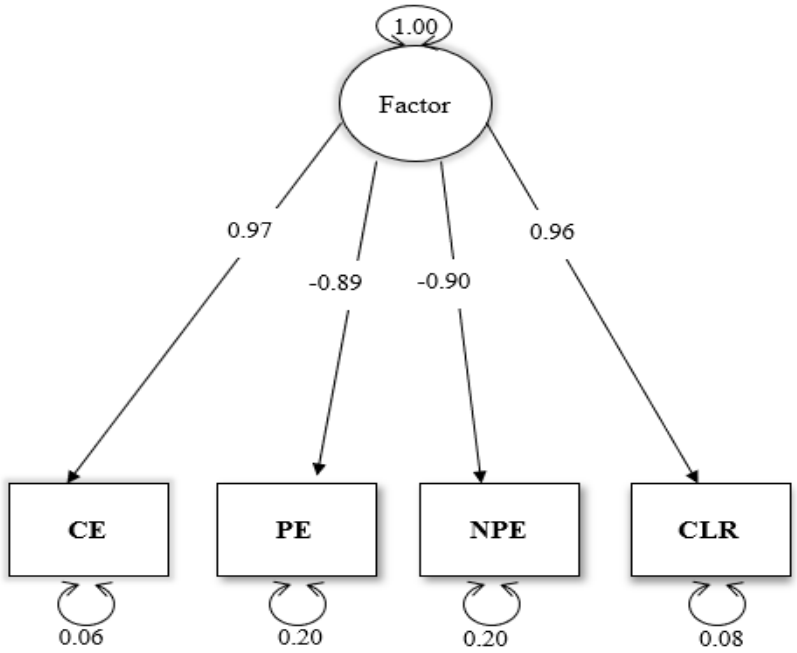

Fig. 1 Best-fitting (most parsimonious) CFA model twin 1. CE-Categorizing Efficiency PE-Perseverative errors; NPE- nonerseverative errors; CLR-Conceptual Level Responses

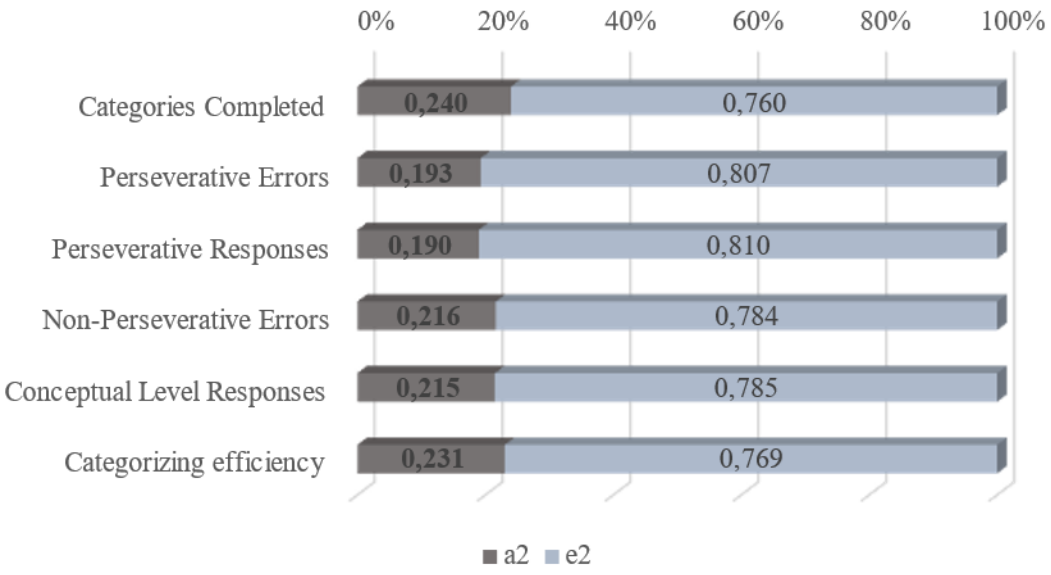

Fig. 2 Variance explained by additive genetic (a2) and non-shared environmental factors (e2) for WCST indexes.

Data availability statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.