Archives of Clinical Neuropsychology 34 (2019) 350–365

Confirmatory Factor Analysis of Neurocognitive Measures in Healthy Young Adults: The Relation of Executive Functions with Other Neurocognitive Functions

Luís Pires[1,2,3,*], Octávio Moura[1,4], Chiara Guerrini[3], Imke Buekenhout[1,2], Mário R. Simões[1,4], José Leitão[1,2]

[1] *CINEICC – Neuropsychology Research and Cognitive and Behavioural Intervention Center, Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal*
[2] *Memory, Language and Executive Functions Lab, Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal*
[3] *Centre for Health and Clinical Neuroscience, Psychology, School of Life Sciences, University of Hull, UK*
[4] *Psychological Assessment Lab, Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal*

*Corresponding author at: CINEICC – Neuropsychology Research and Cognitive and Behavioural Intervention Center, Faculty of Psychology and Educational Sciences, University of Coimbra, Rua do Colégio Novo, 3000-115 Coimbra, Portugal. Tel.: +351-239851450; fax: +351-239851462; *E-mail address*: luismspires@yahoo.com (L. Pires).

Editorial Decision 21 March 2018; Accepted 3 April 2018

Abstract

Objective: The present study aimed to investigate the factor structure of a set of neurocognitive tests theoretically assessing executive functions (EF), verbal abilities (VA), and processing speed (PS). This study extended previous research by analyzing if each test is better explained by the specific factor to which it theoretically belongs or by a more general neurocognitive factor; and also by analyzing the relations between the neurocognitive factors.
Methods: Using confirmatory factor analysis (CFA) we examined the factor structure of nine neurocognitive tests (EF: Working Memory, Tower, Divided Attention, Stroop, and Verbal Fluency tests; VA: Word List and Confrontation Naming tests; PS: Coding and Telephone Search tests) in a nonclinical sample ($N = 90$; 18–33 years old, 76 women). We tested five factor models of neurocognitive functioning: a one-factor model; two models with two-correlated factors; and two models with three-correlated factors.
Results: A three-correlated-factor model, with EF, VA, and PS factors, was the most suitable for our neuropsychological data. The Verbal Fluency test was better explained by the VA factor rather than by the EF factor. The EF factor was correlated with the PS factor, but not with the VA factor.
Conclusions: Most of the neurocognitive measures used in the present study loaded in the expected factors (with the exception of the Verbal Fluency that was apparently more related to VA). EF and PS represent related but separable functions. Our results highlight the need for a careful interpretation of test scores since performance on one test usually requires multiple functions.

Keywords: Cognition; Executive functions; Verbal abilities; Processing speed; Neuropsychological assessment; Confirmatory factor analysis

---

Introduction

Many of the tests currently used in clinical neuropsychology were derived from theoretical models of neurocognitive functions (e.g., Chan, Shum, Toulopoulou, & Chen, 2008). For example, Tower tests (Delis, Kaplan, & Kramer, 2001; Shallice, 1982; Simon, 1975) commonly used in neuropsychological assessment are derived from the supervisory attentional system (SAS) model from Norman and Shallice (1986). According to this model, cognition depends on the activation and interaction of two systems: a contention scheduling system, responsible for routine/automatic behavior; and a supervisory attentional system, responsible for non-routine/controlled behavior. The Tower tests allow the study of these systems since an optimal performance on these tests involves both automatic and control functions such as planning, monitoring and inhibition.

doi:10.1093/arclin/acy040

Even though most neuropsychological tests are theory-driven, it is crucial to analyze their psychometric properties (e.g., the reliability and validity of the test) to improve their clinical utility allowing a better interpretation of the results. Performance on neuropsychological tests often relies on the recruitment of multiple functions and it is critical to define the functions being assessed in each test. This dependence on several functions is especially noticeable in neuropsychological tests assessing executive functions.

Executive Functions (EF) comprise different cognitive capacities such as planning, cognitive flexibility, updating, inhibition, abstraction, decision making, among others (Royall et al., 2002). The role of these diverse functions in orchestrating resources such as memory, attention or language is vital to achieve success in everyday activities and interpersonal relationships (Jurado & Rosselli, 2007). An adequate performance on a test that in theory should assess a specific executive function, depends on the involvement of non-executive functions (nEF) and even on the recruitment of other EF (Royall et al., 2002). For example, in addition to EF abilities, the Stroop Color Word test (Stroop, 1935) requires nEF like reading and naming (MacLeod, 2016), whereas the Verbal Fluency test (Benton, 1968; Newcombe, 1969) entails verbal abilities (e.g., lexical knowledge and retrieval) (Shao, Janse, Visser, & Meyer, 2014). Conversely, many nEF tests also involve EF. For example, in the Word List test (Wechsler, 1997b) the participant must recruit EF, such as the abilities of updating and inhibiting, in order to correctly retain in long-term memory a list of words (Duff, Schoenberg, Scott, & Adams, 2005). EF are also closely related to processing speed (PS). As a basic cognitive function that mediates higher cognitive processes, PS is determinant for EF efficiency. Many EF tests are time-limited and even without a time limit a faster PS can be advantageous since it leads to a more rapid flow of information, enhancing test performance. Regardless of the relationship between PS and EF, PS does not influence all measures of EF in the same way (Salthouse, Atkinson, & Berish, 2003). For example, the Stroop Color Word test, a measure of inhibition, is heavily dependent on PS since a better score is influenced by the time taken to complete the task (MacLeod, 2016). However, a good performance on the Wisconsin Card Sorting Test (WCST; Berg, 1948), a measure of cognitive flexibility, is not dependent on the time taken to complete the task and therefore the PS will be less determinant to task performance (Nyhus & Barcelo, 2009). This complex nature of EF, as multiple control functions acting upon several aspects of cognition, contributed to the current uncertainty about the nature of these abilities.

Currently, EF can be conceptualized as a unitary process (Duncan & Owen, 2000), a set of separable functions (Stuss & Alexander, 2007) or as both unitary and diverse functions (Friedman & Miyake, 2017). In order to clarify the nature of EF, factor analytic studies [exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)] have been conducted to identify constructs underling EF test performance in clinical (Park et al., 2012; Savla et al., 2012) and healthy (Miyake et al., 2000; Testa, Bennett, & Ponsford, 2012) populations.

EFA studies of EF have been conducted to explore the underlying factor structure of a set of EF tests or a set of items within the same EF test. This has generally been done without any a priori hypothesis about the possible relationship between the tests/items or about the nature and number of the underlying factors. For example, Testa and colleagues (2012) used EFA to identify the underling factor structure of 19 EF tests in healthy adults. A principal components analysis with varimax rotation was performed and six independent factors were found: prospective working memory, set shifting and interference management, task analysis, response inhibition, strategy generation and regulation, and self-monitoring or set-maintenance. Their results revealed the diversity of cognitive functions assessed by the EF tests, explored their interrelationship and indicated how each one of these factors explained performance on the EF tests.

CFA are usually conducted to verify an a priori hypothesized factor structure for a set of tests/items based on theoretical assumptions, empirical research (e.g., EFA studies), or both. An example is the study conducted by Miyake and colleagues (2000) in which three EF were analyzed: shifting, updating, and inhibition. Young adults performed nine computerized tasks, three for each EF under study. They compared a three-factor model in which shifting, updating and inhibition were related but separable functions to other possible models (i.e., one-factor model, two-factor model and three-factor models in which the three EF were completely independent). Their results supported a three-correlated factor model showing that the three EF were indeed related but also separable functions. Therefore, CFA seems to be particularly relevant in determining the organization of EF and several CFA studies were conducted for this purpose (see Table 1). Most of these CFA studies were conducted with children (Brydges, Reid, Fox, & Anderson, 2012; Duan, Wei, Wang, & Shi, 2010; Friedman et al., 2006; Lehto, Juujarvi, Kooistra, & Pulkkinen, 2003; Miller, Giesbrecht, Müller, McInerney, & Kerns, 2012; Usai, Viterbori, Traverso, & De Franchis, 2014; van der Sluis, de Jong, & van der Leij, 2007; van der Ven, Kroesbergen, Boom, & Leseman, 2013; Wiebe, Espy, & Charak, 2008) but some CFA studies were also conducted with young adults (Miyake et al., 2000) and older adults (Hull, Martin, Beier, Lane, & Hamilton, 2008).

Most of the factor analytic studies conducted so far focused on defining the underlying factor structure of a set of tests assessing specific cognitive functions like EF. However, in clinical settings comprehensive neuropsychological batteries are often used, assessing distinct cognitive functions that are interrelated (e.g., an EF test encompasses also nEF and at the same time nEF tests rely on EF). Some studies used EFA and/or CFA to explore the underlying factor structure of comprehensive

Table 1. Factor analytic studies of executive functions

Authors/year	Sample	Tests	Factorial analysis	Number of factors	EF factors
Wiebe and colleagues (2008)	*n* = 243 (135 girls) 2 to 6 years old	10 EF tests	CFA	1	-Executive Function
Miller and colleagues (2012)	*n* = 129 (51 girls) 3 to 5 years old	8 EF tests	CFA	1	-Executive Function
van der Ven and colleagues (2013)	*n* = 211 (101 girls), 6 years old	9 EF tests	CFA	2	-Updating -Shifting and Inhibition
Usai and colleagues (2014)	time 1: *n* = 175 (76 girls), 5 and 6 years old; time 2: *n* = 145 (57 girls), 6 years old	6 EF tests 1 Fluid intelligence test	CFA	2	-Inhibition -Working Memory and Shifting
Brydges and colleagues (2012)	*n* = 215 (105 girls) 7 to 9 years old	9 EF tests 1 Fluid intelligence test 3 Crystalized intelligence test	CFA	1	-Executive Function
Lehto and colleagues (2003)	*n* = 108 (48 girls) 8 to 13 years old	9 EF tests	EFA and CFA	3	-Working Memory -Inhibition -Shifting
van der Sluis and colleagues (2007)	*n* = 172 (88 girls), 9 to 12 years old	11 EF tests 2 Fluid intelligence test 2 Crystalized intelligence test	CFA	2	-Updating -Shifting
Duan and colleagues (2010)	*n* = 61 (27 girls), 11 and 12 years old	7 EF tests 1 Fluid intelligence test	CFA	3	-Updating -Inhibition -Shifting
Friedman and colleagues (2006)	*n* = 234 twins 16 to 18 years old	8 EF tests 2 Fluid intelligence test 2 Crystalized intelligence test	CFA	3	-Updating -Inhibition -Shifting
Miyake and colleagues (2000)	*n* = 137 undergraduates	9 EF tests	CFA	3	-Updating -Inhibition -Shifting
Hull and colleagues (2008)	*n* = 100 (80 women), 51 to 74 years old	12 EF tests 1 Fluid intelligence test 1 Crystalized intelligence test	CFA	2	-Updating -Shifting

Note. EF = executive functions; CFA = confirmatory factor analysis; EFA = exploratory factor analysis.

neuropsychological batteries that were administered to clinical (Leonard et al., 2007; Park et al., 2012; Stinnett, Oehler-Stinnett, Fuqua, & Palmer, 2002) or nonclinical (Floyd, Bergeron, Hamilton, & Parra, 2010; Moleiro et al., 2013; Moura et al., 2018; Santos et al., 2015; Siedlecki, Honig, & Stern, 2008) samples. For example, Stinnett and colleagues (2002) examined with EFA the underlying factor structure of the NEPSY, a neuropsychological battery assessing attention/EF, language, sensorimotor functions, visuospatial processing, memory, and learning in children. They found that a structure with only a single neurocognitive factor was the most suitable, with almost all the tests adequately loading on this single factor. This study highlighted the common variance among tests theoretically assessing different cognitive domains. Santos and colleagues (2015) investigated the factor structure of neurocognitive measures assessing general cognitive status, short and long-term memory, inhibition, and verbal fluency in older adults. They found with EFA a two-factor solution with memory and general/executive function. Then, they used CFA in a distinct sample from the same original cohort to confirm that a two-correlated factor model suggested by EFA was indeed the best fit. Recently, Moura and colleagues (2018) studied the factor structure of the Coimbra Neuropsychological Assessment Battery (BANC; Simões et al., 2016), that comprises the assessment of laterality, motor function, orientation, memory, language, and attention/EF abilities in children. Only the tests theoretically assessing memory, language and attention/EF were studied. Several neurocognitive models with three, four and five factors were compared and a three-correlated factor model (memory, language and attention/EF) represented the best fit to their data. Overall, mixed results have been found in respect to the factor structure derived from EFA and/or CFA and the proposed theoretical models. The present study aimed to investigate the factor structure of nine neurocognitive tests theoretically assessing EF, verbal abilities (VA) and PS. We examined three neuropsychological domains (i.e., EF, VA, and PS) that are particularly relevant in the clinical practice. The relationship between these domains is still not well-defined with some EF tests relying on VA or PS abilities. Conversely, the involvement of EF in the performance of VA and PS measures is not fully defined.

This study extended previous research by analyzing if each test is better explained by the specific factor to which it theoretically belongs or by a more general neurocognitive factor; and by analyzing the relations between the neurocognitive factors. In order to analyze different aspects of EF, we included measures of inhibition, working memory (WM), planning, verbal fluency, and divided attention. Inhibition and WM (particularly the updating function of its contents) are considered core EF functions (Miyake et al., 2000). Verbal fluency tasks have also been commonly used as an EF measures and it is clear that a number of EF (e.g., inhibition, updating) contribute to performance in these tasks along with memory and language processes (Henry & Crawford, 2004; Moura, Simoes, & Preira, 2015). The ability to execute complex plans (Allain et al., 2005) and divided attention abilities have also been described as relevant EF (Diamond, 2013). We also included two VA measures, one assessing verbal episodic memory and the other one assessing confrontation naming ability. We are not advocating that these VA measures do not comprise any EF abilities but we selected measures in which the level of EF involvement is certainly reduced in comparison to the one needed in the EF tests (Duff et al., 2005). Because PS may be an important mediator of the selected EF and VA measures (Bryan, Luszcz, & Crawford, 1997; Henninger, Madden, & Huettel, 2010; Lee et al., 2012), we also included two PS measures.

We tested through CFA five models of neurocognitive functioning:

- Model 1: A one-factor model suggesting that neuropsychological tests that in theory assess different cognitive domains are better explained by a single neurocognitive factor, as found for example by Stinnett and colleagues (2002);
- Model 2: A two-correlated factor model with EF and nEF factors, in which the VA and PS measures are included in the nEF and Verbal Fluency (that is theoretically related to both EF and VA) was included in the EF factor. The separation between EF and nEF (such as memory, language, and PS) has been found in different studies (Floyd et al., 2010; Leonard et al., 2007). The inclusion of the Verbal Fluency measure as EF is supported by previous studies (Moleiro et al., 2013; Moura et al., 2018; Santos et al., 2015);
- Model 3: Another two-correlated factor model with EF and nEF factors was tested. PS measures were included in the EF factor, considering PS as a basic function underlying performance in EF tests (Park et al., 2012);
- Model 4: Despite the close relationship between PS and other neurocognitive measures, several studies suggest that PS is separable from EF and nEF (Floyd et al., 2010; Leonard et al., 2007; Moleiro et al., 2013). Therefore, we tested a three-correlated factor model with EF, PS, and VA factors.
- Model 5: Another three-correlated factor was tested. It was identical to Model 4, except for the inclusion of Verbal Fluency in the VA factor, as suggested by previous studies (Floyd et al., 2010; Park et al., 2012; Siedlecki et al., 2008).

We hypothesized that a three-correlated-factor model (Model 4 or Model 5) would best represent the underlying structure of our neurocognitive battery. We expected that EF, VA, and PS would be separable but related functions.

Method

Participants

A sample of 115 young adults volunteered to participate. All the participants were first-year Psychology students at the University of Coimbra (Portugal) and received course credit in return for their participation. They provided written informed consent in accordance with institutional guidelines prior to their inclusion in the study. The study was approved by the ethical committee of the Psychology department. Exclusion criteria comprised current or previous diagnosis of psychological, psychiatric or neurologic disorders, vision or hearing impairment. In addition, during neuropsychological assessment, participants were screened for depressive symptoms with the Beck Depression Inventory II (Beck, Steer, & Brown, 1996) and a cut-off score of 20 points (i.e., moderate depression symptoms) was used to determine exclusion. Due to the presence of moderate depressive symptoms, sixteen participants were excluded from the sample. Nine participants were also excluded due to current psychoactive medication intake (e.g., anxiolytics) and/or the presence of other medical conditions that could interfere with behavioral testing (e.g., Diabetes Mellitus). As a result, data from 90 young adults (76 female; 18–33 years old, $M = 19.77$, $SD = 2.85$; 12–20 years of formal education, $M = 13.10$, $SD = 1.47$) were analyzed. The mean estimated intelligence, as measured by the TeLPI - Irregular Words Reading Test (Alves, Simoes, & Martins, 2012), a Portuguese test similar to the National Adult Reading Test (Nelson, 1982), was 118.33 (range = 108–130), indicating that the sample was well placed within the average range for their age and education levels.

Materials

A comprehensive neuropsychological battery designed to assess EF (i.e., inhibition, planning, WM, divided attention, and verbal fluency), VA (i.e., episodic memory and confrontation naming) and PS was administered to all participants.

Below we present the different measures included in the CFA analysis.

Coding. The Digit Symbol Coding or Coding test (Wechsler Adult Intelligence Scale – Third edition; Wechsler, 1997a) allows the evaluation of psychomotor control, speed, sustained attention, and (incidental) memory. Despite the number of functions (both cognitive and motor) that are involved in this task, the Coding test is considered to mainly reflect PS (Joy, Kaplan, & Fein, 2004), and has been widely used with this purpose in clinical settings (Gonzalez-Blanch et al., 2011; Khanahmadi, Malmir, Eskandari, & Orang, 2013). In this test, the examinee must copy symbols previously associated with numbers, in a predetermined matrix. Thus, the examinee's capacity to make a fast association between symbols and numbers is evaluated in this task. There is a time limit of 2 min. The total score is the number of correct items produced within the time limit, ranging from 0 to 133 points.

Telephone search. This subtest from the Test of Everyday Attention battery (TEA; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1994) involves visual selective attention and PS. In this task the participants have to look at stimuli similar to the yellow pages from a telephone directory, in which they can find different services (i.e., restaurants, plumbers…). These yellow pages include pairs of symbols (e.g., a square and a circle; two stars…), one pair per service, and the participants must search for the pairs in which the symbols are the same (e.g., two circles) and ignore the mismatched pairs (e.g., one star and one circle). The participant is instructed to complete the task as fast and accurately as possible. There is a time limit of 4 min. The total score is the time per target (the mean time needed to correctly identify each one of the targets). This time per target measure is used in the analyses.

Divided attention. This divided attention measure is obtained from the difference between the time per target in the Telephone Search and the time per target when the participant has to perform the same task (a parallel form is used) while trying to execute a second task at the same time. This second task requires the participant to count a series of phone tones played by an audio recorder – the Dual Task Telephone Search subtest from TEA. In this test the instructions emphasize that both tasks are equally important and the participant must complete both tasks as accurately as possible. The time required to complete the task is 4 min. The Telephone Search and Dual Task Telephone Search subtests from the TEA were selected to assess selective and divided attention, respectively. The TEA was mainly developed to offer a clinically valid assessment of attentional processes in individuals from 18 to 80 years old (Evans, Greenfield, Wilson, & Bateman, 2009; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996). Importantly, the TEA provides enhanced ecological validity (i.e., a better predictive power with respect to everyday cognitive functioning) by introducing subtests that closely relate to everyday tasks, in this case searching telephone directories (for a review of ecological validity in neuropsychological assessment see Chaytor & Schmitter-Edgecombe, 2003).

Working memory. This is a composite score obtained from two tests: a verbal WM test (Digit Suppression Test – DST; Beblo, Macek, Brinkers, Hartje, & Klaver, 2004) and a spatial WM test (Block Suppression Test – BST; Beblo et al., 2004). Both verbal and spatial WM have been studied in clinical neuropsychology (Marquand, Mourao-Miranda, Brammer, Cleare, & Fu, 2008; Wood et al., 2003). Beblo and colleagues (2004) highlighted that the most used methods to assess WM in clinical settings (e.g., memory span tasks) were not capable of capturing all aspects of working memory, due to the simplicity of their information-manipulation component. This new set of measures, the DST and the BST, address this problem by introducing the requirement of suppressing some of the incoming information before reproduction. In the DST the participant must repeat every second digit of a sequence of digits orally presented by the examiner, beginning with the first digit. (e.g., 1-7-4 repeats 1-4; 1-5-7-8 repeats 1-7). The trial starts with a sequence of three digits and the length of the sequences increases until 16 digits are reached. There is a total of 28 items, two for each level. The task ends when the participant fails to correctly recall the digits of the two items of the same level (i.e., same sequence length). The total score is the number of accurately remembered digits, ranging from 0 to 28 points. In the BST, the spatial version of the DST, the participants must tap every second block of a sequence of blocks tapped by the examiner, beginning with the first block. Toepper, Beblo, Thomas, and Driessen (2008) found that the BST can provide useful information for the early detection of Alzheimer's disease. In this test the blocks are tapped by the examiner in a 1-s pace. The task ends when the participant fails to correctly tap the blocks in the two items of the same level. The total score is the number of accurately reported taps. The composite measure obtained from verbal and spatial WM tests is used for data analysis.

Stroop. The Stroop test has been widely used in research and clinical settings (Castro, Martins, & Cunha, 2003; Homack & Riccio, 2004; MacLeod, 2016). In this version of the Stroop test (Castro et al., 2003; Trenerry, Crosson, DeBoe, & Leber, 1989) there are four colors (red, blue, pink, and gray) used in four tasks: (i) reading colored words in which the written color word and the ink color in which the word is printed are incongruent (e.g., the word red printed in blue reads “red”); (ii) the Stroop condition – naming the color of the colored words read in (i) (e.g., the word red printed in blue says “blue”); (iii) reading color words printed in black ink (e.g., the word red printed in black reads “red”); (iv) naming the color of colored bars (e.g., when the bar is filled with red ink says “red”). The task contains 112 items and has a time limit of 2 min. In the Stroop Test we used the number of correct items in the reading black words task (W) and the number of correct items in the naming colored bars task (C) to predict the number of correct items in the Stroop condition [$WC' = (W \times C)/(W + C)$]. Then we computed the difference between the number of correct items in the Stroop condition (WC) and the prediction ($WC–WC'$) and used it as a measure of interference.

Tower. This is a subtest from the Delis–Kaplan Executive System (D-KEFS; Delis et al., 2001). There are other variations of the tower test such as the Tower of London (Shallice, 1982) and the Tower of Hanoi (Humes, Welsh, Retzlaff, & Cookson, 1997) but the D-KEFS version allows for a greater difficulty variation, improving the overall psychometrics of the test (Delis et al., 2001). Participants are asked to construct several towers up to a maximum of nine towers. The difficulty level increases from the first tower to the ninth tower (e.g., in the first tower one movement is necessary to complete the tower; in the ninth tower at least 26 movements must be executed to correctly construct the tower). All the items have a time limit (e.g., the first tower has a time limit of 30 s; the ninth tower has a time limit of 240 s). A photograph with the tower to be constructed is presented to the participant, who has to manipulate wooden disks with different sizes and blue color tones on a board with three vertical pegs. They must construct each tower using as few movements as possible and follow two rules: (i) move just one disk at a time, using only one hand; (ii) never place a larger disk on top of a smaller disk. The achievement score, combining the number of correct towers and the overall movement accuracy, was used in the analyses.

Verbal fluency. Phonemic and semantic verbal fluency tests are frequently included in clinical neuropsychological assessment batteries because they assess different neurocognitive functions (e.g., EF, memory, language) and show sensitivity to the cognitive deficits that characterize different clinical conditions (Melrose et al., 2009; van den Berg, Jiskoot, Grosveld, van Swieten, & Papma, 2017). In this study we used a composite Verbal Fluency score obtained from two D-KEFS tests, the Phonemic Fluency test, and the Semantic Fluency test. In the Phonemic Fluency test the participant is asked to generate as many words as possible beginning with a specific letter within 1 min. There were three phonemic categories: P, M, and R. These categories represent the Portuguese language frequency counterparts of the F, A, and S categories used in English language. In the Semantic Fluency test the participant is asked to generate as many words from a semantic category as possible within 1 min. There were three semantic categories: animals; food commonly found in a supermarket; things people do (i.e., action verbs). In both phonemic and semantic tasks the sum of the number of words recalled during the three categories was computed. The composite measure obtained from Phonemic and Semantic Fluency tests was used in the analyses.

Confrontation naming. This is a subtest from the Psycholinguistic Assessment of Language battery (PAL 09; Caplan & Bub, 1990; Portuguese version, PAL-PORT, Festas, Martins, Leitão, & 2007). This task is similar to the Boston Naming test (Kaplan, Goodglass, & Weintraub, 1983), the most popular naming test used in clinical neuropsychology (Harry & Crowe, 2014). Naming difficulties are commonly found in clinical settings and depending on the cognitive processes that are impaired, they can have different clinical presentations, like aphasia, anomia, or agnosia (Laine & Martin, 2006). This task assesses the ability of the participant to access the phonological forms of words from their meanings (activated by black and white pictures). There are 44 pictures corresponding to 44 words from seven semantic categories (animals, fruits, clothes, artefacts; instruments, vehicles and vegetables), with high and low frequency (22 items each) and with long and short length (22 items each). The total score is given by the number of correctly named pictures, ranging from 0 to 44 points.

Word List. This test from Wechsler Memory Scale – third edition (WMS-III; Wechsler, 1997b) assesses verbal episodic memory. It encompasses free immediate recall of 12 words’ lists over four trials, followed by a free short-delayed recall (after the free immediate recall of an interference 12 words’ list) and a long-delayed recall (after a 25 min interval). There is also a delayed recognition test (with 24 words, 12 new and 12 previously presented). The total score is derived from the number of words recalled or accurately recognized. The delayed recall of previously learned word lists is among the most used measures of episodic memory abilities in clinical neuropsychology (Cerami et al., 2017). In this task we only used the percentage of long-delay retention.

Procedure

The testing session lasted approximately 90 min and was always conducted in a well-lit room within the Psychology Department. The neuropsychological tests were administered in random order except for the following order constraints: (i) the long-term interval of the Word List test (25 min) was always filled with the same tests applied in the same order: a semi-structured interview; the Beck Depression Inventory II; and the Coding test. These tests did not include words as stimuli that could cause interference with the learned word list; (ii) the Telephone Search subtest was always administered prior to the Dual Task Telephone Search subtest, following guidelines from TEA; (iii) the Digit Suppression and the Block Suppression tests were always administered together to facilitate instructions comprehension and were applied in a random order; (iv) the Phonemic and the Semantic Fluency tests were always administered together to facilitate instructions comprehension and were applied in a random order.

Data Analysis

Age and level of education are relevant moderators of performance on tests of neurocognitive function (Ardila, Ostrosky-Solis, Rosselli, & Gomez, 2000; Beeri et al., 2006; Lam et al., 2013; Ostrosky-Solis, Ramirez, & Ardila, 2004). A preliminary correlation analysis revealed significant (or marginally significant) correlations between these two demographic variables and the neurocognitive measures (e.g., $r = -.221$, $p < .05$ between age and spatial WM; $r = -.212$, $p < .05$ between age and phonemic verbal fluency). Thus, to control for the influence of age (ranging from 18 to 33 years old) and years of formal education (ranging from 12 to 20 years) on the neurocognitive measures, an adjusted score was created by regressing the raw score of each neurocognitive measure onto age and years of formal education, and then saving the standardized residual score. These standardized residual scores were used in the CFA analysis.

To test the factor structure of the neuropsychological battery, a CFA was performed using IBM SPSS Amos 20 (Arbuckle, 2011). The models tested were estimated through covariance matrices using maximum likelihood estimation. Model fit was assessed through various fit statistics. We reported chi-square (χ^2), two absolute fit indices (SRMR and RMSEA), as well as an incremental fit index (CFI) and a parsimonious fit index (AIC). The χ^2 is known to be extremely sensitive to sample size, meaning that with larger samples, even reasonable models are likely to produce statistically significant chi-square p values (Bentler & Bonett, 1980; Bentler, 1990; Bryant & Yarnold, 1995; Joreskog & Sorbom, 1986). Therefore, the use of other fit indices besides the chi-square is recommended (Byrne, 2005). Hu and Bentler (1999) recommend a CFI of >.95, a SRMR of <.08 and an RMSEA of <.06 to determine good fit. The AIC was reported to compare different models, with smaller values representing a better fit. As suggested by Marsh, Hau, and Wen (2004) these traditional cut-off values should not be used as rules of thumb. Therefore, more stringent cut-off values are recommended for simple models, and less stringent cut-off values are recommended for more complex models (Cheung & Rensvold, 2002; Marsh et al., 2004).

Results

Descriptive Statistics and Correlation Analysis

Descriptive statistics (raw scores) for each neurocognitive test are presented in Table 2. For the Telephone Search test and the Stroop test higher values indicate worse performance. For the Divided attention measure, values closest to 0 indicate better performance. For all the other tests higher values indicate better performance.

The correlations among all the measures are presented in Table 3. There are several significant correlations between the neuropsychological measures, mostly of moderate size (Cohen, 1988). There was a strong and positive correlation between WM and Stroop measures [$r(88) = .501$, $p < .01$], a moderate and positive correlation between the WM and Divided Attention measures[$r(88) = .355$, $p < .05$], and a small and positive correlation between WM and Tower measures [$r(88) = .255$, $p < .01$]. Verbal Fluency was not significantly correlated to any EF measure, but a moderate and positive correlation was found with Confrontation Naming [$r(88) = .316$, $p < .01$] and a small and negative correlation was found with Telephone Search [$r(88) = -.292$, $p < .01$]. A moderate and negative correlation between the Coding and Telephone Search tasks [$r(88) = -.374$, $p < .01$] was found.

Confirmatory Factor Analysis

CFA enabled us to evaluate several a priori models (i.e., based on theoretical considerations and on the results of previous factor analysis studies) in terms of their fit with the observed data. Five theoretical models of neurocognitive functioning were tested through CFA (see Table 4).

In Model 1 (one-factor model) we tested how the nine neurocognitive measures are explained by a single (general) neurocognitive factor (see Table 4). As shown in Table 5, Model 1 did not provide a good fit to the data, with CFI = .674;

Table 2. Descriptive statistics for the neurocognitive measures ($N = 90$)

Test		Dependent measure	*M* (*SD*)	Range
Coding (WAIS-III)		Number of correct codifications	89.50 (13.37)	59–117
Telephone Search (TEA)		Time per target (s)	2.96 (.72)	1.8–6
Confrontational Naming (PAL)		Number of correct named pictures	33.41 (3.55)	22–42
Long-term Recall–Word List (WMS-III)		% of retention after a 25 min interval	89.70 (12.77)	50–110
Divided Attention		Difference between the time per target (s) in Telephone Search and in Dual Task Telephone Search	−1.38 (1.87)	−8.34–3.77
Working Memory	Digit Suppression Test	Number of correct items	12.36 (3.49)	3–18
	Block Suppression Test	Number of correct items	11.09 (3.35)	4–19
Stroop		Interference score	49.09 (10.67)	10–56.50
Tower (D-KEFS)		Achievement score	17.80 (3.59)	10–27
Verbal Fluency	Phonemic Fluency	Sum of total number of words in three conditions: P; M; R	35.46 (10.23)	16–64
	Semantic Fluency	Sum of total number of words in three conditions: Animals; Food commonly found in a supermarket; Verbs	56.97 (10.54)	29–82

Note: WAIS-III = Wechsler Adult Intelligence Scale third version; TEA = Test of Everyday Attention; PAL = Psycholinguistic Assessment of Language; WMS-III = Wechsler Memory Scale third version; D-KEFS = Delis–Kaplan Executive Function System.

Table 3. Pearson's product–moment correlations between neurocognitive measures ($N = 90$)

	1	2	3	4	5	6	7	8	9
1. Telephone Search	—	−.009	−.340**	.003	−.108	−.292**	−.129	−.123	−.374**
2. Divided Attention		—	.181	.189	.355**	−.077	−.083	−.0.90	.135
3. Stroop			—	.181	.501**	.120	−.016	.070	.398**
4. Tower				—	.255**	−.029	.027	−.051	.169
5. Working Memory					—	−.017	.025	.080	.195
6. Verbal Fluency						—	.316**	.193	.111
7. Confrontation Naming							—	.347**	.110
8. Word List								—	.175
9. Coding									—

Note: *$p < .05$ **$p < .001$.

Table 4. Factor models estimated through CFA

Models	Factors	Variables
Model 1 (one-factor model)	Single neurocognitive factor	Working Memory; Tower; Divided Attention; Stroop; Verbal Fluency; Word List; Confrontation Naming; Coding; Telephone Search
Model 2 (two-correlated-factor model)	EF	Working Memory; Tower; Divided Attention; Stroop; Verbal Fluency
	nEF	Word List; Confrontation Naming; Coding; Telephone Search
Model 3 (two-correlated-factor model)	EF	Working Memory; Tower; Divided Attention; Stroop; Verbal Fluency; Coding; Telephone Search
	nEF	Word List; Confrontation Naming
Model 4 (three-correlated-factor model)	EF	Working Memory; Tower; Divided Attention; Stroop; Verbal Fluency
	VA	Word List; Confrontation Naming
	PS	Coding; Telephone Search
Model 5 (three-correlated-factor model)	EF	Working Memory; Tower; Divided Attention; Stroop
	VA	Word List; Confrontation Naming; Verbal Fluency
	PS	Coding; Telephone Search

Note: EF = executive functions; nEF = non-executive functions; VA = verbal abilities; PS = processing speed.

RMSEA (90% CI) = .106 (.064–.147); and SRMR = .109. So, a single neurocognitive factor is not sufficient to explain the variance in most tests.

A two-correlated-factor model, Model 2, was then tested. This model predicts that the neurocognitive measures included in our battery can be explained by two factors: an EF factor, and a nEF factor, including VA and PS measures (see Table 4). Model 2 did not provide a good fit to the data, with CFI = .765; RMSEA (90% CI) = .092 (.044–.135); and SRMR = .101. Another two-correlated factor model (EF and nEF factors) was tested, with PS measures included in the EF factor (Model 3).

Table 5. Goodness-of-fit indices for the five factor models estimated through CFA

CFA models	χ^2	*df*	χ^2/df	CFI	RMSEA (90% CI)	SRMR	AIC
Model 1	53.964, p = .002	27	2.00	.674	.106 (.064–.147)	.109	89.964
Model 2	45.397, p = .011	26	1.75	.765	.092 (.044–.135)	.101	83.397
Model 3	42.853, p = .020	26	1.65	.796	.085 (.034–.130)	.098	80.853
Model 4	34.887, p = .070	24	1.45	.868	.071 (.000–.120)	.089	76.887
Model 5	24.661, p = .424	24	1.03	.992	.018 (.000–.088)	.065	66.661

Note: χ^2 = Chi-square; χ^2/df = Relative/Normed Chi-square; CFI = Comparative Fit Index; RMSEA (90% CI) = Root Mean Square Error of Approximation (90% confidence interval); SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion.

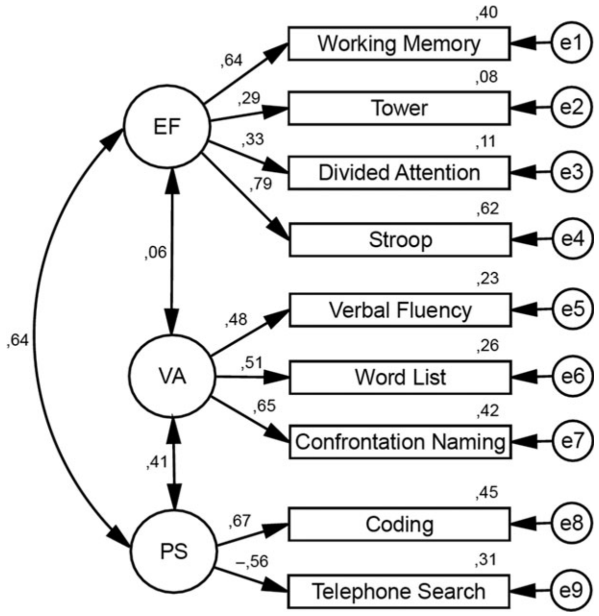

Fig. 1. Model 5 – three-correlated-factor model (standardized solution) with Executive Functions (EF), Verbal Abilities (VA) and Processing Speed (PS). Circles represent the latent variables and boxes represent each observed variable. Values in the middle of two arrow heads lines represent correlation between factors. Values in the arrows pointing from the factors to the observed values represent loadings of each one of the observed values in the corresponding factor. The values above each observed variable represent the variance explained by the factor (e.g., in the Stroop test a loading of .79 indicates that the EF factor explain 62% [(.79^2) = .62] of the variance in the Stroop test' performance).

This model also did not provide a good fit to the data, with CFI = .796; RMSEA (90% CI) = .085 (.034–.130); and SRMR = .098. So, the two-factor models (i.e., Model 2 and Model 3) were unable to adequately explain the data.

Subsequently, we tested two different three-factor models, with EF, VA and PS factors. In Model 4, Coding and Telephone Search were included in the PS factor, whereas Verbal Fluency remained in the EF factor. Despite these changes, Model 4 did not provide a good fit to the data, with CFI = .868; RMSEA (90% CI) = .071 (.000–.120); and SRMR = .089. Finally, another three-correlated factor model was tested with one adjustment relative to Model 4: the inclusion of Verbal Fluency in the VA factor. In Model 4, Verbal Fluency was poorly explained by the EF (λ = .12). Thus, the inclusion of Verbal Fluency on the VA factor makes empirical and statistical sense. Model 5 showed a good model fit, with CFI = .992; RMSEA (90% CI) = .018 (.000–.088); and SRMR = .065.

Taken together, the results from the CFA showed that Model 5 provided the best fit to the data. The factor loadings and correlations among factors of Model 5 are presented in Fig. 1. EF and VA are not related to each other (r = .06), and PS is more related to EF (r = .64) than to VA (r = .41).

Discussion

A large number of factor analytic studies of neurocognitive tests have been conducted to analyze the psychometric properties of neuropsychological batteries assessing specific neurocognitive domains such as the ones assessing EF (e.g., D-KEFS; Delis et al., 2001), attention (e.g., TEA; Robertson et al., 1994), memory (e.g., WMS; Wechsler, 1997b) or language (e.g., PAL; Caplan & Bub, 1990). However, in a comprehensive clinical evaluation, several neurocognitive tests tapping multiple domains (usually from different batteries/scales) are commonly administered.

In the present study, we analyzed the factor structure of neurocognitive tests that assess different cognitive domains and that are widely used in clinical settings. We used CFA to determine if the tests included in our test battery, that in theory should assess EF, VA, and PS, are better explained by a single neurocognitive factor or by an underlying factor structure with two or three related neurocognitive factors. Consistent with the theoretical conceptualization of our test battery, a three-correlated factors model (EF, VA, and PS; Model 5), revealed the best fit to the data.

When we examined local fit, most factor loadings were adequate. However, the EF factor only explained 11% of the variance in the Divided Attention test and 8% of the variance in the Tower test. Higher factor loadings for these two tests were expected, as they are usually associated with EF.

The Divided Attention measure reflects dual task coordination and has been linked also to sustained attention (Robertson et al., 1996). One possible explanation for the low factor loading of the Divided Attention test on the EF factor could be its reduced adequacy to assess dual task coordination in our sample. Indeed, there was little variation in the young adults' performance and a small dual task cost. This could indicate that the task was not challenging enough to activate control processes associated with dual task coordination. Previous studies also did not find a relation between dual tasks and other EF tests (Miyake et al., 2000; Fournier-Vicente, Larigauderie, & Gaonac'h, 2008), thus suggesting that these tasks due to their complexity may be relying on a wide range of EF and nEF functions.

Concerning the Tower test we also expected a higher factor loading on the EF factor. The Tower test is usually interpreted as a measure of planning ability. This capacity to organize behavior in order to achieve a particular goal relies on different but related cognitive processes linked to EF such as shifting of mental sets, updating of WM representations, and inhibition (D'Antuono et al., 2017; Lehto et al., 2003; Miyake et al., 2000). Other studies also did find a smaller loading of the Tower test (Moura et al., 2018) and a small correlation between the Tower test and other EF tests (Floyd et al., 2010; Savla et al. 2012).

One possible explanation for this finding could be the level of reliance on EF entailed by the different strategies used by the participants while performing the Tower test (e.g., perceptual strategy, move pattern strategy, or peg selection strategy; see Goel, Pullara, & Grafman, 2001). Some of these strategies rely on cognitive processes that are more dependent on EF (e.g., maintaining the number of movements made down to the minimum needed to correctly perform the task) while others rely on cognitive processes that are less dependent on EF (e.g., trading efficiency for speed and completing correctly the tower without trying to complete it with the fewest number of movements possible). As suggested by Keith Berg and Byrd (2002), while referring to a different tower test (i.e., Tower of London), it is important to consider different performance measures to fully understand the processes involved in the Tower test. Another possible explanation for the smaller EF loading of the Tower test is the nature of the performance score that we used in our analysis, namely the achievement score. In fact, it could be argued that other performance scores that are obtained for the Tower test reflect more adequately the complex interplay of processes that is evoked by the task. The time needed for the first movement is usually considered a "planning" time in which the participants select the best strategy to accomplish each tower (Koppenol-Gonzalez, Bouwmeester, & Boonstra, 2010). However, variations in the time needed to perform the first movement could be due to factors other than planning ability; for example, the participant could take more time to perform the first movement because he or she is being overcareful or inefficient, inversely he or she could be faster in performing the first movement due to impulsivity. Other measures reflect specific processes that do contribute to planning but that cannot on their own fully account for it (e.g., the number of errors or rule breaks is linked to inhibition but does not reflect other processes relevant to the completion of the Tower test, like updating of working memory representations or shifting abilities). Nevertheless, we acknowledge that given the complexity of the Tower test even the achievement score can be insufficient to tap all the cognitive processes involved in test performance. Future studies could better define the precise nature of the executive processes involved in performing the Tower test, as well as assess the clinical relevance of the performance scores that this test yields.

Finally, the smaller loading of the Tower test on the EF factor could also indicate that the cognitive processes involved in this test are not strongly related to the cognitive processes assessed in the Working Memory, the Divided Attention and the Stroop test, as indicated by the correlation coefficients shown in Table 3.

The EF factor seems to better explain the variance in the Stroop test and in WM test performance. This highlights the key role of inhibition (assessed in the Stroop test) and of the activation of the relevant information in WM to EF. These two functions have been considered core EF (Diamond, 2013; Friedman & Miyake, 2017) and previous CFA studies conducted with EF tests identified factors representing inhibition and WM (Duan et al., 2010; Lehto et al., 2003; Miyake et al., 2000; van der Ven et al., 2013).

Contrary to our initial prediction, variance in the Verbal Fluency test performance was explained by the VA factor and not by the EF factor. Also, no significant correlations were found between Verbal Fluency and the tests loading in the EF factor (i.e., WM test; Stroop test; Tower test and Divided Attention test).

Multiple cognitive functions have been appointed to be recruited during performance on Verbal Fluency tests, including EF (e.g., cognitive flexibility, WM, and inhibition), verbal intelligence, semantic retrieval, and PS (Boone, Ponton, Gorsuch, Gonzalez, & Miller, 1998; Bryan et al., 1997; Henry & Crawford, 2004; Ross et al., 2007). Some studies attempted to isolate the core functions in these tests (Kraan, Stolwyk, & Testa, 2013; Ross et al., 2007; Shao et al., 2014; Whiteside et al., 2016). For example, Whiteside and colleagues (2016) used EFA to examine the underlying cognitive structure of Verbal Fluency tests (including both phonemic and semantic fluency) and found that the language processing is the critical component for these tasks. However, the authors assessed a mixed clinical sample and their study was retrospective, making the generalization of the results difficult. Their version of Verbal Fluency included three categories of phonemic fluency (i.e. words starting with F, A, and S) and only one category of semantic fluency (animals). Also, they included just two EF measures in the study (WSCT and Trail Making Test). Our study addressed some of these limitations by assessing healthy young adults, a different and larger set of EF measures and more semantic categories in the Verbal Fluency test. Our results supported their findings and suggest that Verbal Fluency is more related to VA than EF.

Recently, van den Berg and colleagues (2017) examined Verbal Fluency in a clinical sample and found a close relationship between these tests and language/verbal memory tests similar to the ones included in the present study. In agreement to these findings, our results indicate that Verbal Fluency was significantly correlated to the Confrontation Naming test. In both tests the participants use their lexical access ability in order to retrieve words from the mental lexicon.

This closer relationship of Verbal Fluency to the VA factor rather than the EF factor can contribute to a more effective selection of the tests included in neuropsychological assessment batteries. According to the present study, the cognitive substrate of Verbal Fluency performance is distinct from other EF measures (Whiteside et al., 2016), it follows that we should have some cautions in interpreting the performance of verbal fluency tasks as measuring executive functioning. Language ability is critical for the performance in Verbal Fluency tests (even if we can assume some level of recruitment of EF). The clarification in the construct validity of Verbal Fluency and its separation from other EF tasks is also important to balance the number of verbal and non-verbal tasks when designing a neuropsychological battery (Chan et al., 2008; Shao et al., 2014). The present results should also contribute to a more accurate identification of verbal and non-verbal EF deficits in patients and in the disentanglement between the two functions/dysfunctions

As expected, variance in the performance on the Confrontation Naming test and on the Word List test was adequately explained by the VA factor. Previous studies also found a close relationship between language and memory abilities (Moura et al., 2018).

Verbal Fluency (both phonemic and semantic fluency), Confrontation Naming and Word List measures were all related to the VA factor which indicates the existence of commonalities between the cognitive functions necessary to perform on these tests. Neuroimaging studies have shown the activation of a fronto-temporal network in both Verbal Fluency tests (Melrose et al., 2009), Confrontation Naming test (Bonelli et al., 2011), and episodic/semantic memory tests like the Word List test (Takashima, Bakker, van Hell, Janzen, & McQueen, 2017). Also, temporal lobe epilepsy patients typically show deficits in verbal fluency, naming, semantic and episodic memory abilities (Bonelli et al., 2011; McAndrews & Cohn, 2012), thus suggesting the activation of similar brain networks when these abilities are recruited.

In relation to the PS factor, we found adequate loadings for the two PS measures, the Telephone Search and the Coding tests. Performance on the Telephone Search heavily relied on PS despite being considered a measure of selective attention (Chan, Lai, & Robertson, 2006; Robertson et al., 1996). The Coding test is often used as measure of PS (Joy et al., 2004).

Concerning the correlations among the three factors, we found that PS was indeed related to both EF and VA but surprisingly EF and VA were unrelated.

PS and EF were separable but related domains. The role of PS in the Stroop test was recently highlighted by Naber, Vedder, Brown, and Nieuwenhuis (2016). In their study, the performance on the Stroop task was not only explained by EF. Specifically, both stimulus PS and lateral inhibition explained 40% of the variance in their Stroop task. The strong relation that we found between PS and EF was also expected due to the general nature of the tests that intended to measure PS and EF abilities. Three of the four EF tests are time-constrained: the Stroop, the Divided Attention and the Tower tests. Regarding WM, even if participants were not performing the WM tests under time constraints they still benefited from higher PS. A slower rate of processing reduces the amount of information that can be processed, impairing encoding and retrieval (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002).

Indeed, PS is important to performance in EF tests and vice versa(Cepeda, Blackwell, & Munakata, 2013). In our study, we used two measures of PS in which there is an involvement of EF, the Coding test from WAIS-III and the Telephone Search from TEA. In the Coding test visual–motor coordination and WM are relevant to task performance (Baudouin, Clarys, Vanneste, & Isingrini, 2009). In the Telephone Search, visual selective attention abilities along with PS allow successful task performance (Chan et al., 2006). Thus, this interaction between PS and EF has important implications for clinical neuropsychological assessment: in order to accurately interpret performance on EF tests, a PS measure should always be included in the assessment battery. By disentangling which component is compromised (i.e., whether PS, EF, or both), it is possible to make a more accurate diagnosis and a targeted intervention. Evidence for this close interplay between EF and PS was also recently unveiled in recent research with clinical populations (Liebel et al., 2017; Motes et al., 2018).

We also found a moderate correlation between PS and VA. This is consistent with the view that PS is an important neurocognitive function that mediates the performance of other neurocognitive functions like EF, language, and memory (Floyd et al., 2010; Henninger et al., 2010; Lee et al., 2012; Moleiro et al., 2013; Naber et al., 2016; Salthouse et al., 2003).

In our study EF was not significantly correlated to VA. However, the current conceptualization of EF (Jurado & Rosselli, 2007; Royall et al., 2002) suggests that these functions interact with nEF in order to control the execution of complex activities. Other studies did find EF to be related to performance in Verbal Fluency tests (Kraan et al., 2013; Shao et al., 2014), Confrontation Naming tests (Abrahams et al., 2003) and in verbal memory tests like the Word List test (Duff et al., 2005). One explanation for this discrepancy could be that the EF recruited in these tests are in some manner distinct from the EF assessed in our study. Abrahams and colleagues (2003) suggested that verbal fluency and confrontation naming may recruit a semantic executive system, which is responsible for accessing, maintaining and manipulating semantic representations. None of the EF measures included in our study are closely related to semantic processing. DeDe, Caplan, Kemtes, and Waters (2004) found that online syntactic processing was not related to traditional WM measures and proposed that this language comprehension mechanism could rely on a different WM resource linked to language processing. Overall, these studies suggest the existence of a specific EF system operating in verbal tasks that is autonomous from EF systems linked to traditional EF measures. This conclusion is in line with the recent suggestion that an effective EF assessment should target EF in a set of different modalities, that should be verbal, non-verbal, and also address other EF dimensions, such as emotion and social needs (Diamond & Ling, 2016).

Our findings support a view of the EF system that emphasizes the diversity and relative autonomy within this set of functions. In fact, only two EF tests were strongly related to each other (i.e. Stroop and WM) and adequately loaded in the EF factor, while others EF tests were either poorly related (i.e., Tower and Divided Attention), or unrelated to the EF factor (i.e., verbal fluency). There were also distinct associations of PS to different EF tests used in our study (Salthouse et al., 2003; Salthouse, 2005). Performance on PS measures was related to performance on the Stroop test but not to performance on the WM, Divided Attention or Tower tests. The EF factor was unrelated to the VA factor underlying verbal fluency, verbal episodic memory, and naming abilities. This was not expected and it can suggest the existence of autonomous EF systems (e.g., one linked to linguistic/semantic processing and another, more general, related to inhibition and updating of information processed out of the scope of the lexical and syntactic systems) (Abrahams et al., 2003; DeDe et al.2004). Another important consideration is that various EF (e.g., WM, inhibition) work together to solve complex problems and conduct intricate decision processes (Alvarez & Emory, 2006). The interplay of the various EF requires coordination, and both frontal and non-frontal brain regions (i.e., temporal and parietal) are likely involved in the orchestration of these processes (Stuss, 2011). In light of the present results, the traditional conceptualization of EF found in clinical neuropsychology practices and academic literature as frontal abilities, seems to be insufficient and inadequate.

We acknowledge that our findings have some limitations. Other EF tests commonly used in clinical neuropsychology (e.g., WCST and TMT) could have been included in our battery. Additionally, the reduced size of our sample and its homogeneity may affect the generalizability of our results, namely, in respect to individuals with fewer years of formal schooling and/or below average intellectual functioning. This is because our sample includes young adults that mostly have more than 12 years of formal education. Therefore, even though the estimated IQs in this sample (ranging between 108 and 130, with a mean of 118) are within the average range (the age/education appropriate mean obtained from the TelPI normative data is 123.42 ± 4.087), it can be argued that our participants' intellectual functioning only represents a limited band of the overall range, and that their scores in the overall test battery also reflect this limitation. Indeed, some studies have found that IQ is a significant predictor of performance in several neurocognitive domains (Diaz-Asper, Schretlen, & Pearlson, 2004; Friedman et al., 2006). It would be also particularly interesting to measure invariance of the factor structure of Model 5 across age groups (e.g., middle-aged adults vs. older adults) and clinical populations. In the context of our study we found that a three related factor model was the most adequate model to explain our data but we cannot predict if similar results would be obtained for a clinical young adults' sample or an older adults' sample. The measurement invariance analysis would clarify if the factor structure is equivalent across age and/or clinical conditions (Wicherts et al., 2016). If the relation between the observed variables (i.e., neuropsychological tests) and the latent variables (i.e., factors) included in the model was not invariant, all subsequent between-groups comparisons (e.g., age or clinical groups) may likely be invalid (Moura et al., 2018). The validity of the clinical neuropsychological assessment conclusions would be compromised if this lack of invariance is not properly taken into account while interpreting neuropsychological data.

In sum, we confirmed our a priori assumptions about the general cognitive domains assessed in our neurocognitive test battery. A three-correlated factor model with EF, VA, and PS factors presented the best fit to the data. Most of the neurocognitive measures loaded on the expected factors (the exception was Verbal Fluency). EF and PS were strongly related suggesting that PS is relevant for the performance in EF tasks and vice versa. These findings are relevant for a more reasoned selection of tests for the neuropsychological assessment protocol and for a more accurate interpretation of the test scores.

Funding

This original research has not been published elsewhere and was supported by a fellowship from Fundação para a Ciência e a Tecnologia (FCT), a portuguese public agency that supports science, technology and innovation [FCT; (SFRH/BD/70011/2010/Psicologia)].

Conflict of Interest

None declared.

References

Abrahams, S., Goldstein, L. H., Simmons, A., Brammer, M. J., Williams, S. C., Giampietro, V. P., et al. (2003). Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. *Human Brain Mapping*, *20*, 29–40. doi:10.1002/hbm.10126.

Allain, P., Nicoleau, S., Pinon, K., Etcharry-Bouyx, F., Barre, J., Berrut, G., et al. (2005). Executive functioning in normal aging: A study of action planning using the Zoo Map Test. *Brain and Cognition*, *57*, 4–7. doi:10.1016/j.bandc.2004.08.011.

Alvarez, J. A., & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*, *16*, 17–42. doi:10.1007/s11065-006-9002-x.

Alves, L., Simoes, M. R., & Martins, C. (2012). The estimation of premorbid intelligence levels among Portuguese speakers: The Irregular Word Reading Test (TeLPI). *Archives of Clinical Neuropsychology*, *27*, 58–68. doi:10.1093/arclin/acr103.

Arbuckle, J. L. (2011). *Amos (version 20.0)* [Computer Program]. Chicago: IBM SPSS.

Ardila, A., Ostrosky-Solis, F., Rosselli, M., & Gomez, C. (2000). Age-related cognitive decline during normal aging: The complex effect of education. *Archives of Clinical Neuropsychology*, *15*, 495–513.

Baudouin, A., Clarys, D., Vanneste, S., & Isingrini, M. (2009). Executive functioning and processing speed in age-related differences in memory: Contribution of a coding task. *Brain and Cognition*, *71*, 240–245. doi:10.1016/j.bandc.2009.08.007.

Beblo, T., Macek, C., Brinkers, I., Hartje, W., & Klaver, P. (2004). A new approach in clinical neuropsychology to the assessment of spatial working memory: The block suppression test. *Journal of Clinical and Experimental Neuropsychology*, *26*, 105–114. doi:10.1076/jcen.26.1.105.23938.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.

Beeri, M. S., Schmeidler, J., Sano, M., Wang, J., Lally, R., Grossman, H., et al. (2006). Age, gender, and education norms on the CERAD neuropsychological battery in the oldest old. *Neurology*, *67*, 1006–1010. doi:10.1212/01.wnl.0000237548.15734.cd.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Benton, A. L. (1968). Differential behavioural effects in frontal lobe disease. *Neuropsychologia*, *6*, 53–60. doi:10.1016/0028-3932(68)90038-9.

Berg, E. A. (1948). A simple objective technique for measuring flexibility in thinking. *Journal of General Psychology*, *39*, 15–22. doi:10.1080/00221309.1948.9918159.

Bonelli, S. B., Powell, R., Thompson, P. J., Yogarajah, M., Focke, N. K., Stretton, J., et al. (2011). Hippocampal activation correlates with visual confrontation naming: fMRI findings in controls and patients with temporal lobe epilepsy. *Epilepsy Research*, *95*, 246–254. doi:10.1016/j.eplepsyres.2011.04.007.

Boone, K. B., Ponton, M. O., Gorsuch, R. L., Gonzalez, J. J., & Miller, B. L. (1998). Factor analysis of four measures of prefrontal lobe functioning. *Archives of Clinical Neuropsychology*, *13*, 585–595.

Bryan, J., Luszcz, M. A., & Crawford, J. R. (1997). Verbal knowledge and speed of information processing as mediators of age differences in verbal fluency performance among older adults. *Psychology and Aging*, *12*, 473–478.

Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In Grimm L. G., & Yarnold P. R. (Eds.), *Reading and understanding multivariate statistics* (pp. 99–136). Washington, DC: American Psychological Association.

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, *40*, 458–469.

Byrne, B. M. (2005). Factor analytic models: Viewing the structure of an assessment instrument from three perspectives. *Journal of Personality Assessment*, *85*, 17–32. doi:10.1207/s15327752jpa8501_02.

Caplan, D., & Bub, D. (1990). *Psycholinguistic assessment of aphasia*. Paper presented at the American Speech and Hearing Association Conference, Seattle, USA.

Castro, S. L., Martins, L., & Cunha, L. (2003). *Neuropsychological screening with a Portuguese Stroop test*. Paper presented at the 111th Annual Convention of the American Psychological Association, Toronto, Canada.

Cepeda, N. J., Blackwell, K. A., & Munakata, Y. (2013). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, *16*, 269–286. doi:10.1111/desc.12024.

Cerami, C., Dubois, B., Boccardi, M., Monsch, A. U., Demonet, J. F., Cappa, S. F., et al & Geneva Task Force for the Roadmap of Alzheimer's Biomarkers. (2017). Clinical validity of delayed recall tests as a gateway biomarker for Alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiology of Aging*, *52*, 153–166. doi:10.1016/j.neurobiolaging.2016.03.034.

Chan, R. C. K., Lai, M. K., & Robertson, I. H. (2006). Latent structure of the Test of Everyday Attention in a non-clinical Chinese sample. *Archives of Clinical Neuropsychology*, *21*, 477–485. doi:10.1016/j.acn.2006.06.007.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, *23*, 201–216. doi:10.1016/j.acn.2007.08.010.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*, 181–197. doi:10.1023/B:Nerv.0000009483.91468.Fb.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. B. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183. doi:Pii/S0160-2896(01)00096-4.

D'Antuono, G., La Torre, F. R., Marin, D., Antonucci, G., Piccardi, L., & Guariglia, C. (2017). Role of working memory, inhibition, and fluid intelligence in the performance of the Tower of London task. *Applied Neuropsychology: Adult*, *24*, 548–558. doi:10.1080/23279095.2016.1225071.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system: Technical manual*. San Antonio, TX: Harcourt Assessment Company.

DeDe, G., Caplan, D., Kemtes, K., & Waters, G. (2004). The relationship between age, verbal working memory, and language comprehension. *Psychology and Aging*, *19*, 601–616. doi:10.1037/0882-7974.19.4.601.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. doi:10.1146/annurev-psych-113011-143750.

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience*, *18*, 34–48. doi:10.1016/j.dcn.2015.11.005.

Diaz-Asper, C. M., Schretlen, D. J., & Pearlson, G. D. (2004). How well does IQ predict neuropsychological test performance in normal adults? *Journal of the International Neuropsychological Society*, *10*, 82–90. doi:10.1017/S1355617704101100.

Duan, X., Wei, S., Wang, G., & Shi, J. (2010). The relationship between executive functions and intelligence on 11- to 12-year old children. *Psychological Test and Assessment Modeling*, *52*, 419–431.

Duff, K., Schoenberg, M. R., Scott, J. G., & Adams, R. L. (2005). The relationship between executive functioning and verbal and visual learning and memory. *Archives of Clinical Neuropsychology*, *20*, 111–122. doi:10.1016/j.acn.2004.03.003.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences*, *23*, 475–483.

Evans, J. J., Greenfield, E., Wilson, B. A., & Bateman, A. (2009). Walking and talking therapy: Improving cognitive-motor dual-tasking in neurological illness. *Journal of the International Neuropsychological Society*, *15*, 112–120. doi:10.1017/S1355617708090152.

Festas, I., Martins, C., & Leitão, J. (2007). Avaliação da compreensão e da leitura de palavras na PAL-PORT (Bateria de Avaliação Psicolinguística das Afasias e de outras Perturbações da Linguagem para a População Portuguesa). *Revista Educação: Temas e Problemas*, *4*, 223–239.

Floyd, R. G., Bergeron, R., Hamilton, G., & Parra, G. R. (2010). How do executive functions fit with the cattellñhornñcarroll model? Some evidence from a joint factor analysis of the Delis-Kaplan executive function system and the Woodcock-Johnson III tests of cognitive abilities. *Psychology in the Schools*, *47*, 721–738. DOI:10.1002/pits.20500.

Fournier-Vicente, S., Larigauderie, P., & Gaonac'h, D. (2008). More dissociations and interactions within central executive functioning: A comprehensive latent-variable analysis. *Acta Psychologica*, *129*, 32–48. doi:10.1016/j.actpsy.2008.04.004.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. doi:10.1016/j.cortex.2016.04.023.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*, 172–179. doi:10.1111/j.1467-9280.2006.01681.x.

Goel, V., Pullara, S. D., & Grafman, J. (2001). A computational model of frontal lobe dysfunction, working memory and the Tower of Hanoi task. *Cognitive Science*, *25*, 287–313.

Gonzalez-Blanch, C., Perez-Iglesias, R., Rodriguez-Sanchez, J. M., Pardo-Garcia, G., Martinez-Garcia, O., Vazquez-Barquero, J. L., et al. (2011). A Digit Symbol Coding Task as a screening instrument for cognitive impairment in first-episode psychosis. *Archives of Clinical Neuropsychology*, *26*, 48–58. doi:10.1093/arclin/acq086.

Harry, A., & Crowe, S. F. (2014). Is the Boston Naming Test still fit for purpose? *The Clinical Neuropsychologist*, *28* (*3*), 486–504. doi:10.1080/13854046.2014.892155.

Henninger, D. E., Madden, D. J., & Huettel, S. A. (2010). Processing speed and memory mediate age-related differences in decision making. *Psychology and Aging*, *25*, 262–270. doi:10.1037/a0019096.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*, 284–295. doi:10.1037/0894-4105.18.2.284.

Homack, S., & Riccio, C. A. (2004). A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology*, *19*, 725–743. doi:10.1016/j.acn.2003.09.003.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6* (*1*), 1–55.

Hull, R., Martin, R. C., Beier, M. E., Lane, D., & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology*, *22*, 508–522. doi:10.1037/0894-4105.22.4.508.

Humes, G. E., Welsh, M. C., Retzlaff, P., & Cookson, N. (1997). Towers of Hanoi and London: Reliability and validity of two executive function tasks. *Assessment*, *4*, 249–257. doi:10.1177/107319119700400305.

Joreskog, K. G., & Sorbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood and least squares methods*. Mooresville, IN: Scientific Software.

Joy, S., Kaplan, E., & Fein, D. (2004). Speed and memory in the WAIS-III Digit Symbol – Coding subtest across the adult lifespan. *Archives of Clinical Neuropsychology*, *19*, 759–767. doi:10.1016/j.acn.2003.09.009.

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, *17*, 213–233. doi:10.1007/s11065-007-9040-z.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.

Keith Berg, W., & Byrd, D. (2002). The Tower of London spatial problem-solving task: Enhancing clinical and research implementation. *Journal of Clinical and Experimental Neuropsychology*, *24*, 586–604. doi:10.1076/jcen.24.5.586.1006.

Khanahmadi, M., Malmir, M., Eskandari, H., & Orang, T. (2013). Evaluation of visual information processing speed in depressed people. *Iranian Journal of Public Health*, *42*, 1266–1273.

Koppenol-Gonzalez, G. V., Bouwmeester, S., & Boonstra, A. M. (2010). Understanding planning ability measured by the Tower of London: An evaluation of its internal structure by latent variable modeling. *Psychological Assessment*, *22*, 923–934. doi:10.1037/a0020826.

Kraan, C., Stolwyk, R. J., & Testa, R. (2013). The abilities associated with verbal fluency performance in a young, healthy population are multifactorial and differ across fluency variants. *Applied Neuropsychology: Adult*. doi:10.1080/09084282.2012.670157.

Laine, M., & Martin, N. (2006). *Anomia: Theoretical and clinical aspects* (pp. 37–93). New York: Psychology Press.

Lam, M., Eng, G. K., Rapisarda, A., Subramaniam, M., Kraus, M., Keefe, R. S., et al. (2013). Formulation of the age-education index: Measuring age and education effects in neuropsychological performance. *Psychological Assessment*, *25*, 61–70. doi:10.1037/a0030548.

Lee, T., Crawford, J. D., Henry, J. D., Trollor, J. N., Kochan, N. A., Wright, M. J., et al. (2012). Mediating effects of processing speed and executive functions in age-related differences in episodic memory performance: A cross-validation study. *Neuropsychology*, *26*, 776–784. doi:10.1037/a0030053.

Lehto, J. E., Juujarvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*, 59–80.

Leonard, L. B., Weismer, S. E., Miller, C. A., Francis, D. J., Tomblin, J. B., & Kail, R. V. (2007). Speed of processing, working memory, and language impairment in children. *Journal of Speech Language and Hearing Research*, *50*, 408–428. doi:10.1044/1092-4388(2007/029).

Liebel, S. W., Jones, E. C., Oshri, A., Hallowell, E. S., Jerskey, B. A., Gunstad, J., et al. (2017). Cognitive processing speed mediates the effects of cardiovascular disease on executive functioning. *Neuropsychology*, *31*, 44–51. doi:10.1037/neu0000324.

MacLeod, C. M. (2016). The Stroop Effect. In Luo, M. R. (Ed.) *Encyclopedia of Color Science and Technology*. New York, NY: Springer. doi:10.1007/978-3-642-27851-8_67-1.

Marquand, A. F., Mourao-Miranda, J., Brammer, M. J., Cleare, A. J., & Fu, C. H. Y. (2008). Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport*, *19*, 1507–1511. doi:10.1097/WNR.0b013e328310425e.

Marsh, H. W., Hau, K. T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler’s (1999) findings. *Structural Equation Modeling*, *11*, 320–341.

McAndrews, M. P., & Cohn, M. (2012). Neuropsychology in temporal lobe epilepsy: Influences from cognitive neuroscience and functional neuroimaging. *Epilepsy Research and Treatment*, *2012*, 925238. doi:10.1155/2012/925238.

Melrose, R. J., Campa, O. M., Harwood, D. G., Osato, S., Mandelkern, M. A., & Sultzer, D. L. (2009). The neural correlates of naming and fluency deficits in Alzheimer’s disease: An FDG-PET study. *International Journal of Geriatric Psychiatry*, *24*, 885–893. doi:10.1002/gps.2229.

Miller, M. R., Giesbrecht, G. F., Müller, U., McInerney, R. J., & Kerns, K. A. (2012). A latent variable approach to determining the structure of executive function in preschool children. *Journal of Cognition and Development*, *13*, 395–423.

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. doi:10.1006/cogp.1999.0734.

Moleiro, C., Madureira, S., Verdelho, A., Ferro, J. M., Poggesi, A., Chabriat, H., et al. (2013). Confirmatory factor analysis of the Neuropsychological Assessment Battery of the LADIS study: A longitudinal analysis. *Journal of Clinical and Experimental Neuropsychology*, *35*, 269–278. doi:10.1080/13803395.2013.770822.

Motes, M. A., Yezhuvath, U. S., Aslan, S., Spence, J. S., Rypma, B., & Chapman, S. B. (2018). Higher-order cognitive training effects on processing speed-related neural activity: A randomized trial. *Neurobiology of Aging*, *62*, 72–81. doi:10.1016/j.neurobiolaging.2017.10.003.

Moura, O., Albuquerque, C. P., Pinho, M. S., Vilar, M., Lopes, A. F., Alberto, I., et al. (2018). Factor structure and measurement invariance of the Coimbra Neuropsychological Assessment Battery (BANC). *Archives of Clinical Neuropsychology*, *33*, 66–78. doi:10.1093/arclin/acx052.

Moura, O., Simoes, M. R., & Preira, M. (2015). Executive functioning in children with developmental dyslexia. *The Clinical Neuropsychologist*, *28*, S20–S41. doi:10.1080/13854046.2014.964326.

Naber, M., Vedder, A., Brown, S. B., & Nieuwenhuis, S. (2016). Speed and lateral inhibition of stimulus processing contribute to individual differences in Stroop-Task Performance. *Frontiers in Psychology*, *7*, 822. doi:10.3389/fpsyg.2016.00822.

Nelson, H. E. (1982). *National Adult Reading Test*. Windsor, UK: NFER-Nelson.

Newcombe, F. (1969). *Missile wounds of the brain: A study of psychological deficits*. London: Oxford University Press.

Norman, D., & Shallice, T. (1986). Attention to action. In Davidson R., Schwartz G., & Shapiro D. (Eds.), *Consciousness and self-regulation* (pp. 1–18). New York: Plenum Press.

Nyhus, E., & Barcelo, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: A critical update. *Brain and Cognition*, *71*, 437–451. doi:10.1016/j.bandc.2009.03.005.

Ostrosky-Solis, F., Ramirez, M., & Ardila, A. (2004). Effects of culture and education on neuropsychological testing: A preliminary study with indigenous and nonindigenous population. *Applied Neuropsychology*, *11*, 188–195. doi:10.1207/s15324826an1104_3.

Park, L. Q., Gross, A. L., McLaren, D. G., Pa, J., Johnson, J. K., & Mitchell, M., & Alzheimer’s Disease Neuroimaging Initiative. (2012). Confirmatory factor analysis of the ADNI Neuropsychological Battery. *Brain Imaging and Behavior*, *6*, 528–539. doi:10.1007/s11682-012-9190-3.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1994). *The Test of Everyday Attention*. Bury St. Edmunds: Thames Valley Test Company.

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society*, *2*, 525–534.

Ross, T. P., Calhoun, E., Cox, T., Wenner, C., Kono, W., & Pleasant, M. (2007). The reliability and validity of qualitative scores for the Controlled Oral Word Association Test. *Archives of Clinical Neuropsychology*, *22*, 475–488. doi:10.1016/j.acn.2007.01.026.

Royall, D. R., Lauterbach, E. C., Cummings, J. L., Reeve, A., Rummans, T. A., Kaufer, D. I., et al. (2002). Executive control function: A review of its promise and challenges for clinical research. A report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences*, *14*, 377–405. doi:10.1176/jnp.14.4.377.

Salthouse, T. A. (2005). Relations between cognitive abilities and measures of executive functioning. *Neuropsychology*, *19*, 532–545. doi:10.1037/0894-4105.19.4.532.

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General*, *132*, 566–594. doi:10.1037/0096-3445.132.4.566.

Santos, N. C., Costa, P. S., Amorim, L., Moreira, P. S., Cunha, P., Cotter, J., et al. (2015). Exploring the factor structure of neurocognitive measures in older individuals. *PLoS One*, *10*, e0124229. doi:10.1371/journal.pone.0124229.

Savla, G. N., Twamley, E. W., Delis, D. C., Roesch, S. C., Jeste, D. V., & Palmer, B. W. (2012). Dimensions of executive functioning in schizophrenia and their relationship with processing speed. *Schizophrenia Bulletin*, *38*, 760–768. doi:10.1093/schbul/sbq149.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B*, *298*, 199–209.

Shao, Z., Janse, E., Visser, K., & Meyer, A. S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, *5*, 772. doi:10.3389/fpsyg.2014.00772.

Siedlecki, K. L., Honig, L. S., & Stern, Y. (2008). Exploring the structure of a neuropsychological battery across healthy elders and those with questionable dementia and Alzheimer's disease. *Neuropsychology*, *22*, 400–411. doi:10.1037/0894-4105.22.3.400.

Simões, M. R., Albuquerque, C. P., Pinho, M. S., Vilar, M., Pereira, M., Lopes, A. F., et al. (2016). *Bateria de Avaliação Neuropsicológica de Coimbra (BANC) [Coimbra Neuropsychological Assessment Battery]*. Lisboa: CEGOC-TEA.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology*, *7*, 268–288.

Stinnett, T. A., Oehler-Stinnett, J., Fuqua, D. R., & Palmer, L. S. (2002). Examination of the underlying structure of the nepsy: A developmental neuropsychological assessment. *Journal of Psychoeducational Assessment*, *20*, 66–82. doi:10.1177/073428290202000105.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662.

Stuss, D. T. (2011). Functions of the frontal lobes: Relation to executive functions. *Journal of the International Neuropsychological Society*, *17*, 759–765. doi:10.1017/S1355617711000695.

Stuss, D. T., & Alexander, M. P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*, 901–915. doi:10.1098/rstb.2007.2096.

Takashima, A., Bakker, I., van Hell, J. G., Janzen, G., & McQueen, J. M. (2017). Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. *Brain and Language*, *167*, 44–60. doi:10.1016/j.bandl.2016.05.009.

Testa, R., Bennett, P., & Ponsford, J. (2012). Factor analysis of nineteen executive function tests in a healthy adult population. *Archives of Clinical Neuropsychology*, *27*, 213–224. doi:10.1093/arclin/acr112.

Toepper, M., Beblo, T., Thomas, C., & Driessen, M. (2008). Early detection of Alzheimer's disease: A new working memory paradigm. *International Journal of Geriatric Psychiatry*, *23*, 272–278. doi:10.1002/gps.1873.

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *The Stroop Neuropsychological Screening Test*. Odessa, FL: Psychological Assessment Resources.

Usai, M. C., Viterbori, P., Traverso, L., & De Franchis, V. (2014). Latent structure of executive function in five- and six-year-old children: A longitudinal study. *European Journal of Developmental Psychology*, *11*, 447–462. doi:10.1080/17405629.2013.840578.

van den Berg, E., Jiskoot, L. C., Grosveld, M. J. H., van Swieten, J. C., & Papma, J. M. (2017). Qualitative assessment of verbal fluency performance in frontotemporal dementia. *Dementia and Geriatric Cognitive Disorders*, *44*, 35–44. doi:10.1159/000477538.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*, 427–449. doi:10.1016/j.intell.2006.09.001.

van der Ven, S. H., Kroesbergen, E. H., Boom, J., & Leseman, P. P. (2013). The structure of executive functions in children: A closer examination of inhibition, shifting, and updating. *British Journal of Developmental Psychology*, *31*, 70–87. doi:10.1111/j.2044-835X.2012.02079.x.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Whiteside, D. M., Kealey, T., Semla, M., Luu, H., Rice, L., Basso, M. R., et al. (2016). Verbal fluency: Language or executive function measure? *Applied Neuropsychology: Adult*, *23*, 29–34. doi:10.1080/23279095.2015.1004574.

Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., van Aert, R. C., & van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-Hacking. *Frontiers in Psychology*, *7*, 1832. doi:10.3389/fpsyg.2016.01832.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*, 575–587. doi:10.1037/0012-1649.44.2.575.

Wood, S. J., Pantelis, C., Proffitt, T., Phillips, L. J., Stuart, G. W., Buchanan, J. A., et al. (2003). Spatial working memory ability is a marker of risk-for-psychosis. *Psychological Medicine*, *33*, 1239–1247. doi:10.1017/S0033291703008067.