



**UNIVERSITY OF PADOVA**

**Department of General Psychology\***

**Master's degree in Cognitive Neuroscience and Clinical Neuropsychology**

**Final dissertation**

**Attention and Executive Functions in Digitalized  
Neuropsychological Assessments**

*Supervisor*

**Professor: Mario Bonato**

*Co-supervisor: Vincenzo Livoti & Matteo Vitacca*

*Candidate:*

*Student ID number: 2121861*

Academic Year 2024/2025

<b>ABSTRACT</b> .....	<b>4</b>
<b>CHAPTER I: INTRODUCTION</b> .....	<b>6</b>
<b>CHAPTER II: FOUNDATIONAL FRAMEWORK OF ATTENTION AND EXECUTIVE FUNCTIONS</b> .....	<b>9</b>
<b>CHAPTER III: ATTENTION AND EXECUTIVE FUNCTIONS WORKING TOGETHER</b> .....	<b>14</b>
<b>THE ROLE OF ATTENTION AND EXECUTIVE FUNCTIONS IN NEUROPSYCHOLOGICAL ASSESSMENTS OF     NEUROLOGICAL AND PSYCHIATRIC DISORDERS</b> .....	<b>15</b>
<b>CHAPTER IV: ATTENTION AND EXECUTIVE FUNCTIONS IN PAPER-PENCIL ASSESSMENTS</b> .....	<b>19</b>
<b>CHAPTER V: METHODOLOGY</b> .....	<b>25</b>
<b>DATA COLLECTION AND ANALYSIS</b> .....	<b>26</b>
<b>STRATEGY FOR DATA SYNTHESIS</b> .....	<b>26</b>
<b>ANALYSIS OF SUBGROUPS OR SUBTESTS</b> .....	<b>26</b>
<b>SEARCH PROCESS AND PRISMA ORGANIZATION</b> .....	<b>27</b>
<b>DATA EXTRACTION VARIABLES</b> .....	<b>28</b>
<b>CHAPTER VI: RESULTS</b> .....	<b>29</b>
<b>MOST USED ASSESSMENTS AND COGNITIVE FUNCTIONS ASSESSED</b> .....	<b>29</b>
<b>NORMATIVE DATA AND PSYCHOMETRIC PROPERTIES</b> .....	<b>31</b>
<b>TARGET POPULATION</b> .....	<b>31</b>
<b>ADMINISTRATION, PRESENTATION &amp; ACCESSIBILITY</b> .....	<b>31</b>
<b>CHAPTER VII: DISCUSSION</b> .....	<b>32</b>
<b>CHAPTER VIII: ATTENTION AND EXECUTIVE FUNCTIONS IN DIGITALIZED ASSESSMENTS</b> .....	<b>33</b>
<b>COGSTATE COMPUTERIZED BATTERY (CCB)</b> .....	<b>43</b>
<b>NEUROCOGNITIVE PERFORMANCE TEST (NCPT)</b> .....	<b>55</b>
<b>THE BOSTON COGNITIVE ASSESSMENT (BOCA)</b> .....	<b>59</b>
<b>LIMITATIONS OF DIGITALIZED NEUROPSYCHOLOGICAL ASSESSMENTS</b> .....	<b>63</b>
<b>CHAPTER IX: FUTURE DIRECTIONS</b> .....	<b>67</b>
<b>CHAPTER X: CONCLUSION</b> .....	<b>68</b>
<b>REFERENCES:</b> .....	<b>70</b>

### List of Figures

Figure 1: Diagram based on PRISMA method .....	27
Figure 2: Distribution of Cognitive Functions Assessed .....	30
Figure 3: Psychometric Properties Measured .....	31

## Abstract

Attention and executive functions are critical cognitive processes that enable goal-directed behavior, problem-solving, and adaptive decision-making. These functions are vulnerable to decline across lifespans, with significant impacts observed in normal aging, neurological diseases such as Alzheimer's disease (AD), and psychiatric disorders like schizophrenia and major depressive disorder. Impairments in attention and executive functions not only hinder daily performance but also serve as early indicators of cognitive decline and disease progression. To assess these cognitive processes, neuropsychologists administer neuropsychological paper-pencil assessments to: point out cognitive dysfunction, aid in diagnosis, distinguish strengths and weaknesses, and give recommendations regarding rehabilitation and adjustment (Casaletto & Heaton, 2017). Despite their great impact in the field of neuropsychology, paper and pencil tests come with their drawbacks, like the interpretation of scoring, time-consuming data documentation, norms that do not get updated for decades, practice effects, and delayed diagnosis after the patient has already started to complain about symptoms (Gates et al., 2017). To compensate for the limitations of paper-pencil assessments, digitalized neuropsychological assessments are starting to replace traditional tests with the promises of earlier detection of cognitive deficits related to neurological diseases, psychiatric disorders, neurodevelopmental disorders and other disease that affects the brain. They also offer longitudinal monitoring of cognitive deficits without practice effects, accurate data capturing in tasks (like those that involve reaction times), and automatic documentation and interpretation of scores. This research is a systematic review that aims to identify the most used digitalized neuropsychological assessments. It will focus on the differences between paper-pencil assessments and digitalized assessments in testing attention and executive functions (including advantages and limitations), and discussing reliability and validity of the digitalized assessments used like Cambridge

Neuropsychological Test (CANTAB), Neurocognitive Performance Test (NCPT), NIH Toolbox (NIHTB), Cogstate Computerized Battery (CCB), and Boston Cognitive Assessment (BoCA).

## Chapter I: Introduction

Cognitive efficiency is essential for enabling humans to interact effectively with their environment. However, neurological diseases and psychiatric disorders hinder cognitive abilities in different ways, resulting in different neuropsychological profiles. According to the DSM-5, there are 6 cognitive domains that decline from standard functioning in neurological diseases and psychiatric disorders: language, perceptual-motor function, learning and memory, social cognition, complex attention, and executive functioning (American Psychiatric Association,2013). Among those core cognitive domains, attention and executive functions play an important role in supporting adaptive behavior. Attention refers to the ability to selectively focus on relevant stimuli while filtering out distractions, ensuring efficient cognitive processing (Sternberg,2012). Executive functions, on the other hand, encompass higher-order processes such as working memory (WM), inhibitory control, cognitive flexibility, planning, and problem solving, which are necessary for goal-directed behavior (Rabinovici, Stephens & Possin,2015). While attention provides the foundation for executive control, executive functions regulate and sustain attention across tasks, causing these two domains to be intricately connected. The connection between those two cognitive abilities can be disrupted due several factors including, normal aging, neurological diseases, psychiatric disorders, metabolic disorders, or traumatic brain injuries.

Even though certain neurological diseases are characterized by specific symptoms, attention and executive functions still play a role in the manifestations of those diseases. In neurological disorders like Alzheimer's disease, for example, memory is considered to be the prominent symptom. However, defects in attention and executive functions impede the process of attending, storing, and retrieving relevant information to be stored in memory (Hodges,2013). In frontotemporal dementia, deficits in executive functions affect the regulation of appropriate behavior and the process of decision making (Bang, Spina &

Miller,2015). This is also evident in psychiatric disorders like schizophrenia and major depressive disorder; however, the extent and manifestation of impairments vary. To understand the deficits in the different neuropsychological profiles, neuropsychologists administer paper-pencil assessments that typically involve single or multiple batteries; those batteries address 1-5 of the cognitive domains (Parsons & Hammeke,2014). The main purpose of those assessments is to identify cognitive deficits and aid in differential diagnosis, distinguish strengths and weaknesses in cognitive ability, monitor deficit progression, recommend rehabilitation, and test the effectiveness of treatment (Casaletto & Heaton,2017). To reach this goal assessments have to be equipped with specific criteria.

For neuropsychological assessments to fulfill their purpose, the assessments need to be sensitive to brain-behavior relationships, feasible across different settings, accurately measure what they are supposed to measure, and be consistent across time (Casaletto & Heaton,2017). The process of administering a neuropsychological paper-pencil assessment starts by either choosing a hypothesis-driven approach based on a deficit mentioned by the patient, or a comprehensive approach that targets all cognitive abilities (Parsons & Hammeke,2014). The comprehensive approach allows the distinguishing of every possible pathology that a patient might have. This is important in the cognitive domain of executive functions, where a deficit may be hidden on one task but not another (Parsons & Hammeke,2014). The score of the assessment is then scored based on correct responses. The score is then compared with a set of scores from a group of people that share the same normative data of sex, age, and education. From the data, cutoffs are created to establish validity, specificity, and sensitivity, meaning that the assessment is able to measure cognitive domains that it is supposed to measure, and accurately identify patients with cognitive impairment and patients without cognitive impairment (Hogan,2013). All of these criteria are important in the evaluation of newly established assessments.

Despite their common use, the drawbacks of paper-pencil assessments have been gaining more attention. In a meta-analysis conducted by Lang et al. (2017), it was reported that 60% of dementia was undetected due to the lack of accessible and cost-effective assessments that can detect subtle impairment. Also, the tests that are commonly used to assess attention and executive functions are not dynamic enough to reflect impairments that can happen in everyday life (Lauriane et al.,2020). Moreover, the controlled environment in which paper-pencil assessments are usually administered hinders diagnoses because of outpatient clinics procedures, where people usually go when they have already experienced symptoms of cognitive deficits, and increase practice effects (Lunardini et al.,2020). Furthermore, the administration, scoring, and interpretation consume time from the healthcare provider. For those reasons and more that will be discussed below, it is necessary for neuropsychological assessments to be: easily accessible, regularly updating normative data, precisely measuring cognitive impairments, automatically scored and documented, and decreasing the susceptibility to practice effects. Researchers now see that digitalized assessments can compensate for those disadvantages and match everyday functioning (Biorngirim et al.,2019).

Although digital assessments have existed for decades, the shift toward their widespread adoption has been largely driven by the increasing accessibility of smartphones and other portable devices. As a result, a wide range of samples can be reached, the process of being assessed is not costly, and people do not have to wait to go to outpatient clinics when they feel subjective decline. Those assessments would accurately measure tasks that involve reaction times, hand movements, and processing speed (Chan et al.,2021). Moreover, digitalized assessments would provide long-term monitoring without patients having to go to healthcare facilities where the controlled environment could affect performance (Chan et al.,2021). Also, they would prevent the occurrence of subjective scoring for tasks where

standards are ambiguous. However, digitalized neuropsychological assessments for attention and executive functions are yet to be researched and understood further. As Latendorf et al. (2021) stated, digitalized versions of paper-pencil assessments might measure different aspects and hence need their own normative data.

This systematic review will focus on the cognitive domains of attention and executive functions, highlighting their interdependence and critical role in supporting adaptive behavior. It will identify the most commonly used neuropsychological assessments, examine the significance of accurately assessing these domains, and discuss traditional paper-and-pencil assessments commonly employed in neuropsychological evaluations, along with their inherent limitations, such as susceptibility to examiner bias and sensitivity to cognitive impairment. Additionally, the review will explore several digitalized assessment tools, including, among other, two of the most widely cited the Cambridge Neuropsychological Test (CANTAB) and Cogstate, among others. These tools represent a shift toward more innovative and accessible approaches, offering advantages such as enhanced precision, scalability, and automated data analysis. Furthermore, their limitations will be critically examined. This systematic review aims to provide a comprehensive analysis of digital assessment methods, considering their respective benefits and limitations in the evaluation of attention and executive functions.

## **Chapter II: Foundational Framework of Attention and Executive Functions**

According to Sternberg (2012), when faced with a vast amount of information, attention enables humans to process the part of information that is considered important. This information can come through the senses, memories, or other cognitive processes. Through attention, one is able to judge which aspects of information should be highlighted and which

should not. Furthermore, it allows space to process information, work on it rapidly while being oriented to one's surroundings, and maintaining an alert state and focus (Sternberg,2012). Sternberg (2012) classified four types of attention: detection and vigilance, selective attention, search, and divided attention. First, through detection and vigilance, one can maintain attention when waiting for the identification of a stimulus or signal (Sternberg, 2012). Second, selective attention allows for the processing of specific stimuli while ignoring others (Sternberg,2012). A famous example of this phenomenon is the cocktail party situation where one tries to maintain a conversation with a person while ignoring other conversations from the surrounding crowd (Diamond,2014). The third type is search, which is trying to actively identify a stimulus amidst distractions like the famed illustrated puzzle book, "*Where is Waldo?*". Finally, divided attention allows for the allocation of attention to more than one task at a time (Sternberg,2012).

Attention is also divided based on how information reaches one's awareness: it can be bottom up which refers to alertness to stimuli that come from extrinsic factors, or it can be top-down which refers to alertness to stimuli that come from intrinsic factors (Nasiri & Hakimzadeh,2023). Researchers have also studied where those types of attention are processed in the brain anatomically. For example, it was proposed by Corbetta and Shulman in 2002 that the top-down system is modulated by the dorsal frontoparietal area to allocate attention to the environment, while the bottom-up system is modulated by the ventral frontoparietal area to trigger attention to unexpected stimuli. The dorsal network is composed of the intraparietal sulcus (IPS) and the frontal eye fields bilaterally (Vossel, Geng, & Fink,2014). On the other hand, the ventral network is composed of the ventral frontal cortex (VFC) and the temporoparietal junction (TPJ). These areas activate in response to sudden stimuli that occur unexpectedly (Vossel, Geng, & Fink,2014). According to Shulman (2003), during top-down processes, areas in the ventral network like the TPJ gets inhibited. This

mechanism is important during vigilant attention for the effectiveness of goal-directed behavior. Vossel et al. (2014) proposed that the process of allocating attention is an interplay between the two systems and that their interaction is flexible.

Attention is also believed to rely on a combination of networks that work together, not just isolated areas. That was evident when Posner, Rothbart & Sheese (2014) divided attention into three separate networks: alerting network, orienting network, and executive control. The alerting network is triggered by expected, forthcoming stimuli (Posner, Rothbart & Sheese,2014). It is controlled by the norepinephrine systems that involve the frontal and parietal cortex (Posner, Rothbart & Sheese,2014). The orienting network interacts with the sensory environment, allowing for the prioritization of tasks (Posner, Rothbart & Sheese,2014). Finally, the executive control is responsible for resolving conflict between competing stimuli. This conflict is resolved by inhibiting distractors and prioritizing stimuli related to one's goals. It involves the anterior cingulate cortex, anterior insula, and parts of the mid prefrontal cortex (Posner, Rothbart & Sheese,2014).

The executive control model introduced by Posner (2014) explains the intricate interplay between attention and executive functions, highlighting how these domains work together to support goal-directed behavior. This is an important point in understanding the significance of having neuropsychological tests that assess them together accurately. Furthermore, this critical relationship will be explored further in subsequent sections, particularly in the context of neuropsychological assessments.

Baddeley and Hitch (1974) first labeled executive functions as “central executive” and then they were described later by Lezak (1983) as a way to understand how behavior is expressed (Jurado & Rosselli,2007). Executive functions are considered to be a collection of cognitive abilities that are responsible for goal-directed behavior and are important in

adapting to one's fluctuating environment (Rabinovici, Stephens & Possin,2015). They are divided into: working memory, cognitive flexibility, inhibition, planning, reasoning, and problem solving (Cristofori, Grafman, & Zimmerman,2019). Using executive functions is effortful, and consequently, one must also use attention to equip abilities like planning and problem solving instead of getting into an "autopilot" mode. Luria (1973) also proposed that the process of using executive functions depends on three phases: anticipation, planning, and execution and self-monitoring. He also emphasized the role of the frontal lobes in the management of all of these aspects (Cristofori, Grafman, & Zimmerman,2019).

According to Judo & Rosselli (2007), the first executive function that appears in children is inhibitory control. Diamond (2013) described inhibitory control as one's ability to regulate attention, emotions, thoughts and behaviors to interact appropriately with the environment. Furthermore, without the ability to inhibit oneself, people become more prone to impulsiveness and conditioned stimuli that are not appropriate to the situation (Diamond,2013). Unlike animals, humans do not give away easy to instinct and are more in control of their behaviors. Cansino et al. (2011) stated that inhibition occurs along three phases: restraint, access, and deletion. Inhibition also coincides with selective attention in choosing what to perceive and what not to (Diamond,2013). Moreover, inhibitory control is used in information processing of mental stimuli along with working memory (Diamond,2013). Diamond (2013) also mentioned it is an important part of self-control when it comes to emotional regulation like inhibition of anger or delayed gratification. According to Cristofori, Grafman, & Zimmerman (2019), the process of inhibitory control is largely dependent on parts of the right hemisphere that includes the right frontal gyrus, anterior insula, inferior parietal lobe, and frontal limbic area.

Working memory is the cognitive function that allows one to manipulate and maintain information. According to Diamond (2014), "working memory is critical to one's ability to

see the connection between seemingly unrelated things and to pull apart elements to form an integrated whole” (Diamond, 2014). Moreover, Ku (2018) stated that working memory is the foundation of everyday behavior that includes learning, reasoning, and language comprehension. However, it has a limited capacity, in which: one can act on information that is no longer perceived, or it can either be lost or stored in long-term memory (Ku,2018). Diamond (2014) emphasized that working memory and inhibitory control always coexist together and rarely do they exist alone. When in a situation where one must hold specific information through working memory and inhibitory control, they suppress irrelevant information to make the relevant one guide behavior (Diamond,2014). Cristofori, Grafman, & Zimmerman (2019) were able to conclude from lesion mapping studies that the dorsolateral prefrontal cortex is responsible for WM representations, along with the orbitofrontal cortex, which is engaged when there is more than one cognitive process occurring.

Cognitive flexibility enables humans to generate new ideas based on their changing environments, which promotes individual growth (Cristofori, Grafman, & Zimmerman,2019). It is linked to inhibitory control when it comes to switching from one task and/or idea to another (Cristofori, Grafman, & Zimmerman,2019). The orbitofrontal cortex is considered to be an important area in modulating the alteration in behavior/idea when one is faced with an outcome different than what one had expected (Cristofori, Grafman, & Zimmerman,2019). Other areas implicated in cognitive flexibility are ventrolateral prefrontal cortex, dorsolateral prefrontal cortex, caudate and thalamus, however, researchers are still trying to understand how they work together (Cristofori, Grafman, & Zimmerman,2019). Planning comprises the generation, assessment, and selection of goals in order to perform a task which is mostly mediated by the dorsolateral prefrontal cortex (Cristofori, Grafman, & Zimmerman,2019).

According to Cristofori, Grafman, & Zimmerman (2019), reasoning is the process of abstraction that is involved in the formation of theories, creative ideas, and using analogies to understand the environment. They concluded from lesion and imaging studies that the rostolateral PFC modulates reasoning along with other subcortical regions. However, further research is still needed (Cristofori, Grafman, & Zimmerman,2019). Finally, problem solving entails addressing aspects of a problem in order to find its solution, involving all the cognitive processes previously discussed of the executive functions. A network between the frontal, temporal, and parietal lobes is included in this process as one needs to rely on other cognitive domains, like memory, in order to address a problem (Cristofori, Grafman, & Zimmerman,2019).

Executive functions serve as the central control system of cognition, orchestrating wide range of abilities that assist individuals to think, plan, and adapt to their surroundings. They regulate attention, memory, and problem-solving, ensuring that cognitive resources are allocated efficiently to meet the demands of a given task. Without executive functions, even basic cognitive skills would operate in a broken manner, lacking coordination and direction. By integrating information, suppressing distractions, and guiding behavior toward goal-directed actions, executive functions play a crucial role in everything from simple daily tasks to complex decision-making. Their influence extends across various domains, from emotional regulation to social interactions, making them fundamental to adaptive and flexible thinking in an ever-changing environment.

### **Chapter III: Attention and Executive Functions Working Together**

Attention and executive functions are linked cognitive processes that work together to enable goal-directed behavior and effective self-regulation (Diamond,2014). The executive network consists of regions such as the anterior cingulate cortex, anterior insula, midprefrontal cortex, and underlying striatum, playing a pivotal role in resolving competing

actions and regulating responses in tasks requiring conflict resolution (Posner,2014). This regulation enhances goal-relevant neural networks while suppressing conflicting ones, supported by long connections between executive network and cognitive-emotional areas of the brain, aiding voluntary control and self-regulation (Posner,2014).

For example, selective attention and working memory, two core cognitive functions, are highly interdependent. The limited capacity of working memory necessitates the allocation of attention to selected targets to sustain goal-directed behavior (Ku,2018). Selective attention allows individuals to focus on relevant aspects of the environment while ignoring irrelevant stimuli, consequently optimizing the efficiency of working memory processes (Ku,2018). This interaction is mediated by shared neural substrates, particularly within the prefrontal-parietal network, which supports both the maintenance of relevant information in working memory and the selective focus on external stimuli (Ku,2018).

In another example of how both can work together, attention can both enhance and hold back working memory performance (Ku,2018). On one hand, directing attention toward representations held in working memory can improve task outcomes (Ku,2018). On the other hand, external interference, such as distractions or interruptions, can compromise working memory by either requiring suppression of irrelevant stimuli (distraction) or necessitating a reallocation of cognitive resources to manage secondary tasks (interruption) (Ku,2018). These processes rely on top-down signals from the prefrontal cortex and the medial temporal lobes for effective management and recovery of disrupted representations (Ku,2018).

### **The Role of Attention and Executive Functions in Neuropsychological Assessments of Neurological and Psychiatric Disorders**

The importance of assessing attention and executive functions lies in the fact that there is a high prevalence of neurological diseases. The World Health Organization stated that the dementia population is expected to increase to 78 million by 2030 and 139 million by

2050 (World Health Organization [WHO],2021). With the rising prevalence of dementia, it is important to have neuropsychological assessments that are sensitive to subtle impairments. Especially those that are related to attention and executive functions, which are responsible for everyday goal-directed behavior. Mild Cognitive Impairment (MCI) is known to be the transitional stage from normal aging to dementia progression (Lee & Chan,2022). According to Lee & Chan (2022), 30% of patients with MCI progressed to dementia. Therefore, it is important to monitor the decline of attention and executive functions in order to understand disease progression. Indeed, several studies have been conducted in order to track how attention and executive functions affect disease progression and manifestation of symptoms.

According to Belleville, Chertkow, & Gauthier (2007), when a task of divided attention was administered to MCI participants through the Brown Peterson procedure, they were found to have slight impairment, which is considered a severe impairment in Alzheimer's participants. This indicates that the impairment in divided attention serves as a tracker for the progression from MCI to AD dementia. Also, Saunders & Summers (2010) found that patients with MCI had impairments in not only working memory capacity, but also attentional processing. Moreover, Cansino et al. (2011) emphasized that with aging, interference occurs in tasks because of impairment in inhibitory control that affects selective attention. These results highlight the fact that diseases characterized by certain symptoms possess other defects that trigger the worsening of those symptoms.

Even though Alzheimer's Disease is widely known for its initial memory impairments, Perry and Hodges (1999) confirmed that attention is the first cognitive domain to be affected along with the memory domain. Moreover, Guarino et al. (2019) stated that executive functions are affected in the early stages of the disease due to degeneration that occurs in the prefrontal cortex, especially inhibitory abilities. During activities of daily living, the ability to disengage from one task and shift attention to another is impaired (Guarino et

al.,2019). Consequently, other symptoms that are typical of AD occur, like irritability, mood swings and apathy (Guarino et al.,2019). Moreover, defects in attention and executive functions impede the process of attending, storing, and retrieving relevant information as a failure of attentional processing (Nasiri & Hakimzadeh,2023). It is important to note that the biological changes in AD occur up to 30 years before diagnosis, which sheds light on the importance of detecting subtle impairment in order to be able to stall the disease (Ding, Lee & Chan,2022). These findings are not only limited to AD. According to WHO (2023), the second most common type of dementia after AD is vascular dementia. Cognitive impairments in vascular dementia usually depend on the areas affected, yet O'Brian & Thomas (2015) state that frontal striatal circuits are disrupted, causing a significant impairment in information processing, attention and executive function.

Moreover, Gratwick, Jahanshahi & Foltynie (2015) found that impairments in executive functions are present in Parkinson's disease from the beginning of the diagnosis. Those impairments are manifested in concentration and planning that show up in daily activity and occupational life (Gratwick, Jahanshahi & Foltynie,2015). Also, executive functions have been found to be the prognostic sign of developing Parkinson's disease dementia (Gratwick, Jahanshahi & Foltynie,2015). Furthermore, after dopamine therapy, impairment of inhibitory control and increase of impulsiveness and stimulus-bound behaviors are present due to the increase of activity in the corticostriatal-thalamocortical pathway (Vriend et al.,2014). Frontotemporal dementia is characterized by initial degeneration in the frontal lobes which affects all higher cognitive functions, like planning, problem solving, abstraction, inhibition, and cognitive flexibility (Huey et al.,2009). These symptoms are always evident during neuropsychological assessments when patients display stimulus-bound behavior, preservation, and the inability to shift from one rule to another despite being learned (Wisconsin Card Sorting Task) (Huey et al.,2009). All of those symptoms are not

limited to only neurological diseases. They can also appear as symptoms of several other non-neurological causes.

Attention and executive functions are not only affected in neurological diseases but in psychiatric disorders as well. For example, schizophrenia is characterized by significant cognitive impairments that affect attention, inhibition and processing speed (Krabbendam & Jolles,2003). In a study where the Continuous Performance Test (CPT) was administered, patients with schizophrenia displayed impairment in sustained attention (Krabbendam & Jolles,2003). Moreover, this impairment was significant when there was a constant stimulus presentation or when previous stimuli had to be stored in working memory to be used in prospective tasks (Krabbendam & Jolles,2003). They also showed impairment in selective attention and inhibitory control on the Stroop task (Krabbendam & Jolles,2003). Furthermore, patients with schizophrenia also display perseverative behaviors and lack of cognitive flexibility on executive functions tasks like the Wisconsin Card Sorting Test (WCST) (Krabbendam & Jolles,2003). These abilities can also be affected in different ways depending on the type of psychiatric disorder.

Major Depressive Disorder is also characterized by deficits in attention and executive functions. This was evident in a study conducted Habermann, Pohl & Leplow (2005) where they administered three tests for attention and three tests for executive functions. They found that patients with major depressive disorder had deficits in planning, monitoring, and three attention domains. Those findings were discovered during a remission period, which highlights that those deficits are not state dependent (Habermann, Pohl & Leplow,2005). According to Parsons & Hammeke (2014), patients with depression are more likely to make errors of omission on attention tasks. However, bipolar disorder patients with manic episodes commit errors of commission and lack of response inhibition (Parsons & Hammeke,2014).

According to Marinescu et al. (2014), patients with bipolar disorder exhibit impairment in inhibitory control, working memory, cognitive flexibility, and mental manipulation, all of which are related to the recurrence of episodes and hospital admittance. Maalouf et al. (2010) emphasized the role of neuropsychological assessments in differentiating between bipolar depression and unipolar depression. Attention and executive functions deficits are also present in other psychiatric disorders like anxiety disorder, posttraumatic stress disorder and attention-deficit/hyperactivity disorder (ADHD) (Nasiri & Hakimzadeh, 2023). Moreover, Kim et al. (2014) emphasized that deficits in executive functions are predictors of treatment, prognosis and outcomes of those disorders, highlighting the importance of assessing them through neuropsychological assessments.

To conclude, attention and executive functions are critical cognitive domains that are implicated in numerous diseases, not just neurological and psychiatric. Their centrality to goal-directed behavior and daily functioning highlights the necessity of neuropsychological assessments sensitive enough to detect subtle impairment. With the rising prevalence of conditions such as Alzheimer's disease, Parkinson's disease, and psychiatric disorders like schizophrenia and bipolar disorder, accurately identifying early cognitive deficits in attention and executive functions is crucial. These assessments not only aid in understanding disease progression, but also contribute to effective diagnosis, prognosis, and treatment planning, highlighting their necessary role in clinical and research settings.

#### **Chapter IV: Attention and Executive Functions in Paper-Pencil Assessments**

Paper-pencil neuropsychological assessments usually consist of subscales that measure several cognitive abilities from different domains. Those assessments are either used based on a hypothesis approach depending on the patient's history and current symptoms, or in a comprehensive way to understand a patients' strengths and weaknesses (Parsons &

Hammeke,2014). The scores are based on correct responses which are then compared to a cut-off score. The cut-off score is created to establish sensitivity for the assessment so that it can detect subtle impairment (Hogan,2013). Moreover, cutoffs aid in specificity that helps distinguish patients who have a pathology from those who do not (Hogan, 2013). Normative data are then used to compare the patients' score with a group that shares the demographic variables like age, sex, and education (Hogan,2013). A robust neuropsychological assessment needs to have high validity, meaning it measures what it is supposed to measure, and high reliability, meaning that its results are consistent across time (Hogan,2013). For the purpose of this review, this chapter will focus on the limitations posed by traditional paper-pencil assessments.

Paper-pencil assessments have several limitations that make their use debatable. These limitations are important to highlight so that clinicians and researchers can take appropriate precautions when using them. It has been recognized that paper-and-pencil assessments primarily focus on memory domains, leading to a lack of awareness among patients and their caregivers about forms of dementia involving other types of impairments. Some paper-pencil assessments do not adequately assess executive functions, which are critical for understanding impairments in various neurological and psychiatric conditions (Pardo,2014). For example, disorders characterized by frontal lobe dysfunction, such as frontotemporal dementia, schizophrenia, or bipolar disorder, may not be accurately represented due to the lack of sensitivity to executive impairments. Other paper-pencil assessments like the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) do not have subscales that directly assess executive functions just indirectly, which cannot be quantified or used as a measure (Spencer, Andern & Tolle,2018). Additionally, the limited structure of such assessments often impacts their discriminant validity, reducing their ability to effectively distinguish between conditions like mild cognitive impairment (MCI)

and dementia (Pardo,2014). Despite this limited structure, the standardization of those tests is subject to clinicians' or researchers' subjectivity (Pardo,2014). Moreover, Zelazo et al. (2014) highlighted the fact that existing measures do not tackle lifespan progression of attention and executive functions. Particularly, those that are related to structural changes like the prefrontal cortex (PFC) (Zelazo et al.,2014).

Pardo (2014) highlighted that a significant limitation of paper-pencil assessments is a lack of standardization, which can undermine their reliability. When compared to other neuropsychological assessments over time, not all domains show consistent performance. For example, Truong et al. (2024) examined the validity of the Mini-Mental State Examination (MMSE) through network analysis to assess how well it reflects cognitive performance. Particularly, when compared to comprehensive neuropsychological tests over time. The only domain that was consistent with other neuropsychological assessments over time was the memory domain (Truong et al.,2024). However, attention, calculation, and orientation did not have consistent relationships with other neuropsychological assessments over time (Truong et al.,2024). The reduced validity can affect the longitudinal monitoring of those cognitive abilities. Moreover, this limitation will also manifest in psychiatric populations where cognitive impairments are known to fluctuate. Truong et al. (2024) also emphasized that since the domains in the test contain very few items, they do not capture the complexities of cognitive abilities. As a result, this will be reflected not only in neurological diseases, but in other illnesses like psychiatric disorders. Indeed, Balconi et al. (2022) highlighted that, in the case of psychiatric disorders like addiction, executive functions are overlooked. This also motivated researchers to understand the faults with how attention and executive functions are measured.

Marinez-Pernia et al. (2023) highlighted in their review a set of crucial problems with how executive functions are measured. The first issue was content evaluation, where the

focus is on one or more of the subdomains of executive functions, whereas other subdomains are ignored (Martinez-Pernia et al.,2023). According to Martinez-Pernia et al. (2023), those ignored subdomains are multitasking, contextual control and prospective memory.

Furthermore, they emphasized that there are aspects that are not typically defined as part of executive functions, like emotional regulation and social cognition (Martinez-Pernia et al.,2023). These aspects are usually considered in isolation. Consequently, paper-pencil assessments do not capture real-life situations where there is a combination of subdomains (Martinez-Pernia et al.,2023). The second issue has to do with the specified structure of the assessments, where certain appropriate responses are only allowed (Martinez-Pernia et al., 2023). Martinez-Pernia (2023) highlighted that there should be assessments where the patient evaluates their own performance and if it is suited for real-life situations (Martinez-Pernia et al.,2023). These problems are also evident in the assessments used to measure executive functions, which affect their ability to differentiate between conditions.

Tasks designed to assess planning, organization, attention, memory, and visuospatial skills may not be sufficient for accurately differentiating between various cognitive impairments (Price et al.,2011). For example, in The Clock Drawing Test, performing the task based only on command is not enough to differentiate between types of dementia. Indeed, patients with deficits in executive functions show impairment in the tasks of command and copy conditions, whereas those without deficits do not (Price at al.,2011). Evidence for that was found in a study conducted by Price et al. (2011), where patients with AD showed impairment when they had to recall how to draw the clock and set the time. However, in the copy condition they did not show impairment. On the other hand, patients with Parkinson's disease dementia and Vascular dementia transferred errors from the command condition to the copy one, demonstrating impairments in executive functions (Price at al.,2011). Price et al. (2011) added that the clock drawing task lacks interscorer reliability and that the criterion

for scoring the performance is open to interpretation and subjectivity. Laurine et al. (2020) also highlighted that the task does not take into account the rising use of digital clocks as technology becomes more common.

Besides differentiating between different dementias, neuropsychological assessments need to distinguish between those who have cognitive impairments and those who do not. However, according to Mitchell (2009), paper-pencil assessments are adequate at confirming diagnosis but are not adequate for excluding those who do not have them. Another limitation of paper-pencil assessment is that they lack ecological validity and do not capture daily life functions (Chan et al.,2008). The lack of ecological validity results from the target of each cognitive function at a time, whereas daily life requires the combination of multiple cognitive functions (Chan et al.,2008). Furthermore, the lack of ecological validity hinders the detection of patients with frontal lobe syndrome as they have been found to perform as good as controls on standardized assessments (Chan et al.,2008). According to Chan et al. (2008) most paper-pencil assessments that target executive functions and attention have low test-retest reliability.

Dautzenberg & Beekman (2020) warned about using cut-off scores that are not ecologically valid as they may overestimate assessments' accuracy. They conducted a study in 2020 to evaluate the utility of paper-pencil assessments to distinguish between patients with mild cognitive impairment (MCI), major dementia (MD), and normal cognitive aging (NoCI). The authors recruited NoCI patients - individuals referred for assessment but without cognitive impairment - rather than healthy controls (HC) to ensure a realistic evaluation of the assessment (Dautzenberg & Beekman,2020). Dautzenberg & Beekmann (2020) found that the paper-pencil assessments were less effective in real-world settings compared to idealized case-control studies. The authors demonstrated that case-control studies contrasting extreme groups, such as HC versus MCI or MD, tend to inflate specificity and area under the curve

(AUC) measures (Dautzenberg & Beekman, 2020). They emphasized that in practical scenarios, comparisons like NoCI versus MCI or MD reveal reduced specificity and AUC, necessitating adjustments to cutoff scores for accurate diagnosis (Dautzenberg & Beekman, 2020).

Other than how executive functions and attention are measured, the means of administration affect the validity of the scores as well. According to Lang et al. (2016), the controlled environment where the assessments are administered delays the diagnosis because patients are only assessed after they have displayed symptoms. Also, the administration, scoring, interpretation, and documentation consume time from the healthcare provider and involve their subjectivity. (Amanzadeh et al., 2020). Paper-pencil assessments do not allow for longitudinal monitoring because they are not equipped for weekly or monthly evaluations (Vyshedskiy et al., 2022). As such, the variability in standardized assessments is limited, which often leads to practice effects (Vyshedskiy et al., 2022). There is a need for assessments that can monitor cognitive decline regularly and do not require patients to be frequently present in a medical environment (Ding et al., 2024). Clinicians and researchers should put into consideration that paper-pencil assessments depend on tasks that assume a standard level of literacy and cultural familiarity, causing populations with low educational backgrounds or from diverse cultural settings to be excluded (Pardo, 2014). This can result in systematic biases, leading to underdiagnosis or overdiagnosis in these groups and limiting access to appropriate care and clinical trials (Pardo, 2014).

As cognitive impairments increasingly require nuanced and dynamic diagnostic approaches, it becomes evident that paper-pencil assessments alone are insufficient. Their limitations underscore the need for updated, flexible, and ecologically valid tools that can provide more comprehensive insights into cognitive functioning to enable effective, individual care. A shift towards digitalized assessments may introduce an understanding of

how to measure attention and executive functions in a way that allows for: longitudinal monitoring, differentiation between different cognitive profiles, and some correspondence to everyday life. these aspects are going to be explored in the coming sections to investigate if digitalized assessments do really compensate for the limitation highlighted in paper-pencil assessments.

## **Chapter V: Methodology**

According to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, a systematic review was conducted to explore the digitalized assessments that are available for the assessments of attention and executive functions. Moreover, it was conducted to examine what are their main psychometric properties, their accessibility, their eligibility for self-administration, and their target measures. The leading keywords in the conductance of the systematic review were: Attention, Cognitive assessment, Computer-based, Digital, Executive functions, and Tele-neuropsychology. With those keywords, a search process was conducted within three databases : PubMed, Scopus, and Embase. Both observational studies (cohort, cross-sectional, and case-control), and experimental studies (randomized control studies) were included. Per se, those studies focused on neuropsychological deficits and assessment of attention and executive functions.

The inclusion criteria were: studies involving participants aged 18 years or older, the use of validated digital tools, and assessments that focus on attention and executive functions. On the other hand, the exclusion criteria comprised: assessments that do not evaluate attention and executive functions, tools developed primarily for rehabilitation purposes, and telephone-based assessments. Clinical validity, data accessibility, and availability of validated digital tools were the main focus of this systematic review. This presents a great importance since there is a shift towards digitalized assessments and emerging literature that highlight the potential of their use to make up for drawbacks that have been highlighted in paper-and-

pencil assessments. Consequently, the main outcomes for this review were: psychometric properties of the tools, availability and accessibility of normative data, software licensing or hardware costs, and user experience and sustainability of the assessment methods.

### **Data Collection and Analysis**

For the study selection, two independent reviewers screened titles and abstracts according to the previously mentioned inclusion criteria. After that, full texts of relevant studies were assessed for inclusion. A third reviewer was responsible for resolving any discrepancies and disagreements. Moreover, the selection process was documented and managed using digital spreadsheets which added transparency.

Two independent reviewers were involved in the process of data extraction, along with a third one to verify the extracted data and resolve disagreements. A digitalized spreadsheet was used to collect and organize the following data: psychometric information of each tool, openness of the tool, cognitive functions assessed, number and types of digital tool identified, languages used for each tool, devices used for administration, administration modalities, and duration. Two independent raters assessed the quality and risk of bias on the included studies using the Joanna Briggs Institute Critical Appraisal Tools (JBI), which included checklists fit for each study type.

### **Strategy for Data Synthesis**

A narrative, qualitative synthesis of the extracted data was conducted. Furthermore, the digital assessments were categorized based on: specific assessments of attention and executive functions versus general batteries that include these domains, psychometric characteristics of these tools, if the study used validated tools or was involved in the validation on normative samples, and administration information.

### **Analysis of Subgroups or Subtests**

The included digital assessments were classified further according to: year of

publication, languages available, device required (e.g., tablet or PC), administration type (remote, in-person, hybrid, self-administration), and access modalities (e.g., registration/login required, software, download required).

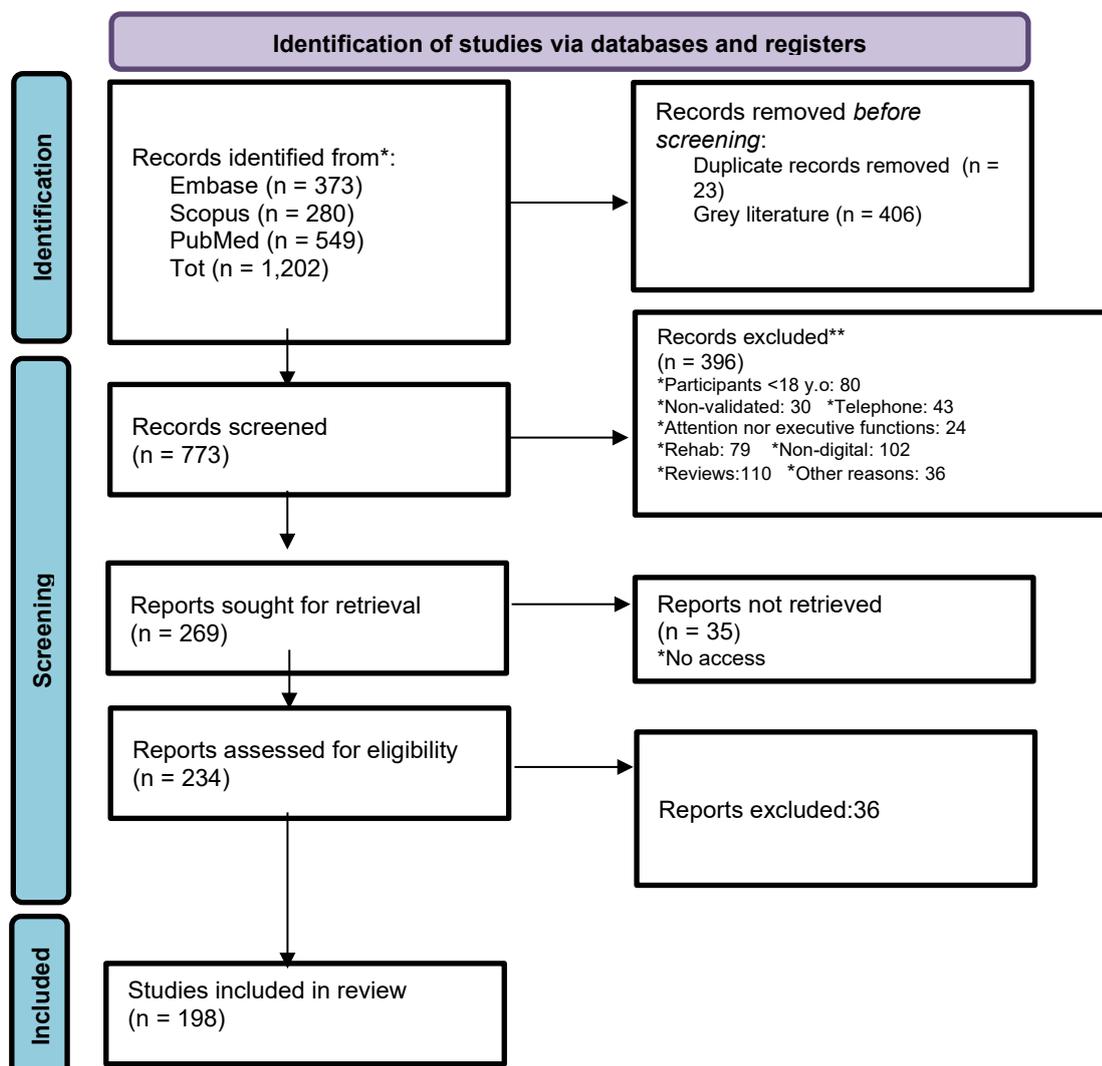
### Search Process and PRISMA Organization

As mentioned previously specific keywords were used in the search process: Attention, Cognitive assessment, Computer-based, Digital, Executive functions, and Tele-neuropsychology. With those keywords, 1,202 articles were identified:

1. Embase: 373 articles identified.
2. Scopus: 280 articles identified.
3. PubMed: 549 articles identified.

**Figure 1**

*Diagram Based on PRISMA Method*



The PRISMA diagram shows the process of elimination and the attainment of the targeted articles. The process started off with 1,202 articles. From them: 23 were duplicates, 406 were methodological, 110 were reviews, and 384 were removed for meeting exclusion criteria. This led to 270 articles that underwent further screening. Of those 270, 38 were removed for fitting the exclusion criteria. Moreover, 34 articles were not included as they had no access. Consequently, 198 articles were sought for retrieval, assessed for eligibility, and are included in this review.

### **Data Extraction Variables**

A set of variables were specified to make a proper identification of targeted articles. Moreover, they were essential in understanding the constituents of the digitalized assessments that target attention and executive functions. Those variables included:

- **Targeted population:** whether they focused on certain type of diseases or healthy people. If they were just healthy populations, it was determined what kind of healthy. For example, students, old adults, athletes, or workers.
- **Cognitive functions assessed:** it was determined if the assessments used were for attention and executive functions, and what aspects of attention and executive functions did they address.
- **Device:** computers, smartphones, tablets, or virtual reality (VR).
- **Administration modality:** remote-supervised, remote-unsupervised, or in-person.
- **Psychometric properties:** if articles reported any form of validity or reliability assessments to support the appropriate and accurate use of the digitalized assessments.
- **Accessibility:** if the assessments are available for free use or have to be subscribed/paid for.
- **Normative data:** if any of the articles using the digital assessments included normative data to support their applicability and interpretability.

## Chapter VI: Results

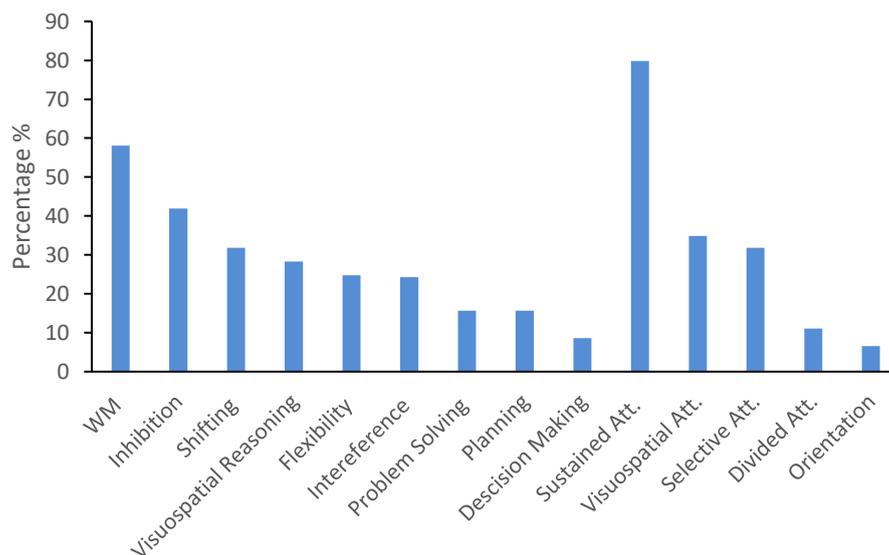
In accordance with PRISMA 2020 guidelines, the systematic review process yielded a comprehensive set of studies addressing or using digital assessments of attention and executive functions. Following an extensive search across three major databases -PubMed, Scopus, and Embase – a total of 1,202 records were initially identified. After duplicate removal and screening for relevance based on the predefined inclusion and exclusion criteria, 198 studies were deemed eligible and included for the analysis. The results presented below reflect a qualitative synthesis of these studies, focusing on key variables such as the most used assessments, assessments that designed specifically for attention and executive functions, psychometric properties that were measured for these assessments, occurrence of validation, presence of normative data, administration modality, and types of attention and executive functions assessed.

### **Most used Assessments and Cognitive functions assessed**

Based on the articles evaluated, it was found that the most used digital neuropsychological assessments were: NeuroTrax ( n= 22), Cambridge Neuropsychological Test Automated Battery (CANTAB, n= 13), Cognitive Drug Research ( CDR, n=13), Cogstate (n=13), and Automated Neuropsychological Assessment Metrics (ANAM, n=10). Other tools used were the Central Nervous System Vital signs (n=8), Wisconsin’s Card Sorting Test (n=7), Cognitron (n=6), and the NIH Toolbox Cognitive Battery (n=5). Collectively, they constituted 97 articles out of the whole data set which accounted for 48.9% of the articles. A total of 86 digitalized assessments were used and among them, only 6 assessments were specifically designed to measure attention and executive functions: CANTAB, Delis-Kaplan Executive Function System (D-KEFS, n=1) , WCST (n=7), Test of Variables of Attention (TOVA, n=2), Paced Auditory Serial Addition Test (PASAT, n= 1), and TAP (n=2).

Data extraction also involved identifying the specific types of attention and executive functions assessed across studies. Executive functions were mentioned to evaluated generally in 83.84% of the articles ( $n = 166$ ), with the most commonly assessed domain being working memory (58%,  $n = 115$ ). This was followed by inhibition (41.92%,  $n = 83$ ), shifting (31.82%,  $n = 63$ ), visuospatial reasoning (28.28%,  $n = 56$ ), flexibility (24.75%,  $n = 49$ ), interference control ( $n = 48$ ), problem solving (15.66%,  $n = 31$ ), planning (15.66%,  $n = 31$ ), and decision-making (8.59%,  $n = 17$ ). Regarding attention, sustained attention was the most frequently assessed subdomain, appearing in 79.80% of the articles ( $n = 158$ ). This is because most articles either describe attention broadly as sustained attention or do not specify the type of attention being measured. Other forms of attention were evaluated less frequently, including visuospatial attention (34.85%,  $n = 69$ ), selective attention (31.82%,  $n = 63$ ), divided attention (11.11%,  $n = 22$ ), and orientation (6.57%,  $n = 13$ ).

**Figure 2**



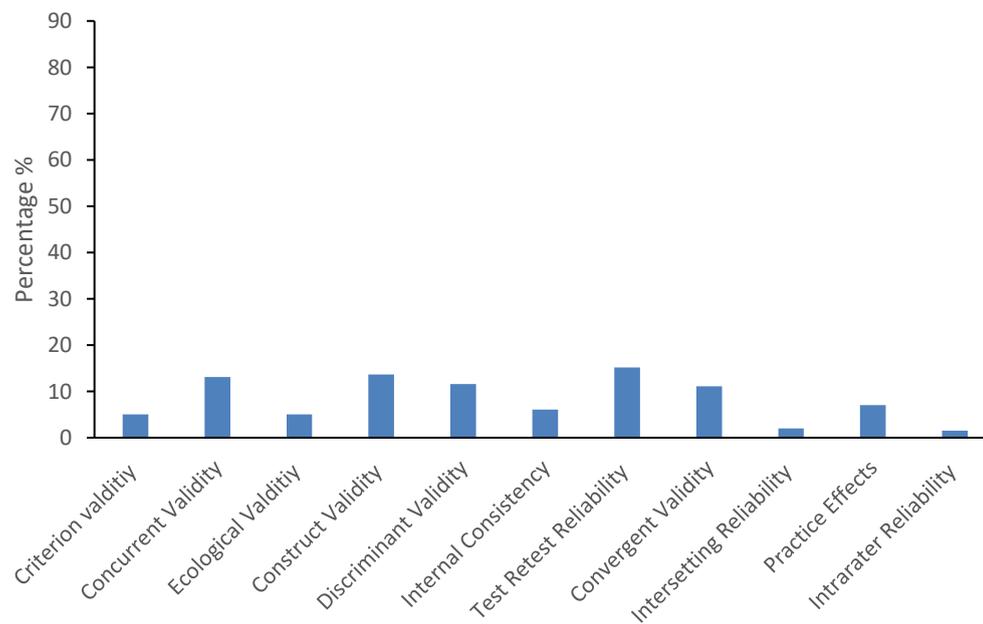
*Distribution of Cognitive Functions Assessed*

## Normative Data and Psychometric Properties

Another aspect explored in the systematic review was whether the included articles reported normative data and psychometric properties, and whether they constituted validation studies. Only 5% (n=10) of the studies were validation studies, and 7% (n=14) reported normative data. While most studies assessed at least one psychometric property of the digital tools, the specific types of validity and reliability examined are detailed in figure 3.

**Figure 3**

### *Psychometric Properties Measured*



### Target population

The target population in the studies included in the systematic review consisted of 25.75% healthy participants (n= 51), 22.72% clinical populations (n= 45), and 32.32% a combination of both healthy and clinical participants (n= 64).

### Administration, Presentation & Accessibility

Most assessments were administered in person (70%, n = 139), with only 7% (n = 14) allowing for self-administration. The remaining studies either utilized both modalities (remote and in-person) or did not mention the mode of administration. In terms of delivery

devices, computers were the most commonly used platform (67.17%, n = 133), followed by tablets (12.12%, n = 24). A smaller number of studies employed screen monitors (2%, n = 4), a combination of smartphone, tablet, and computer (1.52%, n = 3), or virtual reality environments (1.52%, n = 3). Only 1% (n = 2) used smartphones exclusively, while the device used was unreported in the remaining studies. Concerning accessibility, 88% of the studies did not specify access requirements, and 94% did not report whether a subscription was needed

## **Chapter VII: Discussion**

The results show that although a wide range of digital assessments are used in practice, only a small fraction were originally developed to specifically assess attention and executive functions. Most studies assessed working memory and sustained attention, while less focus was placed in other functions like planning, decision-making or selective attention. Consequently, the unbalanced focus on working memory and sustained attention reflect a trend identified by Wild and Musser (2014), who highlighted that executive functions are often operationalized narrowly in digital contexts. Moreover, important psychometric properties such as construct validity and test-retest reliability underreported, raising concerns about standardization and clinical applicability. These findings align with several limitations posed by digitalized assessments that are going to be discussed further in this review.

The high usage of tools like CANTAB and Cogstate align with findings that these assessments are widely used in clinical and research settings due to their availability and ease of use (Green et al.,2019). This usage is also shown because of their versatility with different clinical populations. Another aspect that was highlighted by the results is the fact that 70% of the assessments were administrated in person. This contradicts with idea that those digitalized assessments are often promoted for their ability to be administered remotely. It is also evident in the results that there is a lack of transparency or clarification by the authors of the articles

examined in reporting accessibility of the tools. The following sections will provide a detailed exploration of the assessments identified in the reviewed articles, along with a discussion of their advantages and limitations.

## **Chapter VIII: Attention and Executive Functions in Digitalized Assessments**

As mentioned previously, paper-pencil assessments have several shortcomings, and most do not identify subtle cognitive impairments or differentiate between the different types. With aging populations across several countries, researchers and clinicians are turning towards digitalized neuropsychological assessments that are more sensitive and effective in differentiating impairments. Moreover, the assessments need to be easily accessible to people who are experiencing subjective cognitive decline. Furthermore, aspects of attention and executive functions that are crucial to everyday functioning are not accurately captured in clinical settings. The use of digitalized neuropsychological assessments compensates for some of the limitations of paper-pencil assessments, like ease of accessibility from one's home and reduction of costs. Furthermore, digitalized assessments add standardization and automatic documentation of the results, thus removing clinicians' subjectivity and training requirements.

Still, there is an ongoing research about the assessments' sensitivity and specificity. Furthermore, the quality and validity of their normative data is still debatable. Also, Amanzadeh et al. (2023) examined 32 digital neuropsychological assessments and found that only 38% are new, while the others are just digital versions of already existing paper-pencil assessments. This way of assessment is favorable for clinicians. However, whether the same normative data should be used or whether they are measuring the same constructs is debatable (Amanzadeh et al.,2023). This chapter will discuss the information gathered from

the systematic review about digitalized assessments including their advantages and disadvantages. In addition, a review of some of the most used assessments found in the systematic review and ones that were not used as much will be covered. Moreover, their effectivity to capture decline in attention and executive function will be discussed. These tests are: Cambridge Neuropsychological Test (CANTAB), Cogstate Computerized Battery (CCB), The NIH Toolbox, NeuroCognitive Performance Test (NCPT), and The Boston Cognitive Assessment (BoCA).

### **Advantages**

As previously mentioned, with the growing age of the population, having digitalized assessments creates wide accessibility to a wide range of people without having to go to clinics (Latendorf, Runde & Salminen,2020). As a result, early diagnosis can be made that can help with formulating short-term and long-term plans, unlike clinical settings, where diagnosis is made when symptoms are already significant. Also, the data on digitalized assessment gets automatically stored, which helps in creating a profile for the patient that can be used across different clinical settings. Furthermore, it helps the clinicians when they try to understand patient history. This automatization also helps save time when taking age and education into consideration. In addition, digital assessments are cost effective in a way that does not require the patient to go to a clinic several times and keep paying for repeated services. Another advantage that digital assessments provide is a wide variation in stimuli presentation that reduces practice effects (Latendorf, Runde & Salminen,2020). These qualities overall reduce burden on healthcare professionals.

There are also advantages that affect the patient's experience. For example, self-administration reduces fatigue and other variables associated with the clinical context like the overly controlled environment (Laurine et al.,2020). Furthermore, paper-pencil assessments

require hand-eye coordination, which digital assessments eliminate (Laurine et al.,2020). Another advantage is the standardized stimulus presentation (Laurine et al.,2020). Digital assessments also add items that ask about psychological distress and subjective problems that paper-pencil assessments do not account for (Libon et al.,2023). An additional advantage of digitalized assessments is that the administration time is short, which increases motivation for old adults who are stressed by testing (Chan et al.,2021). Some digitalized assessments are administered through a tablet touch screen that captures complex aspects like time-in-air and pressure, which cannot be captured using paper-pencil assessments (Chan et al., 2021). According to Chan et al. (2021), digitalized assessments give accurate measurement about reaction time, helping in monitoring progression of impairment. When it comes to assessing attention, for example, the digitalized version of the Trail Making Test (TMT) adds the measurements pause, lifts, and timing information. This provides the clinician or researcher additional information to interpret and consider more than the information provided by paper-pencil versions (Fellows et al.,2017).

The favorability of digitalized assessments lies also in the ability to discriminate between those who have MCI and those who have dementia. This was examined in a study by Chan et al. (2021) where they found that twenty-two of the digital assessments that they used had high sensitivity and specificity (0.80). Moreover, eleven had the ability to detect MCI (Chan et al.,2021). They also found that digital assessments were better at capturing information processing and attention, which they mentioned as behavioral indications for MCI (Chan et al.,2021). The authors were able to concluded in their systematic review that digital assessments were better at diagnosing MCI and dementia than paper-pencil ones (Chan et al.,2021). Digital assessments can aid in identifying subtle cognitive impairments which are usually masked by cognitive reserve. This was evident in a study conducted by

Bonato et al. (2012) when patients with right hemisphere strokes were found to be impaired more on digital assessments than on paper-pencil ones.

Ding, Lee, & Chan (2022) stated that measuring executive function digitally could be a potential biomarker for tracking disease progression. They also emphasized that digital assessments show better diagnostic ability than paper-pencil assessments. Moreover, digital assessments are more accessible to patients in remote areas, making the assessments more convenient. Furthermore, they can be used to collect large data samples to use for normative data (Ding, Lee & Chan,2022). As a result, information about demographic variables can be collected to predict and detect MCI and dementia (Ding, Lee & Chan,2022). When already existing assessments become digitalized, several benefits are added. Bjorngrim et al. (2019) conducted a study to compare a digital battery called Minnemera to its paper-pencil counterparts. They found that digitalized assessments that measure aspects of attention and executive functions like the TMT, Corsi Block Test, and Stroop Test show the same effectivity of detecting cognitive impairment (Bjorngrim et al.,2019). When working memory was assessed using the digitalized version of the Corsi Block Test, more benefits were found. This includes reaction time measurement, automated scoring, and less complexity of administration due to the automated presentation of the assessment (Burnetti et al.,2014).

The Clock Drawing Task (CDT), a common tool for assessing executive functions, is now evaluated digitally (dCDT) as in the Core Cognitive Evaluation™ (CCE) (Libon et al.,2023). Consequently, the problem command and copy that was previously mentioned in the limitation chapter for paper-pencil assessments is solved. This allows for more distinctions and less subjectivity in interpretation (Libon et al.,2023). The dCDT was found to differentiate between those who are healthy, with MCI, and those with dysexecutive MCI (Libon et al.,2023). It was also found that calculating the time between each pen stroke was associated with worse performance on tests of working memory (Dion et al.,2020). Rentz et

al. (2021) found that impaired performance on the dCDT was associated with increase in tau and amyloid in patients with AD.

In conclusion, these points highlight benefits digitalized assessments provide that can ease the process of measuring attention and executive functions. Furthermore, they allow for more accurate scores that can aid the clinician in the interpretation process and in understanding different cognitive profiles. In addition to that, these assessments reach a wide range of patients due to the ease of accessibility and administration. The following sections explore the most commonly used digital neuropsychological assessments identified in the systematic review, highlighting their advantages over traditional paper-and-pencil methods, while also considering a few of the least utilized assessments in the review.

### **Cambridge Neuropsychological Test (CANTAB)**

The Cambridge Neuropsychological Test (CANTAB) was found to be one of the most used assessments in the results of the systematic review. It is also designed specifically to target several aspects of attention and executive functions. According to Cormack et al. (2018), the CANTAB has been used 28,000 times since 2012. Moreover, they mentioned that 64.5% of the patients assessed during that period were diagnosed with MCI or dementia, highlighting its effectiveness in clinical settings (Cormack et al.,2018). The CANTAB has also been used in 1,500 peer-reviewed articles, confirming its effectiveness in neuropsychological assessment (Green et al.,2019). It is a web-based assessment that can be administered on computer or touch-screen monitors; however, it is not designed to be used on tablet or smartphones. The assessment is composed of 22 tests that target four domains: memory, social and emotional cognition, attention and psychomotor speed, and executive functions (Cambridge Cognition,2023). Furthermore, the publishers of the assessment grouped some of the tests to form a core battery for specific diagnoses like ADHD, MCI,

schizophrenia, and several others (Torgersen et al.,2012). Because of its common use, it is worth exploring its angle of measuring attention and executive functions.

The Subtests that target executive functions focus on mental flexibility, inhibition, planning, and problem solving (Cambridge Cognition,2023). Those tests include: Cambridge Gambling Task (CGT), Cognition Kit Digit Symbol Substitution (DSST), Intra-Extra Dimension Set Shift (IED), Multitasking Test (MTT), One Touch Stockings of Cambridge (OTS), Spatial Working Memory (SWM), Stockings of Cambridge (SOC), and Stop Signal Task (SST). The attention subtests of the CANTAB focus on one's ability to selectively attend to certain stimuli while ignoring others. Those tests include Match Sample Visual Search (MTS) and Rapid Visual Information Processing (RVP). An ample of research has been conducted to examine the validity and reliability of each of those tests and a loy of different results have been found.

To examine its validation, Torgersen et al. (2012), conducted a study to validate the CANTAB in patients with epilepsy and found that the most difficult level in the SOC correlated with measures of memory and intelligence. This indicates that as the SOC task gets more difficult, other cognitive abilities are equipped in order to complete the task. In another research, Siew et al. (2022) focused on the aging population and conducted a study to create normative data for some of the tests of CANTAB in the Singaporean population. Moreover, they wanted to investigate the concurrent validity of executive functions and attention tests in the CANTAB in comparison with their paper-pencil equivalents (Siew et al.,2022).

Siew et al. (2022) found that the MTT and the RVP did not have good concurrent validity because they did not correlate with their paper-pencil counterparts. Furthermore, Siew et al. (2022) found that when the MTT was compared to its counterpart, the Color Trails Test (CTT), no correlation was found, indicating that they are not matched (Siew et al.,2022).

This was concluded after finding out that the MTT is more complicated than the CTT (Siew et al.,2022). Nevertheless, they found that men tend to perform better on the MTT. Years of education also affected performance on attention and executive functions tasks (MTT, SOC, RVP). The authors explained that those with lower education had less planning time and more impulsiveness (Siew et al.,2022). Consequently, the measures of executive functions on CANTAB show that less years of education are associated with reduced inhibition (Siew et al.,2022). Another finding was that the RVP did not correlate to the Digit Span Forward because they both have different scoring criteria (Siew et al.,2022).

Another study that focused on the construct and concurrent validity of the tests of executive functions in the CANTAB is the study conducted by Kim et al. (2014). In this study, however, the CANTAB was compared to another digitalized assessment called Computerized Neuropsychological Assessment (CNT) (Kim et al.,2014). This study offers a different perspective, as it was conducted on patients with schizophrenia and bipolar disorder (Kim et al.,2014). The subtests examined in this study were the IED, SOC, and SWM. It was found that there was a significant correlation between the IED and negative symptoms of schizophrenia (Kim et al.,2014). Furthermore, there was a correlation between the IED and Wisconsin Card Sorting Test (WCST) among all patients (Kim et al.,2014). It was also found that test-retest reliability showed that the coefficient was higher for the IED ( $r= 0.75$ ) than that of the WCST ( $r=0.65$ ), which shows that the IED has high validity (Kim et al.,2014). Moreover, there was a correlation between the SOC and TMT-B and a correlation between the SWM and the Stroop test. Kim et al. (2014) concluded that the SOC is a challenging task that requires a combination of cognitive abilities including attention, motor function, and inhibition. These results highlight how some tests require the integration of more than one cognitive ability in order to pick out how much the patient has an impairment in that area.

A question is also posed about how the CANTAB can be used remotely without the intervention of a clinician. This was examined by Backx et al. (2020), when conducted a study to examine the use of the CANTAB through web-based administration compared to a normal administration in a clinical setting. The authors found that the CANTAB is applicable to be used unsupervised with good reliability except for tests that include reaction times like the MOT, RVP, ATT, SST and RTI. Meaning most of the tests that target attention and executive functions. (Backx et al.,2020). Another examination has been made by Deluca et al. (2003) to examine how the assessment differentiates between different cognitive profiles by conducting a study that focus on the subtests of the CANTAB that targets working memory, planning, and composition of goal-directed behavior on a normative sample. The authors concluded that the subtests were fit to distinguish age related cognitive decline and difference in gender when it comes to executive functions (Deluca et al.,2003).

It is also important to examine how the consistency of the assessment in a certain population with a specific disorder. Therefore, Chamberlain et al. (2011) conducted a meta-analysis where they reviewed 13 studies that compare the performance of ADHD patients on measures of attention and executive functions to that of healthy controls using the CANTAB. The authors found medium to large effects in inhibition, working memory, and planning (Chamberlain et al.,2011). However, there were smaller deficits in attentional shifting compared to studies that used WCST (Chamberlain et al.,2011). Performance on SST showed impairment that aligned with other studies using other digitalized stop signal tasks (Chamberlain et al.,2011). The SOC also showed similar results in the meta-analysis (Chamberlain et al.,2011). ADHD patients showed impairments of IED, but the impairments were smaller compared to other tests (Chamberlain et al.,2011). Moreover, ADHD patients showed impairments in attention on the RVP test that aligns with neuroimaging studies (Chamberlain et al.,2011). The authors concluded that the CANTAB can effectively

distinguish impairments in facets of attention and executive functions in ADHD patients, which aligns with results of previous research (Chamberlain et al.,2011).

The CANTAB was also found to be a sensitive measure for assessing neurodevelopmental changes related to attention and executive functions in children (Green et al.,2019). This was evident in a study conducted in 2019 by Green et al. (2019) where the CANTAB was used to examine developmental changes among 826 Mexican children (Green et al.,2019). They found that performances on tests of attention and executive functions (SOC, IED, SST, MTS) improved in a linear manner that reflects the developmental course of these abilities (Green et al.,2019). Moreover, the SOC and MTS were found to correlate with Full-Scale IQ, suggesting that the CANTAB measures distinctive cognitive constructs (Green et al.,2019). The CANTAB was found to be suitable for use with Mexican youth as it requires non-verbal application (Green et al.,2019). This shows that the assessment is also culturally versatile and can demonstrate the same usability in different cultures.

Factor analysis is used to identify the cognitive constructs measured by an assessment and whether they are separate or overlap with each other (Lenehan et al.,2015).Lenehan et al. (2015) used factor analysis to identify if tests in the CANTAB measures the same cognitive domains that are measured in traditional neuropsychological assessments, and if they would group together under the same factors. They found that most of the CANTAB tests of executive functions and attention did not load onto specific cognitive domains (MTS, RTI, RVP and SWM) (Lenehan et al.,2015). This showed that they did not align with the traditional assessments (Lenehan et al.,2015). The authors clarified these results by mentioning that traditional assessments measure several cognitive abilities simultaneously (Lenehan et al.,2015). However, the CANTAB is more focused on specific abilities (Lenehan et al.,2015). Lenehan et al. (2015) stated that one should be cautious when using the CANTAB interchangeably with traditional assessments because of the lack of alignment.

However, it is worth mentioning that, in this review, it was mentioned that one of limitations of traditional assessments is their lack of specificity when targeting the different facets of attention and executive functions.

The CANTAB was found to be favored more than paper-pencil assessments because of the standardized administration procedure (Torgersen et al., 2012). Moreover, the tests do not require neuropsychologists to be trained because only short instructions are needed. This makes the CANTAB easily accessible in clinical settings (Torgersen et al.,2012). Furthermore, the difficulty within tests is varied, which lessens ceiling and floor effects (Torgersen et al.,2012). Another advantage of the CANTAB is that all of its tests are non-verbal, so they can also be used with non-English speakers (Torgersen et al.,2012). It also displays good construct validity (Siew et al.,2022). It can also differentiate between those who have AD and MCI (Siew et al.,2022). The CANTAB design reduces anxiety related to the use of computers (Wild & Musser,2014).A disadvantage of the CANTAB that stood out to Wild & Musser (2014) is that some of the measures require the use of past strategies to new levels, which led to practice effects. Speed and accuracy also displayed signs of improvement with repeated testing (Wild & Musser,2014). Consequently, caution is needed when interpreting test results with the application of correction for practice effects. Despite having some limitations, the CANTAB offers more tests to identify decline in attention and executive functions, which other paper-pencil tests do not offer. Moreover, it addresses the complexity of executive functions and attention working together. As such, it can be used in different clinical settings that include aging or dementia, psychiatric and neurodevelopmental disorders. This is also aided by the fact that the Cambridge Cognition website offers different collections of assessments based on the problems addressed such as AD, ADHD, multiple sclerosis, epilepsy, Huntington's disease, and several more. It offers more flexibility in the choices of assessments and the whole battery does not have to be administered if it is not

necessary. It must be noted that when using the CANTAB, one must not compare it to paper-pencil assessments, as both measure cognitive constructs differently.

### **Cogstate Computerized Battery (CCB)**

Cogstate Computerized Battery (CCB) was another assessment that was used repeatedly in the articles included in the systematic review. The CCB has been used in 1200 clinical trials and longitudinal research that covers neurological disease, psychiatric disorders, and other medical conditions (Gates et al.,2021). The CCB has also been included in 1,800 studies (Cogstate,2025). Designed for efficiency, the test takes only 20 minutes to complete and minimizes practice effects, making it ideal for repeated assessments (Maruff et al.,2013). It can also be administered web-based, on computers, or tablets. Additionally, its high sensitivity allows for the detection of subtle cognitive changes, enhancing its value in clinical and research settings (Patel et al.,2017). The assessment include several subtests that can be customized into different batteries for different profiles. This section explores how the CCB specifically measures attention and executive functions through its various subtests.

The CCB consists of 17 subtests, including three that assess executive functions and four that measure attention (Cogstate,2025). The executive functions subtests are the Groton Maze learning test (GMLT), the One back Test (ONB), and the Two back Test (TWB), all of which primarily target planning, cognitive flexibility, and working memory. The attention-related subtests include the Identification Test (IDN), the Psychomotor Vigilance Test (PVT), the Sustained Attention Test (SAT), and the Sustained Attention to Response Test (SART). Notably, the SART also engages response inhibition, which overlaps with executive function domains. These tasks are designed to evaluate aspects such as sustained attention, reaction time, vigilance, and selective attention.

The Groton Maze Learning Test (GMLT) has been examined for its validity and reliability within the CCB, providing insight into its role in assessing executive functions. Pietrzak et al. (2008) compared the test in relation to other tests on the CCB and found that it correlated moderately to other tests. This indicates that the GMLT and other tests from the CCB measure related cognitive abilities (Pietrzak et al.,2008). Error monitoring factor did not correlate with the measures of the CCB, which shows that the GMLT does not capture it (Pietrzak et al.,2008). It was also found that the GMLT provides consistent and reliable measurements across time and different populations (Pietrzak et al.,2008). The GMLT showed that those who performed weakly also struggled on other working memory tests of the CCB (Pietrzak et al.,2008). The authors also found that rule-break errors and total errors correlated moderately with other tests of the CCB (Pietrzak et al., 2008). This showed that more errors on the GMLT indicated weaker attention and poorer working memory (Pietrzak et al.,2008). Consequently, in order to perform well on the GMLT one has to have good working memory and inhibition ability (Pietrzak et al.,2008). The authors conducted factor analysis that distinguished two factors underlying the GMLT: error monitoring and inhibition, and learning (Pietrzak et al.,2008). It was concluded by Pietrzak et al (2008) that the GMLT is an adequate addition to the CCB, as it captures facets of executive functions. It is important also to examine how other tests of executive functions work with the GMLT to identify impairments. Rigoli et al. (2013) found that the ONB test of the CCB can predict fine and gross motor skills for children. This was supported by previous research that emphasizes the relationship between motor skills and executive functions, especially, working memory (Rigoli et al.,2013).

The CCB has demonstrated strong reliability and validity in assessing cognitive deficits across various clinical populations, including schizophrenia (Zhong et al.,2013). This was established by Zhong et al. (2013) when they conducted a study to examine the reliability

and validity of the Chinese version of the CCB in patients with schizophrenia. The patients were required to complete the CCB and the Repeated Battery for the Assessment of Neuropsychological Status (RBANS) (Zhong et al.,2013). For tests of executive functions, like the TWB and GMLT, there was a good test-retest reliability (Zhong et al.,2013). Moreover, for tests of attention like the IDN, scores correlated with the RBANS (Zhong et al.,2013). Additionally, it was found that the Chinese version of the CCB had a Cronbach value of 0.81, indicating good internal consistency (Zhong et al.,2013).

It was found in the study that social emotional cognition tests from the CCB correlated with traditional neuropsychological assessments, as it involves aspects of executive functions like inhibition, decision making and problem solving (Zhong et al.,2013). The authors concluded that the CCB is a good neuropsychological assessment for the detection of cognitive impairments in psychiatric disorders like schizophrenia, which is also supported by previous research (Zhong et al.,2013). Pietrzak et al (2009) also investigated the criterion and contrast validity of the schizophrenia battery of the CCB and moderate to large correlations were found between the measures of attention, working memory and problem solving. This indicated that Cogstate Schizophrenia Battery is a valid assessment for attention and executive function impairment among schizophrenia patients.

In the area of psychiatric disorders, patients with MDD showed large deficits on the IDN, indicating impairments in attention (Davis et al.,2017). However, when years of education were considered, differences between MDD patients and healthy controls disappeared (Davis et al.,2017). This showed that the IDN is affected by years of education (Davis et al.,2017). The IDN test was shown to differentiate between ADHD children with mild impairment and severe impairment in executive functions in a study conducted by Pradej et al. (2024). These findings highlight the sensitivity of the IDN test in detecting

attention deficits across different psychiatric conditions while also emphasizing the influence of education on test performance.

The CCB is not only used for assessing cognitive impairments in neurological conditions but also in diseases that, while not primarily neurological, can still impact brain function and cognitive performance. Petal et al. (2017) examined the criterion and construct validity of the CCB in breast cancer survivors. The authors did so by examining how tests from CCB align with traditional neuropsychological assessments, and by examining if the same differences in CCB tests appear in traditional assessments. The authors found that the IDN correlated with Processing Speed Index (PSI) in the WAIS-IV (Petal et al.,2017). The ONB correlated with WASI-IV PSI and Hopkins Verbal learning Test-Revised (HVLTR) (Petal et al.,2017). However, the ONB did not correlate with its supposed match, the Digit Span Backwards (DSB) (Petal et al.,2017). Consequently, this indicates that it does not substitute the DSB but measures working memory. The authors concluded that CCB tests had limited convergent validity but that may have been due to traditional assessments being auditory, whereas the CCB is a digital visual assessment (Petal et al.,2017).

A factor analysis study examined whether executive function tests in the CCB could be grouped into a single factor (Nordenswan et al.,2020). While a single-factor solution was supported using Summative Scores, the model based on First Test Round Performance provided a better fit (Nordenswan et al.,2020). This suggests that initial performance captures executive function more accurately than total scores. The authors concluded that CCB's executive function tests demonstrate strong validity (Nordenswan et al.,2020).

The Cogstate Brief Battery (CBB) is a reduced version of the Cogstate that contains four tests. Two of those tests are the IDN and the ONB. Maruff et al. (2013) conducted a study using the CBB to differentiate between patients with AD and patients with amnesic

MCI. The study also aimed to create composite scores to identify the specificity and sensitivity of those tests (Maruff et al.,2013). It was found that the impaired performance on the working memory composite (ONB) was associated with poor performance on paper-pencil assessments, like the Clinical Dementia Rating Scale (CDR) (Maruff et al.,2013). Also, impairments in attention were found, however, those with MCI had better performance than those with AD (Maruff et al.,2013). As mentioned previously, impairments on working memory and attention can be used as trackers for the progression of AD (Belleville, Chertkow, & Gauthier,2007). Sensitivity and specificity were found to be optimal for working memory. However, for attention (IDN), the sensitivity was low, and specificity was high (Maruff et al.,2013). The authors found that the composite scores had high test-retest reliability when the assessments have been administered four times over three months, showing no indication of practice effects.

The effectivity of the executive function tests of the CCB was challenged in a study conducted by Hammers et al (2012). In the study, the CCB was used to differentiate between patients with AD, Frontotemporal Dementia (FTD), Dementia with Lewy Bodies (DLB), MCI, and healthy controls (Hammers et al.,2012). It was found that the Cogstate had low ability in differentiating between patients with AD, FTD, and DLB (Hammers et al.,2012). Only one of the tests of executive functions, the ONB was able to differentiate between those who have MCI and healthy controls (Hammers et al.,2012). There was also a modest correlation between the CCB and paper-pencil assessments in the attention domain (Hammers et al.,2012). The authors concluded that the CCB was effective in detecting already-diagnosed dementia, but not for differentiating between its different types (Hammers et al.,2012). Zygouris and Tsolaki (2015) have mentioned in their review that the CCB shows practice effects in healthy controls when they are assessed multiple times in a single day. However, patients with MCI do not show the same practice effects when they are assessed 4

time across 3 hours (Zygouris & Tsolaki,2015). They suggested that the CCB is the only test where practice effects are helpful in distinguishing between those who have MCI and those who do not (Zygouris & Tsolaki,2015).

The use of the CCB is considered advantageous because of its short period of administration. Also, it is translated into 41 languages and has a user-friendly interface that does not require the examiner to be trained (Gates et al.,2021). Another major advantage of the CCB is that it was designed for repeated use over short intervals (Gates et al.,2021). This is shown by the majority of the reaction time tasks that include a variety of values, maximizing the quality of repeated testing and the detection of minimal cognitive changes (Gates et al.,2021). However, it has been indicated by Gates et al. (2021) that this robustness decreases over longer periods of retest by showing low test-retest reliability and practice effects. When compared to paper-pencil assessments, patients with first episodes of psychosis performed better on the CCB. Benoit et al. (2015) justified this by indicating that the CCB demands less motor aspects, which makes the performance easier.

According to the studies mentioned previously, the CCB has demonstrated an effectivity in measuring attention and executive functions. It has also demonstrated effective use across populations of neurological disease, psychiatric, neurodevelopmental and others that are not related to brain functions yet still have an effect on it. Still, the tests do not capture the many facets of executive functions and attention that were mentioned previously. This was shown in the study conducted by Hammers et al. (2012) when it failed to capture the differences between types of dementia. This is a crucial point in neuropsychological assessments and is one of the limitations that paper-pencil assessments are already criticized for. While the CCB makes up for in feasibility for repeated administration and low practice effects, it still fails to capture the complexity of attention and executive functions. Still, some

studies, as mentioned, state that its tests demonstrate construct and criterion validity that is sensitive in detecting impairments of attention and executive functions.

### **NIH Toolbox (NIHTB)**

Even though the NIH Toolbox (NIHTB) was not among the most frequently used assessments, it is important to take a broader look at the types of assessments that are available and if they can offer points that might be overlooked. The NIHTB is a comprehensive battery originally developed for research purposes to assess neurological and behavioral functioning. Developed in 2012 by Dr. Richard Gershon and a team of scientists, the NIHTB was designed to provide brief yet robust measures across four key domains: cognition, sensation, motor function, and emotional health, each requiring approximately 30 minutes to administer (Pilkonis et al.,2013). It is administered through an application that mostly runs on IOS software on iPad (NIH Toolbox,2023) Despite its initial focus on research, there has been growing interest in its clinical applicability, particularly for identifying cognitive impairments. As of recent reports, the NIHTB has been administered over 1.5 million times and cited in more than 300 peer-reviewed publications (NIH Toolbox,n.d.), with Snitz et al. (2020) noting its use in at least 260 studies. For the purpose of this review, the focus will be on the NIH Toolbox Cognition Battery (NIHTB-CB), specifically its subtests related to attention and executive functions. The NIHTB-CB is designed for individuals across the lifespan, from ages 4 to 85, and is now available as a mobile application for tablets and smartphones, increasing its accessibility in both research and clinical contexts (Snitz et al.,2020). While some subtests appear to assess both attention and executive functions simultaneously, this overlap can present a challenge for clinicians aiming to isolate specific impairments. At the same time, it may offer valuable insight into the natural integration of these cognitive abilities, as they often operate in tandem during real-world tasks.

The executive functions subtests in the NIH Toolbox Cognitive Battery (NTCB) include the Flanker Inhibitor Control and Attention Test (Flanker), the Dimensional Change Card Sort Test (DCCS), the Visual Reasoning Test (VR) (Carlozzi et al.,2017). These tasks are designed to assess core executive functions such as inhibition, selective attention, cognitive flexibility, and abstract reasoning. Although working memory is classified under the memory domain in the NTCB, numerous theoretical and empirical studies have emphasized its integral role within executive functioning (Carlozzi et al.,2017; NIH Toolbox, n.d.). In the NTCB, working memory is evaluated using the List Sorting Working Memory (LSWM).

Compared to other assessments reviewed in this study, the NTCB offers a relatively limited selection of subtests that specifically target the multiple dimensions of attention and executive functions. This limitation may pose challenges in studies requiring fine-grained assessment of these domains but may not significantly affect the tool's utility in broader cognitive screening or large-scale research contexts (Carlozzi et al.,2017).

The NTCB supports the theory that executive functions are characterized by a specific course across lifespan, where they peak in their twenties and decrease as age increases. This indicates that it has good sensitivity to age-related changes in cognition (Zelazo et al.,2014). This was evident in a study conducted by Zelazo et al. (2014) where performance on the DCCS showed constant improvement between the ages of 20 and 29. However, the Flanker did not show the same results. The subtests of executive functions also showed good test-retest reliability and low practice effects when they were taken again after a short time, between 1 to 3 weeks (Zelazo et al.,2014). Correspondingly, the authors found that the subtests have good convergent and discriminant validity (Zelazo et al.,2014). Another conclusion that the authors established was that the executive functions subtests of the NTCB correlated with other traditional measures like the D-KEFS Color Word Test (Zelazo et

al.,2014). Finally, the assessment highlighted how executive functions become more separated from general intelligence as children and adolescents grow older (Zelazo et al.,2014). These findings were also confirmed in a research conducted by Akshoomoff et al. (2018). According to the study conducted by Zelazo et al. (2014), the subtests of attention and executive functions show effectiveness when used with healthy adults, but its usability in clinical conditions has yet to be proven.

To understand how the NTCB would work within clinical conditions, several studies have been conducted with different conditions to examine its versatility. For example, the NTCB has been found to differentiate between patients with TBI, stroke, and spinal cord injury through its subtests of attention and executive functions (Carlozzi et al.,2017). Moreover, Garcia et al (2023) investigated the sensitivity of the NTCB in differentiating between those with amnesic MCI, non- amnesic MCI, and healthy controls. The authors found that patients with non-amnesic MCI have frontal deficits that tend to arise on the Flanker subtest (Garcia et al.,2023). This allows for the differentiation between patients with amnesic MCI, non-amnesic MCI and healthy control (Garcia et al.,2023). The authors also emphasized that those with amnesic MCI still have impairments that are beyond memory and can be multi-domain. However, those impairments are not as severe as those that arise in patients with non-amnesic MCI (Garcia et al.,2023). In conclusion, the NTCB was able to differentiate between the three conditions effectively, highlighting its use as a proper cognitive marker.

Across the population of patients suffering from strokes, another study was conducted to examine the construct validity of the NTCB (Carlozzi et al.,2017). Among individuals with strokes, evidence of construct validity was demonstrated by significant performance differences between mild and moderate/severe stroke groups. This was seen especially in composite measures of Fluid and Crystallized Cognition and on tests of executive functions

on the Flanker, and DCCS subtests (Carlozzi et al.,2017). According to Carlozzi et al., (2017), these differences align with known cognitive profiles of strokes, especially deficits in executive functioning. Convergent validity was supported by moderate to strong correlations between NTCB subtests and traditional neuropsychological measures ( $r = .31$  to  $.88$ ), while weaker correlations between different domains supported discriminant validity (Carlozzi et al.,2017).

Across broader research, executive function measures from the NTCB have shown moderate overlap with other cognitive abilities such as memory and attention, suggesting that while these tests are valid, they may not exclusively isolate executive functions. This is made evident in how the assessments clumps attention and executive functions together in all its subtests. The authors explained this by mentioning that it is attributed to the nature of the executive function tasks in the assessment, which are based on simple discrimination paradigms typical of experimental psychology rather than the complex, integrative tasks used in traditional neuropsychological assessments (Carlozzi et al.,2017). Consequently, while the NTCB offers valuable insights into executive function-related processes and serves as a reliable and efficient tool for cognitive assessment, its measures of executive function may capture broader cognitive abilities and lack the specificity of traditional tests designed to evaluate higher-order executive skills (Carlozzi et al.,2017). Moreover, confounding variables were present due to motor impairments, highlighting the need for caution when interpreting results from motor-dependent tasks in clinical populations (Carlozzi et al.,2017). Still, other studies found opposing results, highlighting that the executive functions subtests are effective in differentiating between clinical populations

One of those studies was conducted by Snitz et al. (2020) to differentiate between those with normal aging, subjective cognitive decline, MCI, and AD. The authors did so by examining how patients' results on the assessment would reflect the tau and amyloid

accumulation in the brain. From there on, the authors found that patients' results on DCCS had nothing to do with amyloid accumulation, however, it reflected tau progression. This is plausible because research has supported that amyloid accumulation starts early in the course of AD, then reaches a plateau along with memory impairments. Nonetheless, the disease progression of AD can be identified by examining other cognitive abilities, like attention and executive functions, which occur depending on the progression of tau in the brain. As a result, the authors were able to conclude that the NTCB subtests of executive functions were sensitive enough to differentiate between those who are at the beginning of the disease course and those who are progressing (Snitz et al.,2020).

In a different clinical population of patients with Parkinson's disease, Al Melegy et al. (2024) found that the subtests of executive functions of the NTCB were effective in highlighting that patients with GBA1 mutations and subthalamic nucleus-deep brain stimulation (STN-DBS) have significant impairments in response inhibition and faster reaction times, leading to more error on the Flanker compared to those who do not have GBA1 mutations. The previous studies are important to show that the NTCB is effective in patients with neurological disease, but it is also important to understand its effectivity within psychiatric disorders.

Several studies examine the use of the NTCB in assessing cognitive functioning in populations with psychological disorders. For example, Cassetta et al. (2020) administered the assessment to patients with treatment-resistant psychosis (TRP) and found a correlation between the scores and the severity of positive and disorganized symptoms. Moreover, the NTCB demonstrated strong convergent validity with paper-pencil assessments. It also showed good divergent validity (Cassetta et al.,2020). Another disorder where executive functions and attention work together is Social Anxiety Disorder (SAD) (Toller-Renfroe et al.,2015). This was supported by the theory of attentional control, in which patients with SAD

have deficits in goal-directed behavior that impacts attention to negative stimuli, and failure of inhibitory control (Toller-Renfree et al.,2015). Toller-Renfree et al. (2015) examined this by using the NTCB and found no significant impairments on the subtests of attention and executive functions. As a result, the authors indicated that those subtests are not complex enough to introduce the cognitive load that usually makes those impairments significant. As noted in the study by Carlozzi et al. (2017), the limited complexity of the NTCB's attention and executive function subtests may result in the loss of critical information needed to more accurately understand a patient's cognitive profile.

The use of the NIHTB, like all digitalized neuropsychological assessments analyzed in this review, exhibits mixed findings and introduces both advantages and disadvantages. Starting with the disadvantages, the NTCB was created only for research purposes, which limits its use as an aid in clinical diagnosis. Yet some studies advocate for its effectiveness in clinical use. However, as Carlozzi et al. (2017) mentioned, the assessment does not capture the complexities of attention and executive functions. It only mimics the aspects that are often used in experimental psychology. Furthermore, grouping the measurement of attention and executive functions within the same subtests treats them as broad, undifferentiated constructs rather than as distinct domains comprising multiple facets. As a result, the interpretation of the clinician or researcher is distorted. In addition, overlooking those distinct facets does not introduce the cognitive load that can tell us more information about how a disease or disorder affects the brain, as seen in the study by Toller-Renfree et al. (2015). Even so, the NTCB still managed in some studies to highlight age related differences, distinguish between different types of MCI, and align with neuroimaging research.

The most significant advantage that the NIHTB introduces in general is the fact that it does not only tackle cognition, but also sensation, motor function, and emotional health. Consequently, it underscores the notion that diseases do not affect cognition in isolation but

rather involve a constellation of impairments that impact specific cognitive domains in distinct ways. Despite also those limitations, Fox et al. (2022) documented in their review that it was one of the four most used assessments with a record of 281 publications, of which 111 were for neurological disorders and 39 for psychological disorders. Moreover, those articles have been cited 1000 times from 2015 to 2021 (Fox et al.,2022).

### **NeuroCognitive Performance Test (NCPT)**

Taking a deeper look on assessments that were only used once in the articles included in the systematic review, The NeuroCognitive Performance Test (NCPT) was developed by Lumosity in 2013 as part of their neuropsychological web-based assessment tools. A point that stands out about the NCPT is that it is based on validated paper-pencil assessments (Jaffe et al.,2022). Consequently, it provides a digital manifestation of assessments that clinicians are already familiar with. According to Jaffe et al. (2022), the NCPT has been effectively used for cognitive training, assessment of traumatic brain injury effects, and changes in cognition post-surgical intervention. Moreover, the NCPT can also be self-administered and takes only 20 to 30 minutes to finish (Jaffe et al.,2022).

The NCPT includes several subtests that assess a range of attention and executive function components (Jaffe et al.,2022). Attention is measured through the Divided Visual Attention Test, Digit Symbol Coding, Trail Making A and B, Posner Cueing Task, and the Dual Search Task. These subtests evaluate divided attention, selective attention, vigilance, and visual search. Some tasks, such as the Divided Visual Attention and Dual Search, also engage working memory due to their increased cognitive load and dual task demands. Notably, the digital versions of the Trail Making A and B replicate the classic paper-pencil format and assess processing speed, visual attention, and task-switching abilities. The Posner Cueing Task further measures spatial orienting and response speed. While the variability in

task difficulty across these subtests enhance sensitivity to subtle impairments (Jaffe et al.,2022).

Executive functions are captured through several additional NCPT subtests. Go/No-Go evaluates response inhibition, and Forward and Reverse Memory Span tasks assess working memory using a digital format of the Corsi Block-Tapping Test. The Complex Span Test also targets working memory by requiring participants to simultaneously process and retain verbal and numerical information. Arithmetic Reasoning measures problem solving, while Grammatical Reasoning assesses logical thinking through verbal logic tasks. Progressive Matrices, a digital adaptation of Raven's Matrices, evaluates abstract reasoning. The Scale Balance test, derived from the WAIS-IV Figure Weights subtest, measures quantitative reasoning and relational problem solving (Jaffe et al.,2022). Overall, the NCPT integrates both attention and executive function measures across multiple subtests, many of which inherently involve overlapping cognitive demands. This highlights the interdependence of these domains and their collective importance in assessing goal-directed behavior.

The NCPT has demonstrated effectiveness in measuring attention and executive functions by detecting cognitive impairments in attention and mental flexibility when used to examine the fluctuation of cognitive impairments in patients pre and post operations. This was evident in a study conducted by Bell et al. (2018) where patients showed impairments on Reverse Memory Span ( $P = .0045$ ) and Trail Making B ( $P = 0.118$ ) preoperatively. Furthermore, two weeks post-operation, the NCPT was able to detect improvements. The authors found that performance was low (.338 and .0154), yet patients still showed significant improvement 6 months post-operation (Bell et al.,2018). No practice effects were evident due to the fact that the NCPT is not available publicly for people to use for cognitive training (Bell et al.,2018).

Morrison et al (2015) conducted a study to examine the reliability and validity of the NCPT by examining the datasets of healthy individuals. The study focused on some of the subtests of attention and executive functions which are: Arithmetic Reasoning, Digit Symbol Coding, Forward and Reverse Memory Span, Grammatical Reasoning, Progressive Matrices, and Trail Making A and B. The authors found that each NCPT subtest significantly correlated with another subtest, showing that they share the same aspects in some cognitive abilities (Morrison et al.,2015). The Progressive Matrices Test did not correlate with any other subtest, suggesting that it measures a more unique cognitive ability like non-verbal reasoning (Morrison et al.,2015).

Moreover, factor analysis was further conducted to determine if the previously mentioned subtests group together, and to also examine their construct validity (Morrison et al.,2015). The analysis resulted in four factors, indicating that the subtests measure four unique cognitive abilities (Morrison et al.,2015). Two of those factors are mental flexibility (executive functions), and attention and processing speed (Morrison et al.,2015). The authors concluded that the NCPT demonstrated sufficient validity in assessing different cognitive abilities (Morrison et al.,2015). To examine test-retest reliability, Morrison et al. (2015) administered the assessment to their sample twice, with 78 days between both administrations. There was high test-retest reliability ( $r= 0.848$ ) for young participants and a lower reliability for older participants ( $r= 0.756$ ) (Morrison et al.,2015). Moreover, the authors tested the reliability of each subtest and found that all tests had high reliability except for the matrices test (Morrison et al.,2015). A further investigation also went into the assessment's validity, and it was found that it demonstrated strong validity when compared to paper-pencil assessments. Finally, Morrison et al. (2015) examined if the NCPT can differentiate between patients with AD, MCI, and healthy controls. The results showed that AD patients performed worse than MCI patients, indicating that the NCPT can detect

differences in cognitive decline (Morrison et al.,2015). The authors concluded that the executive functions and attention subtests of the NCPT had adequate test-retest reliability and good construct validity (Morrison et al.,2015).

For further research, Jaffe et al. (2022) presented a study that contains a dataset of five million subtest scores for 750,000 adults who were assessed by the NCPT. Moreover, the research contained the normative data and demographical aspects of each participant (Jaffe et al.,2022). This dataset provided a foundation for researchers to investigate cognitive abilities and impairments across populations, and to understand variabilities (Jaffe et al.,2022). The authors emphasized in their research that the NCPT has a high test-retest reliability, sensitivity to different clinical conditions, and good congruence with other neuropsychological assessments.

In conclusion, the NCPT has been developed to replicate established paper-pencil assessments with the feasibility of remote administration and flexibility of test selections. Moreover, tests can be put together to form a specific battery to measure definite cognitive abilities. A majority of those tests capture several facets of attention and executive functions. As shown in mentioned studies, it has demonstrated good reliability and validity. However, further research is needed to examine whether digitalized versions of paper-pencil tests still measure the same constructs or not. Moreover, is the use of normative data of paper-pencil assessments for the digitalized versions applicable or not? A disadvantage of the NCPT is that the test and publications related to its use are not available to the public to examine and assess. Also, the unavailability of the test prevents its use for training and initiating practice effects. However, this can also be a disadvantage as it hinders its usability in clinical settings and limits it only to research. Future research should examine its clinical utility in differences between neurological and psychiatric disorders.

### **The Boston Cognitive Assessment (BoCA)**

Another assessment that was only used once in the systematic review is the Boston Cognitive Assessment (BoCA) which was created by Prof Andrey Vyshedskiy et al. (2022) at the University of Boston. The aim of the assessment is to monitor cognitive decline to provide prevention and intervention for dementia . The BoCA is a 10-minute self-administered assessment that can be taken from a computer or a smartphone (Vyshedskiy et al.,2022). It evaluates 8 cognitive functions: memory/immediate recall, combinatorial language comprehension and prefrontal synthesis, visuospatial reasoning and mental rotation, executive function/clock test, attention, mental math, orientation, and memory/delayed recall (Vyshedskiy et al.,2022). The BoCA is advantageous because it can be administered on computer, tablets, or smartphones which shows its versatility (Vyshedskiy et al., 2022).

The BoCA includes multiple subscales designed to assess a range of executive functions and attention-related processes (Vyshedskiy et al.,2022). The executive function/clock test evaluates cognitive flexibility, logical reasoning, working memory, inhibition, error monitoring, and visual search by requiring participants to calculate and select the time difference between two clocks across multiple difficulty levels. The mental math subscale targets numeral cognition and mental arithmetic while engaging working memory, problem solving, and inhibition as participants perform increasingly complex addition and subtraction tasks.

Attention is assessed through digit span subscale, in which participants must recall and select digits in either forward or backward order, depending on the difficulty level. This task is designed to measure sustained attention, working memory, and engages recognition, response selection, and inhibition (Vyshedskiy et al.,2022). Notably, BoCA's digital format may emphasize recognition processes more than traditional auditory digit span task. A validation study by Vyshedskiy et al. (2022) reported a strong positive correlation between

performance on the attention and executive function subscales, suggesting a significant overlap in the cognitive processes being assessed.

Other subtests in BoCA also engage aspects of executive function and attention. The visual reasoning and mental rotation measures spatial reasoning, working memory, problem solving, and decision making by requiring participants to mentally rotate objects and match them to a target image (Vyshedskiy et al.,2022). The language and prefrontal synthesis subscale integrates executive functions such as working memory, inhibition, cognitive flexibility, and sustained attention alongside semantic comprehension (Vyshedskiy et al.,2022). It progresses across five levels of increasing difficulty, beginning with basic figure selection and evolving into complex semantic reasoning tasks (Vyshedskiy et al.,2022). Higher-level items involve processing semantically reversible sentences, requiring abstract reasoning and the mental manipulation of hierarchical information. Given the multidimensional nature of these subscales, it is likely that individuals with neurological or psychiatric conditions would exhibit impairments across multiple cognitive domains. The BoCA's integration of executive function and attention components within and across subtests highlights its potential utility in detecting subtle cognitive deficits in clinical populations.

Vyshedskiy et al. (2022) showed in their research that education was not found to predict or affect the BoCA scores. Moreover, the subscales showed good internal consistency of ( $\alpha = 0.87$ ) with all of them contributing to overall score. When subscales were removed the assessments', reliability was not affected, ensuring its robustness (Vyshedskiy et al.,2022). The authors also found that the subscales collectively measure a broad cognitive construct rather than distinct or unrelated abilities. They also conducted a factor analysis and found that a global functioning score combines all subscales, with it accurately capturing the variance in performance across time. It was also found that the variability of the instructions in each

subscale minimizes practice effects (Vyshedskiy et al.,2022). Furthermore, taking the test from a computer or a smartphone did not affect patients' performances (Vyshedskiy et al.,2022).

The BoCA offers convenience in that one does not have to go to a clinical setting to take the assessment (Vyshedskiy et al.,2022). The assessment can be taken from the convenience of one's own home where they can use it for cognitive monitoring and not wait until significant symptoms have emerged (Vyshedskiy et al.,2022). It also reduces examiner bias and performance anxiety all while being cost effective (Vyshedskiy et al.,2022). The assessment also does not have to be downloaded, one can be accessed and conducted from any browser (Vyshedskiy et al.,2022). While there were mentioned technical difficulties faced by older adults, the creators of BoCA addressed those issues by providing training sessions and using large clear buttons in the assessment (Vyshedskiy et al.,2022). Another limitation for the assessment is the lack of supervised testing environment that can result in decreased patient compliance. A solution for that problem is to include eye tracking features in future enhancements (Vyshedskiy et al.,2022). Furthermore, technical problems like internet speed may hinder the patients' accurate assessment (Vyshedskiy et al., 2022).

Ding et al. (2024) conducted a study to assess the use of the BoCA in the Chinese population, as it was never done before. They found that age was negatively correlated with the scores obtained from the assessment. Moreover, the assessment was found to have a good internal consistency of 0.77, good test-retest reliability with a Cronbach alpha of 0.79, and good content validity that ranged between 0.202 and 0.761 (Ding et al.,2024). The authors also carried out a calibration validity between the BoCA and the MMSE where they found a moderate positive correlation of 0.682. Unlike what was stated by the creators of the BoCA, Ding et al. (2024) found that scores from the assessment positively correlated with education. They explained these results by stating that some tasks in the assessment require a level of

literacy that can be easier for those with higher education levels (Ding et al.,2024). Furthermore, patients with prolonged years of education often find ease dealing with technological devices (Ding et al., 2024). Another limitation the authors found for the BoCA was that the slow speed of relaying questions made participants impatient and as a result, they started choosing answers carelessly (Ding et al., 2024).

Padovani et al. (2024) conducted a validation study for BoCA in an Italian population to examine its convergent validity with the MoCA between healthy controls, patients with Alzheimer's disease, patients with MCI (amyloid positive), and patients with MCI (amyloid negative). They found significant differences between groups, indicating that the BoCA has good discriminant validity (Padovani et al.,2024). Moreover, they found that the BoCA has good internal consistency between its subscales with a Cronbach alpha of 0.82 (Padovani et al.,2024). BoCA was also found to be strongly correlated with MoCA scores (Padovani et al.,2024). The authors of the study concluded that the BoCA is a valid tool that can be utilized to identify and monitor cognitive decline (Padovani et al.,2024). Furthermore, they stated that the BoCA is cost effective, and since the results are automatically documented, they are easily sharable and not prone to errors (Padovani et al.,2024). Due to the variability in presented stimuli, the assessment did not show any practice effects (Padovani et al.,2024). Moreover, all patients were able to complete the task by themselves without aid, meaning the lack of supervision needed for the assessment allows for screening larger samples (Padovani et al.,2024). Still, they found that scores are affected by levels of education (Padovani et al.,2024).

Ferguson et al. (2022) assessed the psychometric properties of BoCA by examining the impact of age and education on performance in a sample of 61 cognitively healthy adults aged 19 to 82 years with 12 to 20 years of education. Regression analysis revealed that age significantly predicted total BoCA scores, whereas education did not (Ferguson et al.,2022).

Adjusting for educational levels confirmed age-related differences in total scores and in subscales like delayed memory and language and prefrontal synthesis (Ferguson et al.,2022). These findings highlight the importance of considering age-adjusted scores in BoCA memory and to reduce false-positive identifications of cognitive impairment, emphasizing the need for refined scoring systems for different age groups (Ferguson et al.,2022).

In summary, the BoCA is an effective digital tool designed to assess a broad range of cognitive functions, demonstrating strong validity and reliability across various populations and contexts. Despite minor limitations, such as the influence of educational levels and occasional technical challenges, BoCA's benefits, including convenience, cost-effectiveness, and the ability to monitor cognitive decline remotely, make it a valuable instrument for early detection and intervention. Studies consistently highlight its robustness, discriminant validity, and lack of practice effects, indicating its potential as a reliable alternative for paper-pencil neuropsychological assessments. Still, most of the studies examining the effectiveness of the BoCA focus on the population of AD, creating a need for future studies to examine its efficacy with other types of dementia, and include psychiatric disorders.

### **Limitations of Digitalized Neuropsychological Assessments**

Even though they compensate for several limitations that are evident in paper-pencil assessments, digitalized assessments also possess their own drawbacks. Latendorf, Runde, Salminen (2020) conducted a study comparing the TMT and the Color-Word Interference Test (CWIT) in their digital and paper-pencil formats. This included an examination of differences in administration times. The authors found that digital tablet versions of the TMT and CWIT require a longer time in administration than their paper-pencil counterparts (Latendorf, Runde & Salminen,2020). This was a result of an increase cognitive demand that requires the patient to not only deal with the task, but also, deal with it on a new modality that they are not used to (Latendorf, Runde & Salminen,2020). While computer usage was

associated with less processing and administration time, the increase in cognitive demand in the CWIT meant that more executive functions were required, causing the usage advantage to disappear (Latendorf, Runde & Salminen,2020). Their findings supported the concept of double task burden that is presented in digitalized assessments where patients have an increase of cognitive challenges due to the introduction of a new modality (Latendorf, Runde & Salminen,2020). Even though cognitive demand can identify hidden impairments that are masked by cognitive reserve, patients who are not used to digitalized assessments can still show impairments when they do not have one.

Padovani et al. (2024) emphasized the limited number of studies investigating the construct validity, applicability, and reliability of digitalized assessments. As reflected in the results, only 13.64% of the included articles examined construct validity, 6% assessed internal consistency, and 15% reported on test–retest reliability. Moreover, only 5% were validation studies. In a research conducted by Chan et al. (2021), thirty digital assessments from ones they reviewed had limited validation studies. Also, they highlighted the fact that 36% of the digital assessments they reviewed were new and others were based on paper-pencil assessments. This sheds a light on the scarcity of finding new relevant ways to assess cognitive abilities (Chan et al.,2021). It also highlights another issue that instead of finding new and innovative ways to measure cognitive abilities, researchers and clinicians automatically resort to digitalized assessments of already established paper-pencil ones, thinking that they will compensate for all the limitations.

Laurine et al. (2020) conducted a study to examine user experience and feasibility of digitalized assessments in patients with acquired brain injuries (ABI). They also examined the applicability of paper-pencil norms to digital tests (Laurine et al.,2020). Normative data were not applicable as healthy control showed abnormal performance (Laurine et al.,2020). This brings up the idea that using normative data that were established for paper-pencil

assessments is not applicable for digitalized ones (Laurine et al.,2020). This was also reflected in the results of the systematic review were only 7% included normative data. Certainly as Amanzadeh et al. (2023) stated, digitalizing an assessment changes aspects in how it is conducted, administered, and scored. Moreover, the test can end up measuring the same construct in a totally different way.

Education and age have also been found to affect the administration process of digitalized assessments. For example, Bjorngrim et al. (2019) found that age significantly affects the performance on digital assessments, with younger participants performing better than older ones. Moreover, education affected performance on assessments of executive functions. Showing that, higher education was associated with better inhibition and working memory. Although digitalized assessments enable remote administration and broader patient reach, the presence of a clinician remains essential for detecting subtle cognitive or behavioral impairments that may not be captured through automated testing alone (Latendorf, Runde, & Salminen,2020). As previously reported in the results section, despite the intended utility of these tools for remote use, 70% of the included studies employed digital assessments in face-to-face settings. For example, patients who have high cognitive reserve are known to perform well on assessments of executive functions, yet observable deficits can be identified in behavior and attitudes. In patients with executive dysfunction, deficits such as impaired inhibition and perseveration may be revealed through non-standardized tasks like motor sequencing, which can detect impairments that standardized screening tools may overlook (Keifer & Haut, 2014,p.35).

As effective and as accessible self-administration can be, there is also no control of the distractors present at one's home during completing the assessment. The absence of a clinician can also affect the engagement and motivation to complete the assessment (Casaletto & Heaton,2017). There are small variables that can also affect the administration

of the assessment like whether it is conducted through a managed device or a personal one (Stafforni et al.,2020). Using a managed device is costly and introduces a factor of unfamiliarity, while a personal device lessens test stress and can be adjusted to the patients preferences like screen brightness (Laurine et al.,2020) (Stafforni et al.,2020). Other small variables include the types of resolutions, software and hardware (IOS, Android, PC, Mac) (Stafforni et al.,2020). Casaletto & Heaton (2017) emphasized that since digitalized neuropsychological assessment require continuous upgrades and software fixes, any minimal changes can alter how an assessment is administered or affect how its measured.

While digitalized neuropsychological assessments offer several advantages, they also present challenges. One of the reasons for the shift to digitalized assessments is that paper-pencil assessments do not accurately capture the multiple facets of attention and executive functions. While digitalized assessments have offered a compensation for that issue, the tests still do not reflect how attention and executive functions interact in daily life. The problems mentioned previously by Martinez-Pernia et al (2023) still were not accounted for in digitalized assessments. As such, attention and executive functions are still being addressed as mechanistic abilities. Moreover, the combination of decision making, problem solving, emotional regulation, social interaction, contextual control, and multitasking is still not captured.

Taken together, these findings suggest a number of limitations in digitalized assessments. A prominent limitation is that digitalized assessments may trigger increased cognitive demand, particularly for older individuals or those unfamiliar with digital tools, which can impact performance and administration time. Additionally, the applicability of normative data from paper-pencil tests to digital versions remains questionable, highlighting the need for more validation studies. Factors such as education, age, device type, and software variations further complicate standardization. Moreover, the absence of a clinician

during remote assessments may affect engagement, motivation, and the detection of subtle cognitive impairments. As digital assessments continue to evolve, ongoing research and refinement are essential to ensure their reliability, validity, and usability across diverse populations. The development of tests that emphasizes the importance of attention and executive functions as marker for the progression of cognitive decline is important.

## **Chapter IX: Future Directions**

Since the matter of ecological validity is still up to debate when it comes to digitalized neuropsychological assessments, future directions are heading towards things like virtual (VR) reality and serious games. Virtual reality creates a graphic world to engage the patient in real-life activities (Lee, Kim & Park,2024). These activities immerse the patient in real-life situations where goal-directed behavior is evaluated (Lee, Kim & Park,2024). Consequently, functional impairment caused by neurological disease or psychiatric disorders can be apparent and understood (Martinez-Pernia et al.,2023). In addition, a controlled environment similar to that of traditional assessment is available to gather objective measurements and responses (Lee, Kim & Park,2024). The patient can be put in a situation where they have to do tasks in a three-floor building or follow a list of items at a grocery store. According to Lee, Kim & Park (2024), VR-based assessments for attention and executive functions have shown correlations with other neuropsychological assessments. Nonetheless, Martinez-Pernia et al (2023) highlighted some limitations for VR methods such as the head-mounted systems, high prices, and can cause anxiety to the patient because of unfamiliarity with the technology.

Another approach to assessment that demonstrates high ecological validity is the use of serious games. Martinez-Pernia et al. (2023) described serious games as specialized digital applications designed primarily for therapeutic or diagnostic purposes, rather than for entertainment. These games offer patients opportunities for self-evaluation and have shown effectiveness not only in assessment but also in rehabilitation and the teaching of real-life

skills. Serious games can be used to assess attention and executive functions by immersing individuals in realistic, task-driven virtual environments. For example, Levy et al. (2019) developed a virtual grocery store with seven aisles where patients are required to complete tasks involving time management, planning, searching, and budgeting—engaging multiple executive functions and attentional processes simultaneously, much like real-world scenarios.

Similarly, Jovanovski et al. (2021) created a virtual city setting with various public services that patients interact with to complete complex, goal-directed tasks. Martinez-Pernia et al. (2023) emphasized that these serious games possess adequate psychometric properties while maintaining high ecological validity. The realistic, dynamic nature of these environments promotes exploration, decision-making, and strategy development, offering patients with attention and executive function deficits opportunities to practice compensatory skills. In addition, the use of goal-oriented tasks and meaningful feedback can enhance patient motivation by providing real-world relevance, in contrast to traditional assessments that often lack contextual significance (Martinez-Pernia et al., 2023).

Despite their promise, serious games and virtual reality applications still require further empirical validation and standardization. Moreover, their implementation can be costly and resource-intensive, requiring ongoing technical support and maintenance.

## **Chapter X: Conclusion**

This systematic review aimed to explore and critically evaluate the current field of digitalized assessments used to measure attention and executive functions. These two cognitive domains are central to goal-directed behavior, adaptive functioning, and are highly susceptible to impairment across a range of neurological and psychiatric conditions. Traditional paper-pencil assessments, although long-established in clinical neuropsychology, present several limitations including ecological validity, standardization, and sensitivity to

early and subtle deficits. As such, digitalized assessments have emerged as promising alternatives offering enhanced precision, standardization, scalability, and accessibility.

The systematic review identified 86 digital tools used in the assessment of attention and executive functions, among which only a small number were specifically designed to target these domains. While most tools measured at least one psychometric property, the overall reporting of construct validity, internal consistency, and test-retest reliability was limited. Moreover, normative data were often absent, and only a minority of studies constituted formal validation research. Despite these limitations, the review highlighted that digitalized tools such as CANTAB, Cogstate, NIH Toolbox, NCPT, and BoCA offer a range of advantages, including automated scoring, reduced clinician bias, flexibility of administration, and the capacity for remote use.

However, challenges remain regarding the psychometric precision of these tools, their comparability to traditional assessments, and their ability to distinguish complex cognitive profiles. Additionally, a significant proportion of studies still employed in-person administration, indicating that the shift to remote or self-administered digital testing is ongoing and not yet fully adopted.

In conclusion, digitalized neuropsychological assessments hold significant promise in advancing the assessment of attention and executive functions, particularly in settings that demand scalable, accessible, and time efficient. Nonetheless, more validation studies, standardized reporting of psychometric properties, and normative data collection are necessary to ensure their scientific and clinical reliability. Future research should focus on enhancing the ecological validity of digital tools, tailoring assessments for diverse populations, and integrating digital platforms into routine clinical practice to improve early detection, monitoring, and personalized intervention.

## References:

- Akshoomoff, N., Brown, T. T., Bakeman, R., & Hagler, D. J. (2018). Developmental differentiation of executive functions on the NIH Toolbox Cognition Battery. *Neuropsychology*, 32(7), 777–783. <https://doi.org/10.1037/neu0000476>
- Almelegy, A., Gunda, S., Buyske, S., Rosenbaum, M., Sani, S., Afshari, M., Metman, L. V., Goetz, C. G., Hall, D., Mouradian, M. M., & Pal, G. (2024). NIH Toolbox performance of persons with Parkinson's disease according to GBA1 and STN-DBS status. *Annals of clinical and translational neurology*, 11(4), 899–904. <https://doi.org/10.1002/acn3.52005>
- Amanzadeh M, Hamedan M, Mohammadnia A, Mahdavi A. Digital Cognitive Tests for Dementia Screening: A Systematic Review. *Shiraz E-Med J*. 2023;24(6):e137241. <https://doi.org/10.5812/semj-137241>
- American Psychiatric Association. (2013). *\*Diagnostic and statistical manual of mental disorders\** (5th ed.). American Psychiatric Publishing.
- Arevalo-RodriguezI, SmailagicN, Roqué-FigulsM, CiapponiA, Sanchez-PerezE, GiannakouA, PedrazaOL, Bonfill CospX, CullumS. Mini-Mental State Examination (MMSE) for the early detection of dementia in people with mild cognitive impairment (MCI). *Cochrane Database of Systematic Reviews* 2021, Issue 7. Art. No.: CD010783. DOI: 10.1002/14651858.CD010783.pub3
- Baddeley, A., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.) *Recent advances in learning and motivation* (vol. 8). New York: Academic.
- Backx, R., Skirrow, C., Dente, P., Barnett, J. H., & Cormack, F. K. (2020). Comparing Web-Based and Lab-Based Cognitive Assessment Using the Cambridge Neuropsychological

Test Automated Battery: A Within-Subjects Counterbalanced Study. *Journal of medical Internet research*, 22(8), e16792. <https://doi.org/10.2196/16792>

Balconi, M., Losasso, D., Balena, A., & Crivelli, D. (2022). Neurocognitive impairment in addiction: A digital tool for executive function assessment. *Frontiers in psychiatry*, 13, 955277. <https://doi.org/10.3389/fpsy.2022.955277>

Bang, J., Spina, S., & Miller, B. L. (2015). Frontotemporal dementia. *Lancet (London, England)*, 386(10004), 1672–1682. [https://doi.org/10.1016/S0140-6736\(15\)00461-4](https://doi.org/10.1016/S0140-6736(15)00461-4)

Bell, C. F., Warrick, M. M., Gallagher, K. C., & Baregamian, N. (2018). Neurocognitive performance profile postparathyroidectomy: a pilot study of computerized assessment. *Surgery*, 163(2), 457–462. <https://doi.org/10.1016/j.surg.2017.09.001>

Belleville, S., Chertkow, H., & Gauthier, S. (2007). Working memory and control of attention in persons with Alzheimer's disease and mild cognitive impairment. *Neuropsychology*, 21(4), 458–469. <https://doi.org/10.1037/0894-4105.21.4.458>

Benoit, A., Malla, A. K., Iyer, S. N., Joobar, R., Bherer, L., & Lepage, M. (2015). Cognitive deficits characterization using the CogState Research Battery in first-episode psychosis patients. *Schizophrenia research. Cognition*, 2(3), 140–145. <https://doi.org/10.1016/j.scog.2015.03.006>

Björngrim, S., van den Hurk, W., Betancort, M., Machado, A., & Lindau, M. (2019). Comparing Traditional and Digitized Cognitive Tests Used in Standard Clinical Evaluation - A Study of the Digital Application Minnemera. *Frontiers in psychology*, 10, 2327. <https://doi.org/10.3389/fpsyg.2019.02327>

Bonato M, Priftis K, Marenzi R et al (2012) Deficits of contralesional awareness: a case study on what paper-and-pencil tests neglect. *Neuropsychology* 26:20–36. <https://doi.org/10.1037/a00253069>. O'Halloran

Brunetti, R., Del Gatto, C., and Delogu, F. (2014). eCorsi: implementation and testing of the corsi block-tapping task for digital tablets. *Front. Psychol.* 5:939. doi: 10.3389/fpsyg.2014.00939

Cambridge Cognition. (2023). *Digital cognitive assessments*. Retrieved February 15, 2025, from <https://cambridgecognition.com/digital-cognitive-assessments/>

Cansino, S., Guzzon, D., Martinelli, M., Barollo, M., & Casco, C. (2011). Effects of aging on interference control in selective attention and working memory. *Memory & cognition*, 39(8), 1409–1422. <https://doi.org/10.3758/s13421-011-0109-9>

Carlozzi, N. E., Goodnight, S., Casaletto, K. B., Goldsmith, A., Heaton, R. K., Wong, A. W. K., Baum, C. M., Gershon, R., Heinemann, A. W., & Tulsky, D. S. (2017). Validation of the NIH Toolbox in Individuals with Neurologic Disorders. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 32(5), 555–573. <https://doi.org/10.1093/arclin/acx020>

Carnero-Pardo C. (2014). Should the mini-mental state examination be retired? *Neurologia (Barcelona, Spain)*, 29(8), 473–481. <https://doi.org/10.1016/j.nrl.2013.07.003>

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological Assessment: Past and Future. *Journal of the International Neuropsychological Society: JINS*, 23(9-10), 778–790. <https://doi.org/10.1017/S1355617717001060>

Cassetta, B. D., Menon, M., Carrion, P. B., Pearce, H., DeGraaf, A., Leonova, O., White, R. F., Stowe, R. M., Honer, W. G., Woodward, T. S., & Torres, I. J. (2020).

Preliminary examination of the validity of the NIH toolbox cognition battery in treatment-resistant psychosis. *The Clinical neuropsychologist*, 34(5), 981–1003.

<https://doi.org/10.1080/13854046.2019.1694072>

Chamberlain, S. R., Robbins, T. W., Winder-Rhodes, S. E., Muller, U., Sahakian, B. J., Blackwell, A. D., et al. (2011). Translational approaches to frontostriatal dysfunction in attention-deficit/hyperactivity disorder using a computerized neuropsychological battery. *Biological Psychiatry*, 69, 1192–1203.

Chan, J. Y. C., Yau, S. T. Y., Kwok, T. C. Y., & Tsoi, K. K. F. (2021). Diagnostic performance of digital cognitive tests for the identification of MCI and dementia: A systematic review. *Ageing research reviews*, 72, 101506.

<https://doi.org/10.1016/j.arr.2021.101506>

Chan, R. C., Shum, D., Touloupoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 23(2), 201–216. <https://doi.org/10.1016/j.acn.2007.08.010>

Cogstate. (2025). *Academic research*. Retrieved February 27, 2025, from <https://www.cogstate.com/academic-research/>

Cooley, S. A., Heaps, J. M., Bolzenius, J. D., Salminen, L. E., Baker, L. M., Scott, S. E., & Paul, R. H. (2015). Longitudinal Change in Performance on the Montreal Cognitive Assessment in Older Adults. *The Clinical neuropsychologist*, 29(6), 824–835.

<https://doi.org/10.1080/13854046.2015.1087596>

Corbetta M, Shulman GL. 2002. Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–15.

Cormack, F., Morris, H., Baker, E., Housden, C., & Barnett, J. H. (2018). CANTAB Mobile: Experience of use of a digital memory screening tool in primary care. *Cambridge Cognition*. Retrieved from [https://cambridgecognition.com/wp-content/uploads/2023/05/Cormack\\_AAIC\\_2018\\_Cantab\\_Mobile\\_FINAL2.pdf](https://cambridgecognition.com/wp-content/uploads/2023/05/Cormack_AAIC_2018_Cantab_Mobile_FINAL2.pdf)

Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., Thom, V. M., Nash, K. J., Elhamoui, H., Milligan, R., Patel, A. S., Tsivos, D. V., Wing, T., Phillips, E., Kellman, S. M., Shackleton, H. L., Singleton, G. F., Neale, B. E., Watton, M. E., & Cullum, S. (2016). Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *The Cochrane database of systematic reviews*, 2016(1), CD011145. <https://doi.org/10.1002/14651858.CD011145.pub2>

Cristofori, I., Cohen-Zimmerman, S., & Grafman, J. (2019). Executive functions. *Handbook of clinical neurology*, 163, 197–219. <https://doi.org/10.1016/B978-0-12-804281-6.00011-2>

Dautzenberg, G., Lijmer, J., & Beekman, A. (2020). Diagnostic accuracy of the Montreal Cognitive Assessment (MoCA) for cognitive screening in old age psychiatry: Determining cutoff scores in clinical practice. Avoiding spectrum bias caused by healthy controls. *International journal of geriatric psychiatry*, 35(3), 261–269. <https://doi.org/10.1002/gps.5227>

Davis, M. T., DellaGioia, N., Matuskey, D., Harel, B., Maruff, P., Pietrzak, R. H., & Esterlis, I. (2017). Preliminary evidence concerning the pattern and magnitude of cognitive dysfunction in major depressive disorder using cogstate measures. *Journal of affective disorders*, 218, 82–85. <https://doi.org/10.1016/j.jad.2017.04.064>

DeLuca, C. R., Wood, S. J., Anderson, V., Buchanan, J., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, 25, 242–254.

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

Dion, C., Arias, F., Amini, S., Davis, R., Penney, D., Libon, D. J., et al. (2020). Cognitive correlates of digital clock drawing metrics in older adults with and without mild cognitive impairment. *J. Alzheimers Dis.* 75, 73–83. doi: 10.3233/JAD-191089

Ding, Jiahui & Gao, Xiaoxia & Tang, Ying & Wu, Yehuan & Yuan, Xiaofang & Zhang, Jianan & Shi, Yue & Wang, Ya & Guo, Jing & Fei, Xiao & Zhang, Yi. (2024). A preliminary study on the use of the Boston Cognitive Assessment to assess cognitive function in a Chinese population. 10.21203/rs.3.rs-4821369/v1.

Ding, Z., Lee, T. L., & Chan, A. S. (2022). Digital Cognitive Biomarker for Mild Cognitive Impairments and Dementia: A Systematic Review. *Journal of clinical medicine*, 11(14), 4191. <https://doi.org/10.3390/jcm11144191>

Duff, K., Hobson, V. L., Beglinger, L. J., & O'Bryant, S. E. (2010). Diagnostic accuracy of the RBANS in mild cognitive impairment: limitations on assessing milder impairments. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, 25(5), 429–441. <https://doi.org/10.1093/arclin/acq045>

Fellows, R. P., Dahmen, J., Cook, D., and Schmitter-Edgecombe, M. (2017). Multicomponent analysis of a digital trail making test. *Clin. Neuropsychol.* 31,154–167. doi: 10.1080/13854046.2016.1238510

Ferguson H, Turok N, Gold D, Vyshedskiy A, Piryatinsky I. A-303 Preliminary Analysis of the Influence of Age and Education on the Boston Cognitive Assessment (BoCA). *Arch Clin Neuropsychology*. 2022;37(6):1457.

Garcia, S., Askew, R. L., Kavcic, V., Shair, S., Bhaumik, A. K., Rose, E., Campbell, S., May, N., Hampstead, B. M., Dodge, H. H., Heidebrink, J. L., Paulson, H. L., & Giordani, B. (2023). Mild Cognitive Impairment Subtype Performance in Comparison to Healthy Older Controls on the NIH Toolbox and Cogstate. *Alzheimer disease and associated disorders*, 37(4), 328–334. <https://doi.org/10.1097/WAD.0000000000000587>

Gates, T. M., Kamminga, J., Jayewardene, A., Vincent, T., Quan, D., Brew, B. J., Bloch, M., & Cysique, L. A. (2021). An examination of reliable change methods for measuring cognitive change with the Cogstate Computerized Battery: Research and clinical implications. *Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists*, 36(4), 597–612. <https://doi.org/10.1093/arclin/acia076>

Global Status Report on the Public Health Response to Dementia. World Health Organization; Geneva, Switzerland: 2021. [(accessed on 1 June 2022)]. Available online: <https://www.who.int/publications/i/item/9789240033245>. [Google Scholar]

Goette W. (2020). Reconsidering the RBANS Factor Structure: a Systematic Literature Review and Meta-Analytic Factor Analysis. *Neuropsychology review*, 30(3), 425–442. <https://doi.org/10.1007/s11065-020-09447-3>

Gratwicke, J., Jahanshahi, M., & Foltynie, T. (2015). Parkinson's disease dementia: a neural networks perspective. *Brain : a journal of neurology*, 138(Pt 6), 1454–1476. <https://doi.org/10.1093/brain/awv104>

Green, R., Till, C., Al-Hakeem, H., Cribbie, R., Téllez-Rojo, M. M., Osorio, E., Hu, H., & Schnaas, L. (2019). Assessment of neuropsychological performance in Mexico City

youth using the Cambridge Neuropsychological Test Automated Battery (CANTAB). *Journal of clinical and experimental neuropsychology*, 41(3), 246–256.

<https://doi.org/10.1080/13803395.2018.1529229>

Guarino, A., Favieri, F., Boncompagni, I., Agostini, F., Cantone, M., & Casagrande, M. (2019). Executive Functions in Alzheimer Disease: A Systematic Review. *Frontiers in aging neuroscience*, 10, 437. <https://doi.org/10.3389/fnagi.2018.00437>

Hammers, D., Spurgeon, E., Ryan, K., Persad, C., Barbas, N., Heidebrink, J., Darby, D., & Giordani, B. (2012). Validity of a brief computerized cognitive screening test in dementia. *Journal of geriatric psychiatry and neurology*, 25(2), 89–99.

<https://doi.org/10.1177/0891988712447894>

Hodges, John R, *Cognitive Assessment for Clinicians*, 2 edn (Oxford, 2007; online edn, Oxford Academic, 1 1 Nov. 2013),

<https://doi.org/10.1093/med/9780192629760.001.0001>, accessed 1 Dec. 2024.

Hogan, T. P. (2013). *Psychological Testing: A Practical Introduction* (3rd ed.). New York: John Wiley.

Huey, E. D., Goveia, E. N., Paviol, S., Pardini, M., Krueger, F., Zamboni, G., Tierney, M. C., Wassermann, E. M., & Grafman, J. (2009). Executive dysfunction in frontotemporal dementia and corticobasal syndrome. *Neurology*, 72(5), 453–459.

<https://doi.org/10.1212/01.wnl.0000341781.39164.26>

Jaffe, P. I., Kaluszka, A., Ng, N. F., & Schafer, R. J. (2022). A massive dataset of the NeuroCognitive Performance Test, a web-based cognitive assessment. *Scientific data*, 9(1), 758. <https://doi.org/10.1038/s41597-022-01872-8>

Jovanovski, D., Zakzanis, K., Ruttan, L., Campbell, Z., Erb, S., & Nussbaum, D. (2012). Ecologically valid assessment of executive dysfunction using a novel virtual reality task in patients with acquired brain injury. *Applied neuropsychology. Adult*, *19*(3), 207–220. <https://doi.org/10.1080/09084282.2011.643956>

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: a review of our current understanding. *Neuropsychology review*, *17*(3), 213–233. <https://doi.org/10.1007/s11065-007-9040-z>

Kim, H. S., An, Y. M., Kwon, J. S., & Shin, M. S. (2014). A preliminary validity study of the Cambridge neuropsychological test automated battery for the assessment of executive function in schizophrenia and bipolar disorder. *Psychiatry investigation*, *11*(4), 394–401. <https://doi.org/10.4306/pi.2014.11.4.394>

Krabbendam, L., & Jolles, J. (2003). The Neuropsychology of Schizophrenia. *Biological Psychiatry*, *2*, 631-647.

Ku Y. (2018). Selective attention on representations in working memory: cognitive and neural mechanisms. *PeerJ*, *6*, e4585. <https://doi.org/10.7717/peerj.4585>

Lang L, Clifford A, Wei L, Zhang D, Leung D, Augustine G, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. *BMJ Open* 2017 Feb 3;7(2):e011146 [FREE Full text] [doi: 10.1136/bmjopen-2016-011146] [Medline: 28159845]

Latendorf, A., Runde, L.M., Salminen, T. *et al.* Digitization of neuropsychological diagnostics: a pilot study to compare three paper-based and digitized cognitive assessments. *Aging Clin Exp Res* **33**, 1585–1597 (2021). <https://doi.org/10.1007/s40520-020-01668-z>

Lauriane A. Spreij, Isabel K. Gosselt, Johanna M. A. Visser-Meily & Tanja

C. W. Nijboer (2020) Digital neuropsychological assessment: Feasibility and applicability in patients with acquired brain injury, *Journal of Clinical and Experimental Neuropsychology*, 42:8,781-793, DOI: 10.1080/13803395.2020.180859

Lee, S.-A., Kim, J.-Y., & Park, J.-H. (2024). Concurrent Validity of Virtual Reality-Based Assessment of Executive Function: A Systematic Review and Meta-Analysis. *Journal of Intelligence*, 12(11), 108. <https://doi.org/10.3390/jintelligence12110108>

Lee, YC., Lee, SC. & Chiu, EC. Practice effect and test-retest reliability of the Mini-Mental State Examination-2 in people with dementia. *BMC Geriatr* **22**, 67 (2022). <https://doi.org/10.1186/s12877-021-02732-7>

Lei, L. K. S., Lam, B. Y. H., Lai, D. W. L., Bai, X., Li, J., Zou, Z., & Chan, C. C. H. (2022). Stability of Montreal Cognitive Assessment in Individuals with Mild Cognitive Impairment: Potential Influence of Practice Effect. *Journal of Alzheimer's disease : JAD*, 87(3), 1401–1412. <https://doi.org/10.3233/JAD-220003>

Lenehan, ME, Summers, MJ, Saunders, NL, Summers, JJ and Vickers, JC (2015) Does the Cambridge Automated Neuropsychological Test Battery (CANTAB) Distinguish Between Cognitive Domains in Healthy Older Adults? *Assessment*, 23 (3). pp. 163-172. ISSN 1073-1911

Levy, C. E., Miller, D. M., Akande, C. A., Lok, B., Marsiske, M., & Halan, S. (2019). V-Mart, a Virtual Reality Grocery Store: A Focus Group Study of a Promising Intervention for Mild Traumatic Brain Injury and Posttraumatic Stress Disorder. *American journal of physical medicine & rehabilitation*, 98(3), 191–198. <https://doi.org/10.1097/PHM.0000000000001041>

Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.

Libon, D. J., Matusz, E. F., Cosentino, S., Price, C. C., Swenson, R., Vermeulen, M., Ginsberg, T. B., Okoli-Umeweni, A. O., Powell, L., Nagele, R., Tobyne, S., Gomes-Osman, J. R., & Pascual-Leone, A. (2023). Using digital assessment technology to detect neuropsychological problems in primary care settings. *Frontiers in psychology, 14*, 1280593. <https://doi.org/10.3389/fpsyg.2023.1280593>

Lunardini, F., Luperto, M., Romeo, M., Basilico, N., Daniele, K., Azzolino, D., Damanti, S., Abbate, C., Mari, D., Cesari, M., Borghese, N. A., & Ferrante, S. (2020). Supervised Digital Neuropsychological Tests for Cognitive Decline in Older Adults: Usability and Clinical Validity Study. *JMIR mHealth and uHealth, 8*(9), e17963. <https://doi.org/10.2196/17963>

Maalouf, F. T., Klein, C., Clark, L., Sahakian, B. J., Labarbara, E. J., Versace, A., Hassel, S., Almeida, J. R., & Phillips, M. L. (2010). Impaired sustained attention and executive dysfunction: bipolar disorder versus depression-specific markers of affective disorders. *Neuropsychologia, 48*(6), 1862–1868. <https://doi.org/10.1016/j.neuropsychologia.2010.02.015>

Malek-Ahmadi, M., & Nikkhahmanesh, N. (2024). Meta-analysis of Montreal cognitive assessment diagnostic accuracy in amnesic mild cognitive impairment. *Frontiers in psychology, 15*, 1369766. <https://doi.org/10.3389/fpsyg.2024.1369766>

Maruff, P., Lim, Y. Y., Darby, D., Ellis, K. A., Pietrzak, R. H., Snyder, P. J., Bush, A. I., Szoeker, C., Schembri, A., Ames, D., Masters, C. L., & AIBL Research Group (2013). Clinical utility of the cogstate brief battery in identifying cognitive impairment in mild

cognitive impairment and Alzheimer's disease. *BMC psychology*, 1(1), 30.

<https://doi.org/10.1186/2050-7283-1-30>

Martínez-Pernía, D., Olavarria, L., Fernández-Manjón, B., Cabello, V., Henríquez, F., Robert, P., ... Slachevsky, A. (2023). The limitations and challenges in the assessment of executive dysfunction associated with real-world functioning: The opportunity of serious games. *Applied Neuropsychology: Adult*, 1–17.

<https://doi.org/10.1080/23279095.2023.2174438>

Mitchell A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of psychiatric research*, 43(4), 411–431. <https://doi.org/10.1016/j.jpsychires.2008.04.014>

Morrison, G. E., Simone, C. M., Ng, N. F., & Hardy, J. L. (2015). Reliability and validity of the NeuroCognitive Performance Test, a web-based neuropsychological assessment. *Frontiers in psychology*, 6, 1652. <https://doi.org/10.3389/fpsyg.2015.01652>

Nasiri, E., Khalilzad, M., Hakimzadeh, Z. *et al.* A comprehensive review of attention tests: can we assess what we exactly do not understand? *Egypt J Neurol Psychiatry Neurosurg* 59, 26 (2023). <https://doi.org/10.1186/s41983-023-00628-4>

NIH Toolbox. (n.d.). *NIH Toolbox: Home*. Retrieved April 6, 2025, from <https://nihtoolbox.org/>

Nordenswan, E., Kataja, E.-L., Deater-Deckard, K., Korja, R., Karrasch, M., Laine, M., Karlsson, L., & Karlsson, H. (2020). Latent structure of executive functioning/learning tasks in the CogState computerized battery. *SAGE Open*, 10(3), 1–10.

<https://doi.org/10.1177/2158244020948846>

Northwestern University. (2023). *NIH Toolbox iPad app information*. NIH Toolbox®.  
<https://www.nihtoolbox.org/app>

O'Brien, J. T., & Thomas, A. (2015). Vascular dementia. *Lancet (London, England)*, 386(10004), 1698–1706. [https://doi.org/10.1016/S0140-6736\(15\)00463-8](https://doi.org/10.1016/S0140-6736(15)00463-8)

O'Driscoll, C., & Shaikh, M. (2017). Cross-Cultural Applicability of the Montreal Cognitive Assessment (MoCA): A Systematic Review. *Journal of Alzheimer's disease : JAD*, 58(3), 789–801. <https://doi.org/10.3233/JAD-161042>

Padovani, A., Caratozzolo, S., Galli, A., Crosani, L., Zampini, S., Cosseddu, M., Turrone, R., Zancanaro, A., Gumina, B., Vicini-Chilovi, B., Benussi, A., Vyshedskiy, A., & Pilotto, A. (2024). Validation and convergent validity of the Boston cognitive assessment (BOCA) in an Italian population: a comparative study with the Montreal cognitive assessment (MoCA) in Alzheimer's disease spectrum. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, 10.1007/s10072-024-07775-3. Advance online publication. <https://doi.org/10.1007/s10072-024-07775-3>

Paelecke-Habermann, Y., Pohl, J., & Lelow, B. (2005). Attention and executive functions in remitted major depression patients. *Journal of affective disorders*, 89(1-3), 125–135. <https://doi.org/10.1016/j.jad.2005.09.006>

Parsons, Michael W., et al. *Clinical Neuropsychology [Risorsa Elettronica]: A Pocket Handbook for Assessment*. 3. ed, American Psychological Association, 2014.

Patel, S. K., Meier, A. M., Fernandez, N., Lo, T. T. Y., Moore, C., & Delgado, N. (2017). Convergent and criterion validity of the CogState computerized brief battery cognitive assessment in women with and without breast cancer. *The Clinical neuropsychologist*, 31(8), 1375–1386. <https://doi.org/10.1080/13854046.2016.1275819>

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease. A critical review. *Brain: a journal of neurology*, *122* ( Pt 3), 383–404.

<https://doi.org/10.1093/brain/122.3.383>

Pietrzak, R. H., Maruff, P., Mayes, L. C., Roman, S. A., Sosa, J. A., & Snyder, P. J. (2008). An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. *Archives of clinical neuropsychology: the official journal of the National Academy of Neuropsychologists*, *23*(4), 433–445. <https://doi.org/10.1016/j.acn.2008.03.002>

Pietrzak, R. H., Olver, J., Norman, T., Piskulic, D., Maruff, P., & Snyder, P. J. (2009). A comparison of the CogState Schizophrenia Battery and the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Battery in assessing cognitive impairment in chronic schizophrenia. *Journal of Clinical and Experimental Neuropsychology*, *31*(7), 848–859. <https://doi.org/10.1080/13803390802592458>

Pilkonis, P. A., Choi, S. W., Salsman, J. M., Butt, Z., Moore, T. L., Lawrence, S. M., Zill, N., Cyranowski, J. M., Kelly, M. A., Knox, S. S., & Cella, D. (2013). Assessment of self-reported negative affect in the NIH Toolbox. *Psychiatry research*, *206*(1), 88–97. <https://doi.org/10.1016/j.psychres.2012.09.034>

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Voelker, P. (2014). Developing Attention: Behavioral and Brain Mechanisms. *Advances in neuroscience (Hindawi)*, *2014*, 405094. <https://doi.org/10.1155/2014/405094>

Price, C. C., Cunningham, H., Coronado, N., Freedland, A., Cosentino, S., Penney, D. L., Penisi, A., Bowers, D., Okun, M. S., & Libon, D. J. (2011). Clock drawing in the Montreal Cognitive Assessment: recommendations for dementia assessment. *Dementia and geriatric cognitive disorders*, *31*(3), 179–187. <https://doi.org/10.1159/000324639>

Rabinovici, G. D., Stephens, M. L., & Possin, K. L. (2015). Executive dysfunction. *Continuum (Minneapolis, Minn.)*, *21*(3 Behavioral Neurology and Neuropsychiatry), 646–659. <https://doi.org/10.1212/01.CON.0000466658.05156.54>

Rentz, D. M., Papp, K. V., Mayblyum, D. V., Sanchez, J. S., Klein, H., Souillard-Mandar, W., et al. (2021). Association of digital clock drawing with PET amyloid and tau pathology in normal older adults. *Neurology* *96*, e1844–e1854. doi: 10.1212/WNL.0000000000011697

Rigoli, D., Piek, J. P., Kane, R., Whillier, A., Baxter, C., & Wilson, P. (2013). An 18-month follow-up investigation of motor coordination and working memory in primary school children. *Human movement science*, *32*(5), 1116–1126. <https://doi.org/10.1016/j.humov.2013.07.014>

Saunders, N. L., & Summers, M. J. (2010). Attention and working memory deficits in mild cognitive impairment. *Journal of clinical and experimental neuropsychology*, *32*(4), 350–357. <https://doi.org/10.1080/13803390903042379>

Siew, S. K. H., Han, M. F. Y., Mahendran, R., & Yu, J. (2022). Regression-Based Norms and Validation of the Cambridge Neuropsychological Test Automated Battery among Community-Living Older Adults in Singapore. *Archives of clinical neuropsychology : the official journal of the National Academy of Neuropsychologists*, *37*(2), 457–472. <https://doi.org/10.1093/arclin/acab073>

Shulman GL, McAvoy MP, Cowan MC, Astafiev SV, Tansy AP, D'Avossa G, and others. 2003. Quantitative analysis of attention and detection signals during visual search. *Neurophysiol* *90*(5):3384–97.

Snitz, B. E., Tudorascu, D. L., Yu, Z., Campbell, E., Lopresti, B. J., Laymon, C. M., Minhas, D. S., Nadkarni, N. K., Aizenstein, H. J., Klunk, W. E., Weintraub, S., Gershon, R.

C., & Cohen, A. D. (2020). Associations between NIH Toolbox Cognition Battery and *in vivo* brain amyloid and tau pathology in non-demented older adults. *Alzheimer's & dementia (Amsterdam, Netherlands)*, *12*(1), e12018. <https://doi.org/10.1002/dad2.12018>

Spencer, R. J., Kitchen Andren, K. A., & Tolle, K. A. (2018). Development of a scale of executive functioning for the RBANS. *Applied neuropsychology. Adult*, *25*(3), 231–236. <https://doi.org/10.1080/23279095.2017.1284664>

Staffaroni, A. M., Tsoy, E., Taylor, J., Boxer, A. L., & Possin, K. L. (2020). Digital Cognitive Assessments for Dementia: Digital assessments may enhance the efficiency of evaluations in neurology and other clinics. *Practical neurology (Fort Washington, Pa.)*, *2020*, 24–45.

Sternberg, R. J. (2012). \*Cognitive psychology\* (6th ed.). Wadsworth Cengage Learning

Torgersen, J. , Flaatten, H. , Engelsen, B. and Gramstad, A. (2012) Clinical Validation of Cambridge Neuropsychological Test Automated Battery in a Norwegian Epilepsy Population. *Journal of Behavioral and Brain Science*, *2*, 108-116.  
doi: [10.4236/jbbs.2012.21013](https://doi.org/10.4236/jbbs.2012.21013).

Troller-Renfree, S. V., Barker, T. V., Pine, D. S., & Fox, N. A. (2015). Cognitive functioning in socially anxious adults: insights from the NIH Toolbox Cognition Battery. *Frontiers in psychology*, *6*, 764. <https://doi.org/10.3389/fpsyg.2015.00764>

Truong, Q. C., Cervin, M., Choo, C. C., Numbers, K., Bentvelzen, A. C., Kochan, N. A., Brodaty, H., Sachdev, P. S., & Medvedev, O. N. (2024). Examining the validity of the Mini-Mental State Examination (MMSE) and its domains using network analysis. *Psychogeriatrics : the official journal of the Japanese Psychogeriatric Society*, *24*(2), 259–271. <https://doi.org/10.1111/psyg.13069>

V. Marinescu, M. Vrabie, A. Talasman, EPA-0749 – Assessment of attention and executive function in bipolar affective disorder, *European Psychiatry* Volume 29, Supplement 1, 2014, Page 1, ISSN 0924-9338, [https://doi.org/10.1016/S0924-9338\(14\)78098-8](https://doi.org/10.1016/S0924-9338(14)78098-8).  
(<https://www.sciencedirect.com/science/article/pii/S0924933814780988>)

Vossel, S., Geng, J. J., & Fink, G. R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *The Neuroscientist: a review journal bringing neurobiology, neurology and psychiatry*, 20(2), 150–159.  
<https://doi.org/10.1177/1073858413494269>

Vriend, C., Pattij, T., van der Werf, Y. D., Voorn, P., Booij, J., Rutten, S., Berendse, H. W., & van den Heuvel, O. A. (2014). Depression and impulse control disorders in Parkinson's disease: two sides of the same coin?. *Neuroscience and biobehavioral reviews*, 38, 60–71.  
<https://doi.org/10.1016/j.neubiorev.2013.11.0014>

Vyshedskiy, A., Netson, R., Fridberg, E., Jagadeesan, P., Arnold, M., Barnett, S., Gondalia, A., Maslova, V., de Torres, L., Ostrovsky, S., Durakovic, D., Savchenko, A., McNett, S., Kogan, M., Piryatinsky, I., & Gold, D. (2022). Boston cognitive assessment (BOCA) - a comprehensive self-administered smartphone- and computer-based at-home test for longitudinal tracking of cognitive performance. *BMC neurology*, 22(1), 92.  
<https://doi.org/10.1186/s12883-022-02620-6>

Wild, K. V., & Musser, E. D. (2014). The Cambridge neuropsychological test automated battery in the assessment of executive functioning. In *Handbook of Executive Functioning* (pp. 171-190). Springer New York. [https://doi.org/10.1007/978-1-4614-8106-5\\_11](https://doi.org/10.1007/978-1-4614-8106-5_11)

Wilk, C. M., Gold, J. M., Bartko, J. J., Dickerson, F., Fenton, W. S., Knable, M., Randolph, C., & Buchanan, R. W. (2002). Test-retest stability of the Repeatable Battery for

the Assessment of Neuropsychological Status in schizophrenia. *The American journal of psychiatry*, 159(5), 838–844. <https://doi.org/10.1176/appi.ajp.159.5.838>

World Health Organization. (2023). *Dementia*. Retrieved December 21, 2024, from <https://www.who.int/news-room/fact-sheets/detail/dementia>

Zelazo, P. D., Anderson, J. E., Richler, J., Wallner-Allen, K., Beaumont, J. L., Conway, K. P., Gershon, R., & Weintraub, S. (2014). NIH Toolbox Cognition Battery (CB): validation of executive function measures in adults. *Journal of the International Neuropsychological Society : JINS*, 20(6), 620–629. <https://doi.org/10.1017/S1355617714000472>

Zheng, B., Udeh-Momoh, C., Watermeyer, T., de Jager Loots, C. A., Ford, J. K., Robb, C. E., Giannakopoulou, P., Ahmadi-Abhari, S., Baker, S., Novak, G. P., Price, G., & Middleton, L. T. (2022). Practice Effect of Repeated Cognitive Tests Among Older Adults: Associations With Brain Amyloid Pathology and Other Influencing Factors. *Frontiers in aging neuroscience*, 14, 909614. <https://doi.org/10.3389/fnagi.2022.909614>.

Zhong, N., Jiang, H., Wu, J., Chen, H., Lin, S., Zhao, Y., Du, J., Ma, X., Chen, C., Gao, C., Hashimoto, K., & Zhao, M. (2013). Reliability and validity of the CogState battery Chinese language version in schizophrenia. *PloS one*, 8(9), e74258. <https://doi.org/10.1371/journal.pone.0074258>

Zygouris, S., & Tsolaki, M. (2015). Computerized cognitive testing for older adults: a review. *American journal of Alzheimer's disease and other dementias*, 30(1), 13–28. <https://doi.org/10.1177/1533317514522852>