MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„A Prospective Validation of a Digital Neurocognitive Assessment Battery - A Pilot Study in the Development of The International Neurocognitive Test Profile (INCP) "

verfasst von / submitted by

Magdalena Maria Maier, BSc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc.)

Wien, 2024/ Vienna, 2024

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Psychologie UG2002
Betreut von / Supervisor:	Univ-Prof.Mag.Dr. Claus Lamm
Mitbetreut von / Co-Supervisor:	Assoc.Prof.PD.Dr. Johann Lehrner

Acknowledgments

I would like to take a moment to express my gratitude to those people who have provided me with invaluable support and guidance, enabling to complete my master's thesis.

I am particularly grateful to my external advisor, Assoc. Prof. PD. Dr. Johann Lehrner, who provided me with the opportunity to supervise my master's thesis as part of an internship. He was consistently available and offered constructive feedback, which assisted me in the completion of my thesis. In particular, the clinical practice in the memory clinic of the Department of Neurology (AKH) provided valuable insights and reinforced my motivation to complete my studies.

I am grateful to my internal thesis advisor, Univ.-Prof. Mag. Dr. Claus Lamm, who was always supportive and available when needed.

I would like to express my appreciation to my parents, who supported me unconditionally and always believed in me. I am very grateful to my mother, who spent so much time talking to me by encouraging me and strengthening my personality. I am also very thankful to my father, who consistently provided me with encouragement through his easy-going manner, enabling me to become the person I am today. I consider myself very fortunate to have such great parents.

I would like to express my gratitude to my love, Martin. He is so kind, caring, and humorous, and has consistently provided me with emotional support during this stressful period of time. I am grateful for the late-night talks, and his wonderful ability to cheer me up.

To my friend Vali, with whom I spent the days studying in the library. Thanks for the enlightening talks we had over our lunchtime meals in the park.

Finally, I would like to express my gratitude to my boss, Prof. Dr. Hackenberg, who is my role model and my mentor. She is one of the most impressive women I have ever met. Regardless of time, she always took the time for a good talk. She provided me with so much personal and professional expertise, which has been a great source of inspiration. Her sense of humor made the part-time job a memorable one.

Abstract

Background: As the population ages, the prevalence of dementia is anticipated to triple by 2050. Given the lack of pharmacological treatment, it is considered a priority to develop new technologies for the detection of Alzheimer's disease in its prodromal stage. The purpose of the digital assessment battery (INCP) is to monitor cognitive function. Objectives: The aim of the study was to provide initial insights into the intricacies of variables, the underlying structure, and the discriminatory power of the INCP. Methods: The pilot study had a prospective cross-sectional design. Data from 41 healthy controls and 12 patients with MCI were analyzed. First insights into the validity of the INCP were gained through examining associations within the INCP using Spearman's correlation coefficient. Additionally, domain-specific variables were correlated with a well-established paper and pencil assessment battery. Furthermore, an exploratory factor analysis was conducted, and performance differences across groups were analyzed using Mann-Whitney-U-test. Results: The findings implied convergent validity for the domains of learning and memory, and language. Furthermore, the data indicated discriminatory power for the subtests: FPT, AVT, VVT, and CITY. Conclusion: The study implied that the INCP may be a valid tool for monitoring cognitive function. However, it is essential to acknowledge the limitations of the exploratory approach, emphasizing the need for cautious interpretation. Further research with refined methodology is necessary to reinforce the INCP's position as a valid tool for the assessment of cognitive decline.

Keywords: Alzheimer's disease (AD), mild cognitive impairment (MCI), prodromal stage of dementia, digital neurocognitive assessment, prevention

Zusammenfassung

Hintergrund: Die demographische Alterung führt zu einer prognostizierten Verdreifachung der Demenzprävalenz bis 2050. Angesichts fehlender Behandlungsmöglichkeiten ist die Entwicklung neuer Technologien zur Früherkennung von Demenz von großer Bedeutung. Das Internationale Neurokognitive Testprofil (INCP) wurde entwickelt, um kognitive Funktionen selbstständig zu überwachen. Studien zu den psychometrischen Kriterien der digitalen Testbatterie fehlen weitgehend. Ziel: Im Rahmen einer Zwischenevaluierung untersuchte die vorliegende Studie die Validität des INCP. Es wurden erste Einblicke in die zugrundeliegende Struktur, korrelierende Variablen und die Trennschärfe der digitalen Testbatterie gewonnen. Methode: Diesbezüglich wurde eine Pilotstudie mit einem prospektiven Querschnittsdesign durchgeführt und die Daten von 53 Teilnehmer*innen analysiert, darunter 41 gesunde Kontrollpersonen (HC) und 12 mit leichter kognitiver Störung (MCI). Die Validität des INCP wurde mittels einer Korrelationsmatrix unter Verwendung des Spearman Korrelationskoeffizienten untersucht. Einerseits wurden domänenspezifische Variablen innerhalb des INCP analysiert, andererseits wurde untersucht, inwieweit das INCP mit einer validierten P&P Testbatterie (NTBV) korreliert. Um erste Einblicke in die zugrundeliegende Struktur des INCP zu erhalten, wurde eine explorative Faktorenanalyse durchgeführt. Zusätzlich wurde mit Hilfe des Mann-Whitney-U-Tests untersucht, ob sich die beiden Gruppen (HC, MCI) hinsichtlich ihrer Leistungen in den Subtests unterscheiden. Die Ergebnisse zeigten vielversprechende Korrelationen und Trennschärfe in den Bereichen Lernen und Gedächtnis sowie Sprache. Schlussfolgerungen: Die Studie impliziert, dass das INCP ein valides Instrument zur Erfassung kognitiver Beeinträchtigungen im Forschungskontext sein kann. Das explorative Design schränkt die Interpretierbarkeit und Generalisierbarkeit der Ergebnisse ein. Für den Einsatz des INCP zur Früherkennung von Demenz im klinischen Setting sind weitere Forschungsarbeiten notwendig; psychometrische Kriterien müssen erst etabliert werden, um valides Testen zu gewährleisten.

Schlüsselwörter: Alzheimer-Krankheit (AD), leichte kognitive Beeinträchtigung (MCI), Prodromalstadium der Demenz, neuropsychologische Diagnostik, Prävention

Table of Contents

Introduction

In September 2021, the World Health Organization (WHO) stated that nearly 55 million people worldwide suffer from *dementia*. As the population ages, the prevalence of dementia is forecast to hit 139 million by 2050. It is generally agreed that dementia has tremendous consequences on the personal lives of people affected by dementia, their families, the public healthcare system, and the economy. Its direct impact on the global cost of health, social care, and informal care is estimated at 1.3 trillion US dollars in 2019, equivalent to 0.76% of the global gross domestic product (GDP). These costs are anticipated to increase and may overwhelm the health care system (World Health Organization [WHO], 2021, 2022). In response, the WHO has recognized dementia as a public health priority by supporting awareness, research, and prevention efforts since 2017 (WHO, 2017).

Dementia manifests itself in various forms, with dementia due to *Alzheimer's disease* (AD) being the most prevalent form. Although dementia is a syndrome that mostly affects the elderly, it can impact individuals of any age. Approximately 9% of dementia cases occur in people under the age of 65 (WHO, 2021, 2022).

AD is characterized by a decline in cognition from a previous higher level of function and has a presymptomatic course that can last for several years to decades (Jessen et al., 2014). Jessen et al. (2014) have suggested a three-stage model for the progression of dementia, which is defined by *subjective cognitive decline* (SCD), *mild cognitive impairment* (MCI), and dementia (Jessen et al., 2014).

While a cure for dementia remains elusive and pharmacological treatment options show only modest effects (Chan et al., 2021; Pons et al., 2018), early screening is highly recommended, as detecting dementia in its early stages can facilitate preventive measures, including physical, cognitive, and social interventions, which can positively impact the course of the disease (Rosas et al., 2022; Österreichische Alzheimer Gesellschaft, n.d.). Especially the two stages preceding AD provide great opportunities for early detection and the implementation of prevention strategies (Jessen et al., 2014).

Research has identified several modifiable risk factors that are collectively responsible for about 40% of dementia incidents and could consequently be prevented (Livingston et al., 2020b).

Given the fact that the healthcare system is challenged by the rising prevalence of AD due to a higher life expectancy in an increasingly aging population while healthcare resources are constrained (Petersen et al., 2018), the invention of new technologies that enable

population-level cognitive screening to enhance tailored prevention for at-risk individuals is considered a priority (Alty et al., 2022; Sabbagh et al., 2020).

However, the current clinical practice involves neuropsychological tests that are time-consuming, resource-limited, and expensive (Lehrner, 2021b). An accurate, easy-to-use, and reliable tool that is faster and more accessible to both patients and providers is critical to implementing secondary prevention trials to enable help as quickly as possible (Rosas et al., 2022; Sperling et al., 2011; Wild et al., 2008).

Consequently, further development on digital and self-administered neurocognitive assessment batteries is needed, offering a promising way to shut the dementia diagnosis gap without imposing an excessive load on the healthcare system (Sternin et al., 2019).

Theoretical and Empirical Background

Dementia

Dementia is an umbrella term that encompasses various neurodegenerative or cerebrovascular pathologies and their associated clinical consequences. The underlaying causes of this progressive neuronal disorder, which can ultimately lead to dementia syndrome, are diverse and not yet fully understood. Dementia due to *Alzheimer's disease (AD)* has a progredient onset, with amnestic symptoms being the prototypical clinical phenotype (Klotz & Gelpi, 2021; Knopman et al., 2021).

Accounting for 60-80% of all cases, AD is the leading cause of dementia, followed by vascular dementia (15-20%), dementia with Lewy bodies (7-20%), and other forms of dementia (Österreichische Alzheimer Gesellschaft, n.d.). Mixed forms of dementia are common; thus, AD pathology rarely occurs in isolation (Knopman et al., 2021).

In 1997 the Lancet Commission documented in the "Canadian study of health and aging" the following prevalence rate for dementia: 2.4% of individuals aged between 65-74, 11.2% of those aged 75-84 and 34.7% of those aged 85 and older are affected by the syndrome (Graham et al., 1997). Another study replicated these findings in the European population and found comparable prevalence rates. It is concluded that the prevalence rate of dementia rises significantly with age (Lobo et al., 2000).

Diagnostic Guidelines

Diagnostic guidelines are necessary to establish a common language and a clear understanding to ensure appropriate care for those who are affected (Balogh et al., 2015).

Several well-established diagnostic guidelines for dementia are currently in use. The most frequently clinically applied ones are namely the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V) developed by the American Psychiatric Association (2013) and the International Classification of Diseases, 10th edition (ICD-10) developed by the WHO (1992). These guidelines differ slightly in terms of symptom evaluation. Despite their differences, they collectively serve as an indispensable source for clinicians in the diagnosis of dementia (Wetterberg et al., 2024).

In 2019, the 10th edition of the International Statistical Classification of Diseases and Related Health Problems was replaced by the 11th revision (World Health Organization, 2019).

A study by Wetterberg et al. (2024) compared the DSM-V and the ICD-11 with their earlier editions and found that the selected diagnostic criteria significantly impact the prevalence rate of AD, with DSM-V and ICD-11 having a higher prevalence rate.

It is noted that neither diagnostic guideline captures the preclinical stage of dementia, as SCD is not considered a diagnosis (Jessen et al., 2020). However, the revision of both DSM-V and ICD-11 captures the prodromal stage of dementia (MCI), (Deutsche Gesellschaft für Neurologie [DGN e. V.] & Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde [DGPPN e. V.], 2023).

Furthermore, the research framework proposed by the National Institute on Aging and Alzheimer's Association (NIA-AA) is described as well.

ICD-10/11 Definition

The World Health Organization (1992) defines dementia as following:

> Dementia is a syndrome due to a disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, calculation, learning capacity, language, and judgement. Consciousness is not clouded. Impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation. This syndrome occurs in Alzheimer's disease, in cerebrovascular disease, and in other conditions primarily or secondarily affecting the brain. (WHO, 1992, p. 45)

Dementia, as categorized by the ICD-10, falls within the organic, symptomatic, mental disorder chapter (F00-F09). This classification encompasses various subcategories denoted by codes F00 through F03 and is categorized into four classifications: Dementia in Alzheimer's disease (F00), vascular dementia (F01), dementia in other diseases classified elsewhere (F02), and unspecified dementia (F03). Additionally, other forms of dementia, including vascular dementia, Lewy body dementia, and frontotemporal dementia, are situated in the same chapter.

The ICD-11 still utilizes the term "dementia" but categorizes dementia within the group "neurocognitive disorders" (Wetterberg et al., 2024). The ICD-10 defines memory impairment as mandatory (WHO, 1992), whereas the ICD-11 emphasizes the presence of decline in at least two cognitive domains, with memory being just one of many cognitive domains that can be affected by AD pathology (Wetterberg et al., 2024).

It should be highlighted that the ICD provides a standardized way of classifying and coding diagnoses, but it does not provide any recommendations on the treatment or management of progressive disorders (Lehrner et al., 2009a, p. 376) and is therefore only used in clinical practice (Wetterberg et al., 2024).

DSM-V Definition

The Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V) provides guidelines that are primarily intended for clinicians rather than for research. The DSM-V differentiates between major and mild neurocognitive disorders based on the severity of cognitive deterioration and thereby tries to capture the stages preceding dementia (Hugo & Ganguli, 2014).

According to the DSM-V, dementia is defined as a major neurocognitive disorder that results in cognitive deterioration beyond the typical trajectory of normal aging. This cognitive decline must manifest itself as a regression from a previous higher level of cognitive functioning and must be significant enough to impair social or occupational functioning. Furthermore, the syndrome must affect at least two cognitive domains, including memory, attention, language, perception, problem-solving, and social cognition, one of which must be memory (Hugo & Ganguli, 2014).

Additionally, the DSM-V provides etiological categories for the most common subtypes of dementia, encompassing dementia syndrome due to: Alzheimer's disease, vascular neurocognitive disease, frontotemporal neurocognitive disorder, traumatic brain injury and Lewy body dementia, Parkinson's disease, HIV infection, Huntington's disease, Prion disease, another medical condition, or multiple etiologies.

To meet the DSM-V criteria for a diagnosis of dementia, other causes, such as mental or medical conditions like depression or substance abuse, must be clearly excluded (Hugo & Ganguli, 2014).

NIA-AA Definition

In 2011, the National Institute on Aging-Alzheimer's Association (NIA-AA) established separate guidelines for the diagnosis of dementia (Montine et al., 2012). Contractionary to the ICD-10 and DSM-V guidelines, the NIA-AA guidelines are labeled a "research framework" with a focus on a multidisciplinary approach that defines dementia on a continuum, which progresses in stages (Jack et al., 2018). The framework involves clinical evaluation, neuropsychological assessment, biomarker testing, and recommendations, but shifts the focus to a research approach rather than clinical practice (Jack et al., 2018; McKhann et al., 2011).

According to the NIA-AA guidelines, dementia is defined as a decline in cognitive function, which is severe enough to interfere with daily functioning. This decline must be

greater than what would be expected for normal aging and last at least six months. The guidelines also specify that the decline must affect at least two cognitive domains, and they propose a new diagnostic category, namely "preclinical Alzheimer's disease" to capture the early stages of cognitive decline. Notably, the NIA-AA guidelines enhance a biological construct, incorporating biomarker tests across the disease continuum to improve diagnostic accuracy and outline a scientific framework for disease progression. Furthermore, they emphasize that not all cases of (preclinical) dementia involve abnormal protein deposits, but that this particular pathology defines AD as a unique neurodegenerative disease among other underlying diseases (Jack Jr. et al., 2011; Jack et al., 2018; McKhann et al., 2011).

Neurobiological Background

Alzheimer's disease is a progressive neurodegenerative disease that develops slowly. It is defined by a pathological accumulation of intraneuronal and extracellular protein deposits in the central nervous system. The typical symptoms of AD, like memory loss, behavioral changes, and loss of independence are caused by pathological protein deposits. Those are mainly caused by two proteins, namely β-amyloid-protein and τ-protein, which are also part of the healthy human brain (Lehrner et al., 2009a, pp. 376; National Institute on Aging, 2024).

Key biological processes necessary for neuronal function, and survival are disrupted by pathological modifications of β-amyloid- and τ-protein. The pathological extracellular β-amyloid plaques are made of amyloid β-proteins, which are derived from the precursor protein Amyloid Precursor Protein (APP), which is essential for the neuronal cell membrane. Intracellular neurofibrillary tangles are formed by τ-proteins, which hold the cytoskeleton together. Due to these chemical changes the cytoskeleton becomes unstable. This leads to irreversible neuronal degeneration as microtubules lose their structural integrity and degenerate. As a result, these chemical changes have an impact on the metabolism, communication, and repair mechanisms of neurons, which in turn leads to synaptic and dendritic degeneration, neuronal loss, and atrophy of the brain. Furthermore, AD pathology is associated with chronic inflammation and vascular contributions, which further exacerbate neuronal damage, and cognitive decline (National Institute on Aging, 2024).

AD can be divided into stages based on the typical hierarchical expansion of pathological changes over several decades. In the early stages of AD, pathological alterations usually arise in memory-related brain regions, like the hippocampus. As the disease proceeds,

these pathological changes spread to other brain regions, such as the cerebral cortex, which is leading to severe cognitive and behavioral dysfunction. Consequently, the clinical picture worsens over time (Knopman et al., 2021; Patterson, 2018; Thal & Braak, 2005).

Even though dementia is a syndrome that primarily affects the elderly, research suggests that complex brain changes caused by AD pathology start decades before the onset of clinical symptoms. This emphasizes the importance of focusing on the preclinical stage since early detection and, thus, early intervention, may be the key to managing the disorder's devastating effects (Jansen et al., 2015; Jessen et al., 2014; Knopman et al., 2021).

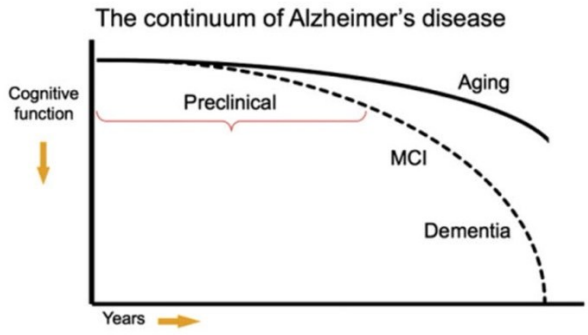
Stages of Dementia

The relationship between memory issues, mild cognitive impairment, and the transition period to dementia is not yet fully understood (Lehrner et al., 2009a).

According to Jessen et al. (2014), dementia evolves in stages, which encompass the preclinical stage, namely subjective cognitive decline (SCD), the prodromal stage, namely mild cognitive impairment (MCI), and the last stage, which is referred to as dementia (see Figure 1).

Figure 1

The Clinical Trajectory of Alzheimer's Disease



Note. The figure shows the overlap between normal aging, preclinical AD, the prodromal stage of MCI, and AD, describing cognitive impairment on a disease continuum, with preclinical AD and MCI depicted as an intermediate state between normal aging and early AD (adapted from Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease by Sperling et al., 2011, p. 283).

Because biomarker evidence is the best predictor for the risk of transitioning from normal mental function to dementia due to AD pathology, it is strongly advised that it be incorporated into the diagnostic process along the disease continuum (Jessen et al., 2014).

Recently, the stages preceding AD have gained significant attention, as early detection and intervention at this point are very promising with high potential for preventing or postponing the onset of the disease in those who are considered especially vulnerable (Jessen et al., 2014; Rabin et al., 2020; Reid & Maclullich, 2006).

Preclinical Stage of Dementia: Subjective Cognitive Decline (SCD)

Currently, prevention strategies target the prodromal stage of dementia (MCI), when irreversible and progressive neuronal loss has already occurred (Jessen et al., 2014).

Sperling et al. (2011) raise concerns about the conventional approach, assuming that interventions in the prodromal stage may be ineffective, highlighting that the key to targeted prevention is during the preclinical phase of AD, when neuropathological changes are modest.

In 2014, Jessen and colleagues proposed a research framework that defines the preclinical stage of dementia, called subjective cognitive decline (SCD). SCD refers to the period of time when cognitive decline is present but clinical symptoms have not yet manifested. SCD is described as a self-perceived decline in cognitive function, which is mostly compensated. Following the framework of Jessen et al. (2014), SCD is diagnosed by using an age, gender, and education-related mean score greater than -1.50 standard deviation in each domain of a neurocognitive assessment battery (Lehrner et al., 2016).

Detecting the preclinical stage of dementia (SCD) would be a useful approach for targeted prevention, as treatment at this stage could preserve cognitive functions at a high level. Unfortunately, it cannot be objectively assessed within a neurocognitive assessment battery; thus, healthy individuals, and individuals with subjective memory complaints do not show any differences in cognitive performance (Jessen et al., 2014).

Therefore, it is well recognized that there is a need for greater standardization in the assessment of the preclinical stage (SCD) in older adults and the development of reliable and valid instruments that can be used across different populations and cultures (Rabin et al., 2015).

With up to 50-80% of individuals aged over 70 years reporting subjective memory complaints, SCD is a common condition among older adults (Jessen et al., 2020). Subjective

memory complaints are not required for the diagnosis of dementia, are not present in all cases of AD, and not all individuals with SCD develop dementia, as it can be a result of normal aging or other conditions such as depression or anxiety (Jessen et al., 2014, 2020).

Nevertheless, prior studies have found that SCD is linked to early AD-related brain changes (Amariglio et al., 2012), such as decreased gray matter volume (Jessen et al., 2006). A study by Mitchell et al. (2014) examined the annual conversion rate and found that the risk of preceding to AD was doubled in people with SCD. These findings contribute to a growing body of evidence implying that SCD may be a potential early indicator of cognitive impairment due to AD pathology (Jessen et al., 2014; Mitchell et al., 2014).

However, Jansen et al. (2015) found contradictory results. The study compared cognitively healthy individuals with people who reported subjective memory complaints. Both groups did not differ in terms of amyloid positivity, suggesting that people with SCD are not at increased risk of progressing to dementia. The authors highlight the need for additional longitudinal research to gain a more profound knowledge of the complex interactions between the preclinical stage and AD-related brain alterations along the disease continuum.

The Prodromal Stage of Dementia: Mild Cognitive Impairment (MCI)

Petersen (2004) states that there is a transitional period between the physiological changes of healthy aging and the diagnosis of early Alzheimer's disease and proposes a framework that introduces the stage of prodromal AD described by mild cognitive impairment (MCI), (Petersen, 2004; Petersen et al., 2000; Petersen et al., 1999).

Since the transition period is very subtle, the distinction between normal aging, MCI, and early AD can be very challenging (see Figure 1). However, unlike SCD, MCI can be objectively assessed, making it an appropriate approach for implementing prevention strategies in at-risk individuals (Petersen, 2004), as neuropsychological or psychotherapeutic interventions can successfully intervene at this juncture and stabilize or fully remit the condition of MCI (Lehrner et al., 2009a, p. 382).

The diagnosis of MCI follows the framework established by Petersen et al. (1999) and is defined as a cognitive decline that is greater than expected for a person's age and education level but does not significantly interfere with daily functioning (Petersen et al., 1999). In other words, individuals who perform 1.50 standard deviations below age, gender, and

education norms in a single domain or in multiple domains in a neurocognitive assessment battery fulfill the criteria for the MCI diagnosis (Lehrner et al., 2016).

As MCI can be objectively assessed it is recognized as a key construct for targeting prevention efforts (Geddes et al., 2020; Myers et al., 2022).

However, Petersen (2004) states in "Mild cognitive impairment as a diagnostic entity" that almost half of the healthy cohort originally investigated by the Mayo Clinic performed 1.5 *SD* below their demographically normed group. Therefore, he draws attention to the fact that clinicians may be challenged on the one hand by high-functioning individuals, who are highly talented and perform statistically normal on a neurocognitive assessment battery, although their performance represents a cognitive decline from a previous higher level of performance, and on the other hand by low-functioning individuals, who perform below their normative group, but this performance does not represent a change in cognitive function. It is therefore important to interpret this quantitative criterion carefully and only in combination with the criterion of subjective memory complaints. By itself, the SCD criterion indicates only an individual change in cognitive performance (Petersen, 2004).

Petersen et al. (1999) report that individuals with MCI may exhibit cognitive impairment comparable to those with AD but do not differ significantly from healthy individuals in terms of their general cognition. Additionally, Petersen (2004) argues that MCI differs from dementia insofar as functional abilities must be essentially preserved in MCI condition (Petersen, 2004).

Petersen (2004) extended the MCI framework by incorporating several subtypes, such as amnestic and non-amnestic MCI, and by differentiating between single and multiple domains. Non-amnestic MCI implies normative memory performance but abnormal performance in other cognitive domains. In contrast, amnestic MCI includes mandatorily abnormal memory function for age. Both need to fulfill the criterion of self-perceived perception of memory deterioration, whereas daily life activities are not impaired. Figure 2 provides the guidelines for the MCI diagnosis.

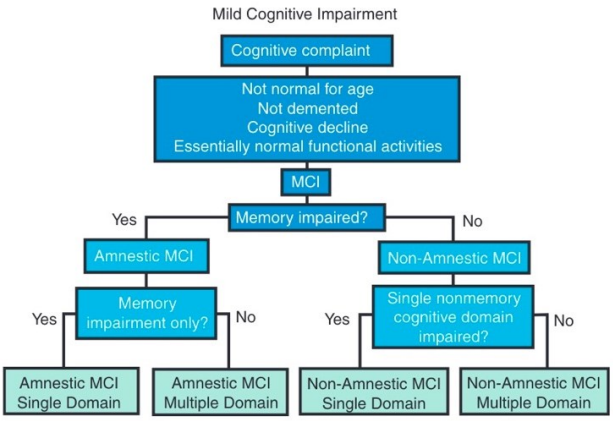
Previous research on the prediction of cognitive impairment due to several subtypes of MCI indicates that patients with amnestic MCI have an 8.6 times higher risk of converting to AD than patients with non-amnestic MCI (Lehrner et al., 2005).

Because MCI may progress to dementia or to another neurodegenerative disease (Petersen & Morris, 2005; Rostdamzadeh & Jessen, 2020), it is highly recommended that

cognitive performance be monitored over a period of 6 to 12 months (Lehrner et al., 2009a, p. 382).

Figure 2

Diagnostic Guidelines for MCI



Note. Flow chart of the diagnostic decision process for the diagnosis of MCI and its proposed clinical subtypes (adapted from mild cognitive impairment as a diagnostic entity by Petersen, 2004, p. 190).

In the general population, the prevalence rate of MCI among people over the age of 60 years is 14-16% (Petersen & Morris, 2005; Petersen et al., 2018), of which approximately 6.5%-6.8% convert to AD annually, compared to 8.1% in the specific population of memory clinics. Within three years, about one-third of patients with MCI convert to AD (Lehrner et al., 2005; Mitchell & Shiri-Feshki, 2009), while almost all patients diagnosed with amnestic MCI proceed to AD within a period of 6 years (Petersen et al., 1999).

MCI can have a variety of causes, such as depression, internal malfunction, sleeping disorder, or drug side effects, but the underlaying cause can also be AD pathology. Advances in biomarker-based screening have enabled the prediction of the likelihood of MCI due to AD and therefore, the identification of MCI patients with and without AD pathology. This progress is considered a “game changer” as it enables the discrimination of low- and high-risk individuals. The presence of abnormal pathology in both amyloid- and tau-markers indicates a 60% chance of progressing to dementia within 3 years and a 90% chance within 5 years (Rostdamzadeh & Jessen, 2020).

In Individuals diagnosed with MCI, researchers have found contractionary prevalence rates for AD pathology. While some prior findings indicated that 40%-50% of people with MCI exhibit AD-like pathology (Petersen et al., 2000; Wagner et al., 2012), a study by Jansen et al. (2015) has reported that patients with MCI have a 20%–30% higher prevalence of abnormal amyloid pathology compared to those with normal objective cognitive performance, as seen in healthy individuals or in individuals with SCD.

Although it is widely recognized that the integration of biomarker-based screening for AD pathology during the prodromal stage of dementia would allow for a proper risk assessment, there are currently no specific guidelines (Rostdamzadeh & Jessen, 2020). Rostdamzadeh and Jessen (2020) strongly advocate for a multidisciplinary approach with scientific guidelines to better evaluate the risk of AD progression, as it has already been proposed by the NIA-AA group, as the prediction of dementia risk is becoming increasingly important for the implementation of preventive strategies.

Dementia

The final stage of the model is called dementia, which is marked by biomarker evidence as well as symptoms of dementia (Jessen et al., 2014).

According to the WHO (2021) dementia manifests itself individually differently, depending on their health condition, the premorbid cognitive functions, and the underlying reasons for the disease.

Once more, the clinical picture of AD develops gradually due to its chronic and progressive nature. Throughout the disease progression, individuals with dementia may also experience changes in mood, behavior, and personality, such as depression, anxiety, apathy, irritability, aggression, and wandering, along with cognitive symptoms (Hugo & Ganguli, 2014).

The symptoms of early-stage dementia are mostly compensated and frequently go unnoticed. The early stage is characterized by forgetfulness, loss of sense of time, and disorientation in familiar surroundings. When middle-stage dementia advances, people may experience confusion while at home, lose track of recent events and people's names, struggle with communication, require help with personal care, and exhibit behavioral changes such as wandering and repetitive questioning. Symptoms in the advanced stage lead to total dependence, inactivity and finally to death (WHO, 2021, 2022).

Prevention of Dementia

There is currently no cure for dementia. Available medical treatments can only manage symptoms and modestly delay progression. Because of the long prodromal period of dementia due to AD, which is estimated to last up to 20 years (Jansen et al., 2015), research has focused on preclinical forms of dementia, the identification of protective and risk factors, and the implementation of preventive as well as health promotion strategies (Clouston et al., 2020).

Prevention considers all actions that are set to avoid disease and its negative consequences with the goal of targeting risk factors, whereas health promotion aims to improve opportunities for health development by enhancing protective factors. The common goal of prevention and health promotion is to improve the individual health as well as the collective health of society.

Prevention can be divided into primary, secondary, and tertiary prevention strategies (Bak, 2023). Primary prevention refers to all public health interventions that aim to prevent disease before it occurs. Secondary prevention strategies intervene when symptoms have already appeared, with the aim of reducing or slowing the disease's progression. For degenerative diseases such as AD, current secondary prevention interventions target the prodromal phase (MCI) of the disease. Tertiary prevention approaches refer to the treatment of manifest disease, with a focus on improving the long-term disease prognosis and reducing disability. Tertiary prevention targets the final stage of the disease continuum (dementia) and includes rehabilitative and palliative interventions (Savica & Petersen, 2011).

Global Dementia Prevention Initiatives

Public health institutions are a component of the healthcare system. Their shared responsibility is to spread objective information about various diseases, including protective and exacerbating factors. Their common goal is to address diseases that have a major global health impact, thereby strengthening healthcare systems and improving the health of society around the world. This common ground involves more than mere research efforts and is imperative (Bloland et al., 2012).

According to Alzheimer's Disease International (ADI) (2021), 75% of individuals affected by dementia remain undiagnosed due to a lack of awareness and the widespread stigma surrounding the condition. Notably, this trend predominantly affects populations residing in low- and middle-income nations. The lack of specialized diagnostic tests

exacerbates the dementia diagnosis gap. Furthermore, ADI reported that these accessibility concerns were further compounded by the COVID-19 pandemic (Barclay & Rees, 2020).

Consequently, ADI advocates for the implementation of culturally sensitive, standardized, and validated digital neurocognitive assessment tools to facilitate lifelong monitoring of cognitive performance. This approach not only enhances a timelier diagnosis but also provides better care and access to treatment worldwide. Additionally, ADI recommends the integration of population-wide annual brain health screenings for individuals aged 50 years and older, along with the surveillance of the preclinical stage of dementia. ADI facilitates prompt and accurate diagnosis and tries to optimize treatment outcomes (Gauthier et al., 2021).

The Global Action Plan on the Public Health Response to Dementia (2017-2025) is a framework established by the WHO. It addresses dementia's growing global burden. It aims to improve the lives of people affected by dementia and reduce the impact of dementia on individuals, families, and societies. Key tenets of this comprehensive approach include enhancing awareness and understanding of dementia to reduce stigma and foster early detection, as well as promoting healthier lifestyles and reducing risk factors associated with the disease. Moreover, the plan emphasizes the need to improve access to accurate diagnosis, evidence-based treatments, and quality care for individuals with AD and their caregivers. Furthermore, the plan underscores the importance of research and innovation to deepen the understanding of dementia and its underlying causes, the development of new treatments, and the refinement of care practices. Lastly, it focuses on the establishment of monitoring mechanisms and the constant evaluation of interventions. Overall, the WHO initiative aims to encourage international cooperation to address the multiple challenges posed by the neurodegenerative syndrome and to improve the effectiveness of public health responses worldwide (WHO, 2017).

The WHO's (2017) designation of dementia as a global health priority has encouraged several research efforts, such as the Horizon 2020 (2021) initiative "LETHE" undertaken by the European Union (EU). The longitudinal project, in which the Department of Neurology of the Medical University of Vienna is also involved, seeks to create a predictive model based on big data for the early detection of dementia using smart technologies. Additionally, the project aims to develop an AI-driven digital profiling methodology to tailor personalized lifestyle interventions that aim to ameliorate cognitive decline (Hanke et al., 2022).

Protective Factors and Risk Factors

Research has shown that lifestyle interventions, such as exercise and cognitive training, can delay or prevent the onset of the progressive disease in people with SCD or MCI with AD pathology (Livingston et al., 2020a; Ngandu et al., 2015).

Experts distinguish between modifiable and unmodifiable risk factors. While age is the most prominent unmodifiable risk factor (Savica & Petersen, 2011), the Lancet Commission has approved 12 modifiable risk factors involving smoking, alcohol abuse, physical inactivity, obesity, an unhealthy diet, high blood pressure, high cholesterol, high blood sugar, depression, social isolation, low educational level, cognitive inactivity, air pollution, hearing impairment, and traumatic brain injury. It is estimated that 30-50% (Hoffmann et al., 2022; Livingston et al., 2020b; Pertl, 2015) of all dementia cases are linked to these prior mentioned modifiable determinants and could consequently be prevented (Hoffmann et al., 2022).

A meta-analysis by Norton et al. (2014) suggests a causal association between proposed risk factors and AD, implying that preventive intervention at the appropriate age is highly effective. From a global perspective, the research points out that reducing the identified lifestyle risk factors by 10% every ten years could significantly decrease the incidence rate by 16.8 million out of 33.9 million by 2050.

Hence, the potential for prevention is high. However, it could be higher in low-income countries, where there is a need to catch up in terms of medical care, education, and nutrition (Livingston et al., 2020b). Due to its high worldwide prevalence, one in five cases of AD is associated with low education, and one in ten cases is related to smoking and depression (Norton et al., 2014). Prior research has shown that the risk factors associated with AD vary by region (Hoffmann et al., 2022) and differ by sex and ethnicity (Nianogo et al., 2022), with women (Nichols et al., 2022) and individuals from minority ethnic groups, especially those who identify as black, Native Americans, and individuals of Hispanic or Latino background, having a higher risk of dementia (Livingston et al., 2020b). Whereas, in Western countries, most cases of dementia are attributed to middle-aged physical inactivity (Norton et al., 2014).

Governments should raise awareness that every individual can reduce their own dementia risk by living a healthy lifestyle and staying cognitively, physically, and socially active throughout life (Livingston et al., 2020b).

Neuropsychological Assessment of Dementia

Neuropsychological assessment aims to clarify differential diagnosis and etiology of the dementia syndrome. It is the basis for further multi-professional treatment measures and is resource oriented (Lehrner et al., 2009b, p. 389).

According to Lehrner et al. (2009b, pp. 388 - 389), the clinical diagnosis of dementia is defined by a multi-stage process that includes:

- Clinical examination of:
  - Personal and medical history,
  - Family and social history,
  - External anamnesis,
  - Neurological, and psychiatric status,
  - Complete blood test,
  - Genetics (Apolipoprotein E) (optional),
  - Cerebrospinal fluid testing (CFU) to detect inflammatory processes, or pathological protein modifications (optional).
- Imaging procedures:
  - CT and MRT: to exclude structural brain changes (e.g. infarction, atrophy). In late-stage dementia due to AD, an internal and external enlargement of the ventricular space, and atrophy of the hippocampus can usually be detected.
  - EEG/SPECT/PET have a high differential diagnostic value in the detection of dementia-related brain diseases, or in other etiologies.
- Neuropsychological assessment:
  - Objective quantification of cognitive performance,
  - Differentiation between normal aging and pathological processes,
  - Differential diagnosis of various mental disorders (e.g. depression),
  - Assessment of daily functioning, independence, and judgment,
  - Assessment of psychopathology and neuropsychological changes along the disease continuum (Lehrner et al., 2009b, pp. 388 - 389).

The S3 guidelines for the diagnosis of dementia recommend that people suspected of having mild dementia or mild cognitive impairment should undergo a detailed neuropsychological assessment, even if brief screening tests such as the MMSE produce

unremarkable findings. Neuropsychological diagnosis should include a standardized assessment of the episodic memory, as the most prominent symptom of AD manifests itself as impaired ability to learn new information. Additionally, episodic memory should be objectively assessed along with other memory, attention, and executive functions, as well as language and visuospatial abilities, and should include a differential diagnosis (particularly depression). Furthermore, the assessment of daily functioning and an evaluation of the premorbid intelligence level is highly recommended as dementia manifests itself differently in individuals based on their premorbid cognitive function level (DGN e. V. & DGPPN e. V., 2023).

Task characteristics for the estimation of premorbid intelligence function are considered valid if scores remain stable along the disease continuum. Reading skills are commonly assessed because they are considered overlearned skills that are maintained at a high level, although the onset of cognitive deterioration has already taken place (Marier et al., 2023).

However, a study by Marier et al. (2023) evaluated a widely used premorbid intelligence task, namely AmNART, in which patients were asked to read irregular words aloud. They found that irregular words reading performance declines along the disease continuum and that it was, among others, strongly associated with reduced hippocampal volume, which is the primary brain area affected in AD. These findings indicate that impaired irregular word reading is an indicator of semantic decline in AD rather than an estimate of premorbid intelligence level. The findings have clinical implications, as it is concluded that the use of irregular word reading in the diagnostic process may lead to an underestimation of premorbid intelligence level and consequently to an underestimation of cognitive impairment. The authors recommend that task characteristics should represent the concept of crystallized intelligence, which remains stable over the course of the disease progression. Given that standardized neuropsychological assessment of cognitive decline is considered critical for determining a diagnosis of dementia along the disease continuum, clinicians should rely on either comparison of the patient's performance with age-, sex-, and education-adjusted normative performance in healthy individuals or on the comparison of current performance with premorbid performance (Marier et al., 2023).

Hugo and Ganguli (2014) provide examples of typical task characteristics for the neuropsychological assessment of cognitive function relevant to dementia diagnosis based on the DSM-V (see Table 1).

Table 1

Neurocognitive Assessment of Dementia: Proposed Task Characteristics based on the DSM-V

Cognitive Domains	**Task characteristics**
Complex Attention	• **Maintenance of attention:** pressing a button upon hearing a tone. • **Selective attention**: counting letters, while listening to letters or numbers • **Divided attention:** tapping a button quickly, while reading a story. • **Processing speed:** completing tasks with a time limit
Executive Function	• **Planning:** solving maze puzzles, interpreting sequential pictures or arranging objects in a sequence. • **Decision making:** engaging in a gambling simulation. • **Working memory:** holding information for a short time, e.g. repeating a list of numbers backward. • **Feedback utilization:** using feedback to adjust performance. • **Inhibition:** Overriding habits to choose the correct but more complex and less obvious solution, e.g., reading the written names of colors rather than naming the color in which they are printed. • **Cognitive flexibility:** Shifting between different tasks or concepts

	e.g., switch between numbers and letters.

Learning and Memory	• **Immediate memory:** Repeating a list of words.
	• **Recent memory:**
	  • **Free recall:** Repeating a list of words.
	  • **Cued recall:** recognizing items from a list.
	  • **Recognition:** identifying whether an item was previously presented.
	• **Semantic memory:** recalling well-known facts.
	• **Autobiographical memory:** recalling personal events.
	• **Implicit (procedural) memory:** recalling skills to conduct procedures.

Language	• **Expressive language:** assessing fluency for words by naming as many words with a given initial letter, or objects, or pictures in a given category (e.g. animals) with a time limit. • **Grammar and syntax:** identifying errors in grammar or syntax. • **Receptive language:** comprehending or defining words.
Perceptuomotor Function	• **Visuoconstructional:** e.g., drawing or copying figures. • **Perceptuomotor:** placing blocks or pegs in appropriate slots.

	• **Praxis:** mimicking gestures such as "salute" or actions such as "use the hammer." • **Gnosis:** e.g., recognizing faces or colors.
Social Cognition	• **Recognize emotions:** identifying emotions e.g., happy, sad, scared, angry faces. • **Theory of mind:** interpreting the thoughts or intentions of characters in stories.

Note. As outlined in Hugo and Ganguli's (2014) paper the table shows typical task characteristics for the assessment of dementia based on the DSM-V criteria, providing a comprehensive framework for clinicians and researchers (adapted from Dementia and cognitive impairment: Epidemiology, diagnosis, and treatment by Hugo & Ganguli, 2014, p. 24).

Digital Neurocognitive Assessment of Cognitive Decline.

Conventional approaches to the diagnosis of AD-related cognitive impairment require specialized memory clinics where at-risk patients must undergo a detailed annual check-up (Perin et al., 2020), which is costly to the health care system and burdensome for those who are affected (Gates & Kochan, 2015; Geddes et al., 2020; Myers et al., 2022). On the one hand, the COVID-19 pandemic and its associated limited access to neuropsychological testing (Myers et al., 2022) as well as the shortage of health care professionals, and on the other hand, the predicted increasing prevalence rate of dementia underline the urgent need for affordable, valid, efficient, and accessible neurocognitive tests for the early detection of organic deterioration of cognitive function (Gates & Kochan, 2015; Geddes et al., 2020).

The advent of the Internet led to the proliferation of new technologies, such as smartphones or portable computers (Geddes et al., 2020). The worldwide widespread use of smartphones across all social classes has opened new paths in neuropsychological diagnostics, especially in the field of remote health care (Pratap et al., 2020).

One example for the first digital neurocognitive assessment batteries is the Cambridge Neuropsychological Test Automated Battery (CANTAB), which was initially invented to evaluate neurodegenerative disorders. CANTAB paved the way for other researchers to

develop digital assessment batteries for the timely detection of cognitive dysfunction (Robbins et al., 1994).

One major advantage of digital testing is that it allows the assessment of many aspects of cognitive performance simultaneously and objectively within milliseconds in a self-administered manner without the need for a highly trained test administrator (Geddes et al., 2020). This fact provides a great opportunity for population-level research on cognition, as it allows for the collection of data from many heterogeneous people regardless of demographics (Belleville et al., 2023).

Moreover, digital tools have been shown to be a valuable component not only in the early detection of cognitive impairment, but also in cognitive training to enhance or maintain cognitive health (Grigoryeva et al., 2024).

Nevertheless, the unsupervised testing environment also presents challenges because the experimenter is unable to control and standardize the environment, which can lead to confounding variables (Belleville et al., 2023). However, within-person studies have shown that performance does not differ significantly when comparing different testing environments (remote vs. face-to-face) in the elderly population (Cyr et al., 2021). Moreover, compliance issues, or technical problems with the application cannot be counteracted by the test instructor, which can consequently result in a loss of motivation and data. As digital testing becomes increasingly relevant, the potential learning effects associated with this technology should be examined in future studies (Gates & Kochan, 2015).

Despite the growing number of digital assessment tools designed to measure cognitive decline in neurodegenerative diseases, there is currently a lack of guidelines on which digital tool to use in clinical practice (Geddes et al., 2020). Therefore, Belleville et al. (2023) conducted a review of the current state of good scientific research in the field of digital neurocognitive testing. The authors emphasize that digital tools have the potential to adress the diagnostic gap but also highlight the lack of documentation of psychometric properties, particularly in terms of construct validity and robust norms (Belleville et al., 2023).

As a matter of fact, the development of a digital cognitive test, which allows for an individualized baseline estimation of cognitive performance and the assessment of subtle but significant cognitive changes over time, is critical to closing the diagnosis gap. However, one of the major challenges is to develop a valid and reliable digital assessment battery that is easy to self-administer, sensitive to subtle changes in cognition, and represents all relevant

cognitive domains affected by AD as defined by the DSM-V (Geddes et al., 2020; Myers et al., 2022).

Addressing this need, the Medical University of Vienna has developed the International Neurocognitive Test Profile (INCP), a digital, remote, and self-administered neurocognitive test battery that aims to screen for dementia in its prodromal stage by facilitating individualized baseline estimation of cognition and by facilitating self-monitoring of subtle changes in cognitive performance. With its 18 subtests and 24 parallel versions, practice effects are minimized. Therefore, the target group is neurologically healthy individuals who describe subjective cognitive decline and who are interested in self-monitoring their cognitive function at home. The subtests are assigned to the different cognitive domains affected by dementia as defined by the DSM-V, including complex attention, executive functions, learning and memory, language, and social cognition. However, there are currently no tasks available that measure perceptual motor skills. The test battery is still under development and is undergoing constant evaluation. Currently, there are no studies available that provide information on the psychometric criteria (Lehrner, 2021b).

Study Aim

The objective of this study is to fill the research gap by conducting an interim evaluation of the validity of the INCP and providing initial insights into the associations between various variables of the INCP. Furthermore, performance on the web-administered test (INCP) was also associated with performance on a validated and standardized paper-and-pencil assessment of equivalent construct (NTBV). The methodology employed follows previous research in this field (Lehrner et al., 2006; Rentz et al., 2016; Stricker et al., 2022).

Main Questions of Research

Is there an association between variables posed in the newly developed web-based neurocognitive assessment battery (INCP)?

- $H1_1$: The tasks for assessing the cognitive domain of complex attention in the INCP (DST) are correlated.
- $H1_2$: The tasks for assessing the cognitive domain of executive function in the INCP (TLT-s) are correlated.
- $H1_3$: The tasks for assessing the cognitive domain of learning and memory posed in the INCP (FPT, FACE, CITY) are correlated.
- $H1_4$: The tasks for assessing the cognitive domain of language posed in the INCP (AVT, VVT, INT) are correlated.

Does the INCP correlate with conventional neuropsychological measures of the same construct (NTBV)?

- $H2_1$: The tasks for assessing the cognitive domain of attention posed in the NTBV (Digit Symbol Subtest, AKT, C.I., TMT B) and the tasks for assessing the cognitive domain of complex attention posed in the INCP (DST) are correlated.
- $H2_2$: The tasks for assessing the cognitive domain of executive function posed in the NTBV (TMT A, Five Point Test, Planning Maze Test, Stroop Test, C.I., PWT) and the tasks posed in the INCP (TLT-s) are correlated.
- $H2_3$: The tasks for assessing the cognitive domain of memory posed in the NTBV (VSRT) and the tasks for assessing the cognitive domain of learning and memory posed in the INCP (FPT, FACE, CITY) are correlated.
- $H2_4$: The tasks for assessing the cognitive domain of language posed in the NTBV (BNT, SWT), and the tasks posed in the INCP (AVT, VVT, INT) are correlated.

Exploratory Questions of Research

Additionally, associations between the social cognition task EFT-s and other subtests of the INCP were displayed, and an exploratory factor analysis was conducted to gain initial insights into whether the internal structure of the INCP accurately represents five (out of six) cognitive domains outlined in the DSM-V that are relevant for the diagnosis of dementia.

Furthermore, the study aims to explore the discriminatory power of the INCP by examining whether the subtests of the INCP can differentiate between neurologically healthy individuals (HC) and those with mild cognitive impairment (MCI).

Methods

The present study is part of the already approved study EK1950/2022 and was conducted in accordance with the Good Scientific Practice and the Good Clinical Practice (GCP) proposed by the Medical University of Vienna. To protect patients' privacy and confidentiality the personal data as well as the signed declaration of consent are stored with restricted access in the Department of Neurology of the General Hospital of Vienna (AKH). The data was pseudonymized, standardized and checked for plausibility. Study subjects didn't receive a direct personal benefit, except for detailed feedback on their cognitive performance. Furthermore, the free digital brain-training game "Lumosity" was recommended to all participants. As the examination period took place during the COVID-19 pandemic, the federally mandated COVID-19 pandemic resources were respected, and no resources that were required for pandemic response were used. There were no risks related to participation in the study. Participants were not exposed to any additional infection risk.

Study Design

In the early phase of clinical development, pilot studies are considered particularly efficient as they allow for flexibility (Arain et al., 2010). A pilot study has an adaptive design (Chow & Chang, 2008) with the aim of evaluating the feasibility of a methodology (Arain et al., 2010) by testing various aspects (Chow & Chang, 2008). Their common goal is to enhance the likelihood of clinical benefits in the main study. Pilot studies are essential to make sure that larger-scaled randomized and controlled trials are both robust, practical, and economically feasible (Arnold et al., 2009).

The INCP is still considered an experimental test battery under development. The lack of prior studies in the field led to limited experience with the INCP. As a result, the present study was a pilot study with a prospective, cross-sectional, explorative design. Consequently, normal case planning was not conducted. The present thesis aimed to gain first insights into the associations between the variables and the underlying structure of the digital neurocognitive test battery, INCP.

Due to technical errors, the composition of the digital test battery was constantly adjusted, resulting in different sample sizes per subtest. Because the study is a feasibility

study, test subjects were asked for feedback, which was constantly incorporated into the assessment procedure to improve operability and usability.

Inclusion & Exclusion Criteria

The exclusion criteria were based on the study of Rosas et al. (2022), which was assessed during the diagnostic interview at the beginning of the examination. Participants were excluded from the study if one or more of the subsequent criteria applied:

- Any known cerebral vascular pathology or severe head injury.
- Any known current mental diagnosis according to the ICD-10.
- Any known diagnosis of dementia according to the DSM-V.
- Any known medical condition that leads to severe cognitive deterioration.
- (Under 50 years of age) The COVID-19 pandemic and its associated issues in recruitment led to the dismission of the original target population because older adults were considered particularly vulnerable. Therefore, exclusion criterion 5 was replaced, and all age groups were included in the present study.

Sample Characteristics and Recruitment

From December 2022 to October 2023, 71 study subjects participated in the present study; 7 study participants didn't meet the inclusion criteria (criterion 4) and 11 participants didn't finish enough subtests and were not considered in the present study (see Figure 3).

Finally, the total sample consisted of 53 participants ($M_{\text{Alter}} = 58.06$; $SD_{\text{Alter}} = 17.64$; 69.8% female subjects). Recruited were neurologically healthy controls ($N_{HC} = 41$) and patients diagnosed with mild cognitive impairment ($N_{MCI} = 12$). According to the central limit theorem, a sample of $N = 30$ is normally distributed. Due to the strict COVID-19 restrictions in the clinical setting, the originally planned sample size could not be realized, which is why the central limit theorem was violated within the MCI group. Because the present research is a pilot study, this is not considered a major problem.

The clinical population (MCI) was recruited from the Department of Neurology at the Medical University of Vienna. The healthy control group (HC) involved caregivers of patients from the Department of Neurology and friends who were asked to participate in the study. Flyers were posted in the AKH and in nearby supermarkets to increase participation in the HC-group.

Assessment Procedure

Participants completed two different assessment batteries at a single time. The examination took approximately 240 minutes and was conducted at the Department of Neurology (AKH). All study subjects received standardized instructions in the same order and were instructed to use headphones to avoid distractions from parallel ongoing examinations. The clinical flowchart (see Figure 3) shows the assessment procedure, which is described in the following sequence:

Each participant provided written consent to participate in the study and was informed of the right to drop out at any time. Additionally, every participant underwent a medical interview to evaluate their physical and mental health. Demographic data was also requested. Participants were required to handle the tablet by themselves. Therefore, listening and reading skills were assessed.

Following this, participants underwent a cognitive screening, a depression screening, and a verbal IQ-test. The cognitive screening determined whether the assessment batteries were submitted in a long or short version.

In a further step, all participants operated on the well-established paper-pencil test (NTBV), which evaluates the cognitive domains of attention, executive functioning memory, and language. Afterwards, all subjects worked on the newly developed digital test (INCP), which encompasses tasks for the assessment of the cognitive domains of complex attention, executive function, learning and memory, language, and social cognition. Currently, there is no empirical research on whether working on the NTBV first leads to learning effects that may be confounded with the performance on the INCP.

While the main question of research was addressed by 53 participants, the exploratory question of research compared NTBV scores and INCP scores across groups (HC and MCI). The initial group assignment method based on the NTBV results following the criteria by Petersen (2004) (using an age-related mean score below -1.50 standard deviation in a single domain or in multiple cognitive domains in a neurocognitive assessment battery (NTBV) was dismissed and replaced by screening results only. In order to be assigned to the HC group, a participant had to score at least 27 points on the MMSE, over 8 points on the VVT 3.0 Screening subtest, and over 85 IQ-points on the WST.

In the clinical population, all participants were included except for those who were preliminary excluded because they didn’t meet the inclusion criteria in the clinical interview.

Participants who were thought to be healthy but did not meet the screening criteria for healthy controls were registered in the clinical population.

One participant who initially met the criteria for the HC group was assigned to the MCI group because of its conspicuous behavior. The recruit seemed to have forgotten that the examination had already taken place and came back two more times. Additionally, the participant complained about SCD. Consequently, the subsample MCI consisted of $N_{MCI} = 12$.

Figure 3

Clinical Trial and Exclusion Process Flow Chart

Timeline	**Assessment Procedure**	**Description of Task Characteristics**	**Exclusion Criteria**	**Sample Size**
1.	Informed Consent	Information about the study and the right to drop out of the study at any time, time to clarify questions, signing the consent form	<u>Exclusion if:</u> signature is missing	
2.	Clinical Interview	Question about mental status, medical condition	<u>Exclusion if:</u> • any known cerebral vascular pathology or severe head injury • any known current mental diagnosis according to the ICD-10 • any known diagnosis of dementia according to DSM-V • any known medical condition that leads to severe cognitive deterioration	$N_{excluded} = 7$
3.	Demographic Data	Age, gender, total years of schooling, highest school education		

<table>
<tr><td>4.</td><td>Cognitive Screening</td><td>MMSE
VVT 3.0
WST
(BDI II)</td><td>Group Assignment
HC-Group
MMSE > 27
VVT 3.0 > 8
WST IQ > 85

MCI-Group
Did not meet the criteria for the HC-group.</td></tr>
<tr><td>5.</td><td>NTBV</td><td>NTBV long version (MMSE > 24)
HAWIE-R DIGIT SYMBOL TEST
AKT
C.I.
TMT B
TMT A
FIVE POINT TEST
PLANNING MAZE TEST
STROOP TEST
PWT
VSRT
BNT
SWT</td><td>NTBV short version (MMSE ≤ 24)
AKT
C.I.
TMT A
PLANNING MAZE TEST
PWT Subtest “F”
VSRT
BNT
SWT Subtest “Animals”</td></tr>
<tr><td>6.</td><td>INCP</td><td>Long Version
VVT
FPT (1)
CITY
BDI II
WST
DST
INT
FPT (2)
TLT-s
AVT
EFT-s
FACE</td><td>$N_{excluded} = 11$</td></tr>
<tr><td>7.</td><td>Delayed Cognitive Screening</td><td>VVT 3.0 Delayed recall task</td><td rowspan="2"></td></tr>
<tr><td>8.</td><td>Feedback</td><td>Feedback Forum</td></tr>
</table>

Instruments

In the following sequence, all instruments used in the examination are described. A detailed overview of all collected variables can be found in the appendix (see Table A1).

Demographic Data

Demographic data were collected to describe the total sample, the subsample, and for further statistical analysis. Age was calculated from the reported date of birth. Gender and years of formal education were self-reported.

Screening

All participants underwent a clinic interview regarding their physical and mental health. If one or more exclusion criteria were applied, subjects were excluded from the study, either preliminarily or retrospectively.

To ensure the absence of severe cognitive impairment (Folstein et al., 1975), cognitive screening tests were implemented. Following the medical health screening, participants were required to complete the Mini-Mental State Examination (MMSE), which was developed by Folstein et al. (1975) and is considered a well-established and economically efficient dementia screening tool (Fok et al., 2023), as it takes approximately 5-10 minutes, including 30 items that aim to screen for impaired orientation, attention, memory, object recognition, ability to follow commands, verbal fluency, and visuo-constructive function (Folstein et al., 1975). The authors highlight that the MMSE is considered a valid (r = .77) and reliable (r = .89) screening tool to ensure the absence of severe cognitive impairment. However, it does not replace a detailed neurocognitive assessment battery, which is necessary for the final diagnosis of dementia (Folstein et al., 1975). The subjects received one point for each correct answer; the score ranged from 0 to 30, with higher scores indicating better performance (Folstein et al., 1975). Based on the MMSE, the long or short version of the NTBV assessment battery was submitted.

A widely recognized and an effective method for the early diagnosis of neurodegenerative disorders is the analysis of visuo-constructive function (Valencia & Lehrner, 2018), therefore, all participants had to operate on the paper-pencil test VVT 3.0, namely the Vienna Visuo-Constructive Test 3.0 (Lehrner, 2021c). In this test, participants were instructed to copy three figures as accurately as possible. As a final task, each study participant completed the VVT 3.0 Delayed Recall task, in which participants were asked to draw the three previously copied figures from the first part once again from memory

(Heidinger & Lehrner, 2020). For this study, only the VVT 3.0 screening was relevant, with scores ranging from 0 to 10, with 10 indicating better performance (Lehrner, 2021c). A prior study examined the psychometric criteria of the VVT 3.0 and stated satisfactory internal consistency with Cronbach's alpha α =.84 for healthy controls and α = .93 for the patient group (Lehrner et al., 2015), another study could confirm these findings (Numrich, 2017).

The Wortschatztest (WST), developed by Schmidt and Metzler (1992), assesses verbal intelligence level and language comprehension, and is widely used to estimate premorbid intelligence level before mild to moderate brain-organic impairment. The WST is a recognition task consisting of 40 items that include one target item and five distractor items, in which participants had to identify the correct German word. The split-half reliability is r = .95 and the internal consistency is Cronbach's Alpha α = .94 (Schmidt & Metzler, 1992). A WST-IQ above 85 points was required for the assignment to the HC-group.

To screen for depressive symptoms, the German version of the Beck Depression Inventory (BDI-II) was submitted (Hautzinger et al., 2006). It consists of 21 items, rated on a four-point Likert scale. An sample item is from 0 = *"Ich bin nicht traurig"* to 3 = *Ich bin so traurig, dass ich es nicht aushalte."* The internal consistency is high with α = .93 for persons diagnosed with depression and α = .90 for healthy controls.

The cognitive screening was used for the group assignment, depression screening, and the demographic data were used for the description of the study participants.

In another step, all participants completed two neurocognitive assessment batteries in the same order, namely the Neuropsychological Test Battery Vienna (NTBV) and the International Neurocognitive Test Profile (INCP).

In the following sequence, both assessment batteries are described.

The Neuropsychological Test Battery Vienna (NTBV)

The valid and well-established assessment battery, namely the Neuropsychological Test Battery Vienna (NTBV) was installed as a comparator assessment battery to examine the construct validity of the INCP. The NTBV was developed at the Medical University of Vienna (Lehrner et al., 2007) to assess cognitive performance commonly affected by dementia. Sensitivity is given and allows for diagnosing dementia in its prodromal stage (Lehrner et al., 2007). The test battery consists of various subtests. Based on cluster analysis the subtests were assigned to one of the various cognitive domains affected by dementia

proposed by DSM-V (Pusswald et al., 2013). The tasks assess the cognitive domains of attention, language, executive function, and memory.

Psychometric criteria are available. Reliability is determined. Cronbach's alpha for the different diagnostic groups ranges from α = .87 to α = .89, with a high internal consistency (α = .86) in the total sample. Consistency over time ranges from .69 to .94, indicating questionable to excellent reliability (Lehrner, 2021a). Construct validity is determined (Rosas et al., 2022).

The NTBV can be submitted in a short or long version, based on the results of the dementia screening tool MMSE. All subtests were administered with the standardized test instructions provided by the test manual (Lehrner, 2021a; Rosas et al., 2022).

In the following sequence, the domain-specific subtests of the NTBV are described.

To assess the cognitive domain of attention, the NTBV submits 5 subtests:

Alters-Konzentrations-Test (AKT). During the AKT, participants were required to identify and strike out semicircles based on predetermined criteria. The task had a time limit and was terminated after 120 seconds.

The score "Time" was the time to complete the task, correctly answered items were summed up as "Right", wrongly answered items were summed up as "Mistakes". The total test score "G" (maximum 20) was calculated by using the following formula:

$$G = \frac{35 + \text{Right} - False}{Time}$$

For the score "Total/Time", the total test score G was divided by the time needed to complete the task.

HAWIE-R: The Digit-Symbol Test. In this test, participants were presented with a sheet of boxes and numbers (1-9), coded with symbols. Participants were instructed to match numbers with corresponding symbols. Subjects had 90 seconds to finish the task. The raw item score was calculated from the number of correctly solved items. Wrong symbols were subtracted from the total sum score, with a higher score indicating a better performance.

The symbol counting task from **The Cerebral Insufficiency Test (C.I.)**, comprises a sheet with three different symbols. The proband was instructed to count all squares (in total 44 squares). The task was cancelled after 1 minute.

The test score was calculated by the time needed in seconds, and for every mistake one second was added.

The Trail Making Test (TMT B). This task involves connecting letters and numbers in an ascending order (1A, 2B, etc.). The time taken to complete the task was measured and was cancelled after 300 seconds. The time needed to complete the task was the sum score, with lower scores indicating better performance. Additionally, the time of TMT A was subtracted from the time of the TMT-B.

The NTBV assesses the cognitive domain of executive function using six subtests, including the following:

The Trail Making Test (TMT A). The TMT-A task is similar to the TMT-B task. The task characteristic was to connect numbers (1-25) in an ascending order as fast as they could. The time needed was recorded, and the task was terminated after 180 seconds, the time needed to complete the task in seconds was the total sum score, with lower scores indicating a better performance.

The Five-Point Test. This task involves a sheet with boxes, each of which has five dots on it. Whitin 3 minutes, the probands were instructed to draw as many different patterns as possible. The number of correct patterns was calculated for the sum score, with higher scores indicating better performance. The number of repetitions was also calculated, with lower scores indicating a better performance.

The Planning Maze Test (NAI). In this task, a labyrinth was presented. The test subject was instructed to find the easiest way out without reaching a dead end. The time was measured, and the task was cancelled after 2 minutes.

The total score was calculated by using the following formula, a maximum of 16 errors was possible.

$$\frac{16 - \text{errors}}{\text{time needed}}$$

The Stroop Test (NAI). The Stroop test consisted of two tasks: the Color-Word-Time-Colors and the Color-Word-Time-Words Subtest. Participants were given a sheet of colors and asked to name them as quickly as possible while the time was measured. The time limit for this task is 60 seconds.

The total score was represented by the time needed (in seconds), with lower scores indicating a better performance.

During the second task, the participants were shown a sheet consisting of written words in various colors and were instructed to name the word rather than the color. The task was cancelled after two minutes.

The time needed (in seconds) and number of errors (maximum 16 errors possible) were needed for the calculation of the total score by using the following formula:

$$\frac{16 - \text{errors}}{\text{time needed "color} - \text{words"}}$$

The C.I. Interference Test. The CI Interference Test involves a sheet with the letters A and B. The test person was instructed to say A instead of B and B instead of A as quickly as possible. The time was measured, and the task was terminated after 1 minute.

The total score was calculated by using the following formula:

$$\frac{34 - \text{errors}}{\text{time needed}}$$

Phonematic verbal fluency test (PWT). An initial letter was presented by the instructor. The task assigned to the participant was to list as many words as possible that started with this initial letter. There were three rounds to the task, each lasting sixty seconds. The sum score was built by the number of correct answers for each round, with a higher score indicating a better performance. Repetitions and false words were subtracted from the total score.

To evaluate language ability, two subtests were administered, which are namely:

Semantic Verbal Fluency Test (SWT). Participants were to recall as many tools, animals, or grocery store items as they could during the three rounds of the task. The task had a time limit set by the instructor, and it ended after three minutes. The sum score was built for each round by summing up the correct answers; mistakes were subtracted from the total score, with a higher score indicating a better performance.

The Boston Naming Test (BNT). In this task, the instructor showed the participants cards with objects and asked them to identify the object by name. The test person was given ten seconds to name each object. The time limit was set to one minute.

The number of correct answers was the total score; mistakes were subtracted, with higher scores indicating a better performance.

To assess the cognitive domain of memory, the NTBV submits one task, including several subtests:

The Verbal Selective Reminding Test (VSRT). The VSRT, with its subtests of immediate recall, total recall, delayed recall, and recognition, was used to evaluate episodic memory.

During the immediate recall subtest, participants were shown cards with 15 grocery items in a specific order and were instructed to remember as many items as possible. The proband was then asked to repeat the grocery list as completely as possible. During the second, third, fourth, and fifth rounds, the instructor presented the items that had not been mentioned in the previous trial. Participants were instructed to recall the entire list during each round. A total of five rounds were conducted, but the task was terminated once the test subject successfully recalled the entire grocery list.

A delayed recall and recognition trial was conducted 20 minutes after the last trial. During the delayed recall task, the participants were asked to reproduce as many items as possible from the grocery list in a random order. In the recognition task, the instructor read a grocery list aloud, and the participant had to say whether each item was on the previously shown grocery list.

The number of correct answers in the first trial is referred to as “Immediate Recall”, with higher scores indicating better performance. The learning performance was calculated by summing up the number of correct answers from the 1st to the 5th trial (“Total Recall”), with higher scores indicating a better performance. The score for “Delayed Recall” was derived from the number of correct answers in the delayed task, with higher scores indicating a better performance. “Intrusions” were considered items that were not part of the grocery list, but the participant named them anyway, with lower scores indicating a better performance. The sum of wrongly remembered items divided by 2 was considered “False-Positive”, with lower scores indicating better performance. The “Recognition” score was calculated by subtracting “False-Positive” from “Recognized Items”, with higher scores indicating a better performance.

The International Neurocognitive Test Profile (INCP)

The International Neurocognitive Test Profile (INCP) is currently under development and under constant evaluation. Normative data and test quality criteria are not available, but they are continuously examined and improved. Because the INCP is constructed in accordance with the DSM-V for the diagnosis of dementia and is similar to another validated neurocognitive assessment battery (NTBV) in this field, construct validity may be implied. Currently, there is no prior research on the INCP's internal consistency. INCP 4.0 consists of 18 subtests, which aim to assess all six neurocognitive domains affected by dementia as proposed by the DSM-V guidelines (see Figure 4), which include language, learning and memory, executive function, complex attention, social cognition, and perceptual-motor function. However, the INCP does not comprise a test that captures perceptual-motor functions, but the development of subtests that assess this cognitive domain is planned (Lehrner, 2021b).

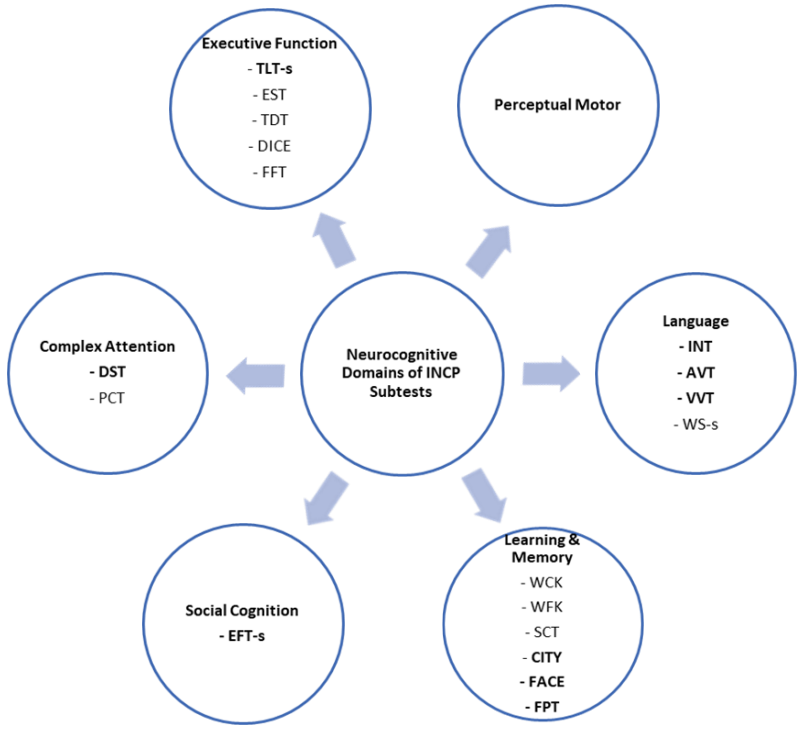
Due to the ongoing development of the INCP, this interim evaluation study examined only 9 out of 18 domain-specific subtests.

The software for the application was developed by psimistriGmbH, which is a spin-off of the Medical University of Vienna (Psimistri, 2022).

The INCP differs slightly from the NTBV, as the NTBV is based on the DSM-IV, whereas the INCP is based on the DSM-V guidelines. Additionally, the NTBV does not include items for the evaluation of the cognitive domain of social cognition. While the INCP measures both learning and memory, the NTBV assesses only memory. Furthermore, the INCP differentiates attention based on complexity.

Figure 4

INCP Subtests for the Evaluation of Cognitive Domains commonly affected by AD proposed by DSM-V



Note. This figure demonstrates the neurocognitive domains proposed by the DSM-V, commonly affected in dementia, and their assigned INCP-subtests. The subtests examined in the present study are highlighted in color (adapted from the International Neurocognitive Profile (INCP) – A web-based neurocognitive assessment battery by Lehrner, 2021b, p.11).

The following section describes the presently submitted INCP subtests.

The cognitive domain of memory and learning is represented by the following subtests:

City Identification Test (CITY). The CITY task evaluates semantic memory by requiring participants to match the capital's name with the corresponding country's name. This multiple-choice subtest consisted of 16 items, and each item encompassed three distractor items. The task took approximately 2 – 5 minutes to complete. The sum score was calculated based on the number of correct responses, ranging from 0 to 16, with a higher score indicating better performance.

Face Identification Test (FACE). The FACE subtest comprised 16 items and was submitted in a multiple-choice format. Photographs of celebrities from the past century were

presented to the participants. They were asked to match the presented photo with the corresponding name. The sum score ranged from 0 to 16, whereby higher scores indicate a better performance. The FACE task took approximately 2-5 minutes to complete and was originally designed to assess semantic memory.

Faces-Pairs-Test (FPT). The Faces-Pairs-Test (FPT) is a task that is intended to assess episodic memory performance and consisted of three subtests. Participants were presented with photos of faces and were instructed to match them based on given rules.

The Faces-Pairs Test -Forced Choice Two Alternative Immediate Recognition (FPT-FCTAIR) presented 20 pairs of items in 2 rounds. In the first round, participants had to learn which two faces go together. In the second round, probands had to correctly match the previously learned pairs of faces. The score was calculated by summing up the number of all correct matches for both rounds. The maximum score is supposed to be 40, with higher scores indicating a better performance. However, the subtest consisted of 8 sample items, so the actual range was 0 - 48.

The Faces-Pairs Test-Forced Choice Six Alternative Delayed Recognition (FPT-FCSADR) was presented after 20 minutes. Participants were required to correctly match the previously learned 20 pairs of faces, including 5 distractor items each. The total score was calculated by summing up all correctly matched pairs of faces. The theoretical maximum score is supposed to be 20, however the subtest consisted of 4 sample items, so the actual range was 0 - 24.

The Faces-Pairs Test-Recognition (FPT-REC) presented 20 previously studied items and 20 new items. Participants were asked to decide whether the faces were familiar or new. The total score was calculated as the sum of all correct answers minus the sum of all incorrect answers. The possible range is 40, but the subtest consisted of sample items, so the actual maximum score was 48.

The FPT “Total Score” was calculated by summing up the total scores of all three subtests: FPT-FCTAIR + FPT-FCSADR + FPT-REC. The theoretical score ranges from 0 - 100. As the FPT still consisted of sample items, the actual scores ranged from 0 to 120, with higher scores indicating a better performance.

The cognitive domain of language was assessed by the following subtests:

Image Naming Test (INT). The task involved pictures of objects and required participants to select the corresponding word. This test was designed to measure language, object naming, word finding, and semantic memory skills. The test took approximately three

minutes to complete, and the score was calculated by summing all correct responses. The sum score ranges from 0 to 107 (7 sample items were included), with higher scores indicating a better performance.

Auditory Vocabulary Test (AVT). This task aims to assess language skills, auditory perception, and comprehension by having participants listen to 109 words. The task included 9 example items and lasted for approximately 3-5 minutes. Half of the spoken words are real German words, while the other half are fictional but familiar in terms of sound and pronunciation. Participants were required to identify the real words by selecting 'YES' or 'NO'. The total score was calculated by summing up the correct answers; the sum score ranged from 0 to 109, with higher scores indicating a better performance.

Verbal Vocabulary Test (VVT). Participants were asked to determine whether a written word was real or fictional by selecting either "Yes" or "No". The test took approximately three minutes to complete and was originally designed to evaluate verbal ability, visual perception, and comprehension. An overall VVT score was calculated by summing up all correct answers, which ranged from 0 to 109 points, with a higher score indicating better performance. The VVT subtest consisted of 9 example items.

In the present study, one subtest was submitted to assess the cognitive domain of complex attention.

Digit Symbol Test (DST). Participants were required to quickly pair symbols and digits. The test consists of three rounds, each lasting one minute. The total duration of the subtest was set to three minutes. The sum score of all correctly paired symbols and digits across the three trials was calculated, ranging from 0 to 180, with a higher score indicating a better performance. The DST subtest was designed to evaluate speed processing, attention shifting, sustained attention, and selective attention.

The cognitive domain of executive function was assessed by the following subtest:

Traffic Light Test (TLT-s). The subtest TLT-s was administered to evaluate executive functioning, inhibition, and cognitive flexibility. The subtest comprises two rounds, each consisting of 30 items. The first round is based on a forward condition, while the second round is based on a reverse condition. Both rounds feature a traffic light that alternates between green and red. During the forward condition, participants were instructed to press the 'Go' button if the light was green and the 'Stop' button if the light was red. In the second round, the instructions were reversed. The task took approximately three minutes. The goal was to press the correct button as quickly as possible.

The sum of all correctly pressed 'Go' and 'Stop' buttons (1 point), incorrectly pressed 'Go' and 'Stop' buttons (-1 point), and unanswered items (0 points) was calculated for each round. The sum score ranges from a maximum of 20 (all correct) to a minimum of -20 (all incorrect) in each round, with a higher score indicating a better performance. The sum scores ranged from 0 – 40.

The INCP provides one test for the evaluation of social cognition, which is described in the following sequence:

Emotion Face Test-short (EFT-s). The task consists of two rounds with 20 items each and encompasses a forward and a reverse condition. It is intended to measure social cognition and emotion recognition. In the first round, emoticons with either a happy or a sad face were presented. Participants were instructed to press as quickly as possible a button, which was saying “Happy” if the emoticon showed a happy face, and a button, which was saying “Sad” if the emoticon appeared with a sad face. In the second round, the conditions were reversed. The total score was calculated by summing up all correctly pressed “Happy”- and “Sad”- Buttons (+1 point), all incorrectly pressed buttons (-1 point), and all unanswered items (0 point) in both rounds. The sum score was calculated for each round. The maximum score for each round was 20 points (all items correct) and the minimum score was -20 points (all items incorrect), with a higher score indicating better performance. There was a time limit. The task took about 3 minutes to complete.

Materials

The letter of consent, the screening tests (MMSE, VVT 3.0) and the NTBV were administered on paper and pencil, with an erasable pen provided by the test instructor. The screening instruments, BDI-II and WST were administered partly with paper and pencil and partly on a tablet. The digital test (INCP) was also administered on a tablet. The digital tests were submitted on an Apple iPad using application software developed by Psimistri-Global Psychometric Test and Intervention Systems. Headphones were provided by the instructor to minimize distractions and ensure standardized testing conditions. They were required only for the AVT subtest.

Statistical Analysis

Statistical analyses were performed using IBM SPSS version 28 and R 4.3.1. The significance level was set at 5%; results with $p \leq .05$ were considered significant.

Initially, the data was checked for plausibility and corrected if necessary. Data was excluded if it was found to be anomalous. However, missing data was not imputed, and reverse items were not recoded. The distribution of the data and descriptive statistics (mean values, median standard deviation, range, skewness, and kurtosis) were analyzed. The data was checked for normal distribution and visualized with histograms and boxplots; this was used to check if there are values outside the possible range. With an eye toward further research, boxplots were used to identify outliers, with the aim of detecting floor and ceiling effects as well as particularly conspicuous participants, who were not excluded due to the exploratory study design. Correlation tables were used to check for unexpected correlations between categorical variables, while t-tests or ANOVA were used for metric variables. Additionally, the preparatory analysis included calculating z-transformations for all INCP subtests. A cross-table displays a comparison of the actual subsample size to the expected sample size when assigning groups following the criteria set by Petersen (2004).

In a second step, hypothesis-testing analyses were implemented.

To evaluate the main question of research, a correlation matrix using Spearman's correlation coefficient (*rS*) and pairwise bivariate scatterplots of all variables of the INCP were calculated, with the aim of gaining first insights into the various associations within the INCP test battery.

Subsequently, domain-specific variables of the NTBV and the INCP were correlated, aiming to model the associations between the NTBV and INCP using Spearman's correlation coefficient (*rS*). The evaluation of whether there is a sufficient correlation between both neuropsychological assessment batteries (NTBV and INCP) was based on the effect size *d* following Cohen (1988, p. 79). A correlation coefficient of $r < .3$ was considered a small effect, $.3 \leq r < .5$ was referred to as a moderate effect; and $r \geq .5$ was considered a large effect. A heatmap was created to visualize the associations.

In an exploratory step, the findings of the main question of research were strengthened through an exploratory factor analysis (EFA), aiming to identify the factors that best describe the underlying structure of the 9 submitted INCP subtests. The EFA was conducted with the data of the overall sample. Factor analysis assumes linear correlations, and positive correlations should exist between variables that are presumably attributable to the same factor (or negative in the case of inverted items). The preconditions were checked, but some were violated. The pilot study design justifies the procedure. The relationship between the variables was not linear. There were outliers in the data and the sample was very

small. However, the variables were metric. A subsequent orthogonal Kaiser- Varimax rotation was used.

In a further exploratory step, the non-parametric Mann-Whitney-U-Test was conducted to gain first insights into the discriminatory power of the INCP. Therefore, the subsample was divided into two groups (HC and MCI) based on the screening results. The prerequisites were checked. Because no participant appeared in two groups, the condition of the independence of measurements was fulfilled. The sample was divided based on the nominally scaled variable "diagnosis" with two characteristics (0 = MCI, 1 = HC). The dependent variable was metric, and the distribution form was approximately the same in both groups (Kolmogorov-Smirnov-Test $p > .05$).

Cronbach's alpha (α) was calculated to assess the internal consistency of some subscales of the INCP, but only for those subtests that provided raw scores (INT, AVT, VVT). The INCP is still considered an exploratory assessment battery, and some of the subtests examined (INT, AVT, VVT) still consists of sample items. The goal was to exclude those items with the lowest internal consistency. The final subtests should consist of 100 items. Based on reliability analysis, the items with the lowest internal consistency were excluded, and reliability analysis was repeated with the shortened scales.

Additionally, item difficulty ($P_i = 1$) was calculated for the items of the subscales INT, AVT, and VVT.

Results

Descriptive Statistics of Overall Sample

The total sample, as displayed detailed in Table 2, consisted of 37 female (69.8%) and 16 male subjects (30.2%), with a mean age of 58.06 years (SD = 17.637) and a range of 22 to 87 years. On average, the sample reported 14.81 years of formal education, with a range of 8 to 26 years. 22 participants reported being retired, 24 reported that they are currently working, and 2 reported being unemployed.

The total sample had a mean BDI II score of 6.12 (SD = 5.735) with a range of 0 to 20 points and an average premorbid intelligence level of 111 WST IQ-points (SD = 10), with a range from 85 to 129 points. The sample achieved an average score of 27.94 on the MMSE with a range from 21 to 30 and an average score of 9.72 on the VVT 3.0 screening test, whereby the scores ranged from 4 to 10 points. In several cases data was missing, which was caused by the exploratory study design.

Table 2

Sample Sizes, Mean Values, Standard Deviations, and Range of Demographic and Screening Variables of the Total Sample

Variable	*N*	*M*	*SD*	*Min*	*Max*
Age	53	58.06	17.64	22	87
Sex	53	.70	.46	0	1
Formal Years of Education	53	14.81	4.17	8	26
MMSE	53	27.94	2.52	21	30
VVT 3.0 Screening	53	9.72	.93	4	10
BDI-II	42	6.12	5.74	0	20
WST	46	111	10.78	85	129

Note. N = sample size, *M* = mean, *SD* = standard deviation, *Min* = Minimum, *Max* = Maximum.
[a]MMSE: possible range 0-30, VVT 3.0 Screening: possible range 0-10, BDI-II: possible range 0-63, WST: possible range 0-139.

The proposed sample was divided into 2 subgroups: "neurologically healthy controls" (N_{HC} = 41) and "mild cognitive impairment" (N_{MCI} = 12). The subsamples are discussed in the following.

Descriptive Statistics of the Subsamples Neurologically Healthy Controls (HC), and Mild Cognitive Impairment (MCI)

Both subsamples showed a preponderance of female participants (~ 70%). It can be noted that, on average, the HC-group comprised younger (M_{HC} = 53.93, SD_{HC} = 16.22) and better educated (M_{HC} = 16.32, SD_{HC} = 3.34) probands. While all study subjects within the MCI-group reported being retired, in the HC-group only 13 persons stated that they were retired, whereas 24 subjects specified that they had a job, and 1 person was unemployed at the time of the examination. The healthy controls scored on average higher on the cognitive screening test (MMSE). While the HC-group reported on average fewer depressive symptoms and scored on average higher on the premorbid verbal intelligence scale WST, the subsamples did not show any significant performance differences on the VVT 3.0 screening test.

For detailed test statistics of demographic data and screening tests for both groups see Table A2 in the appendix. There are different sample sizes for each subtest of the screening tests. This was caused by the explorative study design and the reoccurring technical problems with the tablet application.

Comparison of Actual Group Assignment Method

Due to recruitment issues related to the pandemic situation, the initial group assignment based on the NTBV scores was dismissed and replaced by screening results.).

Table 3 3 displays the deviation of the actual subsample size from the expected subsample size when following the guidelines by Petersen (2004) for the diagnosis of MCI. The actual group assignment was based on screening results, whereas the expected subsample sizes were calculated based on z-scores on each subtest of the NTBV for each participant. Expected N_{MCI} refers to individuals who scored 1.50 standard deviations below age, gender, and education norms in a single domain or in multiple domains on the NTBV and are classified as MCI according to Peterson's criteria.

As shown in).

Table 3 3 the present sample consisted of 41 healthy individuals and 12 individuals with MCI. However, it is shown that the expected subsample size differs from the actual subsample size. Of the 41 participants who were actually classified as healthy individuals in the present study, only 9 participants would meet the criteria for the HC-group, and 32 participants would meet the criteria for the MCI group when following Petersen's criteria based on the NTBV subtests. Of the 32 participants who were actually assigned to the HC group but showed MCI on the NTBV, 11 met the criteria for MCI (single domain), 10 fulfilled the criteria for MCI (multiple domain), 1 participant showed amnestic MCI (single domain) symptoms, and 10 subjects met the criteria for amnestic MCI (multiple domain). 12 patients, who were actually allocated to the MCI-group in the present sample, also showed MCI-symptoms on the NTBV.

To summarize the findings of the comparison of the two different group assignment methods, it is noted that 44 participants are expected to be assigned to the MCI-group, while 9 subjects are expected to be assigned to the HC-group when following the state-of-the-art criteria.

A detailed table is provided in the appendix, which shows the actual *N* per subtest and the number of individuals who deviate from the original group assignment because they scored 1.50 standard deviations below age, gender, and education norms on a single or on multiple subtests on the NTBV for the NTBV assessment battery (see Table A 3 in the appendix) and for the INCP test battery (see Table A 4 in the appendix).

Table 3

Comparison of Actual Subsample Size based on the Screening to Expected Subsample Size based on the NTBV.

	Actual N_{HC}	Actual N_{MCI}
Actual Total *N*	41	12
Expected N_{HC}	9	
Expected N_{MCI}	32	12
Expected $N_{MCI\ (single\ domain)}$	11	1
Expected $N_{MCI\ (multiple\ domain)}$	10	5
Expected $N_{aMCI\ (single\ domain)}$	1	
Expected $N_{aMCI\ (multiple\ domain)}$	10	6
Expected Total *N*	9	44

Note. Actual Total *N* refers to the actual subsample size included in the present sample based on the screening criteria.

Expected N_{HC} corresponds to the number of people who are expected to be part of the HC-group based on the NTBV scores.

Expected N_{MCI} corresponds to the number of people who scored 1.50 standard deviation below age, gender, and education norms in a single domain or in multiple domains on the NTBV and are classified as MCI following Petersen (2004) criteria.

NTBV-Subtests

The descriptive statistics for the NTBV subtests in the total sample and in the subsamples are shown in the appendix (see Table A 5). Reported are sample size, mean value, standard deviation, and minimum and maximum score for each NTBV subtest. The different sample sizes can be attributed to the administration of different versions of the NTBV test battery, depending on the MMSE score.

INCP-Subtests

As the INCP is still considered an experimental test, different versions of the test battery were submitted, resulting in varying sample sizes for each subtest. Due to recurring technical problems, such as programming errors, early test aborts were observable. Furthermore, certain subtests were removed during the assessment procedure. Therefore, only 9 out of the original 18 INCP subtests were examined in the present thesis. Table 4 displays the sample sizes, the sum scores, the average performance, median, standard deviation, and score range for each subtest partialized for the total sample as well as for both subsamples.

To evaluate the normal distribution of the data gathered through the INCP, a Shapiro-Wilk- test was performed. All examined INCP subtests (CITY, FACE, AVT, DST, EFT-s, FPT, VVT, TLT-s, INT) showed an asymmetric distribution (Shapiro-Wilk-Test, $p < .05$). The figures attached display a histogram for the sum score distribution of each subtest, as well as a grouped boxplot showing the median, the 1st, and the 3rd quartiles for each subtest of the total sample, the HC- and the MCI-group.

Associations between INCP subtests and demographic and screening data were quantified using Spearman's' correlation coefficient (*rS*), (see Table A6 in the appendix).

Table 4

Sample Sizes, Mean Values, Standard Deviations and Minimum and Maximum Scores of the INCP Subtests for the Overall Sample, the HC- and the MCI-Group.

Total	Subtests	*N*	*M*	*Mdn*	*SD*	Min	Max
	CITY	41	13.63	14.00	2.44	6	16
	FACE	33	13.62	16.00	3.77	3.50	16
	AVT	44	92.89	93.00	7.58	69	104
	DST	41	175.80	178.00	5.58	156	180
	EFTs	33	34.88	37.00	7.69	-3	40
	FPT	48	81.38	84.00	16.03	49	113
	VVT	45	84.51	86.00	7.27	68	95
	TLTs	42	33.19	38.00	12.26	0	40
	INT	49	98.02	100.00	14.43	1	107
HC-Group	Subtests	*N*	*M*	*Mdn*	*SD*	Min	Max
	CITY	36	14.06	14.50	1.80	10	16
	FACE	30	13.87	16	3.46	5	16
	AVT	39	94.44	95	6.18	76	104
	DST	38	175.68	178	5.74	156	180
	EFTs	31	34.87	37	7.91	-3	40
	FPT	39	86.67	87	12.39	53	113
	VVT	35	86.49	87	6.26	68	95
	TLTs	39	32.77	38	12.63	0	40
	INT	39	97.56	100	16.15	1	107
MCI-	Subtests	*N*	*M*	*Mdn*	*SD*	Min	Max
	CITY	5	10.60	13	4.22	6	14
	FACE	3	11.17	15	6.64	3	15
	AVT	5	80.80	84	6.98	69	86
	DST	3	177.33	178	3.06	174	180
	EFTs	2	35	35	4.24	32	38
	FPT	9	58.44	58	7.28	49	70
	VVT	10	77.60	76.50	6.48	68	88
	TLTs	3	38.67	39	1.53	37	40
	INT	10	99.80	99	2.10	98	104

Note. N = Sample Size, *M* = Mean, *SD* = Standard Deviation, *Min* = Minimum, *Max* = Maximum.

[a] CITY: Possible Range: 0 – 16; FACE: Possible Range: 0 – 16; AVT: Possible Range: 0 – 100, DST: Possible Range: 0 – 180; EFT-s: Possible Range: 0 – 60; FPT: Possible Range: 0 – 100; VVT: Possible Range: 0 – 100; TLT-s: Possible Range: 0 – 60; INT: Possible Range: 0 – 100

A total of 41 probands completed the City Identification Test with a mean score of *M* = 13.63 points (*SD* = 2.44) and a range of 6 – 16 points. The theoretically achievable score range was 0-16. The median of the score was 14, the 1st percentile was 12, and the 3rd percentile was 15,5.

Figure 5

Histogram of the City Identification Test (CITY)

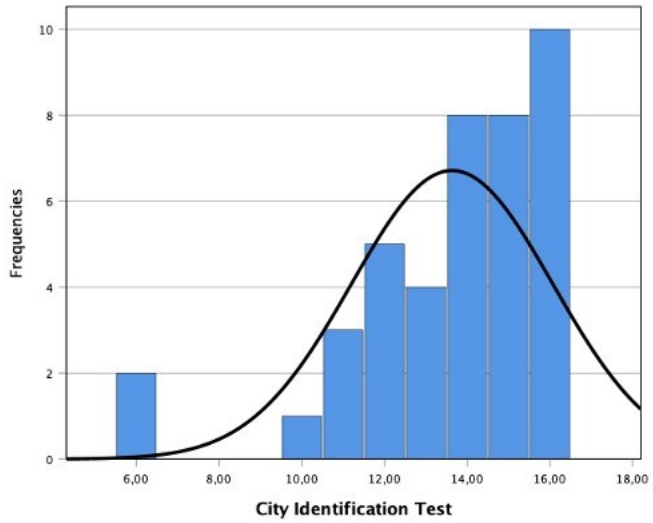
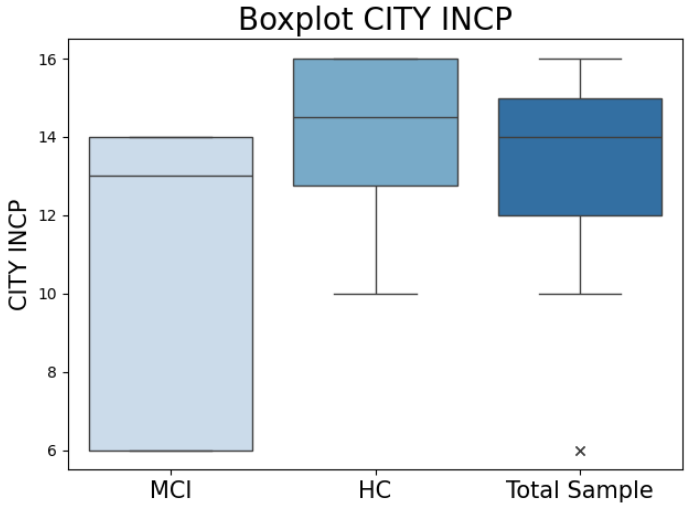


Figure 6

Grouped Boxplot of City Identification Test (CITY)



33 probands completed the Face Identification Test. The possible range of score was 0 – 16 points; the actual range in the overall sample was 3.50 – 16.00, with a mean score of 13.62 points (*SD* = 3.77). The median was 16, the 1st percentile was 11 and the 3rd percentile was 16.

Figure 7

Histogram of the Face Identification Test (FACE)

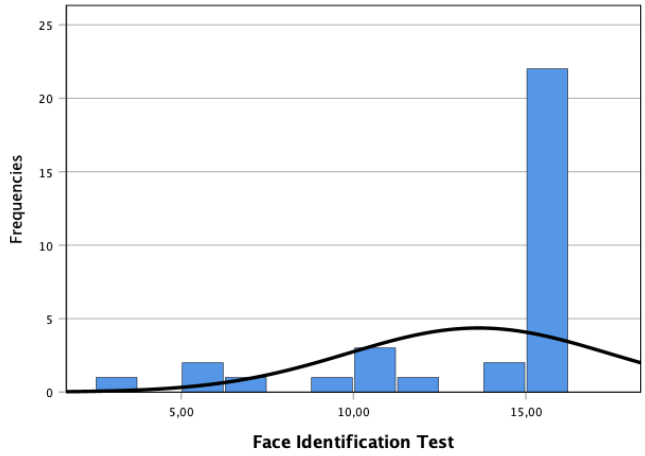
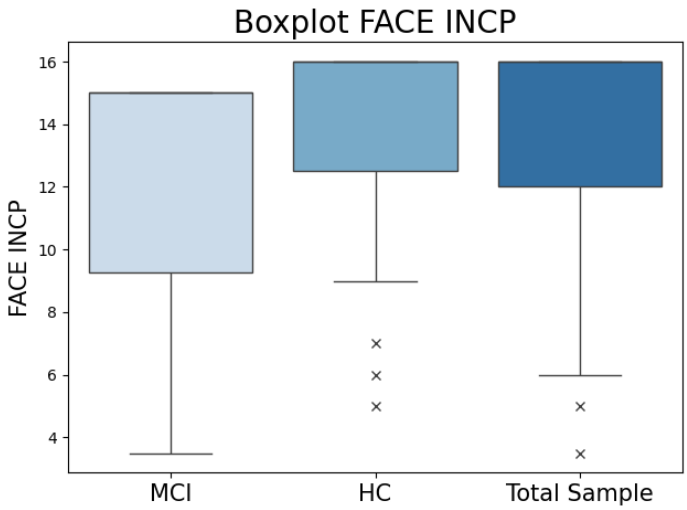


Figure 8

Grouped Boxplot of Face Identification Test (FACE)



In total, 44 study subjects completed the Auditory Vocabulary Test, with a mean score of *M* = 92.89 and a standard deviation of 7.58 points, with a theoretical range of 0 – 100 points. The scores of the total sample ranged from 69 to 104 points. As the INCP is still under development, the AVT comprises 109 items, of which 9 are considered sample items, which should be dismissed in a further development phase. The median was 93, the 1st percentile was 89 and the 3rd percentile was 98.75.

Figure 9

Histogram of the Auditory Vocabulary Test (AVT)

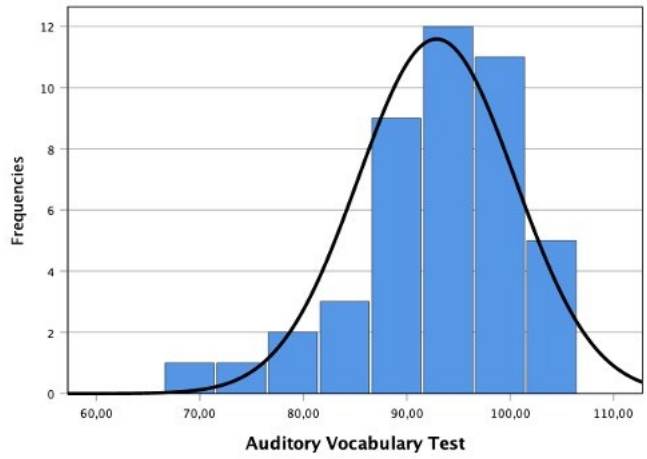
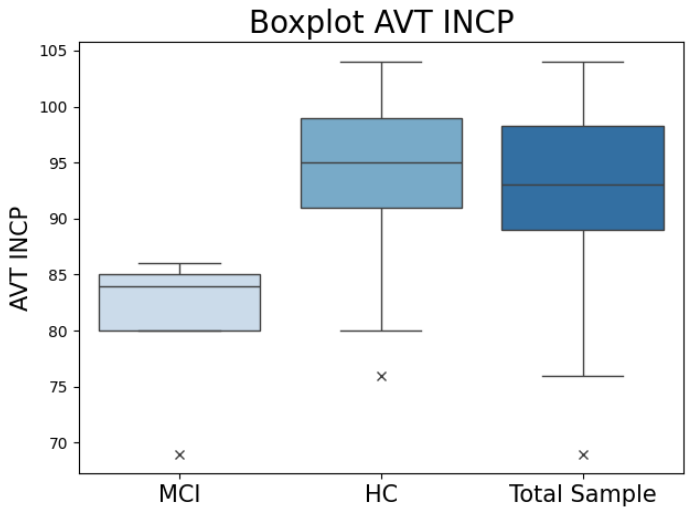


Figure 10

Grouped Boxplot of Auditory Vocabulary Test (AVT)



The DST has been processed by 41 probands, who achieved a mean sum score of *M* = 175.80 (*SD* = 5.58) points. The possible range was 0 – 180 points; the overall sample achieved a minimum of 156 and a maximum of 180 points. The median score was 178, the 1st percentile was 175 and the 3rd percentile was 180.

Figure 11

Histogram of the Digit Symbol Test (DST)

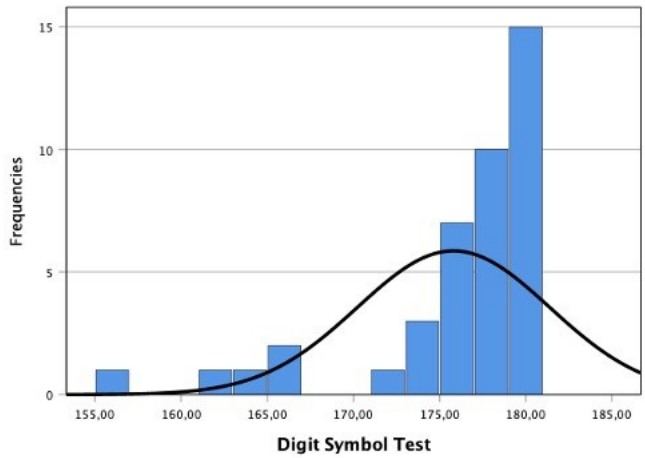
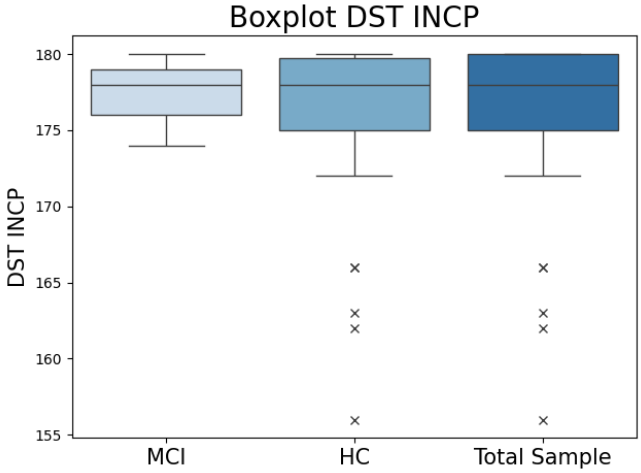


Figure 12

Grouped Boxplot of the Digital Symbol Test (DST)



In total, 33 participants completed the INCP Subtest Emotion Face Test-short. The average sum score of the total sample was *M* = 34.88 points (*SD* = 7.69). The median was 37, the 1st percentile was 33 and the 3rd percentile was 38.5. The actual score ranged from -3 to 40 points. However, the theoretically possible score ranges from 0 to 60 points.

Figure 13

Histogram of the Emotion Face Test-short (EFT-s)

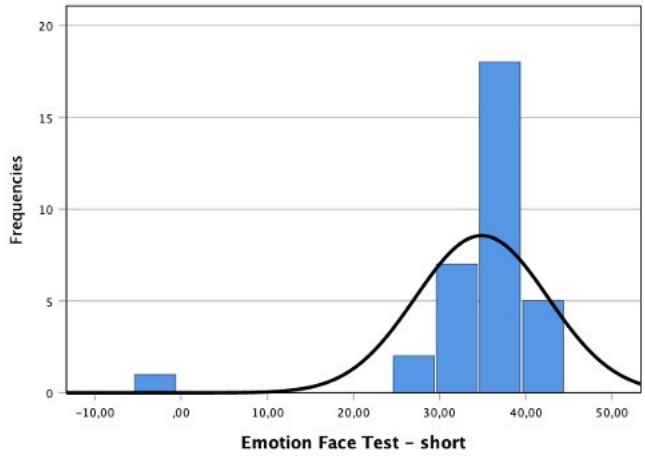
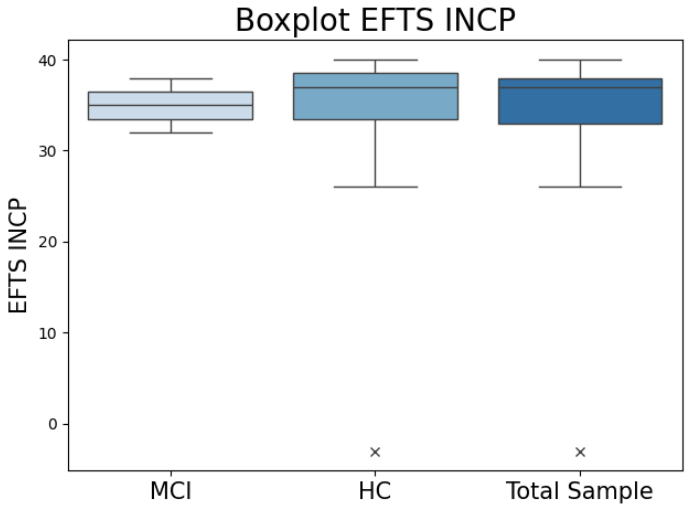


Figure 14

Grouped Boxplot of Emotion Face Test-short (EFT-s)



The Verbal Vocabulary Test consisted of 45 participants. The overall sample achieved an average score of *M* = 84.51 points (*SD* = 7.27) with an actual range of 68.00-95.00 points; however, the possible score ranges from 0 to 100. The median score was 86, the 1st percentile was 80 and the 3rd percentile was 90. As the INCP is still under development, the VVT comprises 109 items, of which 9 are considered sample items, which should be removed in a further development phase.

Figure 15

Histogram of Verbal Vocabulary Test (VVT)

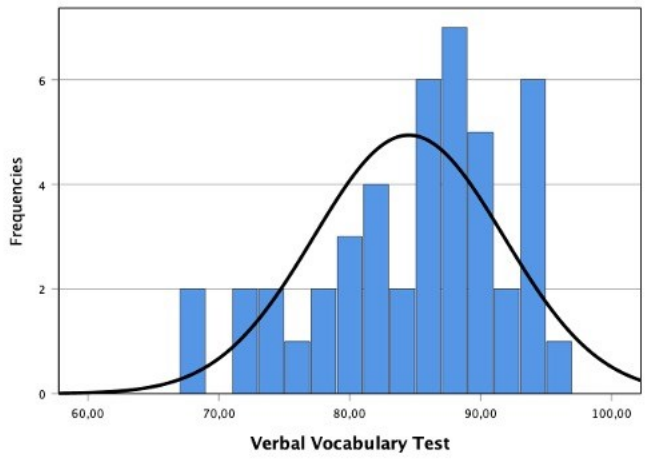
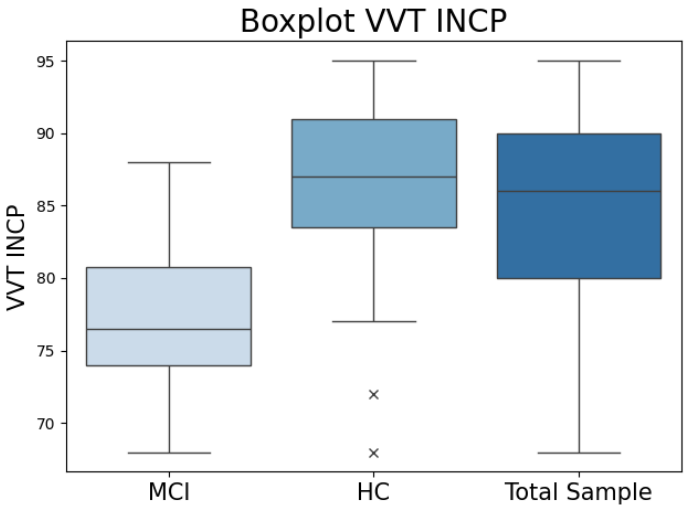


Figure 16

Grouped Boxplot of Verbal Vocabulary Test (VVT)



For the Traffic Light Test-short, the sample consisted of 45 participants. On average, the sample scored 33.19 points (*SD* = 12.26) with a range of 0-40. The possible range was 0-60. The median score was 38, the 1st percentile was 35 and the 3rd percentile was 40.

Figure 17

Histogram of the Traffic-Light-Test-short (TLT-s)

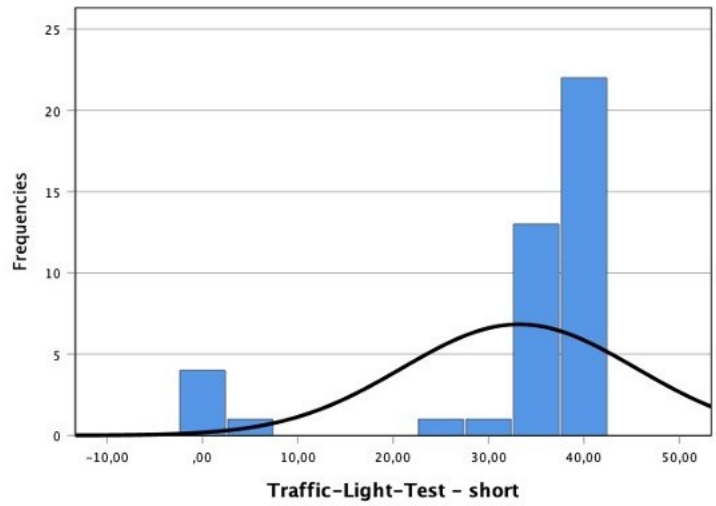
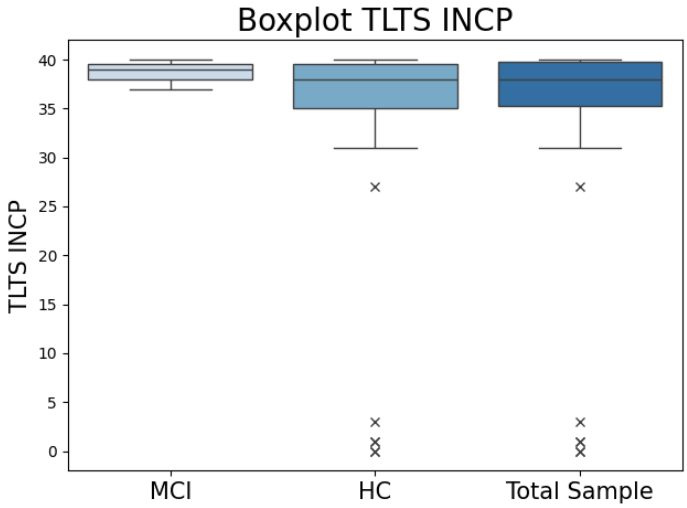


Figure 18

Grouped Boxplot for Traffic Light Test-short (TLT-s)



For the Image Naming Test, the sample consisted of $N = 49$ probands, with a mean score of $M = 98.02$ points ($SD = 14.43$), with a range of 1–107 points; the possible range was 0 – 100. The median of the score was 100, the 1st percentile was 98,5 and the 3rd percentile was 100. The difference between the actual and potential range can be attributed to the fact that the INCP still contains 7 sample items, which should be removed after a thorough item analysis.

Figure 19

Histogram of Image Naming Test (INT)

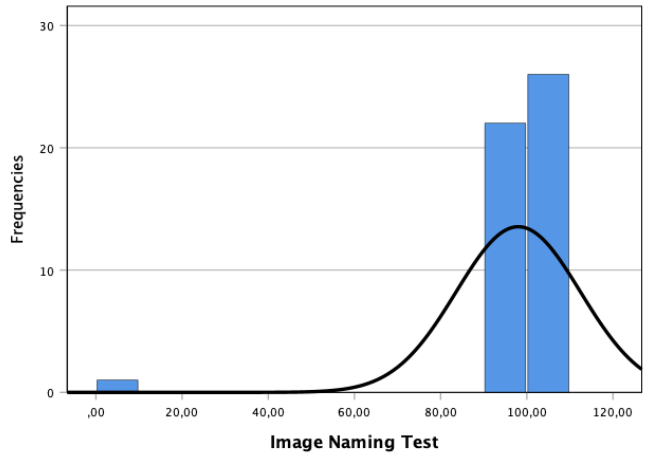
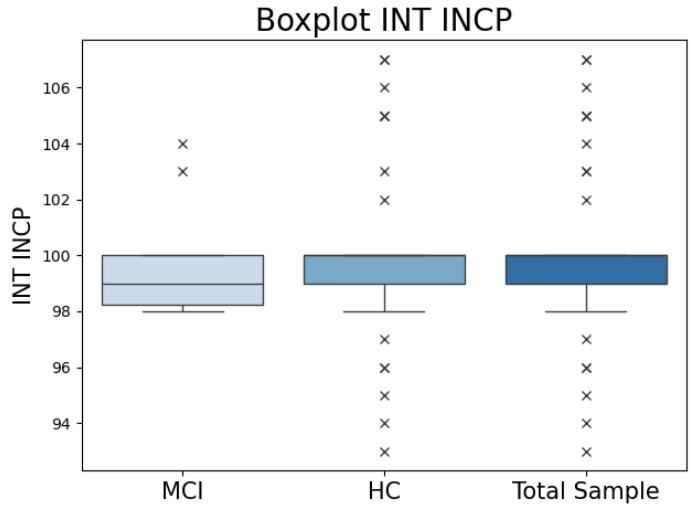


Figure 20

Grouped Boxplot of Image Naming Test (INT)



The Faces-Pairs-Test was administered to 39 participants. The total sample obtained a FPT mean total score of *M* = 86.67 points (*SD* = 12.39), with an actual range of 53-113 points. The median was 84, the 1st percentile = 70 and the 3rd percentile = 93.75.

Figure 21

Histogram of the Faces Pair Test (FPT)

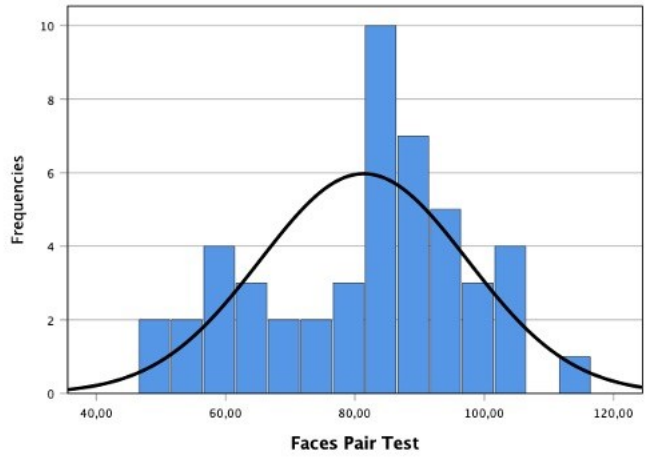
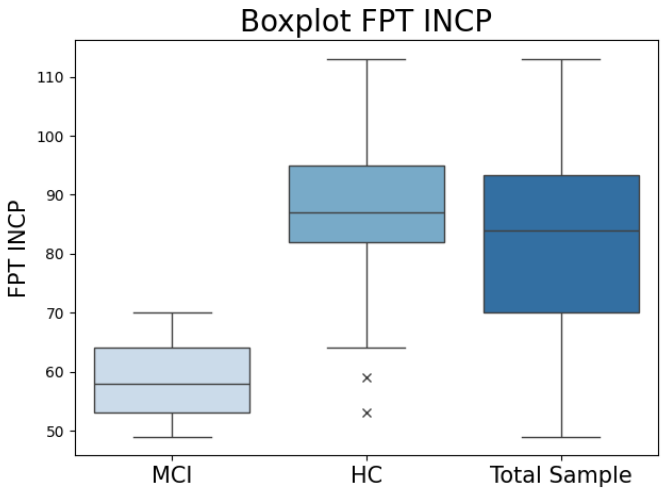


Figure 22

Grouped Boxplot of the Faces Pair Test (FPT)



Main Questions of Research

The objective of this study was to evaluate whether the INCP is a valid test battery for the assessment of cognitive impairment. For this purpose, associations within the INCP subtests were analyzed. In the appendix, a correlation matrix (*rS*) displays associations between domain-specific tasks of the INCP. Unexpected significant correlations are also reported with regard to future research.

In a further step domain-specific variables of the INCP were correlated with another validated and widely used psychometric test battery (NTBV) to gain first insights into the convergent validity of the digital test (INCP). To investigate the associations between the two test batteries, a correlation matrix was created, which includes all domain-specific variables from both tests and is presented in the appendix (see Table A 7).

Due to the violation of the normal distribution and the higher robustness against outliers, Spearman's correlation coefficient (*rS*) was used (Field, 2018).

Following Cohen's (1988) guidelines, a small effect is defined as $r < .3$, a medium effect is considered $.3 \leq r < .5$, and a large effect is referred to as $r \geq .5$. Only significant correlations with moderate to large effect sizes are reported in the following. To enhance visualization, a heat map was created to display the domain-specific correlation matrix between NTBV and INCP for hypotheses 1_7 and 1_8.

Associations within the INCP test battery

Hypothesis 1_1: Cognitive Domain of Complex Attention (INCP). Due to reoccurring technical issues during the assessment procedure, the present study administered only one task (DST) for the evaluation of the cognitive domain of complex attention. As a result, no domain-specific association within the INCP can be reported, leaving hypothesis 1_1 unanswered in this thesis.

However, in terms of further research, it should be noted that the data implied a moderately positive correlation between the INCP subtests DST and the learning and memory task FPT ($r = .46$, $p = .002$).

Hypothesis 1_2: Cognitive Domain of Executive Function (INCP). Due to repeated technical problems, this thesis only submitted one task (TLT-s) for the assessment of executive function. The present study cannot answer the question of whether there is an association between domain-specific tasks of the INCP for the assessment of executive functions.

However, it is important to note that the subtest TLT-s exhibited positive and moderate correlations with two variables outside of the postulated domain, specifically with the learning and memory task FPT (r = .39, p = .011) and with the social cognition subtest EFT-s (r = .39, p = .024).

Hypothesis 1_3: Cognitive Domain of Learning and Memory (INCP). The tasks for assessing learning and memory posed in the INCP (FPT, FACE, CITY) were partially correlated. Thus, this hypothesis can partially be accepted as a large positive correlation was found between the subtests CITY and FACE (r = .53, p = .002).

However, no positive correlation was shown between either the FPT and the CITY subtest or the FPT and the FACE subtest. Contrary to the initial construction of the INCP, the data indicated a moderate negative correlation between the FPT and the CITY task (r = -.37, p = .002).

The section below discusses possible correlations beyond the proposed cognitive domain, with the aim of providing information for further research. The findings implied a moderately positive correlation between the subtest FPT and the executive function task TLT-s (r = .39, p = .011) as well as with the complex attention task DST (r = .46, p = .001).

The results of the CITY subtest indicated moderate positive associations with the language tasks AVT (r = .47, p = .003) and VVT (r = .43, p = .011).

Similarly to the CITY subtest, the FACE subtest showed a moderately positive correlation with the VVT subtest (r = .46, p = .018).

Hypothesis 1_4: Cognitive Domain of Language (INCP). The subtests for the assessment of the cognitive domain of language (AVT, VVT, INT) were partially correlated. A strong positive correlation was found between the AVT and VVT subtests (r = .70, p < .001), but no correlation was found between the INT task and other posed language tasks. Thus, $H1_4$ can be partially accepted.

Further research should consider the following significant domain-independent associations within the INCP assessment battery: Both subtests, AVT (r = .47, p = .003) and VVT (r = .43, p = .011) showed a moderate and positive correlation with the CITY subtest, while the VVT data displayed a moderate and positive correlation with the FACE subtest (r = .46, p = .018).

Associations between Domain-specific Variables of the INCP and Domain-specific Variables of a Construct-like Test Battery (NTBV)

Hypothesis 2_1: Cognitive Domain of Complex Attention (INCP & NTBV). No significant correlation was found between the tasks for assessing attention posed in the NTBV (Digit Symbol Subtest, AKT, C.I., TMT B), and the tasks for assessing complex attention posed in the INCP (DST).

However, the data indicated some associations with variables outside of the proposed cognitive domain. Specifically, there were moderate and positive associations between DST (INCP) and the NTBV scores NAI Labyrinth "Total/Time" (r = .33, p < .05), PWT "b" (r = .44, p = .004).

Additionally, the data showed moderate negative correlations between the DST and the following inverted coded NTBV variables: TMT-A (r = - .36, p = .022), 5-Point Test "Perseveration" (r = - .34, p = .030), NAI-III Color-words-test "Time Words" (r = - .34, p = .028), NA-III – NAI-I Color-word-test "Interference" (r = - .37, p = .016),

Hypothesis 2_2: Domain Executive Function (INCP & NTBV). No significant correlation was found between the executive function assessment tasks in the NTBV (TMT A, Five Point Test, Planning Maze Test, Stroop Test, C.I., PWT) and those in the INCP (TLT-s).

Nonetheless, future research should address the following correlations independently of the cognitive domain of executive function: A moderate positive correlation was observed between the TLT-s and the episodic memory subtest from the NTBV assessment battery, namely the VSRT "Delayed Recall" (r = .36, p = .020), and a moderate negative association was found between the TLT-s and the inverted coded VSRT "False Positive" subtest (r = -.39, p = .011).

Hypothesis 2_3: Cognitive Domain of Learning and Memory (INCP & NTBV). This hypothesis is partially accepted, as the data indicated a partially positive correlation between the tasks posed in the NTBV (VSRT) and the tasks posed in the INCP (FACE, CITY, FPT) for the assessment of the cognitive domain of learning and memory (see Figure 23).

There was no positive correlation found between the CITY and the VSRT subtest, nor between the FACE and the VSRT tasks. Contrary to the initial assumption, the results suggested a moderate negative correlation between the INCP subtest CITY and the VSRT subtest "Delayed Recall" (NTBV), (r = -.33, p = .034).

However, the FPT scores (INCP) showed moderate to large positive correlations with all the VSRT subtest scores (NTBV), including VSRT “Immediate Recall”: (r = .41, p = .005), VSRT “Delayed Recall” (r = .66, p < .01), VSRT “Total Recall” (r = .30, p = .038), and VSRT “Recognition” (r = .34, p = .021).

With an eye toward further research, the following moderately positive associations between the learning and memory task FPT (INCP) and the following NTBV subtests independently of the cognitive domain of learning and memory should be noted: HAWIE-R (r = .47, p < .01), SWT: (r = .34, p = .024), PWT: (r = .31, p = .047), NAI-III “Total/Time” (r = .35, p = .002), and Labyrinth “Total/Time” (r = .38, p = .007).

Moderate negative correlations between the learning and memory task FPT and other inverted coded variables of the NTBV were found, namely: TMT A (r = -.37, p = .009), NAI-I “Color-word-test-Interference” (r = -.31, p = .041), and NAI-I “Time-color”: (r = -.48, p = .001).

Moderate positive correlations were found between the learning and memory task CITY (INCP) and several NTBV subtests outside of the proposed cognitive domain: AKT “Total” (r = .33, p = .036), NAI Labyrinth “Total/Time” (r = .34, p < .05), SWT “Tools” (r = .37, p = .017), PWT “l” (r = .38, p = .015).

Furthermore, the data indicated moderate negative correlations between the learning and memory task CITY and the NTBV subtests VSRT “Delayed Recall” (r = -.33, p = .034), and the inverted coded variable CI Symbols (r = -.33, p = .036).

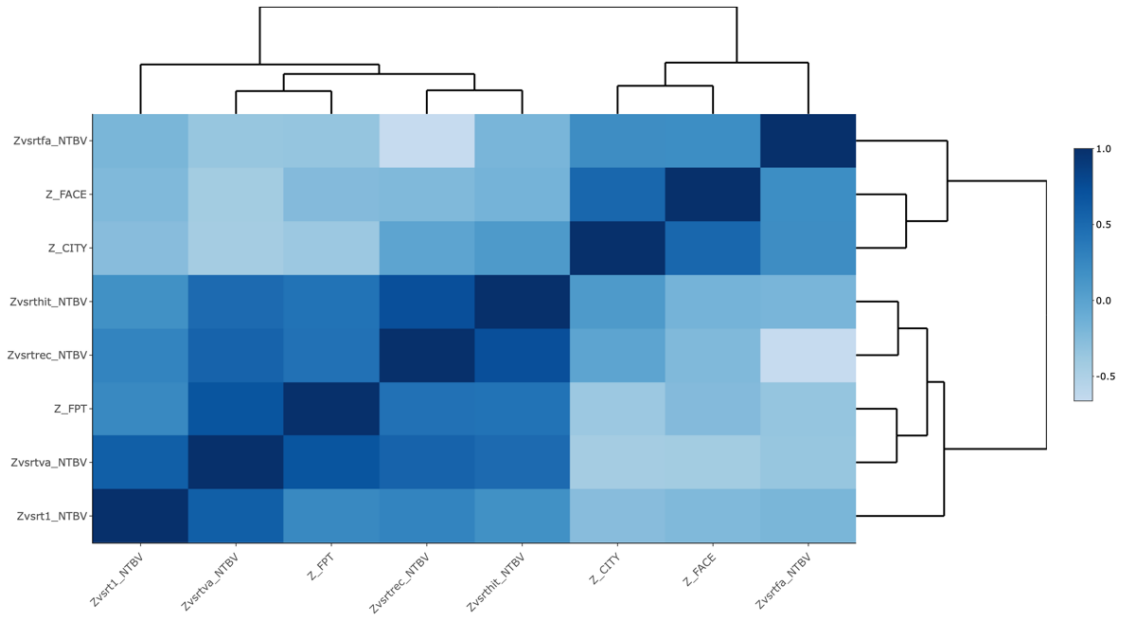
Furthermore, the data indicated several moderate and positive correlations between the learning and memory task FACE (INCP) and other domain variables of the NTBV, namely SWT “Tools” (r = -.37, p = .036), and the inverted coded variables: NAI-I-NAI-III Color-word-test “Interference” (r = .43, p < .05), NAI-III Color-words-test “Time-words” (r = .35, p = .047).

The following paragraph refers to moderate negative correlations, which were found between the INCP task FACE and some variables of the NTBV outside the proposed cognitive domain, which are namely AKT “Total/Time” (r = -.35, p = .049) and the inverted coded variable CI Symbols (r = -.35, p = .046).

Furthermore, large negative correlations were found between FACE and HAWIE-R “Digit Symbol Test” (r = -.53, p = .001) and Labyrinth “Total/Time” (r = -.53, p < .01).

Figure 23

Correlation Matrix: NTBV and INCP Variables for the Assessment of Learning and Memory



Note. The color-coded heatmap illustrates the correlation matrix between domain-specific variables of the NTBV and those of the INCP for the assessment of the cognitive domains of language and memory. The label represents Spearman's correlation coefficient (*rS*), with +1 indicating a perfect positive association, 0 = no association and -1 = a perfect negative association

Hypothesis 2_4: Cognitive Domain of Language (INCP & NTBV). Regarding the cognitive domain of language, positive correlations were found between the tasks posed in the NTBV (BNT, SWT) and the tasks posed in the INCP (AVT, VVT, INT). Therefore, $H1_8$ is partially accepted. Associations between INCP variables and NTBV variables, which are intended to assess the cognitive domain of language, are displayed in the attached heatmap (see Figure 24).

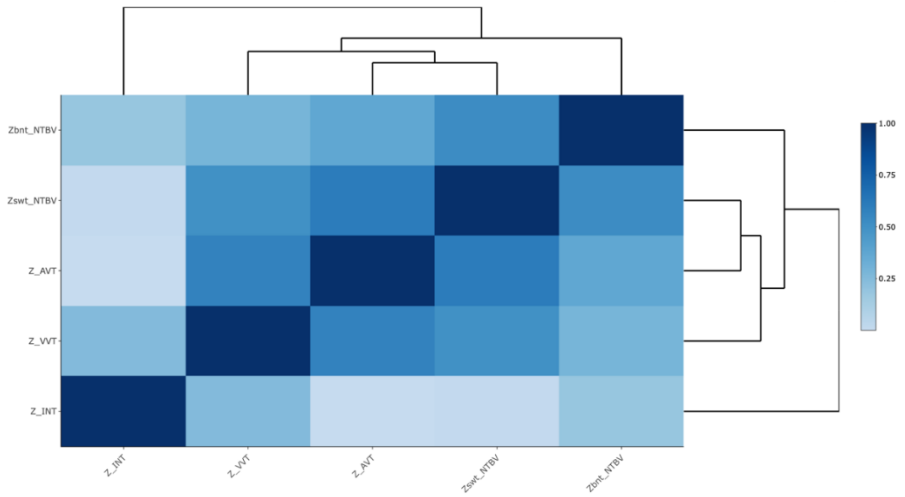
The INCP subtest AVT exhibited a strong and positive correlation with the NTBV subtest SWT ($r = .60$, $p < .001$) and a moderate positive correlation with the NTBV subtest BNT ($r = .50$, $p < .001$).

Similar to the AVT subtest, the VVT data indicated a strong positive correlation with the SWT (NTBV) subtest ($r = .52$, $p < .001$), and a moderate positive correlation with the NTBV subtest BNT ($r = .39$, $p < .001$).

However, no significant correlation was found between the INT subtest (INCP) and domain specific variables of the NTBV.

Figure 24

Correlation Matrix: NTBV and INCP Variables for the Assessment of Language



Note. The color-coded heatmap shows the correlations between domain-specific tasks of NTBV and INCP for the assessment of the cognitive domain language. The label represents Spearman's correlation coefficient (*rS*), with +1 indicating a perfect positive association, 0 = no association and -1 = a perfect negative association.

Further domain-independent associations between language variables (INCP) and other NTBV variables are displayed in the appendix (see Table A 7).

In the following sequence, moderate positive correlations between the language task AVT and other domain independent variables (NTBV) are noted, namely: VSRT "Immediate Recall" ($r = .41$, $p = .006$), VSRT "Recognition" ($r = .35$, $p = .002$), BNT ($r = 51$, $p < .001$), 5-Point Test "Right" ($r = .35$, $p = .021$), NAI-III "Total/Time" ($r = .38$, $p = .001$), and the inverted coded subtest C.I. Interference "Total/Time" ($r = .32$, $p = .036$).

The next paragraph refers to moderate to strong negative correlations between AVT and the NTBV subtests, which are representing other cognitive domains, namely: NAI-III "Total/Time" ($r = -.39$, $p = .011$), and between AVT and other inverted coded variables, namely C.I. Symbols Test ($r = - .57$, $p < .001$), TMT A ($r = - .34$, $p = .022$), NAI-I "Time-color" ($r = - .37$, $p = .014$), NAI-III "Time-words" ($r = - .39$, $p = .009$).

Furthermore, the data indicated moderate to large positive correlations between the language task VVT and other domain-independent variables from the NTBV test battery, namely: NAI-III Color-word-test "Total/Time" (r = .38, p = .017), SWT "Animals" (r = .48, p < .001), SWT "Grocery" (r = .43, p = .006), SWT "Tools" (r = .32, p = .047), SWT "Total "(r = .52, p < .001), PWT "Total" (r = -.35, p = .049), BNT (r = -.39, p = .009).

Moderate negative correlations between the language task VVT and other variables from the NTBV test battery were found between VVT and the inverted coded variables C.I. Symbols (r = - .42, p = .004), NAI-I "Time-color" (r = -.35, p = .030), and NAI-III Color-word-test (r = -.38, p = .008).

Exploratory Analysis

Associations of the Social Cognition Task EFT-s

Currently, the INCP test battery includes one subtest for assessing the cognitive domain of social cognition. The EFT-s showed a moderate and positive correlation with the executive function task TLT-s (r = .39, p = .024).

Additionally, the EFT-s subtest displayed moderate and positive correlations with other domain independent variables of the NTBV, including: VSRT "Delayed Recall" (r = .35, p = .013), VSRT "Immediate Recall" (r = .44, p = .011), VSRT "Recognition" (r = .40, p = .022), NAI-III Color-words-test "Total/Time" (r = .44, p = .010), NAI Labyrinth "Total/Time" (r = .43, p = .013), and the inverted coded variable AKT "Total/Time": (r = .39, p = .024); strong positive correlations were found between EFT-s and VSRT "3" (r = .55, p = .001).

Furthermore, the EFT-s subtest exhibited moderate negative correlations with the following inverted coded domain independent subtest of the NTBV, namely TMT A: (r = -.40, p = .022), TMT B (r = - .38, p = .028), NAI-I "Time Color" (r = -.48, p = .004), NAI-I-NAI-III Color-word test "Interference"(r = -.36, p = .037), and a large negative correlation were shown between EFT-s and the inverted coded subtest 5 Point Test "Perseveration" (r = -.57, p < .001).

Exploratory Factor Analysis (EFA) of INCP- Subtests

The purpose of the exploratory analysis was to determine whether the data structure accurately represents the cognitive domains outlined in the DSM-V, which are relevant for the diagnosis of dementia. As this thesis had an exploratory nature, an exploratory factor

analysis (EFA) was conducted, including all subtests of the INCP. However, some of the preconditions were violated.

The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy yielded a value of .539, signifying a poor level of adequacy for the sample. The Bartlett's test of sphericity was significant ($p < .05$). Kaiser (1990) recommends as the lower acceptable limit to proceed with a factor analysis KMO = .50, indicating that the correlations between the variables in the present sample were marginally large to conduct an EFA.

Only factors with eigenvalues greater than or equal to 1 were considered (Guttman, 1954; Kaiser, 1960). The examination of Kaiser's criterion and the examination of the scree-plot supported the retention of four factors with eigenvalues exceeding 1, explaining 80.41% of the total variance (see Table A8 in the appendix).

Among the factor solutions, the varimax-rotated two-factor solution provided the most interpretable solution, and most variables exhibited a strong loading on a single factor, but also cross-loadings were displayed. The present data revealed that the performance on the 9 INCP subtests examined in this study can be attributed to four underlying factors. The first factor accounted for 31.53% of the variance, the second factor accounted for 23.20% of the variance, the third factor accounted for 15.44% of the variance and the fourth factor accounted for 10.23% of the total variance.

The subtests FACE (.861) and CITY (.851) displayed a single and strong loading on factor 1. However, there were also cross-loadings from other cognitive domains (language) found, the subtest INT showed a strong negative factor loading on factor 1 (-.827), the subtests VVT (.490), and AVT (.529) showed a moderate and positive loading on factor 1. Factor 1 seems to primarily represent a cognitive domain related to memory and language.

The subtest FPT displayed a strong factor loading (.770) on factor 2, as well as the subtest DST, which displayed a moderate loading (.620) on factor 2. As indicated by the loadings of the subtests FPT and DST, the data of factor 2 implied to capture a cognitive domain related to attention and learning.

As indicated by the single strong loading of the subtest TLT-s (.903) on factor 3, factor 3 seems to represent the cognitive domain of executive function. However, there were also moderate cross-loadings from other domains, such as the complex attention task DST (.639), and the episodic memory task FPT (.438).

The subtest EFT-s (.889) showed a strong and single factor loading on factor 4. However, there are some cross-loadings from other cognitive domains, the subscale VVT

(.719) showed a strong loading on factor 4 as well as the subtest AVT, which displayed a moderate positive loading on factor 4, implying that factor 4 represents a cognitive domain associated with social cognition and language.

The scree plot depicting the results of the EFA is presented in the appendix (see Figure B 1).

Comparison of the z-Values of the NTBV-Subtests of the Subsamples HC and MCI

In order to evaluate the exploratory question of research, whether neurologically healthy controls (HC) and people diagnosed with mild cognitive impairment (MCI) show differences in z-values for each subtest of the NTBV, a Mann-Whitney-U test was calculated. The preconditions were checked; effect sizes were calculated for each subtest of the NTBV test battery. There is an overview of descriptive statistics and test statistics in the appendix (see Table A 9).

Comparison of the z-Values of the INCP- Subtests of the Subsample HC and MCI

Similar to a study by Rosas et al. (2022), in the present study a Mann-Whitney-U test was calculated for each examined subtest of the INCP to determine if neurologically healthy controls (HC) and people diagnosed with mild cognitive impairment (MCI) show differences in z-values (see Table 5).

The z-scores for the examined INCP subtests were calculated by using the following formula:

$$\frac{x - \mu}{\sigma}$$

The group assignment was based on the screening criteria.

In a further exploratory step, the procedure was repeated with a different subsample composition. The group assignment (HC, MCI) was based on the NTBV subtests (MCI: 1.50 standard deviations below age, gender, and education norms in a single or multiple domains on the NTBV). The results are displayed in the appendix (see Table A 10).

The preconditions were checked. The data gathered through the INCP was not normally distributed (Shapiro-Wilk test $p < .05$, see Figure 7 – 22).

Effect sizes were calculated by using Pearson's correlation coefficient (r):

$$r = |\frac{z}{\sqrt{\mathrm{N}}}|$$

The interpretation of the results followed the criteria set by Cohen (Cohen, 1988); $r < 0.3$ was considered a small effect, $0.3 \leq r < 0.5$ a moderate effect, and $r \geq 0.5$ was referred to as a large effect. Only subtests with moderate to large effect sizes are reported, but an overview of all INCP subtests is provided in Table 5.

Table 5

Descriptive Statistics and Tests Statistics of the Z-Values of the INCP Subtests. Comparison of Healthy Controls and Mild Cognitive Impairment.

Subtests	Group	*N*	*z*	*Mdn*	*U*	*Z*	*p*	*r*
DST	MCI	3	0.48 (0.48)	178.00	53.00	-.203	.839	
	HC	38	0.48 (0.91)	178.00				
TLTs	MCI	3	0.45 (0.12)	39.00	39.00	-.965	.335	
	HC	39	-0.03 (1.03)	38.00				
CITY	MCI	5	-1.40 (1.68)	13.00	40.50	-2.01	.045	
	HC	36	0.53 (0.69)	14.50				
FACE	MCI	3	0.45 (0.00)	15.00	21.00	-1.67	.095	
	HC	30	0.76 (1.06)	16.00				
FPT	MCI	9	-1.45 (0.45)	58.00	13.50	-4.28	< .001	.62
	HC	39	0.35 (0.77)	87.00				
AVT	MCI	5	-1.18 (.96)	84.00	7.50	-3.33	< .001	.50
	HC	39	0.32 (0.85)	95.00				
VVT	MCI	10	-1.02 (0.88)	76.50	57.00	-3.23	.001	.48
	HC	35	0.41 (0.85)	87.00				
INT	MCI	10	0.03 (0.15)	99.00	175.00	-.509	.611	
	HC	39	0.11 (1.22)	100.00				
EFT-s	MCI	2	0.02 (0.56)	35.00	24.50	-.49	.621	
	HC	31	0.00 (1.03)	37.00				

Note. Mann – Whitney – U - Test. *z* = Median of z-Values with (*SD*), *Mdn* = Median, *U* = Mann-Whitney-U, *Z* = Z-statistic,

[a] Grouping Variable: Diagnosis (based on screening)

[b] *p* = Asymp. Sig. (2-tailed)

For the CITY subtest, a Mann-Whitney-U-Test was conducted to determine differences between the HC-group and the MCI-group. The distributions did not differ between both groups (Kolmogorov-Smirnov $p>.05$). However, there was a moderate difference in median CITY-score between the HC-group (*Mdn* = 14.50) and the MCI-group (*Mdn* = 13.00), *U*= 40.50, *Z*=-2.01, $p = .045$, $r = .31$.

Regarding the subtest FPT, a Mann-Whitney U-test was calculated to determine if there were any differences between the two groups (HC & MCI). The distributions did not differ significantly between both groups (Kolmogorov-Smirnov $p>.05$). However, there was a large and significant difference in the median FPT-score between the HC-group (*Mdn* = 87.00) and the MCI-group (*Mdn* = 58.00), $U= 13.50$, $Z=-4.28$, $p < .001$, $r =.62$.

Regarding the AVT subtest, a Mann-Whitney U-test was conducted to determine if there were any performance differences across groups (HC & MCI). The distributions did not differ significantly between both groups (Kolmogorov-Smirnov $p>.05$. However, there was a large and significant difference in the median AVT-score between HC- (*Mdn* = 95.00) and MCI-group (*Mdn* = 84.00), $U= 7.50$, $Z=-3.34$, $p <.001$, $r = .50$.

A Mann-Whitney-U-Test was performed to determine if healthy controls and patients with MCI show differences in performance on the VVT-subtest. The distributions did not differ significantly between the groups (Kolmogorov-Smirnov $p>.05$. However, there was a moderate and significant difference in the median VVT-score between neurologically healthy controls (*Mdn* = 87.00) and the MCI-group (*Mdn* = 76.50), $U= 57.00$, $Z=-3.23$, $p < .001$, $r = .48$.

Reliability and Item Difficulty Analysis of INCP- Subtests (INT, AVT, VVT)

Several subtests of the INCP still encompass sample items, which will be dismissed after conducting a careful item analysis at a later stage in the development of the INCP.

This thesis performed, in an exploratory step, an item analysis using the item difficulty index (P_i) and Cronbach's alpha (α). All examined subtests had two characteristics (0 = wrong, 1 = right). The data from the overall sample was used to calculate P_i for the items of the subtests INT, AVT, and VVT within the overall sample by using the following formula:

$$P_i = \frac{N_R}{N}$$

Image Naming Test (INT). For reliability analysis, Cronbach's alpha was calculated to assess the internal consistency of the subscale INT, which comprised 107 items. The goal was to dismiss those 7 items, which had the lowest internal consistency.

The internal consistency of the subscale was acceptable, with α = .793. Based on reliability analysis, the following items were excluded: 206, 450, 330, 409, 486, 567, 425, which led to a high internal consistency (α = .834).

To examine which items were solved by 100% of the total sample, P_i was calculated. $P_i = 1$ was displayed by the following items: 12, 22, 39, 42, 100, 116, 150, 175, 182, 185, 187, 227, 267, 312, 371, 388, 434, 444, 447, 456, 457, 483, 531, 568.

Auditory Vocabulary Test (AVT). The subtests AVT consisted of 109 items, of which 9 items are considered sample items.

The item difficulty (P_i) for the items of the subtests AVT was calculated. The following items were solved by 100% of the overall sample: 1.17, 1.57, 2.3, 2.15, 2.17, 2.35, 3.14, 3.29, 3.37, 3.46, 4.1, 4.44, 4.55, 5.19, 5.29, 6.48, 7.36, 8.4, 8.57, 9.23, 10.13, 10.59.

For reliability analysis, Cronbach's alpha was calculated to assess the internal consistency of the subtest AVT. In the overall sample, the internal consistency of the questionnaire was high (α = .839). Based on reliability analysis, the following 9 items were excluded: 9.33, 9.2, 7.51, 9.58, 5.9, 5.4, 8.46, 6.26, 10.22. After the exclusion of the above-mentioned items, reliability analysis was repeated with the remaining 100 items, which lead to a high internal consistency (α = .870).

Verbal Vocabulary Test (VVT). For reliability analysis, Cronbach's alpha was calculated to assess the internal consistency of the subscale VVT, which consisted of 109 items. The internal consistency of the questionnaire was acceptable (α = .796). Based on reliability analysis, the following 9 items were excluded: 5.15, 8.11, 8.14, 8.36, 9.56, 10.6, 10.7, 10.18, 10.52; reliability analysis was repeated with the remaining 100 items, leading to high internal consistency (α = .839).

The following items: 1.13, 1.14, 1.32, 1.55, 2.43, 2.51, 3.15, 3.39, 3.42, 4.11, 4.28, 4.31, 4.57, 5.54, 6.59, 8.25, 8.26, 9.6, 9.42, 10.55 were solved by 100% ($P_i = 1$) of the total sample.

Discussion

Due to the very limited options in the treatment of AD, preventive strategies such as the development of new technologies that detect cognitive decline in its preclinical stage without burdening the health care sector became increasingly critical for managing the global challenges posed by the rising prevalence of dementia. The Department of Neurology addressed the dementia diagnosis gap by developing a remote, self-administered, digital neurocognitive assessment battery. As the INCP still lacks psychometric criteria, the initial focus of the present thesis was to unravel the intricate web of relationships among the various variables of the INCP with the purpose of offering first insight into the INCP's position as a valid tool for identifying cognitive decline in its prodromal stage. The present study employed a methodology to unveil the relationships between domain specific INCP variables. Following this, the discussion extends to a comparative analysis between INCP variables and their convergence with another valid and clinically applied neuropsychological test battery (NTBV). By correlating the performance of these two assessment batteries, the study aimed to discern the extent to which the INCP aligns with the dementia framework proposed by the DSM criteria. The validity of the INCP was further assessed through an exploratory factor analysis, which provided initial insights into its underlying structure and the discriminatory power was evaluated. These steps are crucial for establishing the validity of the INCP and evaluating its potential as a valid tool for clinicians and researchers.

The following section discusses the strength of correlations as well as the identified patterns within a cognitive domain that align with the convergent validity of the INCP. Furthermore, the findings are interpreted and integrated into the state of the research, limitations are discussed, and implications for further research are provided.

Summary of Findings

Cognitive Domain of (Learning) and Memory

This sequence discusses the associations between CITY and FACE (INCP), which are intended to assess semantic memory by using verbal input, and VSRT (NTBV) as well as FPT (INCP), which are intended to evaluate episodic memory through visual input.

Within the INCP test battery, the large and positive association between the CITY- and the FACE-subtests suggests that CITY and FACE assess the same construct. These findings are consistent with a pilot study on the development of the INCP conducted by Maierhofer (2023).

However, the findings indicated a moderate negative correlation between the FPT- and the CITY-tasks. Furthermore, distinct factor loadings highlight differences in their underlying construct.

For further evaluation of the INCP's convergent validity domain-specific variables of the INCP were correlated with another validated and widely used neurocognitive test battery (NTBV). Notably, the data showed negative associations between the subtest CITY (INCP) and the VSRT delayed recall score (NTBV), and no significant correlation between FACE (INCP) and the VSRT subtests (NTBV) was found, suggesting that the examined subtests represent divergent cognitive domains.

Conversely, moderate to large positive correlations were found between the FPT task (INCP) and all the VSRT subtest scores (NTBV). Both subtests are intended to measure episodic memory. These findings corroborate a previous study on the relationship between CITY, FACE, and the NTBV test battery, implying that there is no association between semantic and episodic memory components (Lehrner et al., 2017).

In an exploratory step, an EFA revealed that the CITY- and FACE-scores load highly positively on the same factor (1). These findings underline once more the assumption that CITY and FACE are assessing the same construct.

However, not all results of the EFA were consistent with the initial conceptualization of the INCP test battery, since the FPT subtest showed a low negative factor loading on factor 1 and a highly positive factor loading on factor 2.

The reason for the observed discrepancies may be explained by the different task characteristics. In contrast to the NTBV, which only provides tasks to assess the cognitive domain of memory, the INCP assesses the domain of learning and memory. Since the task characteristics of CITY and FACE are designed to measure semantic memory by recalling previously learned information, it can be assumed that these tasks represent the cognitive domain of memory. The FPT, on the other hand, is designed to measure episodic memory by learning new information and thus may represent the cognitive domain of learning.

Despite the initial assumption of commonality between CITY, FACE, and FPT, these results align with prior research, indicating that episodic memory and semantic memory are distinct processes with independent neural bases (Lehrner et al., 2017).

Although semantic and episodic memory belong to the construct of explicit memory (Lehrner et al., 2006), research yields mixed results for these two memory systems. Episodic memory can be uniquely impaired in healthy aging, whereas semantic memory is only

affected in people with dementia. Hence, the impairment of semantic memory is seen as a hallmark criterion in the diagnosis of dementia (Lehrner et al., 2017; Nilsson, 2003; Nyberg et al., 1996; Vogel et al., 2005). The present results support these scientific insights, as elderly individuals performed worse on the episodic memory task FPT and better on the semantic memory task FACE.

According to a hypothesis, healthy elderly individuals may find semantic memory tasks less frustrating than episodic memory tests. This may be due to their use of internalizing strategies to explain their low performance on the episodic memory task, such as an inability to gain new information. However, within the semantic memory task, they may use externalizing strategies and associate their low performance with a feeling of 'never learned' and therefore not being responsible (Lehrner et al., 2017). The feedback given by the participants during the assessment procedure aligns with this statement. The above-mentioned results are also in line with pilot studies in the development of the INCP (Maierhofer, 2023; Pekez, 2021), which assume that elderly participants are more familiar with celebrities from the past century and therefore performed better on the FACE task (Pekez, 2021).

Also noteworthy are the results for the HC subsample, as the subsample showed a strong positive correlation with the semantic memory task FACE and a moderate negative correlation with the episodic memory task FPT. These findings support the scientific postulate that episodic memory may also be affected by healthy aging, but semantic memory is not.

As previously stated, the evaluated subtests are intended to assess two memory systems by using different testing modii (visual and verbal), implying mixed results. A study in the development of the INCP by Heidinger and Lehrner (2020) supports the present findings, as it highlights the impact of testing mode (visual, verbal) on participants' performance. The study found differential results in a healthy sample when comparing two memory subtests (CK, FK) with differing testing modii (visual, verbal). Furthermore, a recent study underlines once more the importance of utilizing both visual and verbal tests to evaluate memory, as the employment of various testing modalities enhances sensitivity within the identification of MCI by 27% (Oltra-Cucarella et al., 2019).

In conclusion, the observed discrepancies are consistent with recent scientific literature, which recommends the installation of multiple tests for the evaluation of the complexity of the cognitive domain of learning and memory, including different testing

methods and demographic characteristics. The findings imply convergent validity of all evaluated INCP subtests (CITY, FACE, FPT) and indicate that all subtests should be included in the next version of the INCP.

With further studies in mind, the INCP subtest FPT and its associations between other domain variables of the INCP are discussed in the following:

The results indicate a moderately positive correlation between the learning and memory subtest FPT and the subtest TLT-s, which is meant to assess executive function, inhibition, and cognitive flexibility. Additionally, the subtest FPT showed a moderate and positive correlation with the complex attention task DST. The data implied further moderate to large associations between the episodic memory task FPT (INCP) and most of the NTBV subtests, which assess attention and executive function performance. These results are consistent with previous research on the relationship between memory, attention, executive functions, and motivational variables (Lehrner et al., 2017). According to Kontaxopoulou et al. (2017), episodic memory is strongly associated with almost all attentional and executive tests, while semantic memory performance is not.

Additionally, it can be noted that CITY and FACE showed moderately positive associations with language tasks from the INCP test battery and with attention variables from the NTBV test battery.

Cognitive Domain of Language

Within the INCP test battery, the large and positive association between the subtest AVT and the subtest VVT provided initial support for the assumption that both tasks are measuring the same construct.

This finding is further corroborated by the correlations between the INCP subtests AVT and VVT and the subtests of another validated, construct-similar test battery (NTBV), which are as follows: Both subtests (AVT, VVT) exhibited a strong positive correlation with the NTBV subtest SWT and a moderate positive correlation with the NTBV subtest BNT. Given that the NTBV subtests SWT and BNT are validated measures of the cognitive domain of language, these findings reinforce the initial construction of AVT and VVT, implying convergent validity.

However, the results of the EFA add a layer of complexity to this assumption. While the AVT scores indicated a moderate positive factor loading on factor 1, and the VVT scores

show a high positive factor loading on factor 4, their common cross-loading raises questions about the distinctiveness and convergence of the construct they assess.

The observed cross-loadings suggested a shared association, possibly indicating that both AVT and VVT assess overlapping cognitive domains. Future research should delve into the specific elements contributing to this shared variance, considering theoretical frameworks or task characteristics that may explain these observed patterns. Despite these considerations, the overall pattern of findings, including the strong and positive association between the INCP subtests AVT and VVT, coupled with their robust correlations with the NTBV subtests SWT and BNT, provide compelling evidence for the convergent validity of the subtests AVT and VVT.

However, the scores on the INT subtest deviate significantly from the original conceptualization of the INCP. The INT task exhibited a strong and negative factor loading on factor 1. Furthermore, the overall sample analysis reveals neither a significant correlation with the INCP or the NTBV test battery.

Semantic memory impairment is mostly assessed by tasks evaluating object naming (Lehrner et al., 2017), such as it is required on the BNT or INT task. Another study argues, that the BNT involves beside the semantic system other distinct cognitive processes (Marier et al., 2023). Currently, the INT (INCP) is assigned to the cognitive domain of language. As the INCP is still under development and construct validity must be assessed in further studies, one consideration is that the INT, like the BNT, shows associations with sematic memory tasks of the INCP (FACE, CITY). However, the present data could not find evidence for this hypothesis.

This observed discrepancy may be attributed to potential ceiling effects in the overall sample. One assumption is that the present sample mostly consisted of neurologically healthy individuals, with only a few cases of mild cognitive impairment. Supporting this assumption, a study suggested that naming objects, such as those that are required on the INT or BNT, may lack sensitivity to detect cognitive decline in its prodromal stage as the task is too easy (Werheid & Clare, 2007). The authors attempt to explain the latter on the basis of neuronal connectivity, which may be stronger with concrete nouns that describe a whole class of objects, and which are more frequently used in an individual's lifetime (Werheid & Clare, 2007). Remarkably, within the MCI subsample, the scores of the paper and pencil test BNT (NTBV) are strongly and positively correlated with the similarly constructed digital task INT. This finding aligns with existing research, suggesting that the lexical representation of

objects, as demonstrated in BNT and INT, may become more vulnerable to decline in the presence of advanced semantic memory impairment, such as it is seen in MCI or dementia (Joubert et al., 2010).

In conclusion, the findings suggest the inclusion of the INT in the next version of the INCP, as the test battery is not designed for healthy controls but for a clinical sample that is affected by cognitive impairment.

With further research in mind, other correlations within the INCP test battery are discussed in the following.

Moderately and significantly positively correlated are the language subtest VVT and the semantic memory tasks CITY and FACE.

The AVT-subtest is designed to measure language skills, auditory perception, and comprehension skills but exhibits a moderate positive correlation with the semantic memory subtest CITY and with several subtests scores from the episodic memory task (VSRT) of the NTBV.

Further associations were found between the language task AVT and multiple variables from the NTBV test battery, specifically with the executive function domains of planning and nonverbal fluency, as well as with interference and attention variables.

The data from the VVT subtest is also related to multiple executive function variables of the NTBV test battery, namely phonemic word fluency, interference, and attention.

Due to a lack of research, further research is needed to clarify the associations between the subtests AVT, VVT, INT, and other memory, attention, and executive function variables.

Cognitive Domain of (Complex) Attention

As the research adopts an exploratory approach, several initial hypotheses, particularly those related to the cognitive domain of complex attention ($H1_1$), could not be fully analyzed due to technical problems with the application software. Unfortunately, only one subtest (DST) out of two subtests (DST and PCT) could be administered to the participants, leaving questions about the cognitive domain of complex attention unanswered within this thesis.

Nevertheless, the administered complex attention task, DST, revealed a moderate, positive correlation with the memory subtest, FPT. Additionally, associations emerged between DST and other domain variables of the NTBV, specifically between the complex

attention task DST and executive function-planning, nonverbal fluency, and executive function-interference variables from the NTBV.

Exploratory factor analysis unveiled highly positive factor loadings on two factors (2 and 3). The DST shared factor loadings with two other domain-independent variables, FPT and TLT-s, indicating a potential overlap between these variables. This raises questions about the distinctiveness and convergence of the DST and a task measuring episodic memory (FPT), and the TLT-s subtest, which is intended to assess executive functioning, inhibition, and cognitive flexibility.

These results align with recent research highlighting positive associations between memory, attention, and executive function performance (Lehrner et al., 2017).

Nonetheless, further research should delve into specific variables contributing to this shared variance, considering theoretical frameworks or task characteristics that may explain the observed patterns. The limitations encountered in this study underscore the need for continued exploration in understanding the interplay between memory, attention, and executive function within the INCP test battery. Future research should consider the observed relationships and address the unanswered questions posed by the technical challenges encountered during the assessment procedure.

Cognitive Domain of Executive Function

The INCP assesses the cognitive domain of executive function with several subtests (TLT-s, EST, TDT, DICE, FFT). However, in the current study, only one subtest (TLT-s) was submitted to the participants due to reoccurring technical problems with the tablet. Hence, $H1_2$ remained unanswered, and the question of whether the cognitive domain of executive function is accurately represented within the INCP cannot be answered within this thesis.

Regarding the TLT-s subtest, an EFA revealed a highly and single positive factor loading on Factor 3. Additionally, the absence of cross-loadings is considered a strength and underscores the initial assumption that the TLT-s task represents a distinct cognitive domain.

However, further research is necessary to validate this assumption, as the TLT-s data did not show significant correlations with the proposed executive function domain of the well-established neuropsychological assessment battery, NTBV.

Contrary to the initial assumption, correlations were found between TLTs and domain independent INCP variables. Specifically, the TLT-s correlated with the social cognition task EFT-s as well as with the episodic memory tasks FPT (INCP) and VSRT (NTBV).

Future research should explore potential explanations for these unexpected correlations and address the unanswered questions regarding the representation of executive function within the INCP.

Cognitive Domain of Social Cognition

To decrease the likelihood of misdiagnosing MCI in elderly patients, research suggests the implication of multiple tests per cognitive domain. However, the current prototype of the INCP only provides one subtest (EFT-s) and the assessment battery NTBV doesn't provide any subtest for the assessment of the cognitive domain of social cognition. Hence, the question whether the cognitive domain of social cognition is correctly represented within the INCP cannot be answered within the thesis.

With an eye toward further research, the following section discusses some potential correlations.

The social cognition and emotion recognition task EFT-s displayed a moderate positive correlation with the executive function task TLT-s and implied relations with variables of attention, executive function planning, and nonverbal fluency and interference, as well as with the memory domain from the NTBV assessment battery. These findings are coherent with prior studies, implying that social cognition tasks are mostly confounded with other cognitive domains due to their high complexity, such as memory, executive function (Singleton et al., 2023), perception, and language (Forbes & Grafman, 2010).

Neuroimaging has revealed that social cognition is a highly complex process involving the interplay between elementary sensory, implicit, and explicit cognitive processes (Forbes & Grafman, 2010). The EFT-s is intended to measure social cognition and emotion recognition trough simplified emoticons. The simplified test characteristics raise the question of whether the EFT-s accurately represents the highly complex process of social cognition. Furthermore, the EFT-s involves a forward condition and a reversed condition, in which participants are required to press "Happy", when a sad face appears and vice-versa. This raises the question of whether task characteristics may elucidate the observed patterns with executive function variables.

However, an exploratory factor analysis implied a unique and highly positive factor loading on only one factor, with no cross-loadings found. These findings are considered a strength and underscore the initial assumption that the EFT-s represents a distinct cognitive domain.

However, the social cognition task EFT-s showed shared factor loading with another domain-independent variable from the language domain, namely VVT, indicating a potential overlap among these subtests. A study on social cognition argues that the evolution of the social brain may have stimulated the development of language (Forbes & Grafman, 2010). Remarkably, the EFT-s and VVT subtests did not show any significant correlation within the correlation matrix in the overall sample, which raises questions about the distinctiveness and convergence of the social cognition task EFT-s.

Future research should address the discrepancies observed between EFT-s and VVT, as well as the unanswered questions regarding the representation of social cognition within the INCP.

Exploratory Analysis

The exploratory question of the research aimed to examine whether the newly developed INCP subtests could significantly differentiate between healthy individuals (HC) and those diagnosed with mild cognitive impairment (MCI). Employing a comparative approach, the study utilized the nonparametric Mann-Whitney U-Test to analyze performance differences between the two groups. The statistical analysis revealed performance differences between neurologically healthy controls and individuals with mild cognitive impairment for several subtests of the INCP. A moderate effect was observed for the subtests CITY, while a large effect was found for the subtests FPT, AVT, and VVT. These findings support the potential discriminatory power of the INCP.

While these findings provide initial insights into the discriminatory potential of the INCP subtests, it is essential to acknowledge the limitations of the exploratory approach, emphasizing the need for cautious interpretation.

Especially the group assignment method based on the screening criteria is considered a major limitation. To evaluate whether the results differ when group assignment is based on the performance on the NTBV, it is noted that discriminatory potential was also found for the subtests FACE, DST, and EFT-s.

These findings align with prior research, suggesting that the knowledge of world capital identification, as it is required on the CITY subtest, is more sensitive for the prodromal stage of dementia, than the knowledge of famous people, as it is required on the FACE subtest (Lehrner et al., 2017).

To strengthen the discriminatory power of the INCP, future research could explore additional factors, such as demographic characteristics or clinical variables, that may influence performance differences between healthy individuals and those with mild cognitive impairment.

Further research is needed to determine the ability of the INCP to discriminate across the full spectrum of the prodromal stage of AD, such as SCD (non)- amnestic MCI (single or multiple domain), thereby enhancing the precision of the INCP and improving outcomes for individuals at risk.

Limitations and Future Research

As with most clinical studies, the current research has several limitations, which are addressed in the following sequence.

The INCP is currently considered an experimental test battery that is under constant evaluation and revision. Therefore, the selected exploratory approach is indeed a valuable component of the research process in the development of the digital test.

However, the pilot study design poses many challenges, which limit the external validity of the results. Thus, no general conclusions can be drawn (Arain et al., 2010).

The small total sample size, the small sample size across groups, specifically in the MCI group, and the varying sample size per subtest affect the generalizability and the interpretability of the present findings. The different sample sizes per INCP subtest can be explained. Firstly, participants were given two different versions of the test battery (short and long) based on their screening scores, and secondly, technical issues with the application resulted in ongoing adaptations of the test battery composition. However, the pilot study design allows for flexibility and justifies the adaptations made during the examination.

Although dementia is a disease of the aging population, with more women affected and education considered a protective factor, the unbalanced sample, with more female participants and a younger, better-educated, healthy subsample, raises concerns about the representativeness of the findings. Furthermore, the sample is limited to German-speaking individuals, which clearly restricts the generalizability of the findings to multicultural

populations. This emphasizes the need to implement broader inclusion strategies. Additionally, the data was not normally distributed, and homogeneity of variance was violated.

Follow-up research should replicate the present study with a larger sample size while prioritizing the balanced distribution of demographic variables, such as age, gender, and years of education, across experimental groups. Establishing well-matched groups will enhance the validity of the study by minimizing the risk of skewed results due to confounding variables. This will strengthen the generalizability and interpretability of the results.

Due to ongoing Corona-19 restrictions, such as lockdowns, the mandatory use of FFP-2 masks, and a PCR-test obligation, the study was accompanied by several challenges. Especially the recruitment presented several difficulties, leading to a very specific cohort.

The ongoing pandemic restrictions impacted participants' motivation and concentration during the assessment. While the examination lasted for approximately 4 to 5 hours, requiring a high level of concentration, some study participants reported that wearing an FFP-2 mask restricted their oxygen supply, which affected their ability to concentrate. Although the test instructor provided immediate feedback, compliance and motivation issues were observable. The elderly cohort experienced significant fatigue after completing the NTBV assessment battery, which prevented them from continuing with the INCP test battery. As a result, there were many test aborts, which consequentially led to a very small sample size, particularly in the MCI group.

During the recruitment process, it was observed that almost exclusively individuals with subjective cognitive complaints (initial exclusion criteria) registered for the neurologically healthy control group to undergo cognitive function check-ups. This can be interpreted as only highly motivated individuals were willing to accept the high requirements posed by hospitals during the COVID-19 pandemic. As a result, the initial group assignment criteria based on the NTBV results were dismissed and replaced by screening results. Consequently, some participants reported SCD during medical condition screening but still met the new criteria for the assignment to the HC-group based on the screening, although they essentially fulfilled the MCI criteria, which is defined as 1.50 standard deviations below age, gender, and education norms in a single or multiple domains in a neurocognitive assessment battery (NTBV). Despite meeting the criteria for the MCI group, they were assigned to the HC group. The VVT 3.0 screening mean scores in both groups were very high. These findings are in line with the results of a study conducted by Tokaj and Lehrner

(2023), who argue that ceiling effects in the VVT 3.0 screening may limit its sensitivity, potentially leading to high scores in the healthy- and the MCI groups. As a matter of fact, the reliability of grouping based on screening assessments such as MMSE and VVT 3.0 is questionable, as the authors found that none of these screening tests reliably differentiate between healthy individuals and those with mild cognitive impairment (Tokaj & Lehrner, 2023). As a result, the HC-group consisted of 32 individuals who performed below 1.5 standard deviation in a single or multiple domains on the NTBV.

While the pilot study design allows for flexibility and justifies the group assignment method, it raises questions about reliability and validity. Former research should address the acknowledged limitations and be cautious of the risk of inadequate recruitment, which may introduce potential confounding factors. To enhance the interpretability of results, follow-up studies should refine methodology and allocate participants to the MCI-group based on the criteria proposed by Petersen (2004).

Additionally, the comparison of domain-specific subtests between INCP and NTBV is constrained by the different methodologies used for z-score calculations. The NTBV provides normed z-scores using state-of-the-art statistical procedures (GAMLSS) based on age, sex, and education (Pusswald et al., 2013), whereas the INCP z-scores were calculated using a simple z-standardization due to the lack of valid norms for the INCP in its early stage of development.

The lack of valid norms for the INCP limits the comparability of results, emphasizing the necessity for standardized measures.

Field experiences have shown that the design of the digital assessment battery poses challenges for elderly individuals, especially for those with severe cognitive impairment, physical restrictions such as poor eyesight, poor hearing ability, arthritis, or limited experience with digital technology.

Furthermore, the clinical setting introduced environmental distractions, emphasizing the importance of a controlled testing environment in future research.

Although the INCP is designed for remote use, the current study was conducted in a clinical setting due to pandemic restrictions. Therefore, it is not possible to make assumptions about how elderly individuals handle the tablet by themselves without the motivation, technical, and handling support of the test supervisor. This highlights the need to replicate the study in a remote environment.

To enhance the validity assumption of the INCP, subsequent studies should examine the convergent validity of the INCP, by comparing the INCP, administered in a remote setting, with a validated paper-and-pencil neurocognitive assessment battery, such as the NTBV, administered in a face-to-face setting (Kochan et al., 2022).

The limitations encountered in this study underscore the need for continued exploration in understanding the interplay between descriptive variables, INCP subtests, and tasks of a likewise test battery, considering item analysis, theoretical frameworks, or task characteristics that may elucidate the observed patterns. Future research should address the hypothesis, which remained unanswered within this thesis, and further evaluate the associations, distinctiveness, and convergence of domain-specific subtests posed in the INCP.

Despite these limitations, the identified associations and underlying factors provide valuable first-hand insights into the construct validity of the INCP. However, due to the lack of prior studies in the field of the INCP, additional research is needed to clarify the associations between the subtests and the underlying structure of the subtests.

By addressing the limitations highlighted within this thesis, future research should contribute to the development of robust norms for the INCP; valid norms that account for the influences of age, gender, and education can enhance the comparability of cognitive assessments and facilitate more accurate interpretations in both clinical and research settings.

Additionally, incorporating a longitudinal perspective may provide insights into the dynamic nature of cognitive decline along the disease continuum.

Based on the insights gained from this pilot study, future research should address the acknowledged limitations and refine methodology to enhance the INCP's validity and applicability in diverse populations. Further correlation studies are necessary to better understand patterns of cognitive impairment, which will lead to a deeper understanding of the preclinical stages of dementia. This study sheds light on potential solutions to overcome biases and lays the foundation for further studies in the development of the INCP, paving the way for more targeted research in the future.

Conclusion

While this study provides initial insights into the intricacies of variables, the underlying structure, and the discriminatory potential of the INCP, it is essential to acknowledge the limitations of the exploratory approach, emphasizing the need for cautious interpretation.

The evaluation of the cognitive domain of learning and memory revealed both convergent and divergent validity. Notably, a positive correlation between the semantic memory tasks CITY and FACE supported their shared construct, which is consistent with prior research. However, the semantic memory task FPT exhibited a distinct pattern and showed positive associations with all the VSRT subtests, indicating differences in the underlying constructs. The conducted EFA supported the assumption of commonality between CITY and FACE and identified differences in the factor loadings of FPT. The findings are consistent with literature, suggesting divergent neural bases for semantic and episodic memory and the submission of multiple subtests with various testing modii when it comes to the assessment of the complexity of memory. The present findings underscore convergent validity in the domains of learning and memory.

Regarding the domain of language, the strong and positive correlation between AVT and VVT, as well as the substantial associations with the subtests SWT and BNT (NTBV), indicated convergent validity. The EFA revealed common cross-loadings, implying a shared association. However, the INT subtest deviated from the original conceptualization in the overall sample but showed a strong positive correlation with the likewise paper pencil test BNT in the patient group. The results are in line with previous research, indicating that the similarly constructed BNT underlies ceiling effects.

Some of the initial research questions could not be answered within this thesis.

The question of whether the domain of complex attention is accurately represented through the subtest DST remained unanswered. However, the correlation matrix and the conducted EFA raised questions about the distinctiveness and interconnectivity of attention (DST), memory (FPT), and executive function (TLT-s).

Furthermore, this study left the question of whether the cognitive domain of executive function is accurately represented within the INCP unanswered. The EFA indicated a high factor loading of TLT-s on a single factor. However, the correlation matrix suggested unexpected correlations with other domain-independent variables of the NTBV. Future research is needed to explain those unexpected associations.

Consistent with the literature, the examination of the cognitive domain of social cognition revealed potential correlations with memory, executive function, and language variables. The EFA results implied that the subtest EFT-s represent a distinct cognitive domain. However, questions about whether the simplified task characteristic represents the highly complex process of social cognition adequately were raised.

The exploratory analysis provided preliminary evidence supporting the discriminatory potential of some of the INCP subtests, particularly for the language, learning and memory domain.

However, it is crucial to acknowledge the preliminary nature of the present finding. Cautious interpretation is essential, and further research with larger sample sizes and confirmatory approaches is needed to validate these findings.

In summary, this thesis provides valuable insights into the structure of the INCP. The study highlights its strengths, weaknesses, and areas for improvement. The findings underscore the complexity of cognitive assessments and emphasize the importance of a multifaceted approach to capture the intricacies of cognitive impairment in its prodromal stage. The preliminary findings underline that the INCP may be a valid and user-friendly tool for monitoring cognitive function. However, further refinement and validation efforts are essential to solidify the INCP's position as a valid tool for assessing cognitive decline not only in research setting but also in a remote environment. The emphasis on validity lays the groundwork for advancing the implementation of the INCP in the context of dementia research, early diagnosis, and prevention, thereby enabling early intervention for those at risk.

References

Alty, J., Bai, Q., Li, R., Lawler, K., St George, R. J., Hill, E., Bindoff, A., Garg, S., Wang, X., Huang, G., Zhang, K., Rudd, K. D., Bartlett, L., Goldberg, L. R., Collins, J. M., Hinder, M. R., Naismith, S. L., Hogg, D. C., King, A. E., & Vickers, J. C. (2022). The TAS Test project: A prospective longitudinal validation of new online motor-cognitive tests to detect preclinical Alzheimer's disease and estimate 5-year risks of cognitive decline and dementia. *BMC Neurology*, *22*(1), 266. https://doi.org/10.1186/s12883-022-02772-5

Amariglio, R. E., Becker, J. A., Carmasin, J., Wadsworth, L. P., Lorius, N., Sullivan, C., Maye, J. E., Gidicsin, C., Pepin, L. C., Sperling, R. A., Johnson, K. A., & Rentz, D. M. (2012). Subjective cognitive complaints and amyloid burden in cognitively normal older individuals. *Neuropsychologia*, *50*(12), 2880–2886. https://doi.org/10.1016/j.neuropsychologia.2012.08.011

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders,* 5th edition. American Psychiatric Association.

Arain, M., Campbell, M. J., Cooper, C. L., & Lancaster, G. A. (2010). What is a pilot or feasibility study? A review of current practice and editorial policy. *BMC Medical Research Methodology*, *10*, 67. https://doi.org/10.1186/1471-2288-10-67

Arnold, D. M., Burns, K. E. A., Adhikari, N. K. J., Kho, M. E., Meade, M. O., Cook, D. J., & Group for the McMaster Critical Care Interest. (2009). The design and interpretation of pilot trials in clinical research in critical care. *Critical Care Medicine*, *37*(1), S69. https://doi.org/10.1097/CCM.0b013e3181920e33

Bak, P. M. (2023). Prävention und Gesundheitsförderung. In P. M. Bak (Ed.), *Gesundheitspsychologie: Eine Einführung – kompakt, prägnant und anwendungsorientiert* (pp. 103–120). Springer. https://doi.org/10.1007/978-3-662-67181-8_6

Balogh, E. P., Miller, B. T., Ball, J. R., Care, C. on D. E. in H., Services, B. on H. C., Medicine, I. of, & The National Academies of Sciences, E. (2015). The diagnostic process. In *Improving diagnosis in health care*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK338593/

Barclay, L., & Rees, G. (2020). *Age, dementia and the allocation of health resources during and beyond COVID-19*. Alzheimer's Disease International. https://unsdg.un.org/sites/default/files/2020-05/Policy-Brief-The-Impact-of-COVID-19-on-Older-Persons.pdf

Belleville, S., LaPlume, A. A., & Purkart, R. (2023). Web-based cognitive assessment in older adults: Where do we stand? *Current Opinion in Neurology*, *36*(5), 491–497. https://doi.org/10.1097/WCO.0000000000001192

Bloland, P., Simone, P., Burkholder, B., Slutsker, L., & De Cock, K. M. (2012). The role of public health institutions in global health system strengthening efforts: The US CDC's perspective. *PLoS Medicine*, *9*(4), 1001199. https://doi.org/10.1371/journal.pmed.1001199

Chan, W. C., Lee, A. T. C., & Lam, L. C. W. (2021). Exercise for the prevention and treatment of neurocognitive disorders: New evidence and clinical recommendations. *Current Opinion in Psychiatry*, *34*(2), 136. https://doi.org/10.1097/YCO.0000000000000678

Chow, S.-C., & Chang, M. (2008). Adaptive design methods in clinical trials – A review. *Orphanet Journal of Rare Diseases*, *3*(1), 11. https://doi.org/10.1186/1750-1172-3-11

Clouston, S. A. P., Smith, D. M., Mukherjee, S., Zhang, Y., Hou, W., Link, B. G., & Richards, M. (2020). Education and cognitive Decline: An integrative analysis of global longitudinal studies of cognitive aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *75*(7), 151–160. https://doi.org/10.1093/geronb/gbz053

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed).* L. Erlbaum Associates.

Cyr, A.-A., Romero, K., & Galin-Corini, L. (2021). Web-Based Cognitive Testing of Older Adults in Person Versus at Home: Within-Subjects Comparison Study. *JMIR Aging*, *4*(1), e23384. https://doi.org/10.2196/23384

Deutsche Gesellschaft für Neurologie & Deutsche Gesellschaft für Psychiatrie und Psychotherapie, Psychosomatik und Nervenheilkunde. (2023). *S3-Leitlinie "Demenzen", Version 4.0*. https://register.awmf.org/de/leitlinien/detail/038-013

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics (5th* ed.). The Bookwatch.

Fok, R. A., Zehetmayer, S., Brücke, C., & Lehrner, J. (2023). Validity of the Mini-Mental State Examination, the Clock Drawing Test, and the Vienna Visuo- Constructional Test 3.0-Screening for diagnosing Parkinson's disease. *Zeitschrift Für Neuropsychologie, 34*(4), 203-210. https://doi.org/10.1024/1016-264X/a000385

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Forbes, C. E., & Grafman, J. (2010). The role of the human prefrontal cortex in social cognition and moral judgment. *Annual Review of Neuroscience*, *33*(1), 299–324. https://doi.org/10.1146/annurev-neuro-060909-153230

Gates, N. J., & Kochan, N. A. (2015). Computerized and on-line neuropsychological testing for late-life cognition and neurocognitive disorders: Are we there yet? *Current Opinion in Psychiatry*, *28*(2), 165. https://doi.org/10.1097/YCO.0000000000000141

Gauthier, S., Rosa-Neto, P., Morais, J., & Webster, C. (2021). *World Alzheimer Report 2021: Journey through the diagnosis of dementia* (p. 314). Alzheimer's Disease International.

Geddes, M. R., O'Connell, M. E., Fisk, J. D., Gauthier, S., Camicioli, R. & Ismail, Z. (2020). Remote cognitive and behavioral assessment: Report of the Alzheimer Society of Canada Task Force on dementia care best practices for COVID-19. *Alzheimer's & Dementia. Diagnosis, Assessment & Disease Monitoring 12*(1), 12111. https://doi.org/10.1002/dad2.12111

Graham, J. E., Rockwood, K., Beattie, B. L., Eastwood, R., Gauthier, S., Tuokko, H., & McDowell, I. (1997). Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *Lancet, 349*(9068), 1793–1796. https://doi.org/10.1016/S0140-6736(97)01007-6

Grigoryeva, D., Beck, B., & Lehrner, J. (2024). Dementia prevention: Digital assessment and cognitive function monitoring. *Psychologie in Österreich*, *1*, 18–23.

Guttman, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, *19*(2), 149–161. https://doi.org/10.1007/BF02289162

Hanke, S., Mangialasche, F., Bödenler, M., Neumayer, B., Ngandu, T., Mecocci, P., Untersteiner, H., & Stögmann, E. (2022). AI-based predictive modelling of the onset and progression of dementia. *Smart Cities*, *5*(2), 700-714. https://doi.org/10.3390/smartcities5020036

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck Depressions-Inventar (BDI-II) (Revision ed.)*. Harcourt.

Heidinger, T., & Lehrner, J. (2020). Comparing a visual and verbal semantic memory test on the effects of gender, age and education as assessed in a cognitively healthy sample: A pilot study in the development of the international neuropsychological test profile—A tablet-based cognitive assessment. *Neuropsychiatrie*, *34*(3), 140–147. https://doi.org/10.1007/s40211-020-00355-9

Hoffmann, C. M., Nianogo, R. A., Yaffe, K., Rosenwohl-Mack, A., Carrasco, A., & Barnes, D. E. (2022). Importance of accounting for regional differences in modifiable risk factors for Alzheimer's Disease and related dementias: The case for tailored interventions. *Journal of Alzheimer's Disease: JAD*, *89*(2), 563–570. https://doi.org/10.3233/JAD-220278

Horizon 2020. (2021). *LETHE (λήθη) – A personalized prediction and intervention model for early detection and reduction of risk factors causing dementia, based on AI and distributed machine learning*. European Commission. https://doi.org/10.3030/101017405

Hugo, J., & Ganguli, M. (2014). Dementia and cognitive Impairment: Epidemiology, diagnosis, and treatment. *Clinics in Geriatric Medicine*, *30*(3), 421–442. https://doi.org/10.1016/j.cger.2014.04.001

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., Holtzman, D. M., Jagust, W., Jessen, F., Karlawish, J., Liu, E., Molinuevo, J. L., Montine, T., Phelps, C., Rankin, K. P., Rowe, C. C., Scheltens, P., Siemers, E., Snyder, H. M., … Contributors. (2018). NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *14*(4), 535–562. https://doi.org/10.1016/j.jalz.2018.02.018

Jack Jr., C. R., Albert, M. S., Knopman, D. S., McKhann, G. M., Sperling, R. A., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*(3), 257–262. https://doi.org/10.1016/j.jalz.2011.03.004

Jansen, W. J., Ossenkoppele, R., Knol, D. L., Tijms, B. M., Scheltens, P., Verhey, F. R. J., & Visser, P. J. (2015). Prevalence of cerebral amyloid pathology in persons without dementia. *JAMA*, *313*(19), 1924–1938. https://doi.org/10.1001/jama.2015.4668

Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., Rabin, L., Rentz, D. M., Rodriguez-Gomez, O., Saykin, A. J., Sikkes, S. A. M., Smart, C. M., Wolfsgruber, S., & Wagner, M. (2020). The characterisation of subjective cognitive decline. *The Lancet. Neurology*, *19*(3), 271–278. https://doi.org/10.1016/S1474-4422(19)30368-0

Jessen, F., Amariglio, R. E., van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., van der Flier, W. M., Glodzik, L., van Harten, A. C., de Leon, M. J., McHugh, P., Mielke, M. M., Molinuevo, J. L., Mosconi, L., Osorio, R. S., Perrotin, A., … Wagner, M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, *10*(6), 844–852. https://doi.org/10.1016/j.jalz.2014.01.001

Jessen, F., Feyen, L., Freymann, K., Tepest, R., Maier, W., Heun, R., Schild, H.-H., & Scheef, L. (2006). Volume reduction of the entorhinal cortex in subjective memory impairment. *Neurobiology of Aging*, *27*(12), 1751–1756. https://doi.org/10.1016/j.neurobiolaging.2005.10.010

Joubert, S., Brambati, S. M., Ansado, J., Barbeau, E. J., Felician, O., Didic, M., Lacombe, J., Goldstein, R., Chayer, C., & Kergoat, M.-J. (2010). The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. *Neuropsychologia*, *48*(4), 978–988. https://doi.org/10.1016/j.neuropsychologia.2009.11.019

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*, 141–151. https://doi.org/10.1177/001316446002000116

Klotz, S., & Gelpi, E. (2021). Neuropathologie der Demenzen. *Wiener Medizinische Wochenschrift*, *171*, 257–273. https://doi.org/10.1007/s10354-021-00848-4

Knopman, D. S., Amieva, H., Petersen, R. C., Chételat, G., Holtzman, D. M., Hyman, B. T., Nixon, R. A., & Jones, D. T. (2021). Alzheimer disease. *Nature Reviews. Disease Primers*, *7*(1), 33. https://doi.org/10.1038/s41572-021-00269-y

Kochan, N. A., Heffernan, M., Valenzuela, M., Sachdev, P. S., Lam, B. C. P., Fiatarone Singh, M., Anstey, K. J., Chau, T., & Brodaty, H. (2022). Reliability, validity, and user-experience of remote unsupervised computerized neuropsychological assessments in community-living 55- to 75-year-olds. *Journal of Alzheimer's Disease: JAD*, *90*(4), 1629–1645. https://doi.org/10.3233/JAD-220665

Kontaxopoulou, D., Beratis, I. N., Fragkiadaki, S., Pavlou, D., Yannis, G., Economou, A., Papanicolaou, A. C., & Papageorgiou, S. G. (2017). Incidental and intentional memory: Their relation with attention and executive functions. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *32*(5), 519–532. https://doi.org/10.1093/arclin/acx027

Lehrner, J. (2021a). *Neuropsychological Test Battery Vienna (NTBV 1.0)* [Manuscript submitted for publication].

Lehrner, J. (2021b). *The International Neurocognitive Test Profile (INCP) – A Tablet-Based Neurocognitive Assessment Battery* [Unpublished Manuscript].

Lehrner, J. (2021c). *Vienna Visuo-Constructional Test 3.0 Screening* [Manuscript submitted for publication].

Lehrner, J., Bodendorfer, B., Lamm, C., Moser, D., Dal-Bianco, P., Auff, E., & Pusswald, G. (2016). Subjective memory complaints and conversion to dementia in patients with subjective cognitive decline and patients with mild cognitive impairment. *Zeitschrift Für Neuropsychologie*, *27*(2), 85–93. https://doi.org/10.1024/1016-264X/a000175

Lehrner, J., Bodner, T., Dal-Bianco, P., & Schmidt, R. (2009a). Demenzsyndrome. In J. Lehrner, G. Pusswald, E. Fertl, W. Strubreither & I. Kryspin-Exner (Eds.), *Klinische Neuropsychologie, Grundlagen – Diagnostik – Rehabilitation* (2nd ed., pp. 375–388). Springer Verlag.

Lehrner, J., Bodner, T., Dal-Bianco, P., & Schmidt, R. (2009b). Klinische Demenzdiagnose. In J. Lehrner, G. Pusswald, E. Fertl, W. Strubreither & I. Kryspin-Exner (Eds.), *Klinische Neuropsychologie, Grundlagen – Diagnostik – Rehabilitation* (2nd ed., pp. 388–389). Springer Verlag.

Lehrner, J., Coutinho, G., Mattos, P., Moser, D., Pflüger, M., Gleiss, A., Auff, E., Dal-Bianco, P., Pusswald, G., & Stögmann, E. (2017). Semantic memory and depressive symptoms in patients with subjective cognitive decline, mild cognitive impairment, and Alzheimer's disease. *International Psychogeriatrics*, *29*(7), 1123–1135. https://doi.org/10.1017/S1041610217000394

Lehrner, J., Gleiß, A., Maly, J., Auff, E., & Dal-Bianco, P. (2006). Der Verbale Selektive Reminding Test (VSRT). *Neuropsychiatrie*, *20*(3), 204–214.

Lehrner, J., Gufler, R., Guttmann, G., Maly, J., Gleiß, A., Auff, E., & Dal-Bianco, P. (2005). Annual conversion to Alzheimer disease among patients with memory complaints attending an outpatient memory clinic: The influence of amnestic mild cognitive impairment and the predictive value of neuropsychological testing. *Wiener Klinische Wochenschrift*, *117*(18), 629–635. https://doi.org/10.1007/s00508-005-0428-6

Lehrner, J., Krakhofer, H., Lamm, C., Macher, S., Moser, D., Klug, S., Dal-Bianco, P., Pirker, W., Auff, E., & Pusswald, G. (2015). Visuo-constructional functions in patients with mild cognitive impairment, Alzheimer's disease, and Parkinson's disease. *Neuropsychiatrie*, *29*(3), 112–119. https://doi.org/10.1007/s40211-015-0141-2

Lehrner, J., Maly, J., Gleiß, A., Auff, E., & Dal-Bianco, P. (2007). Demenzdiagnostik mit Hilfe der Vienna Neuropsychologischen Testbatterie (VNTB): Standardisierung, Normierung und Validierung. *Psychologie in Österreich*, *4*(5).

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., … Mukadam, N. (2020b). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6

Lobo, A., Launer, L. J., Fratiglioni, L., Andersen, K., Di Carlo, A., Breteler, M. M., Copeland, J. R., Dartigues, J. F., Jagger, C., Martinez-Lage, J., Soininen, H., & Hofman, A. (2000). Prevalence of dementia and major subtypes in Europe: A collaborative study of population-based cohorts. Neurologic Diseases in the Elderly Research Group. *Neurology*, *54*(11 Suppl 5), 4-9.

Maierhofer, B. (2023). *The International Neurocognitive Test Profile (INCP) – A tablet-based neurocognitive assessment battery: a pilot study assessing gender effects* [Unpublished Dissertation]. Medical University of Vienna.

Marier, A., Dadar, M., Bouhali, F., & Montembeault, M. (2023). Irregular word reading as a marker of cognitive and semantic decline in Alzheimer's disease rather than an estimate of premorbid intellectual abilities. *Research Square*, rs.3.rs-3381469. https://doi.org/10.21203/rs.3.rs-3381469/v1

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 263–269. https://doi.org/10.1016/j.jalz.2011.03.005

Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(6), 439–451. https://doi.org/10.1111/acps.12336

Mitchell, A. J., & Shiri-Feshki, M. (2009). Rate of progression of mild cognitive impairment to dementia – Meta-analysis of 41 robust inception cohort studies. *Acta Psychiatrica Scandinavica*, *119*(4), 252–265. https://doi.org/10.1111/j.1600-0447.2008.01326.x

Montine, T. J., Phelps, C. H., Beach, T. G., Bigio, E. H., Cairns, N. J., Dickson, D. W., Duyckaerts, C., Frosch, M. P., Masliah, E., Mirra, S. S., Nelson, P. T., Schneider, J. A., Thal, D. R., Trojanowski, J. Q., Vinters, H. V., & Hyman, B. T. (2012). National Institute on Aging–Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: A practical approach. *Acta Neuropathologica*, *123*(1), 1–11. https://doi.org/10.1007/s00401-011-0910-3

Myers, J. R., Glenn, J. M., Madero, E. N., Anderson, J., Mak-McCully, R., Gray, M., Gills, J. L., & Harrison, J. E. (2022). Asynchronous remote assessment for cognitive impairment: Reliability verification of the Neurotrack Cognitive Battery. *JMIR Formative Research*, *6*(2), 34237. https://doi.org/10.2196/34237

National Institute on Aging. (2024, January 19). *What happens to the brain in Alzheimer's Disease?* https://www.nia.nih.gov/health/alzheimers-causes-and-risk-factors/what-happens-brain-alzheimers-disease

Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., Laatikainen, T., Lindström, J., Mangialasche, F., Paajanen, T., Pajala, S., Peltonen, M., Rauramaa, R., Stigsdotter-Neely, A., Strandberg, T., Tuomilehto, J., … Kivipelto, M. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. The Lancet, 385(9984), 2255–2263. https://doi.org/10.1016/S0140-6736(15)60461-5

Nianogo, R. A., Rosenwohl-Mack, A., Yaffe, K., Carrasco, A., Hoffmann, C. M., & Barnes, D. E. (2022). Risk factors associated with Alzheimer disease and related dementias by sex and race and ethnicity in the US. *JAMA Neurology*, *79*(6), 584–591. https://doi.org/10.1001/jamaneurol.2022.0976

Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., Hamad, H. A., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., Ashraf, T., … Vos, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the global burden of disease study 2019. *The Lancet Public Health*, *7*(2), 105–125. https://doi.org/10.1016/S2468-2667(21)00249-8

Nilsson, L.-G. (2003). Memory function in normal aging. *Acta Neurologica Scandinavica*, *107*(179), 7–13. https://doi.org/10.1034/j.1600-0404.107.s179.5.x

Norton, S., Matthews, F. E., Barnes, D. E., Yaffe, K., & Brayne, C. (2014). Potential for primary prevention of Alzheimer's disease: An analysis of population-based data. *The Lancet. Neurology*, *13*(8), 788–794. https://doi.org/10.1016/S1474-4422(14)70136-X

Numrich, J. (2017). *Visuo-constructional functions in patients with subjective cognitive decline (SCD), mild cognitive impairment (MCI), Alzheimer's disease (AD) and Parkinson's disease (PD)* [published dissertation]. Universität Wien.

Nyberg, L., Backman, L., Erngrund, K., Olofsson, U., & Nilsson, L.-G. (1996). Age differences in episodic memory, semantic memory, and priming: relationships to demographic, intellectual, and biological factors. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *51B*(4), 234–240. https://doi.org/10.1093/geronb/51B.4.P234

Oltra-Cucarella, J., Sánchez-SanSegundo, M., Lipnicki, D. M., Crawford, J. D., Lipton, R. B., Katz, M. J., Zammit, A. R., Scarmeas, N., Dardiotis, E., Kosmidis, M. H., Guaita, A., Vaccaro, R., Kim, K. W., Han, J. W., Kochan, N. A., Brodaty, H., Pérez-Vicente, J. A., Cabello-Rodríguez, L., Sachdev, P. S., & Ferrer-Cascales, R. (2018). Visual memory tests enhance the identification of amnestic MCI cases at greater risk of Alzheimer's disease. International Psychogeriatrics, *31*(07), 997–1006. https://doi.org/10.1017/s104161021800145x

Österreichische Alzheimer Gesellschaft. (n.d.). *Symptome und Diagnose*. Österreichische Alzheimer Gesellschaft. http://www.alzheimer-gesellschaft.at/informationen/was-ist-alzheimer/symptome-und-diagnose/

Patterson, C. (2018). *World Alzheimer Report 2018—The state of the art of dementia research: New frontiers* (p. 48). Alzheimer's Disease International.

Pekez, V. (2021). *Pilot study: The International Neurocognitive Test Profile (INCP) a tablet-based neurocognitive assessment battery: Assessing age effects* [unpublished dissertation]. Universität Wien.

Perin, S., Buckley, R. F., Pase, M. P., Yassi, N., Lavale, A., Wilson, P. H., Schembri, A., Maruff, P., & Lim, Y. Y. (2020). Unsupervised assessment of cognition in the Healthy Brain Project: Implications for web-based registries of individuals at risk for Alzheimer's disease. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *6*(1), e12043. https://doi.org/10.1002/trc2.12043

Pertl, D. (2015). Ursachen, Risiko- und Schutzfaktoren. In *Österreichischer Demenzbericht 2014*. Bundesministerium für Gesundheit.

Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine*, *256*(3), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

Petersen, R. C., Dickson, D. W., Parisi, J. E., Johnson, K. A., Ivnik, R. J., Smith, G. E., Tangalos, E. G., & Kokmen, E. (2000). Neuropathological substrate of mild cognitive impairment. *Neurobiology of Aging*, *21*, 198. https://doi.org/10.1016/S0197-4580(00)82232-0

Petersen, R. C., Lopez, O., Armstrong, M. J., Getchius, T. S. D., Ganguli, M., Gloss, D., Gronseth, G. S., Marson, D., Pringsheim, T., Day, G. S., Sager, M., Stevens, J., & Rae-Grant, A. (2018). Practice guideline update summary: Mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American Academy of Neurology. *Neurology*, *90*(3), 126–135. https://doi.org/10.1212/WNL.0000000000004826

Petersen, R. C., & Morris, J. C. (2005). Mild cognitive impairment as a clinical entity and treatment target. *Archives of Neurology*, *62*(7), 1160–1163, https://doi.org/10.1001/archneur.62.7.1160

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology*, *56*(3), 303–308. https://doi.org/10.1001/archneur.56.3.303

Pons, A., LaMonica, H. M., Mowszowski, L., Köhler, S., Deckers, K., & Naismith, S. L. (2018). Utility of the LIBRA index in relation to cognitive functioning in a clinical health seeking sample. *Journal of Alzheimer's Disease*, *62*(1), 373–384. https://doi.org/10.3233/JAD-170731

Pratap, A., Neto, E. C., Snyder, P., Stepnowsky, C., Elhadad, N., Grant, D., Mohebbi, M. H., Mooney, S., Suver, C., Wilbanks, J., Mangravite, L., Heagerty, P. J., Areán, P., & Omberg, L. (2020). Indicators of retention in remote digital health studies: A cross-study evaluation of 100,000 participants. *NPJ Digital Medicine*, *3*(1). https://doi.org/10.1038/s41746-020-0224-8

Psimistri. (2022). *Psimistri—Global Psychometric Test and Intervention Systems.* https://psimistri.com/index.php/de/

Pusswald, G., Moser, D., Gleiß, A., Janzek-Hawlat, S., Auff, E., Dal-Bianco, P., & Lehrner, J. (2013). Prevalence of mild cognitive impairment subtypes in patients attending a memory outpatient clinic—Comparison of two modes of mild cognitive impairment classification. Results of the Vienna Conversion to Dementia Study. *Alzheimer's & Dementia*, *9*(4), 366–376. https://doi.org/10.1016/j.jalz.2011.12.009

Rabin, L. A., Smart, C. M., Crane, P. K., Amariglio, R. E., Berman, L. M., Boada, M., Buckley, R. F., Chételat, G., Dubois, B., Ellis, K. A., Gifford, K. A., Jefferson, A. L., Jessen, F., Katz, M. J., Lipton, R. B., Luck, T., Maruff, P., Mielke, M. M., Molinuevo, J. L., … Sikkes, S. A. M. (2015). Subjective cognitive decline in older adults: An overview of self-report measures used across 19 International Research Studies. *Journal of Alzheimer's Disease, 48*(1), 63-86. https://doi.org/10.3233/JAD-150154

Rabin, L. A., Wang, C., Mogle, J. A., Lipton, R. B., Derby, C. A., & Katz, M. J. (2020). An approach to classifying subjective cognitive decline in community-dwelling elders. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *12*(1), e12103. https://doi.org/10.1002/dad2.12103

Reid, L. M., & Maclullich, A. M. J. (2006). Subjective memory complaints and cognitive impairment in older people. *Dementia and Geriatric Cognitive Disorders*, *22*(5–6), 471–485. https://doi.org/10.1159/000096295

Rentz, D. M., Dekhtyar, M., Sherman, J., Burnham, S., Blacker, D., Aghjayan, S. L., Papp, K. V., Amariglio, R. E., Schembri, A., Chenhall, T., Maruff, P., Aisen, P., Hyman, B. T., & Sperling, R. A. (2016). The feasibility of at-home iPad cognitive testing for use in clinical trials. *The Journal of Prevention of Alzheimer's Disease*, *3*(1), 8–12. https://doi.org/10.14283/jpad.2015.78

Robbins, T. W., James, M., Owen, A. M., Sahakian, B. J., McInnes, L., & Rabbitt, P. (1994). Cambridge Neuropsychological Test Automated Battery (CANTAB): A factor analytic study of a large sample of normal elderly volunteers. *Dementia (Basel, Switzerland)*, *5*(5), 266–281. https://doi.org/10.1159/000106735

Rosas, A. G., Stögmann, E., & Lehrner, J. (2022). Neuropsychological prediction of dementia using the neuropsychological test battery Vienna – A retrospective study. *Brain Disorders*, *5,* 100028. https://doi.org/10.1016/j.dscb.2021.100028

Rostdamzadeh, A., & Jessen, F. (2020). Früherkennung der Alzheimer-Krankheit und Demenzprädiktion bei Patienten mit leichter kognitiver Störung. *Der Nervenarzt*, *91*, 832–842. https://doi.org/10.1007/s00115-020-00907-y

Sabbagh, M. N., Boada, M., Borson, S., Chilukuri, M., Doraiswamy, P. M., Dubois, B., Ingram, J., Iwata, A., Porsteinsson, A. P., Possin, K. L., Rabinovici, G. D., Vellas, B., Chao, S., Vergallo, A., & Hampel, H. (2020). Rationale for early diagnosis of mild cognitive impairment (MCI) supported by emerging digital technologies. *The Journal of Prevention of Alzheimer's Disease*, *7*(3), 158–164. https://doi.org/10.14283/jpad.2020.19

Savica, R., & Petersen, R. C. (2011). Prevention of dementia. *The Psychiatric Clinics of North America*, *34*(1), 127. https://doi.org/10.1016/j.psc.2010.11.006

Schmidt, K., & Metzler, M. (1992). *Wortschatztest (WST)*. https://www.testzentrale.de/shop/wortschatztest-75222.html

Singleton, E. H., Fieldhouse, J. L. P., van 't Hooft, J. J., Scarioni, M., van Engelen, M.-P. E., Sikkes, S. A. M., de Boer, C., Bocancea, D. I., van den Berg, E., Scheltens, P., van der Flier, W. M., Papma, J. M., Pijnenburg, Y. A. L., & Ossenkoppele, R. (2023). Social cognition deficits and biometric signatures in the behavioural variant of Alzheimer's disease. *Brain*, *146*(5), 2163–2174. https://doi.org/10.1093/brain/awac382

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., Craft, S., Fagan, A. M., Iwatsubo, T., Jack, C. R., Kaye, J., Montine, T. J., Park, D. C., Reiman, E. M., Rowe, C. C., Siemers, E., Stern, Y., Yaffe, K., Carrillo, M. C., Thies, B., Morrison-Bogorad, M., … Phelps, C. H. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *7*(3), 280–292. https://doi.org/10.1016/j.jalz.2011.03.003

Sternin, A., Burns, A., & Owen, A. M. (2019). Thirty-five years of computerized cognitive assessment of aging—Where are we now? *Diagnostics*, *9*(3), Article 3. https://doi.org/10.3390/diagnostics9030114

Stricker, N. H., Stricker, J. L., Karstens, A. J., Geske, J. R., Fields, J. A., Hassenstab, J., Schwarz, C. G., Tosakulwong, N., Wiste, H. J., Jack Jr., C. R., Kantarci, K., & Mielke, M. M. (2022). A novel computer adaptive word list memory test optimized for remote assessment: Psychometric properties and associations with neurodegenerative biomarkers in older women without dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *14*(1), e12299. https://doi.org/10.1002/dad2.12299

Thal, D. R., & Braak, H. (2005). Postmortale Diagnosestellung bei Morbus Alzheimer: Stadiengliederungen der kennzeichnenden Hirnveränderungen. *Der Pathologe*, *26*(3), 201–213. https://doi.org/10.1007/s00292-004-0695-4

Tokaj, A., & Lehrner, J. (2023). Drawing a line? -Visuo-constructive function as discriminator between healthy individuals, subjective cognitive decline, mild cognitive impairment and Alzheimer's disease and predictor of disease progress compared to a multimodal approach. *Neuropsychiatrie*. https://doi.org/10.1007/s40211-022-00455-8

Valencia, N., & Lehrner, J. (2018). Screening for dementia with the Vienna Visuo-Constructional Test 3.0 screening (VVT 3.0 screening). *Neuropsychiatrie*, *32*(4), 196–203. https://doi.org/10.1007/s40211-018-0279-9

Vogel, A., Gade, A., Stokholm, J., & Waldemar, G. (2005). Semantic memory impairment in the earliest phases of Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *19*(2–3), 75–81. https://www.proquest.com/docview/232497608/abstract/50098AD187F34556PQ/1

Wagner, M., Wolf, S., Reischies, F. M., Daerr, M., Wolfsgruber, S., Jessen, F., Popp, J., Maier, W., Hüll, M., Frölich, L., Hampel, H., Perneczky, R., Peters, O., Jahn, H., Luckhaus, C., Gertz, H.-J., Schröder, J., Pantel, J., Lewczuk, P., … Wiltfang, J. (2012). Biomarker validation of a cued recall memory deficit in prodromal Alzheimer disease. *Neurology*, *78*(6), 379–386. https://doi.org/10.1212/WNL.0b013e318245f447

Werheid, K. & Clare, L. (2007). Are faces special in Alzheimer's disease? Cognitive conceptualisation, neural correlates, and diagnostic relevance of impaired memory for faces and names, *Cortex, 43(7)*, 898–906. https://doi.org/10.1016/s0010-9452(08)70689-0

Wetterberg, H., Najar, J., Rydberg Sterner, T., Kern, S., & Skoog, I. (2024). The effect of diagnostic criteria on dementia prevalence – A population-based study from Gothenburg, Sweden. *The American Journal of Geriatric Psychiatry*, *32*(2), 230–243. https://doi.org/10.1016/j.jagp.2023.08.018

Wild, K., Howieson, D., Webbe, F., Seelye, A., & Kaye, J. (2008). Status of computerized cognitive testing in aging: A systematic review. *Alzheimer's & Dementia*, *4*(6), 428–437. https://doi.org/10.1016/j.jalz.2008.07.003

World Health Organization (Ed.). (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines* (10th ed.). World Health Organization.

World Health Organization. (2017). *Global action plan on the public health response to dementia 2017–2025*. World Health Organization. https://iris.who.int/handle/10665/259615

World Health Organization. (2019). *International Classification of Diseases, Eleventh Revision.* https://icd.who.int/browse11. Licensed under Creative Commons Attribution-NoDerivatives 3.0 IGO licence (CC BY-ND 3.0 IGO)

World Health Organization. (2021). *Dementia*. https://www.who.int/news-room/fact-sheets/detail/dementia

World Health Organization. (2022). *Dementia*. https://www.who.int/news-room/fact-sheets/detail/dementia

List of Figures

List of Tables

List of Abbreviation

AD = Alzheimer Disease

ADI = Alzheimer's Disease International

AKT = Alters-Konzentrations-Test

APP = Amyloid Precursor Protein

AVT = Auditory Vocabulary Test

BDI-II = Beck Depression Inventory II

BNT = Boston Naming Test

CANTAB = Cambridge Neuropsychological Test Automated Battery

C.I. = Cerebral Insufficiency Test

CITY = City Identification Test

DSM-IV = The Diagnostic and Statistical Manual of Mental Disorders 4th edition

DSM-V = The Diagnostic and Statistical Manual of Mental Disorders 5th edition

DST = Digit Symbol Test

EU = European Union

FACE = Face Identification Test

FPT = Faces-Pairs-Test

GDP = Global Gross Domestic Product

GCP = Good Clinical Practice

ICD-10 = The International Statistical Classification of Diseases and Related Health Problems, 10th edition

ICD-11 = The International Statistical Classification of Diseases and Related Health Problems, 11th edition

INCP = The International Neurocognitive Test Profile

INT = Image Naming Test

MCI = Mild Cognitive Impairment

MMSE = Mini-Mental State Examination

NIA-AA = National Institute on Aging and Alzheimer's Association

NTBV = Neuropsychological Test Battery Vienna

PWT = Phonematic Verbal Fluency Test

SCD = Subjective Cognitive Decline

SWT = Semantic Verbal Fluency Test

TLT-s = Traffic Light Test- short

TMT-A = Trail Making Test

TMT-B = Trail Making Test

VSRT = Verbal Selective Reminding Test

VVT = Verbal Vocabulary Test

VVT 3. 0 = Vienna Visuo-Constructive Test 3.0

WHO = World Health Organization

WST = Wortschatztest

Appendix A

Tables A

Table A 1

Overview of all Collected Variables

Variables	Type of Items
Demo-graphic Data Age Gender total years of schooling, highest school education 	 In Years, metric Male/female/ divers, nominal Years, metric Ranking, ordinal

Test	**Rounds Duration overall Time limit (per round/overall)**	**Type & Amount of Items**	**Scoring in %/raw**	**Domain/Construct**
Cognitive Screening MMSE	 One Round 10min none	 11 major items: temporal orientation (5 points), spatial orientation (5 points), immediate memory (3 points), attention/concentration (5 points), delayed recall (3 points), naming (2 points), verbal repetition (1 points), verbal comprehension (3 points), writing (1 points), reading a sentence (1 points), and constructional praxis (1 points); Metric	 Total sum of all correct answers, max.= 30points	 Screening: orientation, immediate memory, attention/concentration, delayed recall, language

VVT 3.0	2 tasks 2-3min None	3 items (geometrical figures), copying task, metric.	Max. = 98 points (32p clock, 26p pentagon, 40p cube)	Screening visuo-constructive ability
WST	1 round 10min None	40 items (words), each item consists of 1 target word and 5 distractors, recognition task; metric	Total sum of correct answers	Estimation of premorbid intelligence level
BDI II	1 5-10min none	21 items, 4-point Likert-scale	Sum of all values of the individual statement → compared to a cut-off	Screening: depressive symptoms
Test	**Rounds Duration overall Time limit (per round/overall)**	**Type & Amount of Items**	**Scoring in %/raw**	**Domain/Construct**
NTBV				
Language				
BNT	1 round 1min 10sec per item	15 items (objects), Visual**;** metric	Sum of all correct answers	Language
SWT	3 rounds 3min 60sec per round	Free words calling; Metric	Sum of correctly named words	Language
Learning and Memory				
VSRT (immediate recall) VSRT (total recall)	5 rounds - None 1 round – none	15 items (grocery objects), Visual, Metric -	Sum of all correct answers Immediate Recall = Sum of all correct answers in the 1st run	Memory

VSRT (delayed recall)	1 round 20min after 1 run none		Learning Performance = summing up the number of correct answers from the first to the fifth trial. Delayed Recall = total of correct answers in the task Recognition = recognized items – false positive items	
Complex Attention				
Digit Symbol Subtest	1 round - 90sec.	Symbols, digits, visual, metric		Attention
AKT	1 2min 120	20 items (semicircles), Metric	(35 + Sum of correctly striked out semicircles – wrongly striked out semicircles)/time needed to finish the task	Attention
C.I.	1. 1 round 2. – 3. 60sec	44 items (squares, stars, flowers), Metric	Total time needed in sec.	Executive functioning
TMT B	1 – 300sec	Connecting letters and numbers in an ascending order (1A- 13), Metric	Total time needed in sec. → TMT B – TMT A	Attention
Executive Functioning				
TMT A	1 round – 180sec.	Connecting numbers in an ascending order (1-25), Metric	Time needed = total score	Executive functioning
Five Point Test	1 3min 3min	Patterns, dots, Metric	Sum of all correct patterns	Executive functioning,
Planning Maze Test	1 round – 120	Labyrinth, Metric	(16-errors)/time	Executive functioning
Stroop Test: Color- word time color	1 round – 60sec	Words, Metric	Total time in sec.	Executive functioning

Stroop Test: color-word time-word	1 round None 2min		(36-errors)/time needed	
C.I. symbol	1 round – 60sec.	44 items (symbols), Metric	Time needed (sec) + for every mistake 1 second	Executive functioning
PWT	3 rounds 3min 60 sec for each round	Metric	Sum of all correct answers – categorizations and word(stem) repetition)	Executive functioning
Test	**Rounds Duration overall Time limit (per round/overall)**	**Type & Amount of items**	**Scoring in %/raw**	**Domain/Construct**
INCP				
Language				
Auditory Vocabulary Test (AVT)	1 3-5 min none	Auditory items - words 100 items, Metric	Sum of all correct answers	Language skills, comprehension, auditory perception
Verbal Vocabulary Test (VVT)	1 3 min none	Visual, words, 100 items, Metric	Sum of all correct answers	Language, visual perception, comprehension
Image Naming Test (INT)	1 3 min none	100 items (Pictures, descriptions), Metric	Sum of all correct answers	Language, semantic memory
Memory & Learning				
Story Comprehension Test (SCT)	1 5-10 min none	30 items (stories), Metric	Sum of all correct answers	Memory
Face Identification Test (FACE)	1 3-5 min none	16 items (Photos, names), Metric	Sum of all correct answers	Semantic memory
City Identification Test (CITY)	1 3-5 min none	16 items (Names/words), Metric	Sum of all correct answers	Semantic memory
Faces- Pairs Test (FPT) Faces - Pairs Test – Forced Choice Two Alternative Immediate Recognition (FPT-FCTAIR)	3 3'/1'/1'	20 items(photos) per round (60 in total), metric	Round 1 = sum of all correctly matched pairs over two rounds (with a maximum of 40)	Episodic memory for faces, learning skills

Faces Pairs Test - Forced Choice Six Alternative Delayed Recognition (FPT-FCSADR) Faces - Pairs Test Recognition (FPT-REC)			Round 2 = sum of all correctly formed pairs, with a maximum of 20 pairs) Round 3 = sum of all hits minus the sum of all false alarms (with a maximum score of 40) Total score (FPT) = FPT-FCTAIR round 1 + FPT-FCTAIR round 2 + FPT-FCSADR + FPT-REC (in percent or a total raw score, ranging between 0 – 100, with higher scores indicating better memory performance)	
Complex attention				
Digit Symbol-Test (DST)	3 3 min 60 sec per round	Symbols, digits, Metric	Sum of all correct answers over three rounds	Information processing, attention
Executive function				
Traffic-Light-Test (TLT-s)	2 (1 normal/1 reversed) 3 min 2 seconds per item	Traffic light, words, 30 items per round, Metric	Sum of all correct “go” and correct “stop” in both rounds	Executive functions, inhibition, cognitive flexibility
Social Cognition				
Emotion Face Test (EFT-s)	2 (1 normal/1 reversed) 3 min 2 seconds per item	Happy, sad smiley, 20 items per round, Metric	The sum of all correct answers, of all wrong answers and of all "not responding" are calculated	Emotion recognition, social cognition
Cognitive Screening				

VVT 3.0 Delayed recall task	1 3-5min None	3 items (geometrical figures), free drawing from the memory, Metric	extensive scoring system, evaluation of overall size, alignment, and length of individual lines, together with other criteria, (0 point = performance is insufficient/1 point = performance is sufficient), max. = 98 points (32 clock, 26 points pentagons, 40 points cube)	Visuo-constructive ability

Table A 2

Sample Sizes, Mean Values, Standard Deviations, Range of Demographic Variables and Screening Variables of the Subsamples HC and MCI

HC-Group	Variable	*N*	*M*	*SD*	*Min*	*Max*
	Age	41	53.93	16.22	22	81
	Sex	41	.71	.461	0	1
	Formal Years of Education	41	16.32	3.34	9	26
	MMSE	41	29	1.1	27	30
	VVT 3.0	41	9.93	.26	9	10
	BDI II	37	5.32	5.43	0	20
	WST	41	112	9.73	92	129
MCI-Group	Variable	*N*	*M*	*SD*	*Min*	*Max*
	Age	12	72.17	15.25	30	87
	Sex	12	.67	.49	0	1
	Formal Years of Education	12	9.67	2.10	8	15
	MMSE	12	24.33	2.71	21	29
	VVT 3.0	12	9.0	1.76	4	10
	BDI II	5	12.00	4.8	5	17
	WST	5	97	9.28	85	110

Note. N = sample size, M = mean, SD = standard deviation, Min = Minimum, Max = Maximum.

[a]MMSE: possible range 0-30, VVT 3.0 Screening: possible range 0-10, BDI-II: possible range 0-63, WST: possible range 0-139.

Table A 3

Comparison of Actual Subsample Sizes per NTBV- Subtest according to Screening Criteria to Expected Subsample Size when following Classification of MCI by Peterson (2004)

	HC-Group		MCI-Group	
Subtests	Actual N	$N_{Deviation}$	Actual N	$N_{Deviation}$
AKT Total	41	1	12	2
AKT Time	41	0	10	0
AKT Total/Time	41	2	12	0
AKT Right	41	3	12	0
AKT Mistake	41	0	12	0
HAWIE-R Digit-Symbol-Test	41	2	9	0
C.I. Symbols	41	1	12	0
TMT-A	41	1	12	0
TMT-B	41	0	6	0
SWT Animals	41	4	12	0
SWT Grocery	41	4	6	1
SWT Tools	41	2	6	1
SWT Total	41	4	6	0
PWT – b	41	3	6	0
PWT - f	41	1	12	0
PWT – l	41	3	6	1
PWT total	41	2	6	1
BNT	41	2	12	1
VSRT 1	41	3	11	0
VSRT 2	41	4	11	0
VSRT 3	41	5	11	0
VSRT 4	41	6	11	0
VSRT 5	41	5	11	1
VSRT Total	41	3	11	0
VSRT delayed recall	41	4	11	0
VSRT recognition	41	4	11	1
5 Point Test Right	41	2	6	1
5 Point Test Perseveration	41	0	6	0
Color word test (NAI-I)	41	1	6	1

Color word test (NAI-III)	41	1	6	0
Color word test (NAI-III) Mistakes	41	0	6	0
Color-word-test (NAI-III) Total	41	0	6	2
Color-word-test (NAI-III) Total/time	41	3	6	0
Color-word-test (NAI-III – NAI-I) Interference	41	1	6	0
Planning Maze Test (NAI) Time	41	0	12	0
Planning Maze Test (NAI) Mistakes	41	0	12	0
Planning Maze Test (NAI) Total	41	3	12	0
Planning Maze Test (NAI) Total/time	41	2	12	0
TMTB – TMTA Difference	41	0	6	0
C.I. Interference Time	41	1	12	0
C.I. Mistakes	41	0	12	0
C.I. Interference Total	41	3	12	1
C.I. Interference Total/time	41	1	12	0

Note. Actual N_{HC} refers to the actual sample size per subtest that is classified as healthy in the present sample based on the used screening criteria.

Actual N_{MCI} refers to the actual sample size per subtest that is classified as MCI in the present sample based on the used screening criteria.

$N_{Deviation}$ refers to the number of individuals deviating from the actual N per subtest because they scored 1.50 standard deviations below age, sex, and education norms in a single domain or in multiple domains on the NTBV and would be classified as MCI when following the criteria by Peterson (2004).

Table A 4

Comparison of actual Subsample Size to Expected Subsample Size when following Classification of MCI by Peterson (2004) per INCP-Subtest

	HC-Group		MCI-Group	
Subtests	Actual *N*	$N_{Deviation}$	Actual *N*	$N_{Deviation}$
CITY	36	4	5	0
FACE	30	3	3	0
AVT	39	2	5	1
DST	38	5	3	0
EFTs	31	1	2	0
FPT	39	3	9	0
VVT	35	3	10	0
TLTs	39	5	3	0
INT	39	1	10	0

Note. Actual N_{HC} refers to the actual sample size included that is classified as healthy in the present sample based on the screening criteria used for each subtest of the INCP.

Actual N_{MCI} refers to the actual sample size per subtest that is classified as MCI in the present sample based on the used screening criteria.

$N_{Deviation}$ refers to the number of individuals deviating from the actual *N* because they scored 1.50 standard deviations below age, sex, and education norms in a single domain or in multiple domains on the NTBV and would be classified as MCI when following the criteria by Peterson (2004).

Table A 5

Sample Sizes, Mean Values, Standard Deviations and Minimum and Maximum Scores of the NTBV Subtests of the Overall Sample and the Subsamples HC and MCI

Total Sample	Subtests	*N*	*M*	*SD*	*Min*	*Max*
	AKT Total	51	53.76	2.56	40	55
	AKT Time	51	30.80	12.43	16	70
	AKT Total/Time	51	1.96	0.62	0.69	3.24
	AKT Right	51	19.54	1.63	10	23
	AKT Mistake	50	0.68	1.81	0	12
	HAWIE-R Digit-Symbol Test	50	51.01	17.40	1	85
	C.I. Symbols	53	23.30	9.762	9	60
	TMT-A	53	33.72	18.18	14	122
	TMT-B	47	76.26	54.60	12	300
	SWT Animals	53	23.15	7.22	5	33
	SWT Grocery	47	26.49	6.26	7	42
	SWT Tools	47	11.36	5.44	2	30
	SWT Total	47	61.62	16.30	9	93
	PWT – b	47	13.70	4.61	4	23
	PWT - f	53	11.11	5.10	1	28
	PWT – l	47	12.47	3.64	6	21
	PWT total	47	36.49	11.89	6	62
	BNT	53	13.89	1.88	5	15
	VSRT 1	52	9.60	2.65	3	14
	VSRT 2	52	11.40	3.19	4	15
	VSRT 3	52	12.19	3.33	3	15
	VSRT 4	52	13.33	4.80	1	38
	VSRT 5	52	13.00	3.38	2	15
	VSRT Total	52	58.35	16.18	14	74
	VSRT delayed	52	12.12	3.95	0	15
	VSRT recognition	52	13.88	2.83	0	15
	5 Point Test Right	47	36.15	10.19	4	58
	5 Point Test Repetition	47	3.09	9.47	0	65
	Color word Test (NAI-I)	47	22.11	5.78	2	38
	Color word Test (NAI-III)	47	37.91	13.41	22	90
	Color word Test (NAI-III) Mistakes	47	0.42	.90	0	4

Color-word-test (NAI-III) Total	47	35.68	1.11	32	40
Color-word-test (NAI-III) Total/time	47	3.16	14.73	.39	102
Color-word-test (NAI-III – NAI-I) Interference	47	15.32	10.59	1	53
Planning Maze Test (NAI) Time	53	35.94	22.13	13	120
Planning Maze Test (NAI) Mistakes	53	0.91	1.70	0	8
Planning Maze Test (NAI) Total	53	18.91	27.51	8	215
Planning Maze Test (NAI) Total/time	53	1.11	4.05	0.08	30
TMTB – TMTA Difference	47	45.19	45.07	0	269
C.I. Interference Time	53	21.17	7.38	10	44
C.I. Mistakes	53	0.49	.98	0	4
C.I. Interference Total	53	33.06	3.51	9	34
C.I. Interference Total/time	53	2.01	1.76	0.75	14

HC

Subtests	*N*	*M*	*SD*	*Min*	*Max*
AKT Total	41	54.07	2.49	40	55
AKT Time	41	27.24	8.35	16	60
AKT Total/Time	41	2.13	.54	.92	3.24
AKT Right	41	19.66	.69	17	20
AKT Mistake	41	.59	1.92	0	12
HAWIE-R Digit-Symbol Test	41	54.93	11.59	36	85
C.I. Symbols	41	19.73	6.46	9	40
TMT-A	41	28.63	8.46	14	49
TMT-B	41	69.95	42.83	27	300
SWT Animals	41	25.41	5.03	13	32
SWT Grocery	41	26.68	5.23	15	42
SWT Tools	41	11.85	5.51	2	30
SWT Total	41	63.95	12.73	33	93
PWT – b	41	13.98	4.73	4	23
PWT - f	41	12.02	4.13	4	21

PWT – 1	41	12.61	3.77	6	21
PWT total	41	37.88	11.38	16	62
BNT	41	14.32	1.17	9	15
VSRT 1	41	10.44	1.89	7	14
VSRT 2	41	12.68	1.71	8	15
VSRT 3	41	13.37	2.10	6	15
VSRT 4	41	14.15	1.24	11	15
VSRT 5	41	14.41	1.30	10	15
VSRT Total	41	65.05	6.99	46	74
VSRT delayed	41	13.68	1.68	8	15
VSRT recognition	41	14.78	.49	14	15
5 Point Test Right	41	37.49	9.23	22	58
5 Point Test Repetition	41	1.71	2.14	0	9
Color word Test (NAI-I)	41	21.76	3.92	13	33
Color word Test (NAI-III)	41	35.00	6.81	22	51
Color word Test (NAI-III) Mistakes	41	.32	.76	0	4
Color-word-test (NAI-III) Total	41	35.68	.76	32	36
Color-word-test (NAI-III) Total/time	41	1.048	.20	.67	1.64
Color-word-test (NAI-III – NAI-I) Interference	41	13.44	6.15	4	29
Planning Maze Test (NAI) Time	41	29.76	13.51	13	83
Planning Maze Test (NAI) Mistakes	41	.46	.95	0	3
Planning Maze Test (NAI) Total	41	15.61	.86	13	16
Planning Maze Test (NAI) Total/time	41	.61	.23	.19	1.23
TMTB – TMTA Difference	41	41.32	40.87	10	269
C.I. Interference Time	41	17.88	3.53	10.00	26.00
C.I. Mistakes	41	.20	.64	0	3
C.I. Interference Total	41	33.80	.64	31	34
C.I. Interference Total/time	41	1.97	.42	1.23	3.40

MCI	Subtests	*N*	*M*	*SD*	*Min*	*Max*
	AKT Total	10	52.5	2.59	47	55
	AKT Time	10	45.40	15.94	27	70
	AKT Total/Time	10	1.28	0.46	0.69	1.96
	AKT Right	10	18.60	3.41	10	23
	AKT Mistake	9	1.11	1.17	0	3
	HAWIE-R Digit-Symbol-Test	9	33.17	27.31	1	85
	C.I. Symbols	12	35.50	9.43	25	60
	TMT-A	12	51.08	29.61	15	122
	TMT-B	6	119.33	100.98	12	300
	SWT Animals	12	15.42	8.37	5	33
	SWT Grocery	6	25.17	11.79	7	41
	SWT Tools	6	8.00	3.63	2	13
	SWT Total	6	45.67	28.25	9	82
	PWT – b	6	11.83	3.43	8	16
	PWT - f	12	8.00	6.88	1	28
	PWT – l	6	11.50	2.59	7	14
	PWT total	6	27.00	11.87	6	38
	BNT	12	12.42	2.94	5	15
	VSRT 1	11	6.45	2.81	3	12
	VSRT 2	11	6.64	2.91	4	11
	VSRT 3	11	7.82	3.49	3	13
	VSRT 4	11	10.27	9.92	1	38
	VSRT 5	11	7.73	3.58	2	13
	VSRT Total	11	33.36	16.39	14	61
	VSRT delayed	11	6.27	4.56	0	14
	VSRT recognition	11	10.55	4.94	0	15
	5 Point Test (Right)	6	27.00	12.55	4	40
	5 Point Test (Repetition)	6	12.50	25.81	0	65
	Color word Test (NAI-I)	6	24.50	13.28	2	38

Color word Test (NAI-III)	6	57.83	27.20	31	90
Color word Test (NAI-III) Mistakes	6	1.09	1.47	0	4
Color-word-test (NAI-III) Total	6	35.67	2.58	32	40
Color-word-test (NAI-III) Total/time	6	17.56	41.37	.39	102.00
Color-word-test (NAI-III – NAI-I) Interference	6	28.17	22.41	1	53
Planning Maze Test (NAI) Time	12	57.08	31.95	15	120
Planning Maze Test (NAI) Mistakes	12	2.43	2.67	0	8
Planning Maze Test (NAI) Total	12	30.17	58.26	8	215
Planning Maze Test (NAI) Total/time	12	2.77	8.58	.08	30.00
TMTB – TMTA Difference	6	71.67	66.03	0	178
C.I. Interference Time	12	32.42	5.82	26.00	44.00
C.I. Mistakes	12	1.51	1.24	0	4
C.I. Interference Total	12	30.50	6.88	9	34
C.I. Interference Total/time	12	2.12	3.75	.75	14.00

Note. *N* = Sample Size, *M* = Mean, *SD* = Standard Deviation, *Min* = Minimum, *Max* = Maximum.

Table A 6

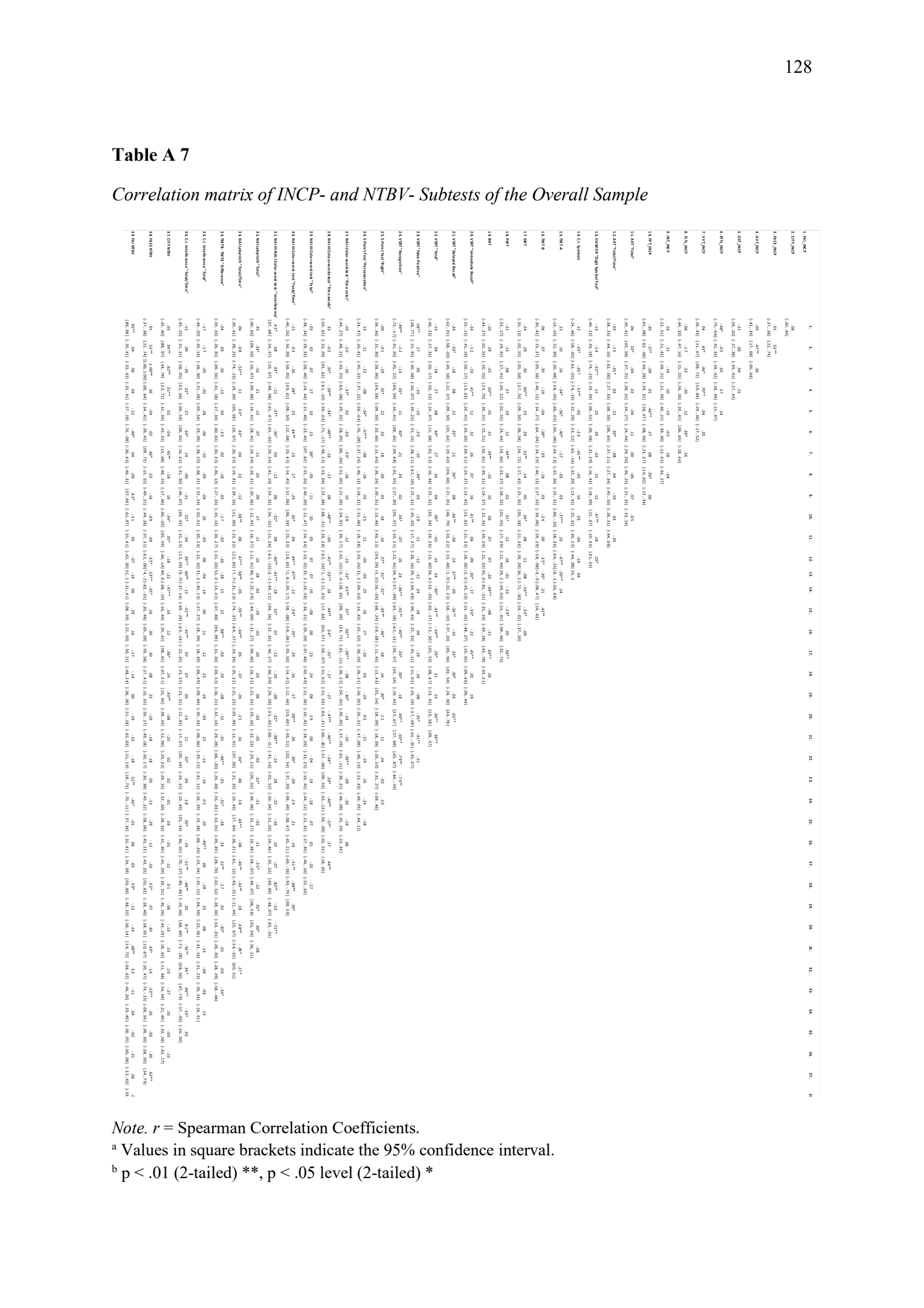
Correlation Table of INCP Subtests and Descriptive-/Screening- Variables of the Overall Sample

	Age	Sex	Occupational status	Formal Years of Education	Diagnosis	MMSE	VVT 3.0 Screening	VVT 3.0 Screening Delayed	VVT 3.0	VVT 3.0 Delayed Recall	Standardab. Check	BDI_II	WST	GDS	CITY_INCP	FACE_INCP	AVT_INCP	DST_INCP	EFTs_INCP	VVT_INCP	TLTs_INCP	INT_INCP	FPT_INCP	FAI_INCP
Age	–																							
Sex	-0,012	–																						
Berufsstatus	-,815**	-0,092	–																					
Bildungsjahre	-,524**	0,022	,404*	–																				
Diagnose1	-,519**	0,037	,394*	,654**	–																			
MMSE	-,483**	0,187	,387*	,616**	,633**	–																		
VVT3.0_Screening	-0,098	0,063	0,168	,373*	,422*	,316*	–																	
VVT3.0_Screening_Delayed	-0,071	-0,067	,385*	,554**	,556**	,397**	,446**	–																
VVT3.0	-,532**	-0,050	0,249	,573**	,558**	,588**	,490**	,437*	–															
VVT3.0_Delayed_Recall	-,340*	0,100	,409*	,536**	,545**	,469**	0,204	-0,036	,551**	–														
Uhr	-,505**	-0,065	0,007	,534**	,491**	,410**	-0,035	,429**	,399*	-0,247	–													
BDI_II	-0,073	-0,178	0,225	,380*	,405*	,333*	0,252	,760**	0,102	,382*	,445*	–												
WST	0,259	0,210	-0,065	-0,202	-0,349	-0,078	-0,166	-0,029	-0,265	-0,179	0,055	-0,183	–											
GDS	0,239	,329*	-0,149	0,094	0,206	0,123	-0,050	0,189	0,140	0,140	-0,148	,711**	-,804**	–										
CITY_INCP	,330*	-0,198	-0,292	0,028	0,246	-0,232	0,046	-0,010	0,252	0,073	-0,136	-0,278	,473*	-,673**	–									
FACE_INCP	,466**	-0,165	0,047	,424*	,509**	0,232	0,161	0,186	0,216	0,132	0,263	-,406*	,416*	-,584**	,535*	–								
AVT_INCP	-0,088	-0,165	0,229	0,209	-0,032	0,208	0,161	0,263	0,216	0,132	0,263	-,380*	,607**	-,595**	,588**	0,322	–							
DST_INCP	-0,156	0,115	0,229	0,209	-0,032	0,208	0,161	0,263	0,216	0,132	0,062	-0,145	0,243	-0,041	0,077	0,148	0,159	–						
EFTs_INCP	-0,298	-0,094	0,228	,382*	0,087	,376*	,391*	,477**	,391*	0,120	-0,144	0,153	,425*	-0,325	0,053	0,119	0,194	,419*	–					
VVT_INCP	0,028	0,104	0,025	,328*	,486**	,340*	,413**	,301*	,306*	,328*	0,273	-0,243	,569**	-,584**	,468**	-0,022	-0,142	0,256	0,104	–				
TLTs_INCP	-0,201	0,042	0,213	,326*	-0,151	0,143	,421**	0,187	0,022	0,026	0,044	0,100	0,051	0,126	-0,316	-0,275	-0,115	-0,085	0,021	0,192	-,334*	-0,068	–	
INT_INCP	-0,037	0,153	0,145	0,201	0,073	0,164	-0,035	0,097	0,033	0,087	0,087	-0,190	0,433	-0,008	-0,169	-0,166	0,184	,380*	,381*	-0,336	-0,060	-0,040	-0,001	–
FPT_INCP	-,507**	0,275	,396*	,745**	,525**	,562**	,443**	,598**	,498**	,560**	0,260	0,062	,319*	-0,008	-0,169	-0,166	0,184	,380*	,381*	,334*	,410**	0,067	–	
FAI_INCP	,394*	0,004	-0,241	-0,297	-,580**	-0,209	-0,120	-0,146	-0,210	-0,095	-0,155	0,265	-0,056	0,377	-0,127	-0,197	-0,134	-0,274	-0,336	-0,060	-0,040	0,175	-0,229	–

Note: r = Spearman Correlation Coefficients;
a p < .01 (2-tailed) *; p < .05 level (2-tailed) *

Table A 7

Correlation matrix of INCP- and NTBV- Subtests of the Overall Sample



Note. r = Spearman Correlation Coefficients.

[a] Values in square brackets indicate the 95% confidence interval.

[b] p < .01 (2-tailed) **, p < .05 level (2-tailed) *

Table A 8

Rotated Component Matrix: EFA of INCP-Subtests

Subtests	1	2	3	4
FACE	.861	-.195	.178	.063
CITY	.851	-.141	-.243	.082
INT	-.827	-.121	-.087	.022
AVT	.529	.498	-.329	.422
FPT	-.329	.770	.438	-.184
TLTs	-.009	-.028	.903	.043
DST	.164	.620	.639	-.026
EFTs	-.141	.133	.087	.889
VVT	.490	-.041	-.078	.719

Note. Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Table A 9

Descriptive Statistics and Test statistics of the Z-Values of the NTBV Subtests. Comparison of Healthy Controls and Mild Cognitive Impairment.

Subtests/z-Values	Group	*N*	*z*	M_{Rank}	R_1; R_2	*U*	*Z*	*p*	*r*
AKT	MCI	12	-0.80 (1.58)	15.20	152.00	97.00	-2.921	.002	.401
	HC	41	0.40 (1.12)	28.63	1174.00				
AKT Total/Time	MCI	10	-0.90 (1.48)	10.10	101.00	46.00	-3.776	< .001	.523
	HC	41	0.30 (1.11)	29.88	1225.00				
TMT B	MCI	6	-1.40 (2.43)	31.42	464.00	78.50	-1.420	.156	
	HC	41	0.30 (1.18)	22.91	967.00				
HAWIE-R Digit-Symbol-Test	MCI	9	-0.90 (1.42)	13.78	124.00	79.00	-2.668	.008	.377
	HC	41	0.40 (0.88)	28.07	1151.00				
TMTB - TMTA	MCI	6	-1.45 (1.98)	30.92	185.50	81.50	-1.323	.186	
	HC	41	0.10 (1.10)	22.99	942.50				
C.I. Symbols	MCI	12	-1.85 (1.18)	44.96	539.50	30.50	-4.588	< .001	.630
	HC	41	0.00 (1.20)	21.74	891.50				
Ø Domain: Attention	MCI	12	-0.85 (1.34)	17.33	208.00	130.00	-2.474	.014	.339
	HC	41	0.20 (0.74)	29.83	1223.00				
PWT - b	MCI	6	0.06 0(.99)	17.50	105.00	84.00	-1.248	.212	
	HC	41	0.20 (1.25)	24.95	1023.00				
PWT - f	MCI	12	-0.85 (0.65)	13.88	166.50	88.50	-3.359	<.001	.461
	HC	41	-0.10 (1.29)	30.84	1264.50				
PWT - l	MCI	6	0.30 (0.69)	20.42	122.50	101.50	-.689	.491	
	HC	41	0.10 (1.18)	24.52	1005.50				

PWT Total	MCI	6	-0.30 (0.88)	13.58	81.50	60.50	-1.995	.046	291
	HC	41	0.10 (1.37)	25.52	1046.50				
Ø Domain: Executive function-phonematic verbal fluency	MCI	5	-0.40 (0.46)	20.10	100.50	85.50	-.602	.548	
	HC	41	0.20 (1.14)	23.91	980.50				
NAI-I (Time colour)	MCI	6	-0.60 (1.59)	30.17	181.00	86.00	-1.185	.236	
	HC	41	0.00 (1.15)	23.10	947.00				
NAI-III (Time words)	MCI	6	-1.45 (1.95)	34.67	208.00	59.00	-2.045	.041	.298
	HC	41	0.40 (0.97)	22.44	920.00				
NAI-III (Total/Time)	MCI	6	-1.25 (1.87)	16.58	99.50	78.50	-1.422	.155	
	HC	41	0.50 (1.03)	25.09	1028.00				
NAI-III – NAI-I (Colour-word-Test, Interference)	MCI	6	1.45 (2.07)	30.83	185.00	82.00	-1.311	.190	
	HC	41	-0.70 (0.94)	23.00	943.00				
C.I. Interference (Time)	MCI	12	-0.75 (1.54)	47.46	569.00	.50	-5.229	< .001	.718
	HC	41	0.50 (0.86)	21.01	861.50				
C.I. Interference (Total/Time)	MCI	12	-0.95 (1.19)	10.08	121.00	43.00	-4.323	< .001	.593
	HC	41	0.40 (0.96)	31.95	1310.00				
Ø Domain: Executive function-Interference	MCI	12	-0.60 (1.21)	17.92	215.00	137.00	-2.323	.020	.318
	HC	41	0.00 (0.61)	29.66	1216.00				
SWT Animals	MCI	12	-1.68 (1.64)	13.13	157.50	79.50	-3.562	.156	
	HC	41	0.60 (1.25)	31.06	1273.50				
SWT Grocery	MCI	6	-0.31 (2.40)	22.92	137.50	116.50	-.209	< .001	.030
	HC	41	0.80 (1.2)	24.16	990.50				
SWT Tools	MCI	6	-0.17 (1.27)	15.75	94.50	73.00	-1.601	.109	
	HC	41	0.30 (1.55)	25.21	1034.00				
SWT Total	MCI	6	-0.40 (2.73)	15.75	94.50	73.50	-1.580	.114	
	HC	41	0.80 (1.32)	25.21	1033.50				
BNT	MCI	12	-1.05 (1.37)	18.08	217.00	139.00	-2.500	.012	.343

	HC	41	-0.70 (.98)	29.61	1214.00				
ØDomain: Language	MCI	12	-1.25 (1.55)	16.08	193.00	115.00	-2.793	.005	.383
	HC	41	0.40 (1.05)	30.20	1238.00				
VSRT Immediate Recall	MCI	11	-0.70 (1.46)	11.18	123.00	57.00	-3.809	< .001	.528
	HC	41	0.80 (1.04)	30.61	1255.00				
VSRT Total Recall	MCI	11	-0.90 (2.16)	7.64	84.00	18.00	-4.656	<.001	.645
	HC	41	1.40 (0.84)	31.56	1294.00				
VSRT Delayed Recall	MCI	11	-0.90 (3.25)	8.86	97.50	31.50	-4.463	< .001	.618
	HC	41	1.00 (0.76)	31.23	1280.50				
VSRT Recognition	MCI	11	-1.30 (1.82)	9.68	106.50	40.50	-4.889	< .001	.677
	HC	41	-0.80 (.89)	31.01	1271.50				
ØDomain: Memory	MCI	11	-0.40 (2.09)	15.77	173.50	107.50	-2.652	.008	.367
	HC	41	0.50 (0.63)	29.38	1204.50				
5- Point Test (Right)	MCI	6	0.15 (2.39)	14.42	86.50	65.50	-1.863	.066	
	HC	41	0.30 (1.01)	25.40	1041.50				
5-Point-Test (Repetitions)	MCI	6	-0.35 (0.70)	29.50	177.00	90.00	-1.086	.277	
	HC	41	-0.40 (0.76)	23.20	951.00				
NAI Labyrinth	MCI	12	-1.10 (1.82)	17.79	213.50	135.50	-2.789	.005	.383
	HC	41	0.40 (0.93)	29.70	1217.50				
NAI - Labyrinth (Total/Time)	MCI	12	-1.60 (1.54)	14.63	175.50	97.50	-3.158	.002	.434
	HC	41	0.30 (1.04)	30.62	1255.50				
TMT A	MCI	12	-0.50 (1.63)	38.67	464.00	106.00	-2.979	.003	.409
	HC	41	0.60 (0.88)	23.59	967.00				
Ø Domain: Executive function-planning & nonverbal fluency	MCI	12	-0.95 (1.26)	18.29	219.50	141.50	-2.223	.026	.304
	HC	41	0.20 (0.61)	29.55	1211.50				

Note. Mann–Whitney–U-Test.

[a] Grouping Variable: Diagnose (based on Screening); *z* = Median of z-Value with (*SD*), M_{Rank} = Mean Rank; R_1 = Sum of Rank Group 1; R_2 = Sum of Rank Group 2, *U* = Mann-Whitney-U, *Z* = Z-statistic, *p* = Asymp. Sig. (2-tailed), *Ø* = in average

Table A 10

Descriptive Statistics and Test Statistics of the z-Values of the INCP Subtests. Comparison of Healthy Controls and Mild Cognitive Impairment, Group Allocation based on z-Values (NTBV).

Subtests	Group	*N*	*Mdn*	*U*	*Z*	*p*	*r*
DST	MCI	34	177.50	62.50	-1.99	< .05	.311
	HC	7	180.00				
TLT-s	MCI	35	37.00	105.00	-.60	> .05	
	HC	7	38.00				
CITY	MCI	34	14.00	54.00	-2.23	< .05	.348
	HC	7	15.00				
FACE	MCI	25	15.00	44.00	-2.55	< .05	.443
	HC	8	16.00				
FPT	MCI	41	83.00	74.00	-2.03	< .05	.293
	HC	7	91.00				
AVT	MCI	37	92.00	60.00	-2.24	< .05	.338
	HC	7	99.00				
VVT	MCI	37	85.00	72.50	-2.25	< .05	.335
	HC	8	89.00				
INT	MCI	42	99.50	120.00	-.80	> .05	
	HC	7	100.00				
EFT-s	MCI	26	36.00	43.50	-2.11	< .05	.367
	HC	7	38.00				

Note. Mann – Whitney – U - Test.

[a] Grouping Variable: Diagnosis (based on NTBV scores)

[b] *Mdn* = *Median*, *U* = Mann-Whitney-U, *Z* = Z-statistic, *p* = Asymp. Sig. (2-tailed)

Appendix B

Figures B

Figure B 1

EFA Screeplot INCP-Subtests