How do individuals with different cognitive performance levels differ in their self-reported limitations after SARS-CoV-2 infection, their well-being and their resting state neural activity?

What is EEG?

Why do I look at resting state and not eeg data while doing cognitive task?

Cognition is essential for everyday functioning. Cognition refers to a range of mental processes related to the acquisition, storage, manipulation, and retrieval of information. It underpins many daily activities, in health and disease, across the age span. Cognition can be separated into multiple distinct functions, dependent on particular brain circuits and neuromodulators. The ability to test, measure and monitor cognitive performance across the lifespan opens up the chance for patients to be identified earlier, access treatments faster, and stay healthy for longer, improving quality of life and reducing costs.
Cognition is defined as “the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses.” The modern word “cognition” actually has its roots back to Latin, the word “cognoscere” which is to “get to know”. Cognitive functioning is therefore critical for day-to-day life, governing our thoughts and actions. Need cognition to help us understand information about the word around us and interact safely with our environment, as the sensory information we receive is vast and complicated: cognition is needed to distill all this information down to its essentials. Cognitive assessment refers to the objective measurement of distinct cognitive abilities, such as working memory, inhibition, cognitive flexibility, psychomotor speed and sustained attention. Cognition can be measured using a variety of methods, each varying in their level of objectivity and sensitivity.

Relation between objective measure of cognition and subjective perception of cognitive functioning reported by …

Patients with … experience a broad range of cognitive impairments affecting domains such as…

(Bland et al. 2024):

Commented [JH1]: Electroencephalography (EEG) is a non-invasive, functional brain-activity2 recording modality with high temporal resolution and relatively low cost.
Luca Pion-Tonachini et al.,2019

Commented [JH2]: https://cambridgecognition.com/what-is-cognition/

Given that subjective and objective cognitive function may be driven by different underlying mechanisms, it is important to explore factors which contribute to objective impairments and subjective feelings of “brain fog” in order to design targeted interventions for individuals living with Post-COVID.

Taken together, understanding the relationship between subjective and objective cognitive impairment and the driving factors underlying these is crucial for the development of tailored rehabilitation programs aimed at improving cognitive function and facilitating recovery in Post-COVID patients. In line with previous literature in other patient groups, we hypothesised that subjective cognitive dysfunction would be associated with increased fatigue and stress whereas objective cognitive function would be most dominantly linked to clinical features of COVID-19.

The present study found that objective and subjective measures of cognitive function were not significantly correlated, and that self-reported experience of cognitive deficits were no longer significant when accounting for heightened stress and fatigue.

Theoretical background

Post-Covid Syndrome

(O’Mahoney et al., 2023)

(Chen et al., 2022)

Cognitive deficits related to Post-Covid Syndrome

Subjective and objective cognitive functions are two distinct measures of cognition. (McWhirter, Ritchie, Stone & Carson, 2020): Cognitive symptoms are common: according to this Review they are present in around a third of the populationm with no clear relation to age.

Subjective

Theoretical background

1. Cognitive Performance as a Key Variable: Introduce cognitive performance as a central factor in your research. Explain its relevance in understanding individual differences in behavior, health, and psychological outcome. Cognitive performance as a key variable that influences

Commented [JH3]: https://www.thelancet.com/pdfs/journals/eclinm/PIIS2589-5370(22)00491-6.pdf

Commented [JH4]: https://academic.oup.com/jid/article/226/9/1593/6569364?login=true

Commented [JH5]: Cognitive Performance as a Key Variable

Cognitive performance refers to an individual’s ability to carry out mental processes such as attention, memory, executive function, and problem-solving. It plays a crucial role in determining how individuals navigate daily life, affecting everything from decision-making to managing complex tasks. Cognitive performance is not a uniform trait but varies significantly between individuals, influenced by factors like age, education, lifestyle, and health conditions. Understanding these variations is essential in psychological and health research because they are often linked to broader outcomes like well-being, self-reported limitations, and brain activity.

Higher levels of cognitive performance have been associated with better problem-solving skills, greater emotional regulation, and improved mental and physical health outcomes. Conversely, individuals with lower cognitive performance may experience difficulties in these areas, leading to reduced well-being and higher levels of self-reported limitations. Therefore, cognitive performance is a key variable when investigating individual differences in how people experience and report limitations in their daily lives, their well-being, and their underlying neural activity.

Cluster Analysis for Grouping Based on Cognition

Given the variability in cognitive performance across individuals, it is important to group people in ways that reflect meaningful differences in their cognitive abilities. One effective method for achieving this is **cluster analysis**, a statistical technique used to classify individuals into groups (or clusters) based on shared characteristics. In this study, cluster analysis will be used to group participants solely based on their cognitive performance scores. This method allows us to objectively identify subgroups of individuals with similar cognitive profiles, which can then be compared in terms of other variables like self-reported limitations, well-being, and resting-state neural activity.

By forming cognitive clusters, this approach provides a way to explore whether individuals with different levels of cognitive performance report their limitations differently, experience well-being in distinct ways, or exhibit variations in neural activity. The use of cluster analysis in this context is particularly valuable because it captures the natural variation in cognitive functioning, rather than relying on predefined categories. This method thus enhances our understanding of how cognitive differences translate into real-world outcomes.

many aspects of life, such as well-being, self-reported limitations, and brain function. Cognitive abilities vary across individuals.

Introduction (catchy start)

Imagine waking up one day feeling disoriented, unable to concentrate, or struggling to remember simple tasks from the day before. For many individuals recovering from COVID-19, this mental cloudiness, often described as "brain fog," is a persistent reminder of their illness. These cognitive challenges, along with fatigue and other lingering symptoms, affect their daily lives long after the infection has passed.

Cognition is essential for navigating the complexities of everyday life. It encompasses a range of mental processes such as memory, attention, and problem-solving, which are critical for acquiring and processing information. Cognitive functions allow us to understand and interact with the world around us by making sense of vast sensory input and distilling it into actionable knowledge. Whether it's remembering a conversation, concentrating on a task, or making decisions, cognition is the foundation that supports these processes.

This thesis focuses on cognitive impairment in individuals experiencing Post COVID-19 Syndrome (PCS), a condition characterized by the persistence of symptoms such as fatigue and cognitive dysfunction months after recovering from the acute infection. While subjective complaints of cognitive difficulties are common among PCS patients, objective measures often reveal discrepancies between what individuals report and their actual cognitive performance. The goal of this thesis is to explore these differences through a cluster analysis approach, aiming to better understand how objective cognitive measures relate to individuals' self-reported symptoms, their overall well-being, and their neural brain activity.

First, the dopaminergic system, closely related to the basal ganglia, is introduced, as it is centrally involved in reward-based learning and movement generation. Then the concepts of reward, prediction, and reinforcement learning are explained and linked to the dopaminergic system. Next, Parkinson's disease is described as it is a disease affecting the dopaminergic system within the basal ganglia. As this thesis is part of an ongoing research project at the Danish Research Center for Magnetic Resonance (DRCMR), first the overall research questions will be mentioned and then the idea of this thesis will be introduced in details, which is to conceptually show how Bayesian cognitive modeling can be used to

model reward-based learning and decision-making in participants diagnosed with parkinson's disease and in healthy control participants.

Theoretical Background

As of September 2024, over 760 million confirmed cases of coronavirus disease 2019 (COVID-19) have been documented by the World Health Organization (WHO) globally, leading to approximately 6.9 million deaths. The actual numbers are likely to be much higher due to underreporting. COVID-19 is an infectious disease caused by the SARS-CoV-2 virus (WHO, 2021). While most patients fully recover, some experience persistent symptoms such as fatigue, shortness of breath, cognitive dysfunction, and other symptoms that generally have an impact on everyday functioning (WHO, 2021). These remaining effects, referred to as Post-COVID-19 Condition or Syndrome (PCS), usually occur three months after the initial infection with the SARS-CoV-2 virus and last for at least two months with no other explanation. Approximately 10-20% of people infected with SARS-CoV-2 meet the criteria for PCS (WHO, 2021).

Cognitive Impairment in PCS

Cognitive impairment is one of the most frequent symptoms of PCS (Davids et al., 2021; WHO) and is therefore of high interest. These impairments are characterized by confusion, memory difficulties, disorientation, and trouble concentrating, which are referred to as experiencing "brain fog" by affected individuals (Bland et al., 2024; Kwan et al., 2024). Around 22% of individuals diagnosed with PCS experience COVID-related cognitive impairment, according to a meta-analysis by Ceban et al. (2022). This finding is based on data from 43 studies, 31 of which used subjective assessments and 12 that employed objective measures. Notably, studies using objective assessments of cognitive function reported significantly greater proportions of individuals with impairment (36%) compared to those relying on subjective modes of ascertainment, which identified 18% as cognitively impaired.

This shows, that subjective and objective measures of cognitive function represent two distinct approaches to assessing cognition. Subjective assessments rely on self-reported experiences and perceptions (Stewart, 2012), while objective assessments use standardized tests and tasks to evaluate cognitive performance in various functional areas. Several studies have illustrated these discrepancies between subjective and objective measures further. In fact, most studies have reported higher rates of cognitive impairment through subjective cognitive complaints than through objective test results (Schild, Scharfenberg, Kirchner et al., 2023). For instance, in a study by Schild, Goereci, Scharfenberg et al. (2023) among 52 patients who

self-reported cognitive impairment after SARS-CoV-2 infection, objective cognitive screening tests confirmed impairment in only 25%, while extensive neurological assessment indicated impairments in 60% of these patients. Moreover, Schild, Scharfberg, Kirchner, et al. (2023) reported that 88% of patients reported persistent self-reported cognitive impairment, with approximately a 40% discrepancy between the subjective reports and objective test results at both follow-up visits, underscoring the discrepancies between patients' self-reports and objective neuropsychological test results. Bland et al. (2024) observed that there was no significant relation between objective and subjective measures of cognitive function, implying that self-reports of "brain fog" may not be reflected by objectively measured cognitive dysfunction.

Subjective cognitive deficits in everyday situations are predicted by elevated anxiety and fatigue levels more than by objective cognitive performance (Zamarian et al., 2024). This lack of alignment highlights the complexity of cognitive impairment and raises questions about which additional factors may influence individuals' perceptions of cognitive difficulties. Recent research has addressed these questions by examining how psychological symptoms influence subjective cognitive and objective cognitive impairment.

Zamarian et al. (2024) discovered that subjective cognitive deficits in everyday situations can be better explained by elevated anxiety and fatigue levels than by objective cognitive performance. In addition to anxiety (Almeria, Cejudo, Sotoca, Deus & Krupinski, 2020; Brück et al., 2019; Costas-Carrera et al., 2022; Hill et al., 2016; Zamarian et al., 2024) and fatigue (Bland et al., 2024; Delgado-Alonso et al., 2023; Zamarian et al., 2024), sleep disturbances (Zamarian et al., 2024) and depressive symptoms (Almeria et al., 2020; Brück et al., 2019; Costas-Carrera et al., 2022; Hill et al., 2016; Zamarian et al., 2024) have been found to be associated with subjective but not objective cognitive impairment (Henneghan, Lewis, Gill & Kesler, 2022). Objective cognitive function, on the other hand, was found to be related to perceived stress (Bland et al., 2024).

These findings highlight the intricate and often discordant relationship between subjective and objective cognitive performance, as well as their complex interactions with psychological factors such as anxiety, fatigue, sleep disturbances, and depressive symptoms. This complexity raises important questions about how these elements interact, particularly in the aftermath of SARS-CoV-2 infection. Understanding these dynamics is crucial for developing effective and personalized rehabilitation programs that aim to improve individuals' perceived cognitive function and assist in their recovery.

Cognition

Cognition is defined as “the mental action or process of acquiring knowledge and understanding through thought, experience, and the sense” (Cambridge Cognition, 2015). Cognition is essential for everyday functioning and refers to a range of mental processes such as the acquisition, storage, manipulation, and retrieval of information (Cambridge Cognition, 2015).

Cognitive impairment. Mild cognitive impairment will be explained in this section.

EEG findings in MCI. Reduced delta power during resting state EEG has been identified in patients with MCI (Liddell et al., 2007). Furthermore, in the study, individuals with MCI demonstrated a significant positive correlation between delta power and immediate memory recall. Liddell et al. (2007) proposed that these findings suggest that delta power may be linked to memory decline in MCI, indicating that it could serve as a sensitive indicator of prodromal or early cognitive decline. However, other studies have shown increased delta power in MCI patients compared to healthy controls, particularly in frontal and centroparietall regions (Adler, Bramesfeld & Jajcevic, 1999; Moretti, Zanetti, Binetti & Frisoni, 2012). A decrease in beta power has been found in individuals with mild AD (Hogan, Swanwick, Kaiser, Rowan & Lawlor, 2003).

Fatigue

Fatigue, alongside cognitive impairment, is the most commonly reported symptom of PCS (WHO, 2021). As mentioned above, subjective perceptions of cognitive performance can be influenced by fatigue. Therefore, a closer examination of fatigue will follow to differentiate between the concepts of fatigue and subjective cognitive impairment.

EEG findings in Fatigue.

EEG findings in PCS patients. Electroencephalography (EEG) is a non-invasive, objective method for assessing neuronal activity and has proven to be a valuable tool in identifying neurophysiological dysfunctions in individuals with cognitive impairment (Koenig, Smailovic & Jelic, 2020; Kubota, Gajera & Kuroda, 2021). Because of this, EEG studies have become increasingly relevant for investigating individuals with COVID-19 and PCS, as they reveal changes in brain neural activity that correlate with fatigue and cognitive deficits in these patients (Antony & Haneef, 2020; Appelt et al., 2022; Cecchetti et al., 2022; Furlanis et al., 2023; Kopańska et al., 2022; Kubota, Gajera & Kuroda, 2021; Pasini et al.,

2020; Pastor, Vega-Zelaya & Abad, 2020; Roberto, Espiritu, Fernandez & Gutierrez, 2020; Wojcik et al., 2023).

Furlanis et al. (2023) found that two-thirds of the 20 participants presenting brain fog were characterized by unexpected abnormal EEG patterns. Ortelli et al. (2023) found that lower performance on cognitive tasks, particularly those assessing executive function, was associated with changes in brain activity in PCS patients.

There are different types of analyses used to evaluate EEG patterns of PCS patients, ranging from common power spectrum and event-related potentials (Cecchetti et al., 2022; Furlanis et al., 2023; Kopańska et al., 2022) to more sophisticated approaches, such as intrinsic mode functions and avalanche analysis (Appelt et al., 2022; Wojcik et al., 2023). However, in this thesis, a power analysis will be conducted, specifically examining delta and beta frequency.

Delta Power in PCS patients. Delta frequency (0.5-3 Hz) is typically absent during the waking state of healthy adults and is associated with deep sleep (Schandry, 2016). Ortelli et al. (2023) reported significant differences in the delta frequency band between PCS and healthy controls, with PCS patients displaying diminished activity compared to healthy controls. Lower delta power was associated with worse cognitive functioning. However, findings regarding delta power in PCS patients are not consistent. For instance, Kopańska et al. (2022) found a decrease in delta in the left hemisphere, similar to Ortelli et al. (2023), but also observed an increase in delta activity in the right hemisphere. In another study of 20 PCS patients, a delta-slowing pattern was revealed in nine of them (Furlani et al., 2022). Furthermore, the relative delta power values in this cohort were higher compared to those reported in the literature for healthy individuals. Similarly, Pastor et al. (2020) demonstrated a significant encephalopathic pattern in PCS patients characterized, among others, by an increase in generalized delta activity.

Beta Power in PCS patients. Beta frequency (14-30 Hz) is typically present when individuals are awake and mentally or physically active, or under psychological stress (Schandry et al., 2016). While Ortelli et al. (2023) found no significant differences in beta frequency bands, Kopańska et al. (2020) reported increased beta2 activity in both hemispheres and elevated beta1 activity in the left hemisphere in PCS patients.

EEG findings conclusion. Those EEG findings discussed above are mainly based on subjective perceived cognitive impairment rather than objective measures of impairment. To illustrate this, the findings of Ortelli et al. (2023) provide relevant insights. The PCS group had a significantly lower MoCA score and higher fatigue score (assessed with the self-

Commented [JH6]: Why beta and delta power?

Commented [JH7]: Other literature

Commented [JH8]: Different reference

evaluation scale measuring perceived fatigue (FSS)), than the control group. However, the global cognitive score assessed with the MoCA was still considered normal, implying that, overall, PCS patients did not have clinically significant cognitive impairment.

Notably, there was no differentiation possible between EEG patterns associated with cognitive impairment and those related to fatigue. This raises an interesting opportunity to examine beta and delta power during resting state in two groups defined solely by objective cognitive measures, allowing for a clearer understanding of the relationship between EEG patterns and cognitive functioning in patients with PCS. This approach would allow potential abnormalities in EEG to be more directly linked to objective cognitive impairment rather than subjective cognitive impairment, which might be influenced by psychological factors, such as fatigue.

Aim of study

The study aims to explore the differences among groups that differ significantly in their objective cognitive performance levels following SARS-CoV-2 infection. This investigation is crucial given the widespread cognitive impairments reported in individuals with PCS and their profound impact on everyday functioning and quality of life. Due to the inconsistent findings in EEG patterns in beta and delta power in patients with PCS, but also in patients with MCI, there is a need for further investigation of this aspect. Specifically, the research will address the following research question: How do individuals with different cognitive performance levels differ in their self-reported limitations after SARS-CoV-2 infection, their well-being, and their resting state neural activity?

By examining the correlations between objective cognitive assessments and self-reported cognitive impairments, as well as the influence of psychological factors, this study aims to provide insights into the complex relationship between cognitive functioning and psychological health in individuals with, and without PCS.

Hypotheses

Following SARS-CoV-2 infection, two distinct groups of individuals will be identified based on objective cognitive assessments, showing significant differences in performance levels between the groups. Suggesting, that one group performs significantly better or worse than the other group. Individuals with objectively assessed lower cognitive performance will report higher levels of self-reported cognitive limitations compared to the group that performed better. Individuals, that have self-reported cognitive impairment, but were not recognized as

Commented [JH9]: Das sieht alles schon gut aus, aber versuch mal deine Hypothesen richtig hervorzuheben, du kannst ruhig richtig Abstand lassen und sagen

Hypothese 1:

Hypothese 2:

Hypothese 3a: Gruppe B hat weniger Fatigue als Gruppe A
Hypothese 3b: Gruppe B hat weniger Depression als Gruppe A
Hypothese 3c: Gruppe B hat weniger Angst als Gruppe A

und so weiter
WARUM du diese Hypothesen aufstellst, arbeitest du direkt vorher heraus, quasi als Übergang vom theoretischen Hintergrund in deine Fragestellung hinein. Und in den Methoden erklärst du dann WIE du die jeweiligen Hypothesen testest. Also die Hypothese, dass mehr Leute sich subjektiv eingeschränkt fühlen, in der Gruppe die objektiv schlecht abgeschnitten hat, müsstest du dann vielleicht mit einem Chi-quadrat Test überprüfen zum Beispiel

low performers in objective cognitive assessment may have higher fatigue, anxiety, and depression scores than all other individuals. Concerning the delta frequency, a decreased delta power, in patients with objective cognitive lower performance compared to the better performers is expected, suggesting, that abnormal delta power is correlated to cognitive impairment. However, abnormal delta power may also be related to fatigue, suggesting that decreased delta power could be observed in patients with subjective cognitive impairment who do not exhibit lower objective cognitive performance. This would imply that their perceived cognitive limitations might be a symptom of fatigue rather than actual cognitive deficits. Concerning the beta frequency, an increase, in patients with objective lower performance compared to the better performers is expected, suggesting, that abnormal beta power is correlated to cognitive impairment. However, abnormal beta power may be (as delta power) also related to fatigue, suggesting that increased beta power could be observed in patients with subjective cognitive impairment who do not exhibit lower objective cognitive performance.

How do individuals with different cognitive performance levels differ in their self-reported limitations after SARS-CoV-2 infection, their well-being and their resting state neural activity?

Idea: Part What is cognition in methods!?
Why beta and delta and not theta and alpha?
Why is closed eye condition the one I chose?

Methods

Study Design Participants

The current data were collected as part of a larger research project (EPOC), which investigates neurophysiological parameters identified from neuropsychological paradigms using a high-resolution stationary laboratory EEG to reflect cognitive impairments and fatigue. The primary goal of the EPOC study is to find EEG parameters that can serve as neurophysiological markers for progression- and therapy-evaluation concerning cognitive functions in PCS.

Participants for the EPOC study were recruited from COVIDOM, a population-based, prospective multi-centre study to investigate Post-COVID Syndrome (PCS) within the German

Commented [JH10]: What does it stand for? Do I need to name it?

Commented [JH11]: EEG Post-Covid

Commented [JH12]: In brief, the experiment consisted of neuropsychological tests (TMT, n-back, PVT, Oddball, RTE), restingstate recording and questionnaires regarding fatigue, sleep quality, depression and anyiety.
The study was conducted at the University Medical Center Schleswig-Holstein (UKSH), Campus Kiel.

Commented [JH13]: Or COVIDOM study?

Commented [JH14]: Well I say in sentence after, that it is a study

Commented [JH15]: Do I need to mention full name or is COVIDOm enough?

National Pandemic Cohort Network (NAPKON). COVIDOM participants had been recruited through public health authorities in Kiel, Berlin, and Würzburg. Patients were assessed between November 15, 2020, and September 19, 2021, at University Medical Center Schleswig-Holstein, Campus Kiel, and University Hospital Würzburg in Germany (Bahmer et al., 2022; Horn et al., 2021; Schons et al., 2022).

The participants were included based on the following criteria: A polymerase chain reaction (PCR) confirmed SARS-CoV-2 infection at least 6 months before study visit, a primary residence in one of the three study regions, age ≥ 18 years at the time of recruitment (Berlin) or infection (Würzburg, Kiel). Participants with acute reinfection of SARS-CoV-2 at the time of the scheduled study visit were excluded (Horn et al., 2021).

In the EPOC study, a subset of individuals from Schleswig-Holstein who participated in COVIDOM was selected, constituting of those with PCS and a control group without PCS. As EPOC is still ongoing at the time of writing, the analysis was conducted based on a preliminary subset of 79 participants (mean age 48.52, range 22–78, female 48, male 31, diverse 0, years of education mean 15.27 years min 9 to 24 years) withPCS (49 participants, age mean 50.29 years min 22–78, F=32 M=17, education: mean 15.04 9–23) and withoutPCS (30 participants, age mean 45.63 years min 22–77, f=16,m=14 d 0, education mean 15.63 min 10–24). The study was conducted at the University Medical Center Schleswig-Holstein (UKSH), Campus Kiel. Participants did not receive payment/financial compensation for their participation. Transportation and parking costs were reimbursed.

Ethics statement

The study was approved by the Ethics Committee of the medical faculty of the Christian-Albrechts-University of Kiel, Germany (record identification: D 446/23). In accordance with the Declaration of Helsinki, informed written consent was obtained from all participants.

Procedure/Study Design

In brief, the experiment consisted of neuropsychological tests, assessing cognitive domains such as working memory, attention, preprocessing speed, cognitive flexibility, executive functions, and multisensory integration, EEG recordings, and questionnaires assessing

Commented [JH16]: Might be bit to specific

Commented [JH17]: COVIDOM is a population-based cohort study of polymerase chain reaction (PCR) confirmed cases of
SARS-CoV-2 infection, recruited through public health authorities in three German regions (Kiel, Berlin, W€urzburg) between November 15, 2020 and September 29, 2021

Commented [JH18]: With und without als Tabelle. Gesammte Stichprobe im Text beschreiben

Commented [JH19]: Table?

Commented [JH20]: Table here -> see Lara

Commented [JH21]: Here, or should I inser it after the inclusion criteria?

Commented [JH22]: All participants gave their written
133 informed consent and were compensated for their participation. The experiment was
134 carried out in accordance with the ethical standards of the Declaration of Helsinki and
135 was approved by the local ethical committee.

Commented [JH23]: The study was conducted monocentrically at the UKSH, Campus Kiel. A maximum duration of 3 hours was allocated for each participant's visit. The eligibility of participants was determined based on their medical history and MoCA tests, which had been previously conducted as part of the COVIDOM study. Participants were then asked to complete questionnaires on demographic data (e.g., age, education) and psychological and neurological conditions. This initial part of the examination took up to 20 minutes. If deemed eligible for the study, all participants underwent neuropsychological testing. This assessment covered cognitive domains such as working memory, attention, processing speed, cognitive flexibility, executive functions, and multisensory integration. First, three neuropsychological tests were conducted without EEG (TMT, n-back task, PVT). These tests took up to 25 minutes in total. Subsequently, EEG was recorded while two additional neuropsychological tests were performed (Oddball task, RTE). At the end, a resting-state EEG was recorded, during which the participant was not required to perform any task. Finally, three short questionnaires on fatigue, depression, anxiety, and sleep were completed. This last part of the examination took up to 2.25 hours. The data were stored in a pseudonymized manner. The study information disclosed that there was no commuting accident insurance for this study. Data collection was scheduled to be completed by the end of September 2023.

Commented [JH24]: Participants first completed demographic and psychological questionnaires, followed by neuropsychological testing to assess cognitive domains such as working memory, attention, and executive functions. EEG recordings were conducted during specific tasks, with a resting-state EEG recorded at the end. The entire session lasted up to three hours, with data stored pseudonymousl

Commented [JH25]: Need to add, that two were there to place EEG cap, and else one person. Here, or at EEG part?

fatigue, sleep quality, depression, and anxiety. All data were collected at the University Medical Center Schleswig-Holstein (UKSH), Campus Kiel.

Participants first filled out a questionnaire on demographic data (e.g., age, education) and psychological and neurological conditions, followed by neuropsychological testing to assess cognitive domains such as working memory, attention, preprocessing speed, cognitive flexibility, executive functions, and multisensory integration. The first test administered was the Trial Marking Test. Following this, the EEG cap was placed, and participants completed a series of other neuropsychological tests, starting with the redundant target effect (RTE), followed by an oddball paradigm, an n-back task, and lastly the psychomotor vigilance task (PVT). Electroencephalographic activity was recorded continuously throughout these tests. Finally, resting state was measured, 5 minutes with eyes open and 5 minutes with eyes closed. During EEG recordings, participants were seated comfortably and instructed to minimize movement, and to focus on a fixation cross displayed on the screen in front of them to reduce eye movements, while the light was turend off. After completing the resting state measurement, the EEG cap was removed, and participants filled out three questionnaires assessing fatigue (FACIT-F), sleep quality (PSQI), depression (HADS-D), and anxiety (HADS-A). The participants got then the chance to wash their scalp/hair. In all, the experiment took up to 3 hours.

Since this thesis focuses on behavioral data obtained from the TMT, n-back and PVT, as well as EEG resting state data, and data from the questionnaires, the RTE, and the oddball task will not be further explained. In addition, the MoCA score was measured in the previous COVIDOM study.

Cognitive tasks

MoCA

The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) is a widely used, validated screening tool originally designed to detect mild cognitive impairment (MCI) (Hauffe, 2024). It assesses several cognitive domains, including visuospatial skills/ability, executive function, naming, memory (short-term and delayed recall), working memory, attention and concentration, language, abstraction, and orientation (Freitas, Simões, Alves & Santana, 2013; Hobson, 2015; Kang et al., 2018; Nasreddine et al., 2005). The MoCA has a

Commented [JH26]: Not sure if nessecary, I like to give an overview but I'm repeating myself

Commented [JH27]: Already wrote that

Commented [JH28]: Where does it fit better? Before? Or here?

Commented [JH29]: Then several neuropsychological tests to assess cognitive domains such as working memory, attention, preprocessing speed, cognitive flexibility, executive functions, and multisensory integration, followed, starting with the Trial Marking Test A and B.

Commented [JH30]: Or: The first tests used were the Trial Marking Tests A and B.
Or should it be The first test used was the Trial Marking Test A and B?

Commented [JH31]: Better way to start this sentence?

Commented [JH32]: Do I need to provide more information here? Which gel etc. ; Two people...

Commented [JH33]: On what? Should I say specific on participants head?

Commented [JH34]: Kind of screen and lights off

Commented [JH35]: Lilly: 27-Zoll-Bildschirm

Commented [JH36]: The or an?

Commented [JH37]: Should I already introduce the abbreviation here?

Commented [JH38]: Maybe a bit informal. Can't think of a better word right now

Commented [JH39]: Bei dem Ruhe-EEG sollen die Versuchspersonen lediglich auf einen Fixationspunkt gucken. Nach 5
Minuten teilt ein Audiosignal der Versuchsperson mit, dass sie nun für die nächsten 5 Minuten die
Augen schließen sollen. Mit einem letzten Audiosignal wird der Versuch beendet.

Commented [JH40]: Resting state was measured next.
Or
Resting state was then measured

Commented [JH41]: Keep their gaze (is that better?)

Commented [JH42]: What kind of screen? Need to as Christian and add information

Commented [JH43]: I'm repeating myself. Is that okay in this case?

Commented [JH44]: Delete sentence?

Commented [JH45]: And will be of interest in this thesis or something like that maybe? Or does it not need further explaination?

Commented [JH46]: Das ...

Commented [JH47]: I like to start with the full name, but maybe dump since I already used it before... or should I ...

Commented [JH48]: literature

Commented [JH49]: Is that correct written? Or wrong?

total possible score of 30 points, with a score of ≥ 26 considered normal (Nasreddine et al., 2005. Cognitive performance on the MoCA is influenced by sociodemographc factors such as age and education (Kang et al., 2018; Larouche et al., 2016). To account for educational backgroundm the MoCA test manual specifies that one additional point is added for individual with ≤ 12 years of formal education, allowing for a maximum score of 30 points (Nasreddine et al., 2005). Additionally, to address variations in performance related to age, normative data for the MoCA are available for precise interpretation of scores (Larouche et al., 2016).

TMT Part A and B

Originally, developed as part of the Army Individual Test Battery (AITB) in 1944, the Trail Marking Test (TMT) was later integrated into the Halstead-Reitan Battery (Reitan & Wolfson, 1985; Tombaugh, 2004). It is now one of the most popular and widely used neuropsychological assessments, included in most test batteries (Tombaugh, 2004). Its widespread use is supported by strong evidence of its validity (Arbuthnott & Frank, 2000; Sánchez-Cubillo et al., 2009). The TMT assesses cognitive processing speed and executive functioning (Lezak, 1995; Mitrushina et al., 2005; Reitan, 1992; Sánchez-Cubillo et al., 2009; Strauss et al., 2016; Tombaugh, 2004), as well as visual search/scanning, and mental flexibility (Sánchez-Cubillo et al., 2009; Tombaugh, 2004).

The TMT consists of two parts: Part A (TMT-A), a number-connection task, and Part B (TMT-B), a number-letter alternation task. Both parts were administered (in this study) according to the guidelines provided by Strauss et al., 2006.

In TMT-A, participants were instructed to connect consecutively numbered circles from 1 to 25 on an A4 page by drawing lines between them, aiming to complete the task as quickly and accurately as possible. In TMT-B, the task becomes more complex (Gaudino, Geisler & Squires, 1995). Participants were instructed to draw lines alternating between numbered circles from 1 to 13 and lettered circles from A to L in sequential order (e.g, 1 to A, to 2, to B, etc.) on an A4 page. The aim, again, was to complete the task as quickly and accurately as possible.

Commented [JH50]: See methods1209 for more information on TMT!

Commented [JH51]: The Trail Making Test (TMT) was originally developed as part of the Army Individual Test Battery (1944) and is one of the most commonly used tests in neuropsychological practice due to its high sensitivity in diagnosing brain impairment (Armitage, 1946; Lewinsohn, 1973; Reitan, 1958; Spreen & Benton, 1965).

Commented [JH52]: Der Trail Making Test (TMT, Reitan, 1958) ist ein neuropsychologisches Testverfahren zur Erhebung von Aufmerksamkeitsstörungen und exekutiven Dysfunktionen. Der Test besteht aus zwei Abschnitten (TMT A und TMT B). Bei dem TMT A müssen Zahlen in aufsteigender Reihenfolge verbunden werden. Im TMT B müssen Zahlen und Buchstaben alternierend verbunden werden, die Buchstaben in alphabetischer Reihenfolge.

Commented [JH53]: I think it's fine to write either search or scanning. Which is better?

Commented [JH54]: Visual search, perceptual/motor speed, speed of processing, working memory, and general intelligence are among the most frequently cited constructs thought to contribute to TMT performance (Sánchez-Cubillo et al., 2009).

The administration of the TMT began with TMT-A, followed by TMT-B. For each part, participants were first given an example to familiarize themselves with the task. After completing the example, they proceeded to the actual test. If participants made a mistake, the experimenter immediately pointed it out, and the participant was required to correct it before continuing. The experimenter timed each part, with the time of completion for each part representing its direct score. In addition to the direct scores, the difference between TMT-B and TMT-A (TMT-B – TMT-A) was calculated.

Commented [JH55]: Why?

N-Back Part A and B

Commented [JH56]: Figure!

The n-back task (Kirchner, 1958) has become a widely used tool in neuroscience for assessing working memory (Jaeggi, Buschkuehl, Perrig &Meier, 2010; Pelegrina et al., 2015). N-back tasks are continuous-recognition measures, that present sequences of stimuli (Kane, Conway, Miura & Colflesh, 2007). In these tasks, participants must determine whether a given stimulus matches one that was presented "n" trials before. In this study, participants completed two blocks of the n-back task. A 1-back task and a 2-back task. Reaction time, hits, misses, and false alarms were recorded for analysis.

Commented [JH57]: The N-back task is used extensively in literature as a working memory (WM) paradigm and it is increasingly used as a measure of individual differences. However, not much is known about the psychometric properties of this task and the current study aims to shed more light on this issue. We first review the current literature on the psychometric properties of the N-back task. With three experiments using task variants with different stimuli and load levels, we then investigate the nature of the N-back task by investigating its relationship to WM, and its role as an inter-individual difference measure. Consistent with previous literature, our data suggest that the N-back task is not a useful measure of individual differences in WM, partly because of its insufficient reliability. Nevertheless, the task seems to be useful for experimental research in WM and also well predicts inter-individual differences in other higher cognitive functions, such as fluid intelligence, especially when used at higher levels of load

(Jaeggi, Buschkuehl, Perrig & Meier, 2010) extensively used in literature as a working memory paradigm. Increasingy used as a measure of individual differences.

Commented [JH58]: Looking at the data relating the N-back task to other measures of WMC, the N-back task does not seem to be a useful measure of individual differences in WMC, due to its low reliability. Nevertheless, it is a very useful tool for the experimental investigation of WM processes because it allows load to be manipulated in a very simple, straightforward way. Further, there is converging evidence that N-back performance can well predict individual differences in Gf and other higher cognitive functions, at least when used at higher levels of load

(Jacola et al., 2014) In particular, we found N-back reaction time during fMRI scanning to be most strongly correlated with measured intelligence, as opposed to clinical measures of working memory

Commented [JH59]: See: Conway et al., 2005; Kane & Engle, 200

Introduced by Kirchner (1958). Widely used tool for assessing working memory in the field of neuroscience. N-back has grown in use in neuroimaging techniques and progressively adopted (Pelegrina et al., 2015). Reliability varies greatly depending on the study (Van Leeuwen et al., 2007 to Friedman et al., 2008). More complex levels, 2-back and 3-back produce higher reliability coefficients (Jaeggi et al., 2010 stolen from (Pelegrina et al., 2015).). Reason: high variety of WM tasks and different contents that they include. N-back relies on recognition processes. Working memory task. Working memory is a system that enables one to actively maintain and regulate a limited amount of task-relevant information (Baddeley & Logie, 1999) (stolen from Pelegrina et al., 2015) Working memory is traditionally measured using processing-and-storing dual tasks that involve performing a cognitive task

Commented [JH60]: Our findings of limited validity between N-back performance and clinical measures of working memory

Commented [JH61]: Our results provide evidence for reliability of N-back accuracy during fMRI scanning; however, reliability of reaction time data is affected by order of task presentation. Data regarding construct validity are inconsistent and emphasize the need to consider clinical utility of behavioral measures in the design and interpretation of functional neuroimaging studies.

while certain information has to be maintained in memory (Pelegrina et al., 2015). N-back does not involve a processing-and-storing dual task and in which individuals are not asked to recall any information but to recognize it (Pelegrina et al., 2015). This study employed an N-back continuous recognition task to assess working memory function.

N-back continuous recognition task in which participants must decide whether a stimulus was previously presented in certain conditions. “n” trials back. Letters were shown. Each item is compared to item presented immediately before, prior letter.
In this study, participants completed two blocks of the n-back task: a 1-back task followed by a 2-back task, with a paus between blocks for additional instructions. The task was programmed using … and presented on a 27-inch computer screen.
In both conditions, participants were shown a series of 60 linguistic stimuli, consisting of 16 different consonants (**B, C, D, F, G, H, J, K, M, Q, R, S, T, V, X, Z**) presented individually in the center of the screen. Each block contained 20 target trials and 40 non-target trials. A trial began with a 250 ms fixation period (a red dot was shown on screen, for the participant to fixate), followed by a 150 ms black screen. The stimulus letter then appeared for 500 ms, succeeded by a variable inter-trial interval of 180 to 220 ms (black screen). Total trial duration ranged from 1080 to 1120 ms.
For the 1-back task, participants were instructed to press the spacebar when the current letter matched the previous one. For example, in the sequence “B, C, C, D,” participants were supposed to respond to the second “C” as it matches the previous letter. In the 2-back task, they were instructed to press the spacebar when the current letter matched the letter presented two trials prior. For instance, in the sequence “B, B, D, F, D,” participants should press the spacebar when the second “D” occurred, as it matches the letter presented two trials before. The response window was limited to the 500 ms stimulus presentation period. Reaction time, hits, misses, and false alarms were recorded. In total, the experiment took around 5 minutes. During the experiment, the light was turned off.

Red circle as fixation point. Two blocks, first 1-back task, sencond a 2-back task (condition) (two positions prior). Between those two blocks is a pause and instructur explaines the new task again, before starting the second block. One trial consist of a fixation period of 250 ms (red dot is displayed), followed by a black screen shown for 150 ms. The stimulus in form of a

Commented [JH62]: Look at numbers again here

Commented [JH63]: Is it clear what I mean with that or should I specify it a bit more?

single letter is presented for 500 ms on screen. Inter-trial interval (1800-2200 ms) Black screen. Reaction time window 0.5 seconds (stimulus presentation). In total 2.3 to 2.7 seconds per trial. Press, when detecting a target, a target is. 1-back task: Participants press spacebar when current letter matches the previous one. 2-back task, participants press spacebar when current letter matches the letter presented two trials ago/before. Reaction time, missing and fals hits measured. Response accuracy and response time was measured. In each condition 60 stimuli presented. 20 targets in both conditions. In total the experiment took around 5 minutes. One can easily manipulate the level of difficulty depending on greater or lesser memory loads as well as the timing of stimulus presentation. N-back task programmed by …. In … (program name) Was used in this study. Task applied consisted of two levels: 1-back and 2-back. Items were letters, that is, linguistic material. The following consonants were used as stimuli in all blocks M, C, K, S, R, V, H, X, T, Q, J, B, Z, G, D, F . Presented one by one in center of the screen. 20 target trials and 40 non-target trials, right? Each stimulus appeared on screen for 500 ms, followed by a screen that remained black for another 180 ms to 220 ms. Response time, number of hits, correct rejection, misses, false alarms, and non-responses were recorded. Age-related changes that might be related to different cognitive processes involved in updating information in working memory (Pelegrina et al., 2015).

PVT

The Psychomotor Vigilance Task (PVT) is a widely used reaction time test originally developed in 1985 to measure sustained attention (Drummond et al., 2005). Sustained attention and cognitive functions, particularly in contexts involving fatigue and sleep deprivation.

Reaction time task

(Drummond et al., 2005) Originally developed in 1985 as a measure of sustained attention. Studies have demonstrated its sensitivity to sleepiness in clinical and experimental settings. (Molina, Sanabria, Jung & Correa, 2019)

The monotonous and unpredictable target presentation in the PVT makes subjects highly prone to lapses of attention. Moreover, the PVT has minimal learning effects, minimizing the variability due to participants different abilities and experience (Basner and Dinges, 2011) (stolen from Molina et al., 2019)

(Jakobsen, Sorensen, Rask , Jensen & Kondrup, 2011) Reaction time was related to cognitive function in both healthy subjects and patients. it is valid.

Commented [JH64]: If I calculated right shortest time: 3.75 minutes, longest 5.8 without instructions

Commented [JH65]: Give example

Commented [JH66]: This sentence is not true in my case

Commented [JH67]: Der Psychomotor Vigilance Test (PVT), die Oddball Aufgabe, der Redundant Target Effect (RTE) und die n-back Aufgabe sind Reaktionsaufgaben, die an einem Computer im EEG-Labor durchgeführt werden. Bei dem PVT müssen die Versuchspersonen so schnell wie möglich einen Knopf drücken, sobald ein Timer aus roten Zahlen erscheint. Sobald die Reaktion stattgefunden hat, stoppt der Timer 14 und zeigt somit die Reaktionszeit der Versuchsperson an. Die Oddball Aufgabe besteht aus einem Zielreiz, einem Distraktorreiz und einem Standardreiz. Die Aufgabe der Versuchsperson ist es, bei Erscheinen des Zielreizes so schnell wie möglich einen Knopf zu drücken. Beim RTE kann ein visueller oder auditiver Reiz erscheinen oder auch beides gleichzeitig. Sobald einer der Reize erscheint (unabhängig von der Bedingung), soll so schnell wie möglich ein Knopf gedrückt werden. Die n-back Aufgabe besteht aus zwei Teilen. Auf dem Bildschirm erscheinen Buchstaben. Im ersten Teil sollen die Versuchspersonen per Knopfdruck reagieren, wenn ein Buchstabe erscheint, der mit dem vorherigen Buchstaben übereinstimmt. Im zweiten Teil sollen die Versuchspersonen per Knopfdruck reagieren, wenn ein Buchstabe erscheint, der mit dem vorletzten Buchstaben übereinstimmt.

Commented [JH68]: figure

Commented [JH69]: Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Beh Res Meth Instr Comp 1985;17:652-5.

Commented [JH70]: Dorrian J, Rogers NL, Dinges, DF. Psychomotor vigilance performance: a neurocognitive assay sensitive to sleep loss. In: Kushida C, ed. Sleep Deprivation: Clinical Issues, Pharmacology and Sleep Loss Effects. New York: Marcel Dekker, Inc; 2005:39-70.

Commented [JH71]: is a straightforward and reliable tool for measuring fatigue in humans.

Commented [JH72]: Might be interesting, when comparring with fatigue score

Commented [JH73]: Research on the relationship between EEG signals and PVT performance showed that an increment in delta power on frontal and occipital …

Commented [JH74]: It is interesting to note that frontal and parietal areas have been often related to sustained attention and, despite the scarce extant evidence with this …

This study employed a 5-minute version of the PVT, which has been established as a valid alternative to the traditional 10-minute PVT-192 (for assessing fatigue) (Lamond et al., 2008). The dynamic stimulus appeared as a red number, counting up in milliseconds, representing the participant's reaction time. Participants were instructed to respond immediately, when the stimulus/red number occured, by pressing the spacebar. Between trials, a white fixation cross was displayed on a black screen for a variable interval ranging from 2 to 10 seconds. Participants were required to maintain their gaze on this fixation cross. Each trial concluded when a response was made. Following each response, the participant's reaction time was displayed on the screen for 50 ms as feedback before the next trial began.
After receiving instructions, participants underwent a training block of 8 trials to familiarize themselves with the task. Following the training, participants proceeded to the main experiment, which consisted of 50 stimulus presentations. The light was turned off during the experiment.

Commented [JH75]: After receiving instructions, participants underwent a training block of 8 trials to familiarize themselves with the task. Following the training, participants proceeded to the main experiment, which consisted of 50 stimulus presentations.

or should that be here, before describing the experiment in details?

Commented [JH76]: The PVT was run on an Intel Core 2 Duo PC and a 17" CRT screen
with a 60 Hz refresh rate, using E-Prime software (Schneider et al.,
2001). The target stimulus was a black circle with a red edge (diameter:
9.15 ° of visual angle at a viewing distance of 50 cm).

Commented [JH77]: (Basner & Dinges, 2011)

Resting state

Do I even need to write something here?

Maybe to investigate delta and beta frequence eyes closed condition and why?

Commented [JH78]: Bei dem Ruhe-EEG sollen die Versuchspersonen lediglich auf einen Fixationspunkt gucken. Nach 5
Minuten teilt ein Audiosignal der Versuchsperson mit, dass sie nun für die nächsten 5 Minuten die
Augen schließen sollen. Mit einem letzten Audiosignal wird der Versuch beendet.

Questionnaires

PSQI

The Pittsburgh Sleep Quality Index (PSQI), developed by Buysse, Reynolds, Monk, Berman, and Kupfer in 1988 is a self-rated questionnaire that assesses sleep quality and disturbance over the past month. The PSQI is the most commonly used sleep health assessment tool in both clinical and research settings. Its reliability and validity have been consistently demonstrated in multiple studies (e.g. Carpenter & Andrykowski, 1998; Manzar et al., 2018; Mollayeva et al., 2016). The questionnaire consists of 24 items in total, 19 of which are self-repoRTEd by the patient and 5 of which require input from a room or bed partner. Only the 19 self-repoRTEd items are used for the quantitative evaluation of sleep quality, as perceived by the patient (Buysse et al., 1989; Manzar et al., 2018). The response formats across the items vary, including the recording of usual bed and wake times, number of hours slept, minutes taken to fall asleep, as well as forced-choice Likert-type responses (Buysse et al., 1989). The items are categorized into seven components, which are sleep quality, sleep

latency, sleep duration, habitual sleep efficiency, sleep disturbance, use of sleeping medications, and daytime dysfunction, for each component given a score. Together, these component scores generate a global sleep quality score ranging from 0 to 21, with scores >5 indicating poor sleep quality (Buysse et al., 1989; Hinz et al., 2017).

HADS

The self-assessment Hospital Anxiety and Depression Scale (HADS) was originally developed by Zigmond & Snaith in 1983 to identify the presence of anxiety and depression states among patients in nonpsychiatric hospital clinics. HADS is an extensively used, reliable, and valid instrument to measure anxiety and depression, not only in psychiatric, and clinical patients (Herrmann, 1997) but in general populations (Bjelland, Dahl, Haug & Neckelmann, 2002; Herrero et al., 2003; Spinhoven et al., 1997). The questionnaire assesses anxiety and depression symptoms during the past week, excluding symptoms also related to physical disorders, e.g., headache, dizziness, or insomnia (Bjelland et al., 2002; Hinz & Braehler, 2011; Zigmond & Snaith, 1983). The scale consists of 14 items, divided into a 7-item anxiety (HADS-A), and a 7-item depression subscale (HADS-D). Both subscales are rated on a four-point Likert scale, giving subscale scores ranging from 0 to 21 (Zigmond & Snaith, 1983). There is no universally accepted cut-off score for the HADS (Herrero et al., 2003; Spinhoven et al., 1997). In this study, the cut-off point was set to eight, indicating elevated/caseness anxiety and depression for scores ≥8, following the recommendations by Zigmond and Snaith (1983), as well as Bjelland et al. (2002) and Herrero et al. (2003).

Zigmond and Snaith (1983), as well as Bjelland, Dahl, Haug, and Neckelmann (2001) and Herrero et al. (2003) recommended a cut-off score of 8 for both subscales,while a score of ≥8 indicates caseness.

(Bjelland, Dahl, Haug & Neckelmann, 2001)

In this study the German Version (HADS-D) of the scale was used.

FACIT-F

The 13-item Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale (FACIT-F; Version 4) was used to assess self-reported fatigue and its impact on daily activities and functions (Cella, Lai, Chang, Peterman & Slavin, 2002; Yellen et al., 1997) during the last 7

Commented [JH79]: Though in some studies the original cut-off scores (8+ and 11+) failed to be optimal,
until now there is no convention concerning the use of other cut-offs.
- Hinz & Brähler 2011

Commented [JH80]: Symptoms relating also to physical disorder, such as dizziness, headache or
insomnia were not included

Commented [JH81]: Zeitform

Commented [JH82]: Time interval (one week)

Commented [JH83]: Or maybe literature already here? And then maybe again after cancer-related fatigue!?

days. While it was originally developed for cancer-related fatigue (Cella et al., 2002; Yellen et al., 1997), it has been shown, that the FACIT-F is a reliable and valid measure of fatigue across various health conditions (Cella et al., 2002), making it a widely used tool for both clinical practice and research (Cella et al., 2022; Butt et al., 2013; Montan, Löwe, Cella, Mehnert & Hinz, 2018; Tinsley, Macklin, Korzenik & Sands, 2011). This questionnaire utilizes a five-point Likert scale, with total scores ranging from 0 (severe fatigue) to 52 (no fatigue). Based on general population data, scores ≤30 indicate clinically significant fatigue (Piper & Cella, 2010).

Cluster Analysis

Cluster analysis is an explorative statistical method used to organize objects, data points, or observations into homogeneous groups, known as clusters, based on similarities (Ketchen & Shook, 1996). The goal is to achieve high homogeneity within groups (intragroup homogeneity) and high heterogeneity between groups (intergroup heterogeneity) (Bacher, Pöge & Wenzig, 2010; Backhaus, Erichson, Gensler, Weiber & Weiber, 2011). In this study, the behavioral cognitive data from the PVT, TMT, MoCA, and n-back task will be utilized as cluster variables, aiming to identify two clusters that differ in their cognitive performance levels suggesting, that one group may perform better or worse than the other.

Data Preprocessing. All participant data was imported from a TSV file (participants.tsv) into R (using read.delim()), where preprocessing and analysis of the data was conducted. Specific variables of interest were selected and stored as subset including demographic information, cognitive test scores and clinical/questionnaire measures. The MoCA variable was converted to a binary variable: scores ≤25 indicate cognitive impairment, while scores >25 indicate no impairment. Missing values were added for nback task, if participant mentioned that they did not understand the task. Rows with missing values in cone of the relevant cognitive test variables (pvt….) were removed. Missing values in the relevant cognitive test variables (PVT reaction time, n-back miss 1, n-back miss 2, TMT a time, TMT b time (need to mention earlier)) were detected/checked is.na() and rows (participants) with missing values in key cognitive variables were removed using drop_na(), to ensure complete cases for analysis. 9 rows have been delted because of missing values, leaving the dataset with 70 participants. Winzorising was used to replace outliers by capping extreme values beyond 1.5 times the interquartile range (IQR). Outliers are identified and winsorized for the before mentioned cognitive variables. A function winsorize_variable is defined to perform winsorization, replacing values beyond 1.5 * IQR with the values at Q1 – 1.5 * IQR or Q3 + 1.5 * IQR.

Commented [JH84]: Again literature here? Not quite sure where exactly it should be

Commented [JH85]: Again „uses“ maybe there is a better/other word I could use instead

Commented [JH86]: Do I need to mention the literature here again?

Commented [JH87]: Check checklist from thesis_2711 to see if I mentioned everything important

Commented [JH88]: Later deleted

Commented [JH89]: Maybe: Why did I chose those?

Commented [JH90]: Further down it‘s written in much more details

Commented [JH91]: Happened in past!

Commented [JH92]: Quote (program number, etc.)

Commented [JH93]: Indicate or indicated?

Commented [JH94]: Potential cognitive impairment?

Commented [JH95]: Indicate or indicated?

Commented [JH96]: Align with what I wrote in MoCA section. Maybe literature here

Commented [JH97]: If comment said..

Commented [JH98]: For how many participants

Commented [JH99]: How many?

Commented [JH100]: Literature?

Commented [JH101]: See Lilys BA

Winsorized versions of variables were created with _w suffix. 4 outliers were detected for PVT reaction time and winsorized. 2 outliers for TMT_a_time. 4 outliers TMT_b_time. A custom function winsorize_variable() was implemented to replace extreme values beyound 1.5 times the interquartile range (IQR) with the nearest non-outlier values (Q1 - 1.5 * IQR or Q3 + 1.5 * IQR). New variable TMT_diff was calculated as the difference between TMT B and A (B-A).

Variable transformation

To account for the influence of age on cognitive performance, participants were divided into age groups, and z-scores were calculated within each group to adjust the data.

Two functions were defined to categorize participants into age groups. Age groups were created. Age groups added to the dataset. Participants were divided into four distinct age groups, 18-34 years (12 participants), 35-49 years (19 participants), 50-64 years (33 participants), and 65-80 years (6 participants). Age groups orientated from TMT norms. Why decided for this age groups? A separate categorization was used specifically for the TMT difference score. Also 4 age groups but 18-24 years (2 participants), 25-54 years (34 participants), 55-64 years (28 participants), and 65-80 years (6 participants). Why different age groups for TMT difference and all other variables?

After age groups were created, mean and standard deviation for each cognitive variable (PVT and TMTa, TMTb and TMT difference) were calculated within each age group. The function calculate_z_scores_individual() was used to compute z-scores for each participant based on age group norms, adjusting for age-related differences in cognitive performance. Creating new variables with _z suffix. Additionally n-back miss scores (miss 1 and miss 2) are standardized using the scale() function, creating new variables with _s suffix. When looking at the n-back means in each age groups, no age related trent could be detected, therefore no z-score was calculated for the n-back values, instead was standardized. N-back miss scores were standardized using the scale() function to ensure comparability across variables.

The final cleaned and processed dataset was saved for further use as a file named "clean_data.Rdata" in the cluster analysis. Saved as an R data file named clean_data.Rdata.

Cluster Analysis/hierarchical clustering. A hierarchical cluster analysis was performed on the preprocessed test data to identify clusters among participants. Hierarchical cluster analysis was performed on the preprocessed cognitive test data to identify (potential) subgroups/clusters within the participant pool/among participants. Which packages were loaded? Tidyverse, dplyr, dendextend, ggplot2, gridExtra, purrr, vroom. This approach was

Commented [JH102]: Why?

Commented [JH103]: Is that detailed enough or do I need to provide even more information? Or maybe an example?

Commented [JH104]: Those I need to mention

Commented [JH105]: Which age groups and why those age groups

Commented [JH106]: Do I need to mention the specific age group means for pvt and tmt?

Commented [JH107]: Probably table here

Commented [JH108]: Why no age groups for this variable?

Commented [JH109]: Prove might be needed here

Commented [JH110]: Since it's explorative investigated k-means and hierarchical. Why then hierarchical? Better fitting

Commented [JH111]: Why hierarchical and not a different one? I also did k-means but didn't give that much meaning. But I should probably still mention that

Commented [JH112]: That would mean that I already gave an introduction on cluster analysis

Commented [JH113]: What is my goal? But I guess I will explain that earlier, so no worries

employed to uncover/reveal patterns in cognitive performance of the participants. Cognitive variables were compared as winsorized, standardized, as well as the original scores were compared.

Only include if good reason to think they will define the clusters. First all cognitive variables, after reduced, why?

At first PVT reaction time, TMT A, TMT B, n-back miss 1 and n-back miss 2, MoCA were used as variables within the clustering. Explorative approach. Using n-back let to bad clustering results. Because everyone performed quite bad. Therefore n-back was excluted from the cluster analysis. TMT difference was not used, since it would double information (not independent). Too high correlation. Why did I choose the variables I chose? Where do I need to write that? Why not used MoCA?

As the hierarchical

The first step in the hierarchical cluster analysis process was to compute a distance matrix. In this approach Euclidean distance method wa used Proximitätsmaß/Distanzmaß bevorzugt bei WARD. All values continuous numerical values that is why used euclidean. Tried different kind of linkage methods and than decided which one performed best, based on. Why? Ward methods was choosen as the algortithm. Specified linkage method via method argument. Dendrogram was build by plotting hierarchical cluster object with hclust. Created desired number of clusters. Cut_mean <- cutree (hclust_median, k = 2) In consideration of the research question and the two groups, withPCS and withoutPCS, a two-cluster solution was chosen. K-means/elbow measure suggested 4 clusters. Therefore a 4-cluster solution was also looked at. As validation, the 2-cluster solution was compared to their self repoRTEd group assignments (withPCS or withoutPCS. Do I here need to mention, how many people are in which cluster? Or is that already result? To visualize cluster on dendrogram abline function used. Stability tested, different proximity measures have been used. But with eucledean best result. Also different algorithms have been tested. Also non-hierarchical clustering was compared to the hierarchical clustering (k-means). Why did I do the analysis with ward and not with k-means? What was my decision there? For the selected number of cluster, three additional analyses were performed using the complete, sigle, and weighted-average linkage methods. The agreement(Übereinstimmung) was assessed using the adjusted Rand index (Hubert and Arabie, 1985). Change of algorithm and alteration of

Commented [JH114]: Might need to mention that earlier

Commented [JH115]: Irrelevant or masking variables should be excluded if possible

Commented [JH116]: Quadriert?

Commented [JH117]: How detailed? Do I also need to write how I cut the three?

Commented [JH118]: Maybe for now write and maybe later than delete

Commented [JH119]: I think in results

number (which numbers where tested?) of clusters was variated. For adjusted rand index: library(fossil) adjusted rand index calculated with rand.index function.

PVT reaction time, TMT a, TMT b, TMT b-a, MoCA, and n-back scores were included. Euclidean distance was used as a distance matrix, and Wars's method was selected for clustering, as it is widely used in practice and known for its effectiveness in identifying distinct clusters (Backhaus et al., 2011). It is considered a reliable algorithm, provided that the variables are on a metric scale, are uncorrelated, and do not contain outliers (Wentura & Pospeschill, 2015). As stated earlier, cluster analysis is an explorative method used to identify patterns in data. However, in this study, the approach is only semi-exploratory, as the number of clusters to be generated was predetermined based on prior knowledge. A two-cluster solution was explored, as that aligns with the self-reported groups (with PCS, and without PCS), thereby allowing for good comparisons between the cluster solution and the self-reported groups. One could also consider this as a confirmatory cluster analysis (Bacher et al., 2010). The stability of the clusters was tested by comparing different proximity measures and algorithms using the adjusted Rand index (Hubert & Arabie, 1985). **Verwendete Kriterien zur Bestimmung der Clusteranzahl**

Inhaltlich, da zwei verschiedene Gruppen. Later 4 because of k-means

DurchgefühRTE Stabilitätsprüfung

Für ausgewählze Clusteranzahl noch drei weitere Analysen mit dem Complete-, Single- und Weighted-Average-Linkage gerechnet. Die Übereinstimmung wurde mittels des adjustieRTEn Randindex (Hubert und Arabie 1985) beuRTEilt.

Auch: Wechsel des Algorithmus und Veränderung Gruppenzahl (Ein Cluster bleibt gleich)

DurchgefühRTE Validitätsprüfung

Zur Validitätsprüfung wurde auf Variablen Z1, Z2 usw. zurückgeriffen

Statistical analysis. The two clusters were compared in several aspects. First, the two clusters were compared in their cognitive performance levels to validate whether significant differences exist between clusters. Clusters were then compared across demographic variables and results in questionnaires. Of particular interest was to examine how those two clusters differ from or align with the self-reported perception of cognitive performance level. To investigate differences between objective and subjective cognitive performance levels, comparisons occured not only between two clusters but also within the clusters between the subjective groups with PCS and without PCS. Additionally, to maximize the insights from the

Commented [JH120]: One cluster stayed the same, no matter which cluster numer was chosen.

Commented [JH121]: Didn't I already mention that?

Commented [JH122]: Why?

Commented [JH123]: Are those metric scales?

Commented [JH124]: Check for correlation?

Commented [JH125]: Outliers were removed

Commented [JH126]: Here or earlier?

Commented [JH127]: Also 4 cluster solution sine k-means said 4 cluster solution is better

Commented [JH128]: Results on stability here or in results?

cluster analysis, the with PCS groups in cluster 1 was compared to the with PCS group in cluster 2, and similarly for the without PCS groups. A t-test was used for these comparisons. Effect size and cohens d were also compared (need to check why)

The clusters were compared in several expects with each other. In demographical variables (sex, age, and years of education), in the used variables for clusteranalysis. But also in their other cognitive variables (PVT, TMT, n-back, MoCa). Also results in the scores from questionnaires were compared. Not only were the two groups compared between each other, but also within comparasion took place. WithPCS and withoutPCS within one cluster were compared. Also withPCS and withoutPCS were compared between clusters (that means, withPCS in Cluster 1 was compared to withPCS in Cluster 2 to clarify). All comparisons were tested by t-test. T-test robust to….. Data is not normal distributed. That was tested by… cat function was used.

The two clusters where compared

Alongside the comparisons of demographic, cognitive data, and questionnaire results, the clusters were also examined for their EEG resting state patterns.

EEG Recording and Analysis

For each group (withPCS and withoutPCS), 5 minutes of resting state with eyes open and 5 minutes of resting state with eyes closed were recorded using high-density EEG. EEG signals were recorded using a 128-channel EEG cap (128Ch Standard Brain Cap for actiCHamp Plus, Easycap GmbH, Wörthsee, Germany) with electrodes positioned in an equidistant layout, connected to an actiCHamp Plus Amplifier (Brain Products GmbH, Gilching, Germany). The sampling rate was 1000 Hz with an amplitude resolution of 0.1 µV. Electrolyte gel was applied to improve conductivity between skin and electrodes, ensuring impedances remained below 20 kΩ. Eye movements and changes in the resting potential of the retina (EOG activity) were monitored using two EOG electrodes placed below each eye, with impedances also maintained below 20 kΩ. In addition, a ground electrode was positioned on the forehead, and a reference electrode was positioned on the tip of the nose. Impedances for both the reference and ground electrode were kept below 5 kΩ.

Preprocessing

Commented [JH129]: Why t-test? And why is that robust? Why am I allowed to use that one?

Commented [JH130]: The analysis of this study was focused on the condition eyes open only.
All code described in the following sections can be found in this public GitHub repository: https://github.com/LGodbersen/Masters-thesis.

Commented [JH131]: EEG signals were collected using the 128-channel actiCap System (Brain Products GmbH, Munich, Germany) with electrodes positioned in an equidistant layout, with a sampling rate of 500 Hz and amplitude resolution of 0.1 µV.

Commented [JH132]: As alternative

Commented [JH133]: Here or earlier? Where should I mention how many Versuchsleiter!?

Commented [JH134]: Or just: EOG activity was recorded using two dedicated EOG electrodes, placed below each eye.

Commented [JH135]: Matlab is widely used by the EEG community and enabed us to use well-established Matlab-based EEG toolboxes that provide robust functions for computing functional, connectivity measures (Avila et al., 2023) - they followed a pragmatic approach towards preprocessing and adopted a simple, established, and automatic workflow in EEGLAB proposed by Pernet et al. And originally developed for ERP data. Adapted this pipeline to resting-state data and detail the seven preprocessing steps below.

Commented [JH136]: Loading the data
1. Line noise removal
2. High pass filtering and bad channel rejection
3. Re-referencing
4. Independent Component Analysis and automatic IC rejection
5. Interpolation of removed channels
6. Bad time segment removal
7. Data segmentation into epochs

Data preprocessing/analysis was performed using the FieldTrip toolbox (Fieldtrip-20240504; Oostenveld, Fries, Maris & Schoffelen, 2011) and the EEGLab toolbox (v2024.0; Delorme & Makeig, 2004) in Matlab (v24.1.0.2578822 (R2024a) Mathworks Inc., 2024, MathWorks® https://de.mathworks.com) on Windows.

The participants' EEG data were organized in BIDS (Brain Imaging Data Structure) format (Gorgolewski et al., 2016; Pernet et al., 2019). BIDS is a community standard that ensures homogeneity in the organization and description of raw neurocognitive/brain-derived/neuroscientific data, enabling efficient data sharing, minimizing errors, and supporting completely automated analysis workflows (Gorgolewski et al., 2016; Pernet et al., 2019; Truong, Robbins, Delmore & Makeig, 2023). The resting state EEG data, organized according to this standard, were identified and imported into MATLAB using the FieldTrip Toolbox. A trial defining function was built to select the data from the eyes-open condition for subsequent processing. This resulted in approximately 300 s per participant.

Filtering and Resampling

A finite impulse response (FIR) windowed-sinc (firws) filter, designed with a hamming windowed sinc function and implemented in the FieldTrip toolbox, was used for both high-pass and low-pass filtering of the continuous data. For high-pass filtering, a cut-off frequency of 0.1 Hz was applied to eliminate very low frequencies (drift) (Keil et al., 2013). This cut-off was based on the findings of Delorme (2023) and Winkler, Debener, Müller and Tangermann (2015), where filtering at 0.1 Hz or higher significantly improved data quality compared to no filtering. Filters above 0.1 were not used due to (I will come up with something).

Prior to applying low-pass filtering, the data was downsampled from 1000 Hz to 250 Hz, to reduce computational load while preserving sufficient temporal resolution for subsequent analysis. A cut-off frequency of 45 Hz was then used to eliminate high-frequency noise and mitigate potential 50 Hz line noise (Delorme, 2023). Finally, the data underwent re-referencing using the Common Average Reference (CAR) technique to remove the influence of the reference and improve signal quality (Ludwig et al., 2009). As the name implies, an average of the recordings from all electrode sites was computed and used as the reference (Ludwig et al., 2009; Offner, 1950). The function reref() was used (to perform this step). The data was converted into the EEGLAB data structure for further processing.

Commented [JH137]: Not only preprocessing

Commented [JH138]: Analyses are performed using Matlab 2020a (The Mathworks, Inc.3) on Windows or Mac OSx with the Statistical and Machine Learning Toolbox installed, along with EEGLAB4 (v2020.0) and its BIDS tool5 (v3.5) and LIMO EEG6 (v3) plugins—both of them available through the EEGLAB plugin manager.

Pernet et al. 2021

Commented [JH139]: EEGLAB (Delorme and Makeig, 2004) is the most commonly used platform for EEG data analysis (Hanke and Halchenko, 2011; Martínez-Cancino et al., 2020) and all steps proposed can also be reproduced from the user interface

Commented [JH140]: How to cite that correctly?

Commented [JH141]: Relevant?

Commented [JH142]: EEG (70 channel Easycap2) and ECG data were extracted from the binary MEG.fif files that combined MEG, EEG, ECG channels, event markers were time corrected (−34 ms) and electrode positions re-oriented to fit the head coordinate system. Out of the 19 participants, participant 1 was removed because of channels digitization errors leading to 18 participants. Pernet et al., 2021

Commented [JH143]: Important for empty dataset

Commented [JH144]: What describes EEG and MRI data best together?

Commented [JH145]: Where to find?

Commented [JH146]: Correct like that?

Commented [JH147]: EEG data were high-pass filtered using a finite impulse response (FIR) filter designed with a ...

Commented [JH148]: Some authors argue against high-pass filtering (or restrict the applicable high-pass cutoff to ...

Commented [JH149]: .The FIR filter was chosen for its stability and linear phase characteristics. ...

Commented [JH150]: Subsequently, high-pass filtering was administered. Since filters above 0.1 Hz can lead to ...

Commented [JH151]: Since filters above 0.1 Hz can lead to ...

Commented [JH152]: Do I need literature?

Commented [JH153]: I think Lara changed that. Before more. Why not 0.1? Need to check voice messages

Commented [JH154]: Avila et al., 2023

Commented [JH155]: After artifact removal, the **CAR** technique was applied again to re-reference the data, ...

Commented [JH156]: True?

Commented [JH157]: Data is re-referenced to the average reference with the function pop_reref(). Avila et al

Due to empty dataset from one participant, the participant was excluded, leaving the dataset with 69 participants (something like that. But where should I write that?)

Artifact removal

After the initial filtering and resampling, the preprocessing pipeline continued with detecting and removing artifacts. First, large artifacts - including the removal of flat-line channels, noisy channels, and short-time bursts of noise - were removed from the data using the EEGLAB pop_clean_rawdata() function with specific parameters. Channels with flat lines for more than 5 seconds were removed (FlatlineCriterion = 5), based on the default recommendation (for this parameter) by Pernet et al. (2021). This ensured the exclusion of “dead” or disconnected channels, thereby improving data quality. Channels were further excluded if their signal could not be predicted from a randomly selected subset of the remaining channels for at least 85% of the recording time (ChannelCriterion = 0.85), to remove those that were highly dissimilar from the rest of the channels (Gil Ávila et al., 2023; Pernet et al., 2021). The euclidean distance metric was used to calculate the similarity between channels. Data segments with abnormally high amplitude bursts, exceeding 100 SD compared to neighboring segments, were eliminated (BurstCriterion = 100), as such extreme bursts are considered unlikely to reflect brain signals (Chang et al., 2018). The default BurstCriterion is set to 20, but it may be adjusted if the default setting results in rejecting too many data segments. Some people recommend setting the threshold to 100 (EEGLAB, "Automated Pipeline Tutorial", 2024), which aligns with the optimal cut-off range of 10 to 100 suggested by Chang et al. (2018). Therefore, a mild threshold of 100 was chosen here, as it still effectively removes large-amplitude artifacts while retaining valuable data (Chang et al., 2018). Time windows where more than 40% of the channels were marked as noisy were removed (WindowCriterion = 0.4), to ensure the quality of the remaining data. A more lenient threshold of 0.4 was chosen over the default of 0.25 to retain more data (even if it is potentially noisier). How many “bad” (excessively noisy) channels were detected or removed in this process?

Again the data is re-referenced to the average reference (CAR), this time using the EEGLAB function pop_reref() (Gil Ávila et al., 2023).

Secondly, Independent Component Analysis (ICA) (Bell and Sejnowski, 1995; Hyvärinen, 2013, Jung et al., 2000; Lee et al., 1999; Palmer et al., 2008) was performed on the data, to

Commented [JH158]: One participant with empty dataset (KA14HH)

Commented [JH159]: „No „bad“ (excessively noisy) channels were detected or removed in this process“ after running pop_clean_rawdata EEGLAB plug-in (Truong, Robbins, Delmore & Makeig,)

Commented [JH160]: One participant with empty dataset (KA14HH)

Commented [JH161]: Data was cleaned

Commented [JH162]: Delete?

Commented [JH163]: Langer Strich, wie funktioniert das?

Commented [JH164]: But they suggested 0.8 which is the default

Commented [JH165]: The Burst Criterion was set to 100, meaning that artifact-bursts that had a higher than 100 SD amplitude compared to neighbouring segments were removed. Bursts of such high amplitude
are thought not to reflect brain signals. The Window Criterion was set to 0.4, meaning that a
data window where over 40% of all channels have been marked as noisy/bad, would be excluded.
This helped detecting periods in time that contained a lot of artifacts simultaneously.

Commented [JH166]: Seems weird but that’s how it is written in tutorial

Commented [JH167]: Maybe just delete people?

Commented [JH168]: Criterion for removing time windows that were not repaired completely

Commented [JH169]: Or bad

detect and reject further artifacts, such as eye or muscle movements (Makeig et al., 1995). ICA was performed with the "runica" algorithm and function pop_runica() with the extended InfoMax method. The runica algorithm was employed with the extended InfoMax method. using the pop_runica function implemented in EEGLAB. To avoid rank deficiency, the number of components was set to one less than the total number of channels (Kim, Luo, Chu, Cannard, Hoffman & Miyakoshi, 2023). This approach decomposes the EEG signal into independent components, potentially separating artifacts from neural activity. Due to the non-deterministic nature of the ICA algorithm, its results vary across repetitions. That is, every repetition of the ICA algorithm leads to small differences in the reconstructed time series after removing artifactual components (Gil Ávila et al., 2023). The resulting ICA weights, which represent the transformation matrix for this decomposition, were saved in a separate file.

Automatic component rejection was implemented using ICALabel (Pion-Tonachini, Kreutz-Delgado & Makeig, 2019), as automatic artifact rejection is preferred over the manual one to ensure standardization (Miljevic et al., 2022). Artifactual components are automatically classified by the ICLabel classifier (Pion-Tonachini et al., 2019). Thresholds were set at probabilities of 0.8 (80%) for muscle-related components (Pernet et al., 2021) and 0.5 (50%) for eye-related components. Components exceeding these thresholds were flagged and automatically removed using the EEGLAB function pop_subcomp(). By default, only components whose probability of being "muscle" is higher than 80% were subtracted from the data (Pernet et al., 2021). The two EOG channels (31 and 32) were removed from the dataset. The cleaned dataset was then checked for consistency using eeg_checkset().

Thirdly and finally, an additional artifacts removal step was implemented to address any remaining problematic channels. This process involved a statistical approach to identify outlier channels based on their signal characteristics. The standard deviation and mean were calculated for each channel across all time points. Then, overall mean values for these standard deviations and means were computed across all channels. Thresholds were established at 2.5 standard deviations above and below the overall mean, creating an acceptable range for channel activity. Channels with standard deviations falling outside this range were identified as outliers. These outlier channels were then removed from the dataset using the function EEGLAB function pop_select(), further refining the EEG data

Commented [JH170]: stolen

Commented [JH171]: In recent years, several EEG data preprocessing pipelines have been developed and published by different laboratories. To our knowledge, there has been no systematic review of these, nor is it quite clear what measures should best be used for fair comparison. It is best that EEG researchers acquaint themselves with the problems involved in adequate data preprocessing and test for themselves the particular pipeline they use or construct for this purpose. Because there is no agreed-upon standard (K. A. Robbins et al., 2020), we hesitate to promote a particular approach (see https://osf.io/8brgv/ for an example of an automated pipeline). We prefer to involve ICA decomposition in this process as it is shown to perform well in identifying and separating out several classes of non-brain source signals typically mixed in the scalp data (eye movements, scalp muscle activities), as well as identifying major spatially localizable effective brain sources that represent much of the brain's (largely cortical) contribution to the scalp data.

T… et al.

Commented [JH172]: Truong, Robbins, Delmore & Makeig, 2023

Commented [JH173]: This was supported by applying ICLabel (Pion-Tonachini et al., 2019), a neural network classifier trained on a large body of expert-labeled IC data, to automatically classify components as representing brain sources or as any of several classes of non-brain sources (see discussion below). Applying ICLabel before removing line noise correctly identified two strong line noise ICs but also caused ICLabel to classify effective brain sources still containing substantial noise noise as most likely representing line noise rather than brain source activity. After running zapline-plus on the scalp data and then applying the AMICA weights to the cleaned data to obtain the IC time courses (activations), ICLabel, now applied to the IC scalp maps and line-noise cleaned …

Commented [JH174]: Truong, Robbins, Delmore & Makeig, 2023

Commented [JH175]: How many good channels were left in both groups?

Commented [JH176]: Detecting and excluding bad channels is important (Delorme, 2023). After the ICA, the data were checked again to reject more possible bad channels or epochs. For each participant, all …

Commented [JH177]: Detecting and excluding bad channels is very important (Delorme, 2023).

Commented [JH178]: This resulted in, on average 105.9 good channels in the with PCS group (SD = 13.4, Range = 75 - 120) and on average 106.3 good channels in the without PCS group (SD = 9.4, Range = 90 - 121). This cor- …

quality. This step ensures that channels with unusually high or low variability, which might represent persistent artifacts or malfunctioning electrodes, are excluded from subsequent analyses.

Interpolate bad channels

Channels removed in the previous step were interpolated using the EEGLAB function pop_interp() with the default spherical splines method (Perrin, Pernier, Bertrand & Echallier, 1989), ensuring a consistent number of channels across participants (Gil Ávila et al., 2023). Interpolated channels were inserted into the original channel order. On average …. % of the channels in each group were interpolated.

Epoch length and number

Lastly, the continuous data are segmented into epochs with the function eeg_regepochs() implemented in the EEGLAB toolbox. By default, data are segmented into 2-second epochs (Gil Ávila et al., 2023), however longer epochs might be desirable for … to increase frequency (Gil Ávila et al., 2023).EEG data for each participant were segmented into into 5-second nonoverlapping epochs. This function will then output the new epoch EEG as a dataset on EEGLAB (Bonello, Garg, Garg & Audu, 2018).

For the restingstate delta power longer epochs are preferred, since delta contains slower frequencies
and the longer the epoch, the higher the resolution. This is why in the preprocessing
for the delta and beta power, the data was cut into 5 s epochs.

Power Analysis

Statistical Analysis

T-test for two group comparasing. For four groups different test.

Results

Cluster Analysis

Which two clusters? Which four clusters? How are they different from each other? T-Werte what is interesting to tell? Table? Stabilty of cluster, validity,…

Clinical data(?)

Might be interesting to compare the reported groups vs “my” groups

EEG Power

Delta (frontal ROI)

Beta (central ROI)

Correlations with clinical data(?)

Commented [JH179]: ?

Commented [JH180]: 7. Data segmentation into epochs. Lastly, the continuous data are segmented into epochs with the function
pop_epoch(). By default, data are segmented into 2-second epochs with a 50% overlap. Although longer epochs might be desirable for the Alpha Peak Frequency estimation to increase frequency resolution, short epochs favor the reliability of functional connectivity measures40,41
. Thus, we propose 2-second epochs to establish a balance between frequency resolution, stationarity of the signal, and reliability of the later extracted features. Fifty percent overlap was chosen to provide a smooth estimation of the power spectra and mitigate the loss of signal due to tapering42
. Epochs containing a discontinuity (e.g., because a segment containing an artifact was discarded) are
rejected automatically. Data segmentation was adapted from Pernet et al.14
, which focused on event-related data.
Alvin et al

Commented [JH181]: For the restingstate delta power longer segments are preferred, which is why in the prepro-cessing for the delta and beta power, the data was cut into 5 s epochs. The maximum would be
60 good epochs here. The preprocessing and age matching resulted in on average 40.7 good

Commented [JH182]: ‘Epoching’ of EEG data refers to extracting sections (epochs) of the data of equal duration time

Commented [JH183]: If I want to do the whole thing in steps

Commented [JH184]: Not in zotero yet

Commented [JH185]: Maybe . Freeman 2004

Commented [JH186]: Power spectrum. Power spectra are computed with the FieldTrip function ft_freqanalysis between 1 and

Commented [JH187]: Donoghue et al. (2020) summarize that evaluating a change in absolute power could potential

Commented [JH188]: The differences between the groups for the hypotheses concerning both power and connectivit

Commented [JH189]: Might be part of the results from cluster Analysis. To compare

Commented [JH190]: For the mean delta and beta power, the assumptions were also investigated. Normality assumption

Commented [JH191]: The with PCS group exhibited a slightly lower average relative delta power in the frontal

Commented [JH192]: Figure 2. Results of power analysis in the A delta (0.6 - 4 Hz) and B beta (14 - 30 Hz)

Commented [JH193]: The with PCS group exhibited higher average relative beta power in the central ROI (M =

Commented [JH194]: The frontal delta power correlated significantly with the HADS-D, ρ(44) = .31, p = .042,

Sex differences

Discussion

Clusters

Cluster 1 always says the same, the other one that gets split up

Delta power hypothesis

Beta power hypothesis

The current study examined the ….

TMT is a valid test (Sánchez-Cubillo et al., 2009).

(Kane, Conway, Miura & Colflesh, 2007) N-back tasks are continous-recognition measures, that present stimulus sequences.

Commented [JH195]: Descriptively, especially in the beta band, there seem to be sex differences, see Figure B2 in Appendix B. Especially for the connectivity but also slightly in the beta power, the values of the female group appear to be elevated compared to the male group. This seems to happen in both groups but might be even more pronounced in the with PCS group. For the delta power, there are two interesting aspects. First, the delta power for males in the with PCS group seems descriptively lower than for females in the with PCS group, but it secondly seems lower than for males and females in the without PCS group.

Commented [JH196]: Our result of no difference in delta power at frontal ROI does not support either the GABAhypothesis for lower delta power in PCS-pts as proposed by Ortelli et al. (2023) nor the brainstem-hypothesis for higher delta power in PCS-pts as suggested in Cecchetti et al. (2022) and Yong (2021b). Kopa´nska et al. (2022) did not find any significant differences in delta power either. Their control were the qEEG values of the same participants prior to infection and not an independent control group. This pre-post design could potentially provide a more robust framework for understanding the disease's impact, since a baseline is assessed and subsequent changes could be attributable to the disease. This might point towards no change in delta power being plausible. It is however not trivial, that we found a significant positive correlation of delta power with the HADS-D as well as a negative correlation with the MoCA and a close to significant negative correlation with the FACIT-F. These findings would indicate that having pronounced frontal delta power coincides with more depressive tendencies, more fatigue and less cognitive functioning. This is in direct contrast to Ortelli et al. (2023), who found their score of the …

Commented [JH197]: There was no clear direction in which the with PCS group was supposed to deviate from the without PCS group in terms of beta power. Our data support this further, since there was no significant difference between the two groups in our central ROI. Descriptively, the frontal and occipital regions appeared to differ between the groups, but the permutation test did not show significance. The higher frontal beta power in the with PCS group could seem like an artifact or like it may be driven by one individual. In relative beta …

Commented [JH198]: However, caution must be taken when trying to generalize the present results to clinical populations since patients may be using compensatory strategies to complete the test (Jefferson et al., 2006 ; Spikman et al., 2001).

Might be good for discussion

Commented [JH199]: They see the n-back task bit critically