University of Zurich^{UZH}

Master Thesis

Department of Psychology

Faculty of Arts and Social Sciences at the University of Zurich

Investigating the Neuronal Basis of Learning Processes and Memory Formation in Children and Adolescents with Various Pychiatric Disorders

Author: B.Sc. Oliver Zingg

16-924-094

Supervisors:Prof. Dr. Nicolas Langer, M.Sc. Dawid Strzelczyk

Chair of Methods of Plasticity Research: Prof. Dr. Nicolas Langer

01.06.2023

Contents

List of Figures

List of Tables

Abstract

Memory and learning are very important in everyday life. In individuals with psychiatric disorders, these cognitive abilities are often compromised. Especially in childhood and adolescents, such deficits can negatively impact many outcomes later in life such as, for example, academic success or health. Identifying learning problems can be challenging in cases where individuals struggle to provide behavioral responses. Furthermore, behavioral responses provide little information about underlying cognitive processes that are relevant to learning and memory. Event-related potentials (ERP), on the other hand, provide a near real-time measure of brain activity occurring in learning settings. Previous studies have linked an ERP component (P300) with successful learning using a visual sequence learning paradigm (VSLP) that allows tracking gradual memory formation on a behavioral and neurophysiological level. Through learning in VSLP, subjective expectancy of upcoming stimuli is systematically modulated, which in turn is suggested to influence P300 amplitude. This thesis extended the previously established link between P300 amplitude and successful learning to a sample of children and adolescents with various psychiatric disorders. More specifically, P300 amplitude was shown to predict learning success within subjects. Furthermore, P300 amplitude decreased over the course of learning regardless of sequence repetition. Associations between the degree of decrease in P300 amplitude and learning-related transdiagnostic constructs were explored through a network analysis. The network yielded no direct associations with the degree of decrease in the amplitude of P300. In summary, these results highlight P300 as a neurophysiological marker of learning and its potential clinical value.

1 Introduction

1.1 Memory and Learning

Memory and learning are concepts that are closely related. When it comes to defining these concepts, no clear consensus has emerged, but most people have an implicit understanding of these terms and know how to apply them (De Houwer et al., 2013; Klein, 2015). Some definitions of memory and learning are kept as general as possible to ensure the inclusion of all different disciplines that investigate memory and learning (Barron et al., 2015; Klein, 2015). For example, a broad definition of memory is that it reflects organically based activity where information is made available through perception or introspection and thus alters the neural machinery (Klein, 2015). A similar broad working definition of learning is that it reflects "processing of information derived from experience to update system properties" (Barron et al., 2015, page. 405). Other definitions are more pragmatic and related to specific subdisciplines such as cognitive psychology (see, Barron et al., 2015). For example, definitions found in cognitive psychology are that learning reflects the act of storing information in memory (Medin & Ross, 1992), or that memory is the product of learning and the process of retention and retrieval (Crowder, 1976). Difficulties in defining learning and memory in an all-inclusive way may reflect the complexity of these constructs and the limitations that still exist when it comes to understanding these concepts. However, with memory being a heavily studied construct in psychology[1], much insight has already been gained. For example, it is commonly accepted that new information passes through different stages and ends up as permanent memory in long-term memory (Aben et al., 2012). These stages are encoding (i.e., initial learning), storage (i.e., maintaining information), and retrieval (i.e., access stored information) (Melton, 1963). Over the years, it became clear that memory is not a unitary system, which led to a taxonomy of memory systems with a main distinction

---

[1]For example, Google Scholar search with ["memory"] yielded 1'720'000 results

between memory about facts and events (i.e., declarative) and unconscious memory such as, for example, skill and habit learning (i.e. non-declarative)(N. J. Cohen & Squire, 1980; Squire, 1986, 2004). What started out as a philosophical inquiry into memory through introspection (Kahana, 2012) has led to an interdisciplinary field that incorporates research from psychology, neuropsychology, biology, and neuroscience (Kahana, 2012; Polster et al., 1991; Squire, 2009).

1.1.1   Beginning of Experimental Study of Human Memory

One of the first scientific endeavors in studying human memory came from Hermann Ebbinghaus in 1885 (Ebbinghaus, 1885; Kahana, 2012). Arguing for experimental methods with highly controlled conditions, Ebbinghaus constructed learning material of syllables, which he used in experiments to study his own memory performance. Focusing on controlled experiments rather than studying memory in its complex manifestation in daily life, Ebbinghaus established the foundations for modern experimental study of memory (Kahana, 2012). Years later, after building on numerous experimental studies, a highly influential theory of memory (i.e., modal model by Atkinson and Shiffrin (1968)) was postulated (Malmberg et al., 2019). As described in Atkinson and Shiffrin (1968) the modal model theory views memory as having three structural components: sensory register, short-term store (STS) and long-term store (LTS). The sensory register is where incoming stimuli are first stored, but decay after a period of several hundred milliseconds[2]. The second component is the STS, where a limited amount of information can be stored. It was assumed that the stored information decays over a time period considerably longer than that of the sensory register. Finally, the long-term store is where unlimited information is stored relatively permanently. It was proposed that information could be transferred bidirectionally from one memory component to the next. Information transfer occurs through control processes. These processes can be described as strategies

---

[2]In subsequent versions of the model, the sensory register is combined with the short-term store to form a single component (Atkinson & Shiffrin, 1971)

that individuals use to encode, store, and retrieve information from the different memory stores. Short-term store when cognitive processes are active is labeled working memory (WM) (Atkinson & Shiffrin, 1968) and plays an important role in learning (A. Baddeley, 2010).

1.1.2 Working Memory

Following criticism (see., Raaijmakers, 1993), A. D. Baddeley and Hitch (1974) proposed a theory of working model to replace STS from the modal model. Their model is composed of two domain-specific stores: the phonological loop (verbal information) and the visuospatial sketch pad (visual-spatial information), a central executive that coordinates the processes of the storage systems, and an episodic buffer, which holds semantic information associated with both domain-specific stores (A. Baddeley, 2000; A. D. Baddeley & Hitch, 1974). Through this model, the concept of working memory became famous (Aben et al., 2012) and led to increased interest in the research of WM with different theories being proposed (Cowan, 2017; Oberauer et al., 2018). With the increase in research, the term WM has received different meanings depending on the theoretical framework in which it was used in (see., Cowan, 2017). Therefore, its distinction from concepts such as short-term memory became blurred and may even be seen as the same (see., Aben et al., 2012). However, most frameworks describe the basic feature of WM as a system for the limited short-term maintenance of information (Eriksson et al., 2015) that can be accessed for a variety of abilities such as learning, reasoning, and preparation for action (A. D. Baddeley & Hitch, 1974). It is suggested that a network of frontoparietal brain areas is activated during WM tasks, where the parietal regions are believed to be involved in manipulating and updating information in WM, and the frontal regions serve to monitor and organize memory processes (J. D. Cohen et al., 1997; Koenigs et al., 2009; van Asselen et al., 2006).

1.1.3   Development of Working Memory and Its Importance for Learning

Working memory is involved in all deliberate cognitive operations (Oberauer et al., 2018) and is especially important in maintaining task-relevant information during cognitive tasks (A. D. Baddeley & Hitch, 1974; Daneman & Carpenter, 1980). It is well known that the capacity to store or manipulate information in WM is limited, but the exact capacity or the mechanistic functioning of WM that cause the limited capacity are still being discussed (see., Cowan, 2010; Oberauer et al., 2016). Individuals differ in WM capacity, and these differences are often associated with differences in many high-level cognitive abilities (Cowan et al., 2005; Daneman & Carpenter, 1980; Miyake et al., 2001; Unsworth & Engle, 2007; Unsworth et al., 2014) and academic performance (Gathercole et al., 2003; Gathercole et al., 2004). The limitation of WM capacity is even more pronounced in childhood, while capacity has been shown to increase linearly from age 4 to 15 in cross-sectional studies (Brady et al., 2009; Gathercole et al., 2004). Furthermore, the development of WM capacity is marked by a large degree of variability between individuals (Gathercole, 1998). This variability may be due in part to the many aspects of working memory that seem to develop during childhood and adolescents, as described by Cowan (2014). First, it is suggested that the capacity to store items itself grows throughout childhood and adolescents (Andrews & Halford, 2002). Second, it has been argued that increased knowledge contributes to the observed increase in working memory performance. For example, Chi (1978) showed that children with knowledge of the chess game remembered the chess configurations better than adults without knowledge of the chess game. Lastly, processing speed (e.g., Kail & Salthouse, 1994) and the availability of strategies are discussed as contributing to improving working memory performance during childhood and adolescents (Cowan, 2014). Although it is not yet clear which factor is primarily driving the observed differences in working memory performance, all should be considered in the context of learning and education (Cowan, 2014). For example, complex learning material could be presented in a way that is accessible with limited working

memory capacity, in a way that facilitates the use of strategies, or presented in the context of previous knowledge of the children. The large corpus of research has shown the crucial link between working memory and learning. For example, children with learning disabilities or low academic performance often have low scores on working memory tests (Sabol & Pianta, 2012; St Clair-Thompson & Gathercole, 2006). Additionally, successful participation in a classroom environment requires processing and storing of new information. Higher-level cognitive abilities involved in processing allow individuals to form plans, initiate actions toward those plans and persevere until their completion, are commonly described as executive functions (EF) or as central executive (referring to the working memory model of A. D. Baddeley and Hitch (1974)). Prominently postulated EF abilities are shifting (e.g., shifting between task, operations, or mental states), updating (e.g., monitoring and encoding relevant information, updating content held in WM by replacing irrelevant content) and inhibition (e.g., deliberately inhibiting automatic responses) (Miyake et al., 2000). These abilities have been shown to strengthen during childhood and adolescents (Best & Miller, 2010). The prefrontal cortex (PFC) is a brain area that is often suggested to play a major role in networks associated with EF (Diamond, 2002; Fiske & Holmboe, 2019). PFC is suggested to be one of the final brain regions to fully mature in humans and to have delayed maturation with neurodevelopmental disorders such as attention deficit hyperactivity disorder (Kolk & Rakic, 2022). Poor performance in executive function tasks where the processing and storage of verbal material is tested has been linked to underachieving in English and mathematics (Gathercole & Pickering, 2000). EF furthermore promotes successful switching and evaluation of new strategies deployed during tasks (Bull & Scerif, 2001), as well as staying on a task and keeping track of one's goal (Kane & Engle, 2003), which is essential in learning settings (Cowan, 2014). Individuals with psychological disorders characterized by deficits in EF and WM are often found to struggle in academic settings (Lonergan et al., 2019; Martinussen et al., 2005; Semrud-Clikeman, 2005; Willcutt et al.,

2005), further highlighting the importance of these cognitive functions for learning and academic performance. Obtaining academic skills such as reading and mathematics is crucial for many outcomes later in life, including monetary income, physical and mental health, and longevity (Calvin et al., 2017; Kuncel & Hezlett, 2010; Wrulich et al., 2014). Recently, cognitive abilities (e.g., WM and EF) and academic achievement have been considered to influence each other in a bidirectional way, that is, schooling facilitates the development of cognitive abilities, and increased cognitive abilities improve academic achievement(Peng & Kievit, 2020). The above mentioned findings indicate the importance of identifying individuals struggling in academic settings and understanding their struggles with learning and memory formation.

1.1.4 Studying Processes Underlying Memory Formation

Studying the underlying processes involved in memory formation was done mainly with mathematical modeling of cognitive processes that are of theoretical importance (Kahana, 2012; Schall, 2004). It is difficult to measure these processes directly, although past studies have used measures such as response time to infer upon processes underlying memory formation (e.g., Garner et al., 1956). With the ability to monitor neurophysiological processes in near instantaneous fashions, an opportunity has emerged to directly map neural and mental processes (Schall, 2004). Relying solely on behavioral data, such a direct link between mental and neural processes is difficult to make, as behavioral data reflect the final output of many cognitive processes (Luu et al., 2009) and can be confounded by numerous biases (Webb et al., 1999). Neurophysiological measures such as electroencephalogram (EEG) have been shown to be appropriate measures for assessing learning (Tinga et al., 2019) and could provide novel insights about the underlying processes that contribute to task performance and possibly explain differences in performance(e.g., Adam et al., 2015; Kim et al., 2014). Additionally, monitoring processes during learning in an online fashion can also be used to evaluate interventions with the aim of facilitating the learning process (e.g., napping vs. playing mobile games)

(Chiang et al., 2018). Thus, incorporating such neurophysiological measurements could provide an opportunity for understanding why some individuals struggle with learning and identifying interventions that facilitate learning.

1.2 Electroencephalogram (EEG) and Event-related Potentials (ERP)

EEG refers to electrical activity recorded from electrodes placed on the head surface. This electrical activity is generated by synaptic excitations of a large population of mainly pyramidal neurons with the same polarity and orientation in the cerebral cortex (Luck, 2014; Teplan, 2002). This summation of many postsynaptic potentials can be measured with a precision of milliseconds, which is a big advantage of EEG (Light et al., 2010). There are multiple methods that allow one to extract useful information from the EEG signal. The computing of event-related potentials (ERP) is one of such methods and enables the extraction of brain activity associated with mental processes that occur in response to experimentally defined events (Luck, 2014; Sur & Sinha, 2009). Such events are, for example, the onset of a stimulus. ERPs are calculated from EEG recording by time locking the signal to an event of interest resulting in EEG epochs that can be averaged in a next step (Luck, 2014). As described in Luck (2014), the assumption behind averaging these EEG epochs is that it results in random event-unrelated positive or negative fluctuations in voltage being canceled out. Therefore, the assumption is that brain activity that is consistently followed after the event remains after averaging, while random fluctuations that are not related to the event are canceled out. The resulting ERP waveform is described in terms of a sequence of positive and negative voltage peaks relative to a reference electrode. These peaks are characterized by amplitude and latency and have traditionally received the greatest interest in research (Luck, 2014; Luck & Kappenman, 2012). Early appearing peaks in the ERP waveform are often described as being exogenous or sensory, since they necessarily appear after an event and vary with physical parameters of an event (e.g., stimulus) (Luck, 2014; Sur & Sinha, 2009). Later appearing peaks in the waveform are seen as cognitive or endogenous, as they are related

to high-order cognitive processes (Luck, 2014; Luck & Kappenman, 2012; Sur & Sinha, 2009). It should be noted that the distinction between exogenous and endogenous peaks is considered helpful but problematic because peaks can be influence by endo- and exogenous factors (Luck, 2014). A highly studied peak is P300, first described by Sutton and colleagues after finding a larger positive deflection around 300 ms after presenting stimuli with unpredictable modality (i.e., visual or auditory) compared to stimuli with predictable modality (Sutton et al., 1965).

1.2.1 P300

P300 or P3b refers to a positive deflection in amplitude of the ERP waveform that typically occurs 250-500 ms after the onset of an event (Polich, 2007). Although the time at which the peak occurs (i.e., latency) depends on age, stimulus, modality, task conditions, and more (Polich & Kok, 1995). This peak in relative amplitude was first described by Sutton et al. (1965) by demonstrating that stimuli with unpredictable modalities (i.e. visual or auditory) elicited larger positive amplitudes around 300 ms after presentation compared to stimuli with predictable modalities. The authors concluded that P300 reflects a correlate of stimulus uncertainty (or expectancy) and that this effect was independent of modality. Later, N. K. Squires et al. (1975) found that presenting deviant auditory stimuli in a series of frequent auditory stimuli without instruction to react to the deviant stimuli elicited a P300-like peak with more of a frontal maximum, which was referred to as P3a to distinguish it from the more parietal located P3b (from here on referred to as P300). P3a is believed to reflect an automatic orientation response to interesting information and is sometimes referred to as novelty P3 (Escera et al., 2000; Polich, 2007). It is difficult to pinpoint the neural sources that generate the P300 signal because it is an inverse problem. An inverse problem refers to estimating unknown parameters given some observation or data where no unique solution is possible, rather there can be an infinite number of combinations of parameters that fit the observed data equality well (Baillet, 2013). In the context of EEG these unknown parameters are

"anatomical location and current-flow orientation in space and amplitude variations in time" of neural sources generating the EEG signal (Baillet, 2013, page. 1). Still, research studying individuals with lesions has suggested that the temporal-parietal junction contributes to the P300 signal and with lesion in that area, P300 amplitudes are greatly reduces (Knight et al., 1989; Verleger et al., 1994).

1.2.2   Modulating Factors of P300

Since its first observation, a lot of research has been conducted, specifying under which circumstances P300 would be elicited or altered, and exploring the functional significance behind this peak. It has been suggested that P300 is an endogenous correlate of cognitive processing of stimulus information rather than reflecting brain activity elicited by the physical parameters of the stimulus presented (Hopfinger & West, 2006; Lacey & Lacey, 1980; Picton, 1992; Pritchard, 1981). It should be noted that physical changes in stimuli can elicit changes in P300 amplitude, but these changes depend on whether participants detect or have prior knowledge of these changes (Donchin et al., 1978). In subsequent research, oddball paradigms were often used to reliably elicit P300 amplitudes. Typical oddball paradigm tasks are where participants have to discriminate between a frequent stimulus and infrequent stimulus, the latter prompting a response from the participant (e.g. covert counting) (Donchin et al., 1978). Additionally, variants of the oddball paradigm with single stimulus or three stimulus variants have been robustly shown to elicit P300 amplitudes (Conroy & Polich, 2007; Walsh et al., 2017). A consistent finding in oddball paradigms has been the inverse relationship between the amplitude of P300 and the probability of which each target stimulus appears (Donchin et al., 1978). Duncan-Johnson and Donchin (1977) found this effect across the entire probability range from 0.10 to 0.90, but crucially, this depends on the subjective probability assigned to the event by the participant (Donchin et al., 1978; Pritchard, 1981). Furthermore, only task-relevant stimuli elicited a large P300 amplitude (Duncan-Johnson & Donchin, 1977; Polich, 2007). For example, Donchin and Cohen

(1967) created an experimental setting in which a flash of light randomly appeared superimposed on a background of a circle randomly alternating with a square. One group was instructed to attend to the background, while the other group was instructed to attend to the flash of light. Only the attend condition (i.e., background or flash of light) elicited a P300 amplitude, suggesting the importance of attention and task relevance for the P300 amplitude (Donchin et al., 1978; Pritchard, 1981). Early on, Donchin (1981) mentioned that the characteristics of P300 are very similar to those of orienting reflex (OR), specifically, task relevant and improbable events (or stimuli) tend to elicit OR. OR refers to an immediate behavioral or neurophysiological response to changes in the environment and was first postulated by Pavlov (1927). As described in Rushby et al. (2005), Sokolov (1963) later proposed that with repeated presentation of a stimulus, an according cortical representation (neuronal model) is formed and new stimuli that do not match this representation elicit an OR. Relating P300 and OR, Donchin (1981) suggested that the process of updating one's neuronal model is manifested in the amplitude of P300, furthermore, directly linking it to memory and learning processes through surprise and orienting reflex. Subsequent studies have aimed to test the context updating theory and its prediction that "events that elicit a P300 are remembered better than events that do not elicit a P300" Donchin (1981, page. 509).

1.2.3   P300 and Memory

In a seminal study, Karis et al. (1984) investigated whether P300 during encoding is related to subsequent recall performance as predicted by the context updating theory. The authors found that for participants using simple learning strategies (i.e., mnemonic), the amplitude of P300 predicted later recall performance, but for participants using more complex strategies to organize the learning material, P300 did not predict later recall but rather a frontal positive slow wave did (Karis et al., 1984). This effect of rehearsal strategy on the predictability of subsequent recall of P300 was further confirmed in a within-subject design manipulating the instructed rehearsal strategies, leading the authors

to suggest future studies to consider paradigms "that narrow the range of variables that can affect recall to variables that are also likely to affect the amplitude of the P300" so that the relationship of P300 and memory can be better understood (Fabiani et al., 1990, page. 300). Johnson Jr. et al. (1985) used a study-test paradigm to test the association of P300 with long-term memory acquisition. In this study, participants learned a list of words that was repeated four times with a test phase after each repetition. The test phase consisted of a recognition test in which learned words were mixed with an equal number of random new words (each test phase included different new words). The authors found that the amplitude of P300 increased during the study and test phases with an accompanying increase in recognition performance and that later recognized words elicited larger P300 compared to unrecognized words, although this difference did not reach significance (Johnson Jr. et al., 1985). The difficulty of the task can directly influence P300 amplitude (Courchesne, 1978; Kok, 2001) and was mentioned by Johnson Jr. et al. (1985) as a possible reason for not finding a significant difference between P300 and later recognized words, as the list of words was rather long. However, in many studies, the use of P300 has been found to predict the subsequent memory performance of words (see., Paller et al., 1988). Using a study-test paradigm with repeated presentation, as was done in Johnson Jr. et al. (1985), could provide valuable information when connecting P300 with memory processes, as repeating the same sequence strengthens memory traces, which is reflected in improved recall performance (Hintzman & Block, 1971). This is further mentioned by Rugg and Doyle (1994) who suggested that changes in neural response to repeated presentation of stimulus indicate the formation of an internal representation of that stimulus. In light of the context updating theory (Donchin, 1981), one would expect P300 to decrease with repetition, since learned material becomes increasingly expected and the internal representation formed of the sequence needs less updating. The findings of Johnson Jr. et al. (1985) that the amplitude of P300 increased with repetition are incongruent with this assumption, although other studies using repeated presentation

have shown that P300 decreases as events become expected (Courchesne, 1978) and is also a consistent finding from oddball paradigm. Furthermore, the P300 amplitude is considered to be influenced by the sequence of immediately preceding stimuli (K. C. Squires et al., 1976) and since the sequence in the recognition phase differed from the sequence in the study phase, this may have influenced P300 amplitudes. Using a list of words to study memory could lead to complexities that systematically vary with age, for example, the use of strategies during encoding (Schneider et al., 2004). Since P300 amplitude has been linked to memory encoding that promotes successful memory rather than directly to memory itself (Fernández et al., 1999; Polich, 2007), using a list of words to learn as in Johnson Jr. et al. (1985) may provide more variability in encoding strategies like chunking (Nogueira et al., 2015) or imagery (Zahedi & Abdi, 2012) between individual items as well as between participants with varying age. Incorporating paradigms with highly simplified stimuli could minimize variables like encoding strategies that possibly vary between items and are known to influence P300 amplitude (e.g., Karis et al., 1984). Steinemann et al. (2016) used a visual sequence learning paradigm (see Chapter2) in which participants learned a sequence of highly simplified stimuli with minimal semantic value (i.e. flashing circles) through repetition with regular behavioral recall. With stimuli only differing in their spatial locations, this paradigm allowed differentiating between endogenous variability in encoding or memory formation and interstimulus variability within and between subjects (Steinemann et al., 2016). Furthermore, Steinemann et al. (2016) mentioned an additional strength of this paradigm is to track the memory traces within individual stimuli as they gradually transit from unknown over newly learned to fully committed to memory. The authors hypothesized that stimuli fully committed to memory become expected during subsequent presentation, which in turn would manifest itself in lower surprise-regulated P300 amplitudes (Steinemann et al., 2016). This hypothesis could be confirmed as well as predicting successful learning with P300 amplitude, which led the authors to discuss the potential of

their findings as a diagnostic tool for learning disabilities, although more studies are needed in such target populations (Steinemann et al., 2016).

1.3 Clinical Application

ERP components provide useful information in clinical settings as near real-time measures of covert information processing, such as encoding processes (D. Friedman & Johnson Jr, 2000). For example, Riggins and Scott (2020) noted that with people whose expressive language or motor skills are impaired, the ability to rely on ERP components would be valuable. Additionally, it was suggested that P300 could in the long run provide information on the identification of individuals from vulnerable groups at risk of developing childhood disorders (Duncan et al., 2009). A study exploring a similar objective was conducted by Willner et al. (2015). Within a socioeconomically disadvantaged population, the authors found an association between higher P300 amplitudes and adaptive learning-related behaviors (e.g. sit still, listen without being distracted, less frustration) in kindergarten. These behaviors, in turn, were related to an increase in academic performance from kindergarten to 1st grade (Willner et al., 2015). Identifying individuals at risk could facilitate early interventions, as experiencing academic underachievement can increase the risk of developing externalizing behaviors and substance dependence (Crum et al., 1998; Hinshaw, 1992). Attention-Deficit/Hyperactivity Disorder (ADHD) is a disorder characterized by impaired attention, hyperactivity, or impulsivity (J. M. Swanson et al., 1998) and is one of the most common disorders in children and adolescents with a prevalence estimated at 5.3% (Polanczyk & Rohde, 2007). ADHD has often been shown to have an impact on academic success (Frazier et al., 2007), with deficits in working memory (WM) being one possible reason (Kim et al., 2014; Martinussen et al., 2005). Meta-analyses investigating P300 differences in ADHD revealed lower amplitudes and longer latencies compared to a control group and further discuss P300 as a potential biomarker of ADHD, although substantial heterogeneity in effect sizes and other factors limit this conclusion (Kaiser et al., 2020;

Szuromi et al., 2011). These meta-analyses only examined P300 in GoNoGo paradigms, although it would be valuable to assess P300 in learning and memory tasks. Especially, since ADHD has a high cooccurrence with specific learning disorders (SLD) (Boada et al., 2012). SLD is used to describe individuals with significant difficulties in learning one or more academic domains (e.g., reading, writing, mathematics) (Pham & Riviere, 2015). The diagnosis of SLD has been described as controversial, and a traditional criterion is a discrepancy between performance on standardized achievement tests and IQ. This criterion has been criticized for its low validity and reliability as diagnostic criteria (Fletcher et al., 2005). Response-to-intervention (RTI) is an alternative approach to identify individuals at risk marked by inadequate response to instructions or interventions (Vaughn & Fuchs, 2003). This approach promotes a move from a deficit model to a risk model, as in early detection of individuals at risk rather than waiting for academic deficits to occur, and also focuses on ongoing progress-monitoring as well as detection in early stages like kindergarten (Vaughn & Fuchs, 2003). P300 could play a potential role in such identification and guide interventions as found in Wiecki et al. (2015). Still, since Wiecki et al. (2015) used a paradigm which did not allow for a direct connection between P300 and learning rather this link was mediated through learning-related behavior. It may be important to use paradigms in which individuals are instructed to learn using repeated presentations to create a direct link between P300 and learning or memory formation. As mentioned, such a paradigm was used by Steinemann et al. (2016) which allowed P300 to directly be connected to successful learning. Strzelczyk et al. (2022) extended these findings in a population of healthy older adults. It is still unclear whether this relationship holds in a population of children and adolescents with psychiatric and learning disorders. As Vaughn and Fuchs (2003) mentions, more research is needed to provide validated methods for testing responsiveness to instructions. P300 amplitudes are associated with encoding processes of successful learning (Polich, 2007) and have been shown to help improve attention and behavioral performance with a P300 based

neurofeedback training (Arvaneh et al., 2019). Therefore, in the future P300 could potentially be helpful as tools for monitoring cognitive processes and responsiveness to intervention or instructions even in communication-impaired individuals (Connolly et al., 2000). Multiple studies have suggested the utility of P300 in academic settings to inform early intervention or disentangle individual differences in cognitive performance from impacting factors such as test anxiety (Privitera, 2022; von der Embse et al., 2018; Willner et al., 2015). Still, none of the above mentioned studies have used paradigms where explicit learning over multiple repetitions was required (see., Privitera, 2022). More research with paradigms assessing learning is needed while considering variables that have been shown to influence P300 amplitude (Cui et al., 2017; Duncan et al., 2009; Polich & Kok, 1995; Riggins & Scott, 2020; Sur & Sinha, 2009; van Dinteren et al., 2014). For example, P300 has often been found to be altered in psychiatric disorders (Polich, 2004; Sur & Sinha, 2009) but has been described as unspecific and not relating to a single diagnosis (Duncan et al., 2009). Moreover, P300 amplitude have been shown to have high heterogeneity within a diagnosis (e.g., Cui et al., 2017; Kaiser et al., 2020; Martinussen et al., 2005). This is not surprising since heterogeneity within diagnosis is a known phenomenon (Astle et al., 2022; Kaiser et al., 2020).

1.4 Research Gap

As previously mentioned, P300 has the potential to inform about the cognitive processes that underlie successful learning without relying on behavioral reports and in near real-time fashion, which could be helpful in studying children and adolescents who struggle with learning (Riggins & Scott, 2020; Steinemann et al., 2016; Strzelczyk et al., 2022; Tinga et al., 2019). In healthy young and old adults, a direct link between P300 amplitudes and successful learning has already been established (e.g., Steinemann et al., 2016; Strzelczyk et al., 2022). Visual sequence learning paradigm (VSLP) in which participants learn a sequence of highly simplified stimuli (i.e., flashing circles) with little semantic content has been a main contributing factor in establishing this link. However, it

is unclear whether the same link exists in a population with psychiatric and learning disorders. Therefore, this thesis aims to close this gap by investigating whether P300 as a neural correlate of successful learning can be found in individuals with psychiatric and learning disorders using the visual learning paradigm in combination with event-related potentials. This thesis is closely aligned with a study by Strzelczyk et al. (2022) that extended the findings of Steinemann et al. (2016) to a sample of healthy older adults. With P300 amplitude being sensible to stimulus-bound subjective degree of surprise (or expectation) bound to stimulus (Donchin, 1981; Duncan-Johnson & Donchin, 1977; Mars et al., 2008; Sutton et al., 1965), this component provides an opportunity of linking P300 to gradually increasing knowledge of the sequence in VSLP over the course of successful learning (Steinemann et al., 2016). Given that participants were instructed to learn a sequence of spatial locations (see Visual Sequence Learning Paradigm) it was expected that the knowledge of the presented sequence (i.e., Knowledge Index: see equation 1) increases over the course of learning:

H1) Behavioral performance increases over the course of learning.

During childhood and adolescents, cognitive abilities increase with age; therefore, it was predicted that behavioral performance increases with increasing age during childhood and adolescents:

H2) Behavioral performance increases with age during childhood and adolescents.

On a neurophysiological level, the amplitude of P300 was expected to decrease monotonically over the course of learning, from stimuli being unknown, to newly learned and fully committed to memory (i.e., known). Similar patterns in P300 decrease were expected in all age ranges, but amplitudes and latency differ as a function of age. Hence, the fourth hypothesis was that in children and adolescents P300 amplitudes latency decreases with age:

H3) P300 amplitude decreases monotonically over the course of learning;

H4) P300 amplitude decreases with increasing age during childhood and adolescents.

Finally, it was investigated whether neurophysiological measures could predict successful learning within individuals. Therefore, it was hypothesized that increased behavioral performance is associated with decreased P300 amplitudes:

H5) Increase in behavioral performance is associated with a decrease of P300 amplitude.

In a next step, it was explored whether P300 amplitude decreases over the course of learning are associated with transdiagnostic constructs. This investigation was exploratory in nature, without concrete hypotheses.

2   Methods

2.1   Data Acquisition

Data was acquired via the 1000 Functional Connectomes Project and its International Neuroimaging Data-sharing Initiative (FCP/INDI) (see Data Citation) (Alexander et al., 2017). Data processing and analysis was performed with MATLAB R2021b (The MathWorks, Inc., Natick, Massachussets, United States) and RStudio 4.2.2 (R Core Team).

2.2   Participants

The sample used in this thesis was obtained from the Child Mind Institute Healthy Brain Network (HBN) Biobank. HBN recruits with a community-referred strategy in New York City with the primary goal of generating a data set that represents the wide range of heterogeneity and impairment that occurs in developmental psychopathology. Upon completion of the study, participants receive comprehensive diagnostic evaluation reports, referral information, up to three in-person feedback sessions, and modest financial compensation (Alexander et al., 2017). Of the 2'091 participants for whom relevant data used in this thesis were available, 497 were excluded (see figure 1). Thus, the final sample consisted of 1'594 participants with an age range of 5.04 to 21.90 years of age. The mean age was 10.77 with a standard deviation of 3.44 (see table 1 for demographics). Many of the participants received or already had at least one diagnosis (n = 1'437) while 157 participants had no diagnosis (see table 2). The most frequent diagnosis can be found in table 2.

```
┌─────────────────────────────────────┐        ┌──────────────────┐
│ Healthy Brain Network Biobank (HBN)  │        │    Exclusions    │
│                                      │        └──────────────────┘
│           N = 2′999                  │
└─────────────────────────────────────┘
```

Healthy Brain Network Biobank (HBN)

$N = 2'999$

Exclusions

Full VSLP Data

$N = 2'091$

Bad EEG Rating ($N = 224$)

$<50\%$ artifact free trials ($N = 177$)

No demographics ($N = 91$)

>10 years old with seq. length 7 ($N = 5$)

Total Sample Size:

$N = 1'594$

$Trials = 58'060$
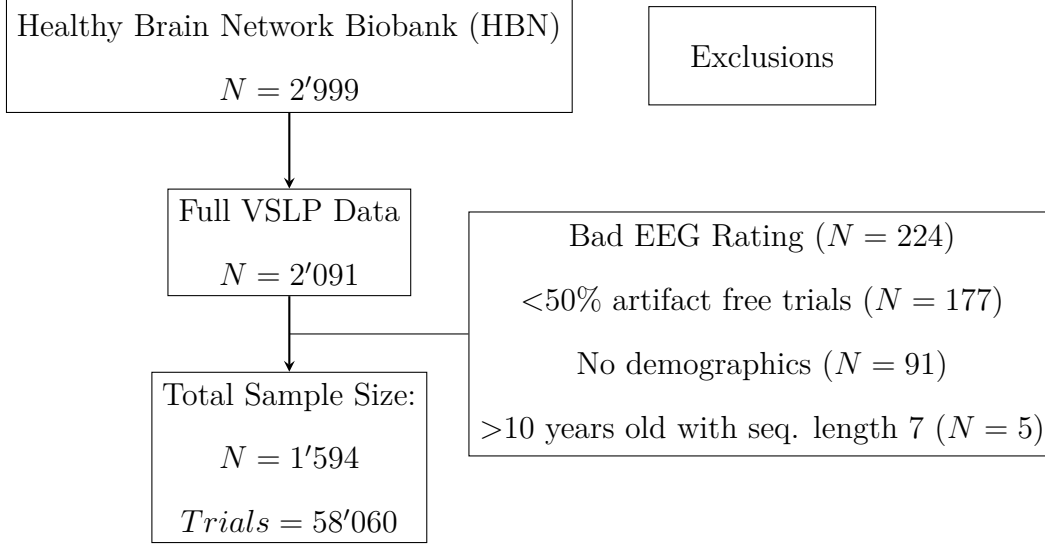
Figure 1

Sample Size and Exclusion Criteria.

2.3 Visual Sequence Learning Paradigm

The visual sequence learning paradigm (VSLP) (Moisello et al., 2013) was one of many tasks applied in the HBN study (see Alexander et al., 2017). During VSLP neurophysiological measures were simultaneously recored, which together with behavioral task performance provided the main data for this thesis. In the VSLP participants had to learn a fixed sequence of visual stimulus positions presented one by one. The possible locations were constantly marked by circular outlines with the target stimuli (i.e., to be learned locations) being filled white circles. These possible locations lay equidistant around a ring of fixed eccentricity. The target stimuli appeared for 200 ms each with 1300 ms between them (i.e. inter-stimulus interval). During the presentation, participants were asked to fixate on a central fixation point. To account for floor effects, the length of the sequence was shorter, and the number of possible locations was less for most participants under nine years. Specifically, this meant a sequence of 8 target stimuli with 6 possible locations. For older participants, the sequence consisted of 10 target stimuli and 8 possible locations. In total, the same sequence was presented five times. After each

sequence presentation, participants were asked to recall the locations in a test phase by clicking on them with a computer mouse. The test phase had no time restriction and no feedback was given, whether the correct locations were recalled (Langer et al., 2017). Multiple studies with an interest in memory formation and learning processes have applied VSLP (e.g., Steinemann et al., 2016; Strzelczyk et al., 2022). Repeated presentation of the same sequence enables the ability to observe learning progress during the recall or test phase. Since VSLP uses highly simplified stimuli, the factors that could influence learning success are minimized (Steinemann et al., 2016).

2.4 Behavioral Data

Individual target stimuli during each sequence presentation were categorized based on the response of the participants in the test phase. Newly learned category (NL) was assigned when the location of the target stimulus was correctly recalled for the first time. Known (K) categorized target stimuli that were correctly recalled at least two consecutive sequence repetitions. Unknown category (UN) classified target stimuli that were not recalled correctly. Finally, the forgotten category (F) was assigned to target stimuli that were NL or K in previous sequence repetitions, and an incorrect recall was made in the current repetition. If a forgotten target stimulus was correctly recalled in a subsequent presentation, the category would change from F to NL. If in a subsequent presentation, a previous forgotten target stimulus is recalled incorrectly again, it becomes unknown. On average, the forgotten category is assigned 9.7% per participant (see table 3). Furthermore, for every sequence presentation, a knowledge index (KI) was calculated to track the progress of cumulative knowledge of the sequence. The KI was calculated for each sequence presentation of a given participant and was defined as:

$$\text{Knowledge Index}_{(P,sr)} = \frac{nK_{(P,sr)} + nNL_{(P,sr)}}{nT_{(P,sr)}} \tag{1}$$

where P is participant, sr is sequence repetition, K is known category, NL is newly learned and T is the total artifact free trials (see 2.6). The KI differed from the one used

in Steinemann et al. (2016) since the paradigm was slightly changed (see Supplementary Material). Another behavior performance measure used was the learning index (LI). LI is the proportion of newly learned stimuli in a sequence repetition for a given participant and was defined as a ratio of NL stimuli to the total number of stimuli (again only considering artifact free trials):

$$\text{Learning Index}_{(P,sr)} = \frac{nNL_{(P,sr)}}{nT_{(P,sr)}} \tag{2}$$

2.5 EEG Data Acquisition

The HBN study, described in Alexander et al. (2017), used a 128-channel EEG geodesic hydrocel system by Electrical Geodesic Inc. (EGI) with a sampling rate of 500 Hz to record the EEG data. Cz (vertex of the head) was the recording reference and the impedance of each electrode was kept below 40 kOhm.

2.6 EEG Data Preprocessing

Since the EEG data was provided in raw format by HBN, the data had to be prepared for further analysis. Data preprocessing was performed using Automagic 2.6, a MATLAB based toolbox for reliable and objective preprocessing in an automatic way (Pedroni et al., 2019). Automagic used the PREP pipeline for large-scale EEG analysis (Bigdely-Shamlo et al., 2015) to detect noisy or outlier channels. Four criteria are used for detection: extreme amplitudes (deviation exceeding a z-score of 5.00), low correlation with any other channel (correlational threshold: 0.4), lack of predictability by other channels, and unusual high frequency noise (deviation of signal-to-noise ratio exceeding a z-score of 5.00). More detailed descriptions of these processes can be found in Pedroni et al. (2019) and Bigdely-Shamlo et al. (2015). Next, automagic handled sweating and power line artifacts by performing a high-pass filter (0.05 Hz) using *pop_eegfiltnew* from EEGLAB (Widmann & Schröger, 2012) and a zapline filter (50.00 Hz) using *nt_zapline* (de Cheveigné, 2020, NoiseTools,). Further artifacts caused by, for example, muscle movement were identified and removed using Independent component Label (ICLabel)

(Pion-Tonachini et al., 2019).

Following automagic preprocessing procedures, a 45 Hz low-pass filter was applied using *pop_eegfiltnew* because only ERP signals are of interest and are mainly below 30 Hz (Luck, 2014). Next, EEG data was re-referenced to the average of all electrode sites with the *pop_reref* function of EEGLAB. Using the average as a reference instead of a single electrode site minimizes artifacts linked to that specific electrode site (Dien, 1998). In a second step, the EEG signal was segmented into trials using triggers that marked the appearance of target stimuli (i.e., stimulus onset). The duration of segments was 900 ms (100 pre stimulus onset to 800 ms post stimulus onset). At this point two data sets were created, one for visualization purpose and a second data set used to extract P300 amplitudes for statistical analysis. For the first data set, the segmented EEG signals were corrected for baseline using the *pop_rmbase* function and a time window of 100 ms pre stimulus onset. The segments of each electrode were compared to an amplitude threshold of -+ 90 microvolts to mark further artifacts. If at any point during a segment this threshold was exceeded, then the trial is marked as artifact and was removed from both data sets. Second, participants who had more than 50% of their total trials marked as artifacts were excluded from both data sets (see figure 1). The cutoff point of 50% was chosen because having a higher cutoff point, for example 75% would have led to over 1'000 subjects being excluded. For the second data set (used for P300 amplitude extraction), the EEG segments of a cluster of centro-parietal electrodes were averaged together, while discarding EEG segments from other electrodes. This cluster consisted of the following electrodes: E54, E55, E61, E62, E78, E79 located at the centro-parietal area. Additionally, in this data set no baseline correction was applied. Rather, baseline was included in linear mixed effects models, following the approach of allowing the amount of baseline correction to be determined by statistical model as suggested by Alday (2019).

2.7 Electrode Sites Selection

The same electrode cluster was selected as in a previous study that used the same paradigm in young and old adults (Strzelczyk et al., 2022). In a subsequent step, the grand average scalp topographies of all trials between 150 and 550 ms with a 50 ms step were plotted for different age groups (see figure 8 in Supplementary Material). By visual inspection of these plots, the electrode cluster was considered appropriate in capturing the P300 amplitude for all age groups. Averaging the signal from an electrode cluster where ERP effects are expected has been discussed as a temporary solution to possible differences in the scalp distribution of ERP in young children (Brooker et al., 2020).

2.8 P300 Extraction

To extract the P300 amplitude, participants were grouped by age into six bins of two years with the exception of participants 17 years and older, which formed a seventh group with a bin size of 5 years. Reasoning behind the seventh group was to decrease the signal-to-noise ratio, as there were fewer participants in this age range. A time window used to calculate the P300 amplitude was defined by averaging all trials in each age group, resulting in seven grand average ERP waveforms. Using all trials to select the time window (i.e., collapsed localizers) minimizes biases, since differences in conditions are not driving the selection of the time window (Luck, 2014; Luck & Gaspelin, 2017). The P300 peak was defined as the most positive amplitude in the time range typical for P300 (i.e., 300 - 500 ms post stimuli onset). Using the latency of the defined peak, a time window was calculated by adding and subtracting 50 ms. In a next step, these time windows were used to calculate the P300 amplitude in each single trial by averaging the voltage in the age appropriate time window (i.e., mean amplitude). Using mean amplitude helps against peak distortion or bias caused by signal noise (Luck, 2014).

2.9   Conceptual Replication

The first part of this thesis aimed at conceptually replicating findings from Steinemann et al. (2016) and Strzelczyk et al. (2022) . A conceptual replication is described as an attempt to test a fundamental idea or hypothesis behind the original study, but can vary in aspects such as experimental operationalization and design, independent and dependent variables and sample population (Crandall & Sherman, 2016). One main difference from the studies mentioned above was the different sample population. Whereas Steinemann et al. (2016) investigated neuronal correlates of learning process and memory formation in a sample of healthy young adults and Strzelczyk et al. (2022) compared a sample of healthy young adults with healthy old adults, this thesis focused on a sample of children and adolescents with various psychiatric disorders.

2.9.1   Statistical Analysis

The main analyses were performed with linear Mixed-Effects Models (LME). LME is a regression-based statistical method that is useful in analyzing hierarchically grouped data (Bickel, 2007). Event-related potentials are naturally grouped within individuals; therefore, it has often been advocated to apply LME in the analysis of ERP data (Volpert-Esmond et al., 2021; Vossen et al., 2011). To estimate the models, lme4 package (Bates et al., 2015) in R was used with LmerTest (Kuznetsova et al., 2017) providing p values and the step function used for model selection. The step function follows a data-driven approach of iteratively deleting nonsignificant effects from a specified model. A step-down approach was used, in which the starting point is a full model that is compared to a reduced model using a likelihood-ratio test in an iterative process until the model with the best fit is found. This algorithmic process first determines the structure of random effects and proceeds with the structure of fixed effects (Kuznetsova et al., 2017). Specifying the model as input for the step function was done by keeping the fixed effects structure as general as possible (i.e., allowing most fixed effects to interact with each other) and keeping the random part of the model parsimonious since having many

random effects led to converging problems during model selection. The model formulas are expressed using the Wilkinson notation (Wilkinson & Rogers, 1973).

2.9.2 Behavioral Analysis

First, learning performance was analyzed descriptively by plotting behavioral measures (i.e., KI: 1 and LI: 2) over repetitions. More specifically, the group averages of KI and LI with their standard errors were plotted across repetitions to display learning effects and potential age effects. Next, two LME (3) were specified with KI and LI as dependent variables (continuous: 0-1), repetition (ordinal: 0-4), age (continuous: grand mean centered) and sequence length (categorical: 7 and 10, 10 as reference category) as independent variables. The random effects structure consisted of a random intercept for every participant and a random slope for repetition including their correlation. To help with interpretation of interaction effects as well as having a meaningful intercept, the variables repetition and age were linearly transformed: repetition was subtracted by one, age was mean centered. Therefore, the intercept represented the first repetition and age equal to zero represents an average-aged person. This becomes important for the interpretation of the main effects of explanatory variables that have a significant interaction with age or repetition as these main effects should be interpreted together with the interaction as a system (Hox et al., 2017). Overall, linear transformation does not change the model fit and residuals, and so models with or without linear transformation are equivalent only differing in what part of the model is emphasized (Hox et al., 2017). On the other hand, the random variance of the intercept can be influenced by transformation when a random slope is included in the model (Hox et al., 2017).

$$KI \text{ or } LI \sim Repetition * Age * SequenceLength + (1 + Repetition | Participant) \quad (3)$$

2.9.3   Neurophysiological Analysis

2.9.3.1   P300 Amplitude Over Sequence Repetitions. In the context of VSLP, the assumption was that learning a sequence increases its expectancy. As it is suggested that the P300 amplitude decreases as a function of expectancy, changes of P300 amplitude over repetition were analyzed. More specifically, it was hypothesized that over the course of leraning, P300 amplitude decreases. This hypothesis was approached through plots and LME. In the first step, the average P300 amplitude per sequence repetition was calculated for each participant (i.e., mP300). Second, to visualize possible age effects, mP300 was averaged within each age group and plotted across repetitions with error bars reflecting standard errors. Next, a linear mixed effects model was specified (4) as input for the step function. The specification was as follows: mP300 as dependent variable (continuous: 0-1), repetition (ordinal: 0-4), gender (categorical: male and female, with male as reference category), age (continuous: grand mean centered), sequence length (categorical: 7 and 10, 10 as reference category) and mBase (continuous) as independent variables. mBase was computed by averaging the baseline values of the trials in a given repetition. For every participant, a random intercept was included.

$$mP300 \sim Repetition * Age * Gender * SequenceLength + mBase + (1|Participant) \quad (4)$$

2.9.3.2   Predicting Successful Learning. To further test the possible connection between P300 amplitude and knowledge index, a LME was specified (5) to investigate whether the differences in KI are explained through differences in mean P300 amplitude. The model was specified with KI as dependent variable (continuous: 0-1), mP300 (continuous), age (continuous: grand mean centered), gender (categorical: male and female, with male as reference category) and mBase (continuous) as independent variables. The model included an intercept as a random effect.

$$KI \sim mP300 * Age * Gender + mBase + (1|Participant) \quad (5)$$

2.9.3.3   P300 Amplitude Across Learning Categories. In a final step, differences in P300 amplitude across learning categories (i.e., Unknown, Newly Learned, Known, Forgotten) were analyzed to directly tie the amplitude of P300 to the behavioral response at a single trial level. Since learning categories differentiate between first accurate recall of a spatial location (i.e., Newly Learned) and subsequent accurate recall (i.e., Known), these categories provide additional information which can be used to refer to how well a spatial location is learned or to what degree it is expected. Connecting P300 amplitude to learning categories on a single trial level is important, as only analyzing the P300 amplitude as an average of a sequence repetition can be influenced by habituation (Ravden & Polich, 1998). Topographical plots were created for the learning categories Unknown, Newly Learned and Known for four randomly selected age groups. For visualization, baseline corrected data set was used and all trials of each learning category were averaged together. Potential age effects were checked by plotting the average P300 amplitude and ERP's of learning categories across age groups the same age groups. In a next step, a LME model was specified (6). The model consisted of P300 as dependent variable, learning categories (categorical: UN,NL,K,F with NL as reference category), age, gender and base as independent variables, and a random intercept for each participant.

$$P300 \sim LearningCategories * Age * Gender + base + (1|Participant) \qquad (6)$$

Since LME only compares each learning category to the reference category, contrasts between the learning categories were computed using the emmeans package in R Studio while correcting for multiple comparisons using Bonferroni correction (Lenth et al., 2019).

2.10   Exploratory Analysis

The second part of the thesis explored whether successful learning in VSLP and the degree of decrease in the amplitude of P300 over repetitions are associated with transdiagnostic constructs related to learning.

2.10.1   Knowledge Index and P300 Decrease

For every participant, an average of their repetition based knowledge index was calculated. Second, a measurement reflecting the degree of a participants increase in expectancy of sequence was created. For this, a linear regression with P300 amplitude as dependent variable and repetition as independent variable was specified for each participant. The resulting slope coefficient of repetition was used in the exploratory part.

2.10.2   Questionnaires and Tests

Since this part is exploratory in nature, no theoretical framework guided the inclusion of questionnaires. Instead, questionnaires that were filled out by most participants were considered and subscales of these questionnaires that have been found to be relevant for either learning, working memory test performance, or P300 amplitude were included. If available, standardized scores were used; otherwise, raw scores were included.

2.10.3   SWAN

The Strengths and Weaknesses of ADHD symptoms and Normal behavior scale (SWAN J. Swanson et al., 2001)) provided scores for attention problems. SWAN has a bidirectional questionnaire design assessing problems but also strength of symptoms relevant in ADHD with a 7-point scale with endpoints: far below average and far above average (Alexander et al., 2020). This questionnaire has been applied in multiple areas such as genetic studies, pharmacotherapeutics, neuropsychology, and has been shown to have good psychometric properties (for review see Brites et al., 2015). Of interest in this thesis was the subscale inattention, as impaired attention has been reported as one of the most ubiquitous clinical symptoms (Mirsky & Duncan, 2001) and plays an important role in maintaining relevant information (Chun, 2011). Therefore, the scores of the inattention subscale were included in the analysis (i.e., raw scores).

2.10.4 WISC-V

Scores from the fifth edition Wechsler intelligence scale for children (WISC-V Wechsler, 2014) were included in the exploratory analysis. More specifically, two primary index scores (i.e., processing speed index and working memory index). For psychometric properties and a more detailed review of this test, see Wechsler (2014) and Meyer and Reynolds (2018). Working memory has been stated as an important factor in learning and memory (Cowan, 2014) and is prominently associated with P300 in the theoretical framework of context updating (Donchin & Coles, 1988; Lenartowicz et al., 2010). Therefore, working memory index was included in the exploratory analysis. Processing speed index as second variable was included since it has been shown to be affected in multiple psychiatric disorders and is linked to academic performance (Mayes & Calhoun, 2007).

2.10.5 Conners 3

Subscale Learning Problems of the Conners 3 self report (Conners, 2008) was used to explore whether behavioral or neurophysiological measures of learning are associated with learning problems. The psychometric properties of Conners 3 have been reported to be good (Conners, 2008).

2.10.6 Child Behavior Checklist

Five subscales were included from the parent report version of the Child Behavior Checklist questionnaire (CBCL) (Achenbach, 1991). More specifically, standard scores of the subscales Withdrawn/Depressed, Rule Breaking Behavior, Attention Problems, Anxious/Depressed, and Aggressive Behavior were included. Following an approach done in previous studies, the latter three subscales (i.e., Attention Problems, Anxious/Depressed, Aggressive Behavior) were summed together to arrive at an index for deficits in emotional regulation: CBCL-AAA profile (Biederman et al., 2012; Donfrancesco et al., 2015). Emotional dysregulation has been stated as a main

transdiagnostic construct affecting multiple disorders (Sloan et al., 2017) and was therefore included to explore its effect on knowledge index and P300 amplitude decrease.

2.10.7 Social Responsiveness Scale

The subscales Social Motivation Problems and Social Cognitions of the Social Responsiveness Scale (Wigham et al., 2012) were included in the exploratory analysis. The validity and reliability of this questionnaire have been reported to be good (Wigham et al., 2012). Including these variables followed the notion that, especially in school settings, social aspects can influence learning and performance (Wentzel & Wigfield, 1998).

2.10.8 Network Analysis

Altered P300 has been associated with many psychiatric disorders (Polich, 2004; Sur & Sinha, 2009) and seems to reflect a general vulnerability (Sur & Sinha, 2009) rather than related to a particular disorder (Duncan et al., 2009). Often the findings of altered P300 are mixed when investigating a single disorder (Salgari et al., 2021). Additionally, P300 has been shown to be sensitive to cognitive dysfunction (Polich & Kok, 1995; Potter & Barrett, 1999) which itself has been discussed as a transdiagnostic phenomenon in psychopathology moderated by factors such as motivation or emotion (Abramovitch et al., 2021). This suggests that it may be of importance to study P300 in a transdiagnostic manner. Recently, network analysis has been shown to be a promising tool in studying transdiagnostic phenomena (e.g., Astle et al., 2022; Borsboom, 2022; Chavez-Baldini et al., 2021).

To explore the question whether the selected transdiagnostic variables provide information on behavioral (i.e., Knowledge Index) and neurophysiological (i.e., P300 decrease) measures obtained from the visual sequence learning paradigm, a network model was estimated. As described in Isvoranu et al. (2022), network models are statistical structures with nodes (i.e., variables) and connecting edges (i.e., estimated statistical relations). Most often these statistical relations are estimated with partial correlations in

an undirected manner using Pairwise Markov Random Fields models (PMRFs). In PMRFs the absence of a connection (i.e., edge) between two nodes indicates that these variables are conditional independent after controlling for all other variables (i.e., parietal correlation). The presence of an edge reflects a conditional dependency, with the partial correlation coefficient being the strength of this relationship. A subclass of PMRF models used in this analysis is the Gaussian graphical model (GGM, Lauritzen, 1996) which is used when the variables in the network (i.e., nodes) are continuous. With PMRFs, the predictive relationships between variables in a network model are visualized and would be comparable to computing many exploratory regression analyses (Isvoranu et al., 2022).

2.10.8.1 Procedure. In a first step, selected variables were plotted to inspect their distribution (see figure 10 in Supplementary Material). Three highly skewed variables (i.e., subscales from CBCL) were transformed with a nonparanormal transformation (Liu et al., 2009) using the function *huge.npn* from the huge package (Zhao et al., 2012). This transformation approach uses the cumulative distribution of the data and maps each point to a z-value of a standard normal distribution via the cumulative distribution of the standard normal distribution (Liu et al., 2009). The estimate of the model was made using the *estimateNetwork* function of the bootnet package (Epskamp & Fried, 2015). There are different estimation (and model selection) methods that come with the bootnet package, and the *ggmModSelect* method was chosen for this analysis. Different estimation methods can produce different results. At this point, however, no clear method is to be preferred and one has to consider the research question and sample size while choosing the estimation method (Isvoranu & Epskamp, 2021). Although it has been observed through simulation studies that with a medium sample size (e.g., N = 1000), *ggmModSelect* produces estimates with good specificity (i.e., proportion of true absent edges selected in the estimated network) and sensitivity (i.e., proportion of true edges selected in the estimated network) in detecting edges that connect different domains in the network structure (i.e., bridge edges). The algorithmic

approach of *ggmModSelect* is described in Isvoranu et al. (2022). First, to limit the search space for the model, *ggmModSelect* estimates various partial correlation matrices with different regularization parameters [3] using glasso (J. Friedman et al., 2008). Next, the resulting models are re-estimated without using regularization and the model structure with the lowest Bayesian information criterion (BIC) is selected. Edges in the selected model are included or removed in an iterative process to once again minizime BIC (Isvoranu et al., 2022). The final model consisted of a partial correlation matrix that was then visualized using the qgraph package (Epskamp et al., 2012). Lastly, the accuracy of edges were estimated using bootstrapped confidence intervals. This was accomplished with non-parametric bootstrapping (bootnet package) in which observations are resampled with replacement and results in 1000 new data sets. When using unregularized estimates of edges (e.g., ggmModSelect), the bootstrapped confidence intervals are reported to be valid as null hypothesis tests (Epskamp et al., 2018).

---

[3]optimizing likelihood function while penalizing for model complexity to a varying degree (Isvoranu et al., 2022)

3   Results

3.1   Behavioral Results

3.1.1   Demographics

The basic demographics of the sample are displayed in table 1.

Table 1

Basic Demographics

	N	Age				Gender	
		Mean	**SD**	**Min**	**Max**	**M/F**	**%**
SL7	541	7.39	0.99	5.04	9.93	390/151	72/28
SL10	1'053	12.50	2.92	8.15	21.90	762/291	72/28
Total	1'594	10.77	3.44	5.04	21.90	1'152/442	72/28

Note. SL7 = Sequence Length 7. SL10 = Sequence Length 10.

3.1.2   Diagnoses

To broadly characterize the sample, table 2 shows the number of participants who received no diagnosis, a single diagnosis and those participants with multiple diagnoses. Furthermore, the most frequent diagnoses are displayed.

Table 2

Most Frequent Diagnoses

General	N
No Diagnosis	158
One Diagnosis	644
Comorbidity	793
Most Frequent Diagnoses	
Attention-Deficit/Hyperactivity Disorder	902
Anxiety Disorder	517
Specific Learning Disorder	331
Autism Spectrum Disorder	251

3.1.3 Visual Sequence Learning Task Performance

As depicted in table 3, participants that learned the shorter sequence had on average 27.42 (4.61) artificial free trials, while participants assigned to the longer sequence had on average 41.05 (6.95) artificial free trials. Across sequence repetitions, participants with SL7 had on average 11.58 (6.02) trials categorized as unknown, 6.04 (2.35) as newly learned, 7.06 (6.03) as known, and 2.74 (1.85) as forgotten. A similar pattern was found in participants with SL10 where, on average, the highest proportion of trials was categorized as unknown followed by known and newly learned.
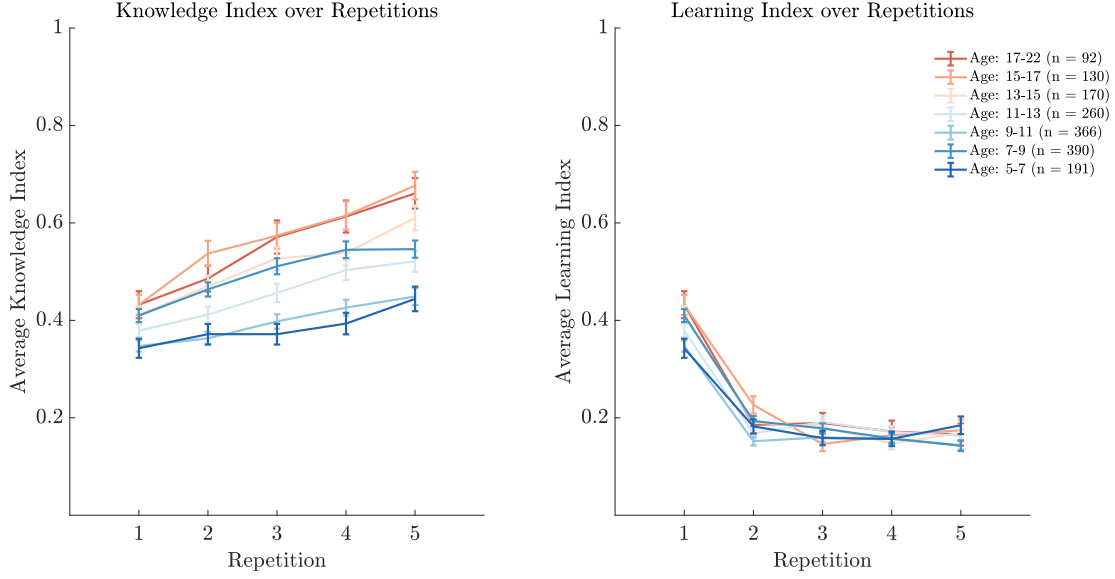
Table 3

Average Trials and Learning Categories per Participant

	Trials	Learning Categories			
		UN	NL	K	F
SL7	27.42 (4.61)	11.58 (6.02)	6.04 (2.35)	7.06 (6.03)	2.74 (1.85)
		43.03 %	22.00 %	25.01 %	9.96 %
SL10	41.05 (6.95)	18.08 (8.90)	8.80 (3.28)	10.26 (8.54)	3.93 (2.60)
		44.69 %	21.40 %	24.30 %	9.61 %
Total	36.42 (8.99)	15.87 (8.61)	7.86 (3.26)	9.17 (7.92)	3.52 (2.43)
		44.13 %	21.61 %	24.54 %	9.73 %

Note. UN = Unknown, NL = Newly Learned, K = Known, F = Forgotten.
Average Trials and Learning Categories per participant with standard deviation.

3.1.3.1   Knowledge Index. It was hypothesized that behavioral performance increases over the course of learning and that this performance increases with age during childhood and adolescents. To test these hypotheses, first behavioral measures KI and LI were plotted over repetitions for each age group. Second, linear mixed-effects models were fitted for inferential testing. Figure 3 shows the average knowledge and learning index across repetitions for each age group (error bars are standard errors). Plotting KI over repetitions suggested that, on average, behavioral performance (i.e., KI) increased over repetitions and that older participants had higher behavioral performance (i.e., KI). The average learning index was highest in the first repetition, with no obvious age differences visible.

Figure 2

Knowledge Index and Learning Index over Repetitions

Providing inferential evidence was provided with a linear mixed effects model (results in table 4). The best-fit model of knowledge index (7) had a fixed effects structure consisting of repetition (REP), age, sequence length (SL) and three interactions. Random effects structure included random intercept for participant, a random slope for repetition and their correlation:

$$KI \sim REP + Age + SL + REP \, x \, SL + REP \, x \, Age + Age \, x \, SL + (1 + REP | Participant) \quad (7)$$

As mentioned in the section Behavioral Analysis, variables age and repetition were linearly transformed to help with interpretation of interaction (Hox et al., 2017). Therefore, in the resulting model, the intercept reflected the average KI for average aged participants with a sequence length of 10 in the first repetition (i.e., expected KI when all explanatory variables are zero). Because a random slope (i.e., repetition) was included, the estimated variance of the intercepts was influenced by the linear transformation and thus reflected the variance of intercepts for average aged participants with SL10 in first repetition (i.e., conditioned on all explanatory variables being equal to zero) (Hox et al.,

2017).

A significant main effect of repetition ($\beta = 0.033$, CI = $[0.03 ; 0.04]$, p <0.001) indicated that for participants of average age with a sequence length (SL) of 10, KI increased with repeated presentation of sequence. This increase was moderate by age ($\beta = 0.005$, CI = $[0.003;0.006]$, p <0.001) and SL ($\beta = 0.016$, CI = $[0.005;0.028]$, p = 0.006). Therefore, the model suggested that the effect of repetition on KI increased with age. Moderation of sequence length on the effect of repetition suggested that, for participants of average age, the effect of repetition on KI is greater in SL7 compared to SL10. This moderation should be interpreted with caution since no actual data is observed at this described age point (i.e., there are no participants of average age who had an SL of 7). Second, age significantly explained the differences in KI in the first repetition suggested by a main effect of age ($\beta = 0.015$, CI = $[0.01; 0.02]$, p<0.001). This indicated that KI in the first repetition was higher in older participants. The effect of age on KI in the first repetition was stronger in participants with sequence length 7 as suggested by a significant interaction of age and sequence length 7 ($\beta = 0.016$, CI = $[0.005; 0.028]$, p = 0.006). Random components of the model revealed a large variance in effect of repetition. Exploring the random slopes estimated by lmer, revealed that 868 participants had a negative slope.

Taken together, the model suggests that participants increase their knowledge of the sequence over repetitions. Additionally, age effects are revealed by the model with older participants having an increased effect of repetition on KI. A second age effect suggested by the model is that KI in the first repetition is on average higher for older participants.

Table 4

Model Output: Increase in Knowledge Index over Repetitions

Variable	Beta	SE	CI	t-value	p-value
Intercept	0.363	0.007	0.35 - 0.38	50.08	<0.001
Rep	0.033	0.003	0.03 - 0.04	12.83	<0.001
Age	0.015	0.002	0.01 - 0.02	6.67	<0.001
SL (7)	0.291	0.037	0.22 - 0.36	7.90	<0.001
Rep x Age	0.005	0.001	0.003 - 0.006	5.75	<0.001
Rep x SL (7)	0.016	0.006	0.005 - 0.028	2.76	0.006
Age x SL (7)	0.048	0.009	0.03 - 0.07	5.23	<0.001
Variance Components	**SD**	**Goodness of fit**			
Participant	0.14	Log Likelihood	-461.09		
Rep	0.04				
Cor(Participant, Rep)	0.53				
Residual	0.21				

Note. Rep = Repetition, SL7 = Sequence Length 7, SL10 = Sequence Length 10. Intercept represents average aged participants with SL10 at first repetition.

3.1.3.2   Learning Index. Learning Index was a second behavioral index of learning and is mentioned here for the sake of completeness even though it was not a focus in this thesis. For the learning index, no random effect structure could be estimated. Therefore, the best-fit model was a linear multiple regression model (8) and results are displayed in table 5).

$$LI \sim REP + Age + SL + REP \, x \, Age + REP \, x \, SL + Age \, x \, SL + REP \, x \, Age \, x \, SL \quad (8)$$

A significant main effect of repetition on learning index was found (β = -0.041, CI = [-0.05; -0.04], p<0.001). This suggested that on average the learning index decreased with

Table 5

Model Output: Decrease in Learning Index over Repetitions

Variable	Beta	SE	CI	t-value	p-value
Intercept	0.286	0.006	0.28 - 0.30	50.09	<0.001
Rep	-0.041	0.002	-0.05 - -0.04	-18.33	<0.001
Age	0.011	0.002	0.01 - 0.02	6.41	<0.001
SL (7)	0.201	0.029	0.14 - 0.26	6.80	<0.001
Rep x Age	-0.003	0.001	-0.004 - -0.002	-4.14	<0.001
Rep x SL (7)	-0.078	0.012	-0.10 - -0.05	-6.45	<0.010
Age x SL (7)	0.031	0.007	0.02 - 0.05	4.21	<0.001
(Rep x Age) x SL (7)	-0.014	0.003	-0.02 - -0.01	-4.56	<0.001

Note. Rep = Repetition, SL7 = Sequence Length 7, SL10 = Sequence Length 10. Intercept represents average aged participants with SL10 at first repetition.

repetition for participants of average age with sequence length 10 and decreased even further with increasing age (β = -0.003, CI = [-0.004; -0.002], p<0.001). The three way interaction of repetition, age, and sequence length 7 suggested that the decreasing effect of age on repetition is even stronger for participants with sequence length 7 (β = -0.014, CI = [-0.02; - 0.01], p<0.001). Furthermore, older age was associated with a higher learning index in the first repetition in SL10 (β= 0.011, CI = [0.01; 0.02], p<0.001) and even more so in participants with SL7 (β = 0.031, CI = [0.02; 0.05], p<0.001).

3.2   Neurophysiological Results

Thus far it was shown that the estimated linear mixed effects models suggested that participants on average increase their knowledge of the sequence over the course of learning and that older participants tended to learn greater proporation of the sequence (i.e., KI). In a next step, it was investigated how the neurophysiological measure (i.e., P300 amplitude) changed over the course of learning.

3.2.1   P300 Amplitude over Repetitions

Firstly, the mean P300 amplitude in each age group was plotted over sequence repetition (see figure 3). This plot gave the first impression that P300 amplitude decreased over repetitions and that there were potential age effects involved. Plotting ERP of repetitions for four randomly selected groups is depicted in figure 4 and showed the decrease in amplitude around 400 ms across repetitions.
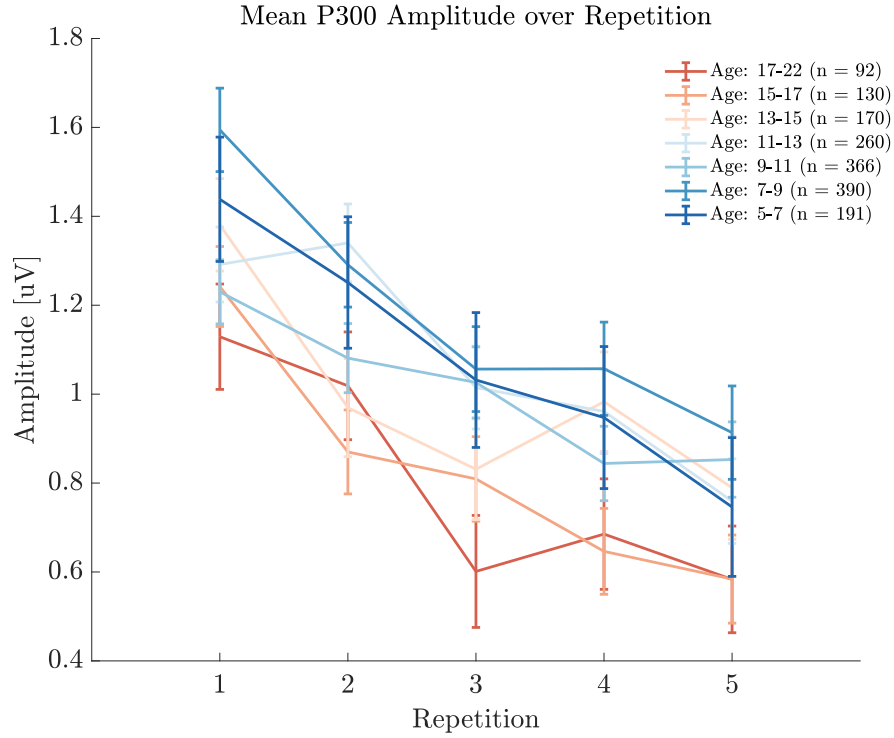
Figure 3

Decrease of Mean P300 Amplitude over Repetitions
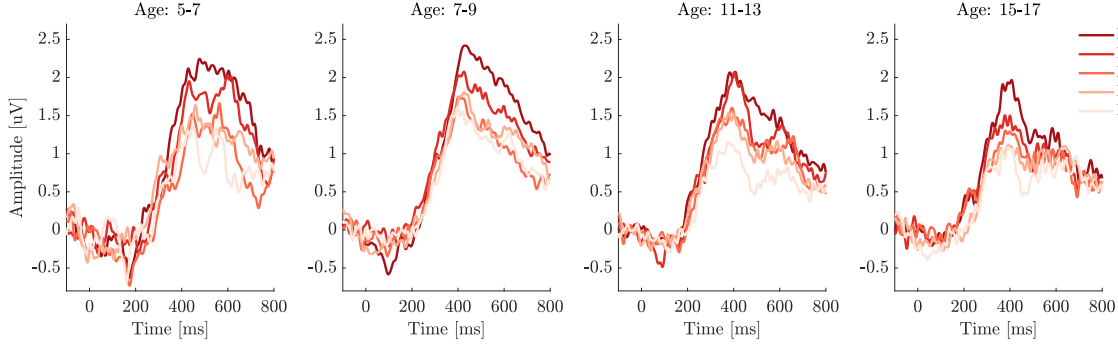
Figure 4

ERP over Repetitions Visualized for Four Groups

Second, a mixed-effects model was selected by the step function to provide inferential evidence for changes in mean P300 amplitude over repetition. The selected model was:

$$mP300 \sim REP + Age + Gender + SL + mBase + REP\,x\,Age + REP\,x\,SL + (1|Participant). \tag{9}$$

In average aged participants with sequence length 10, mean P300 amplitude decreased significantly on average with repetition (β = -0.104, CI = [-0.12; -0.09], p < 0.001). This decrease was significantly stronger in older age (β = -0.008, CI = [-0.011; -0.005], p<0.001). However, age did not explain the differences in the mean P300 amplitude in the first repetition (β = -0.006, CI = [-0.03; -0.52], p = 0.605). A significant interaction of repetition and sequence length 7 (β = -0.091, CI = [-0.12; -0.07], p<0.001) indicated that the effect of repetition is further moderated by SL, meaning that average aged participants in SL7 had a stronger decrease compared to SL10. This interaction should be viewed with caution since no actual data was observed in this range (i.e., no average aged participants had SL of 7). Mean P300 amplitude was significantly higher in the first repetition for participants with sequence length 7 compared to participants with SL10 (β = 0.209, CI = [0.04; 0.40], p = 0.017). On average, female participants had lower P300 amplitudes compared to male participants (β = -0.162, CI = [-0.29; -0.04], p = 0.010).

Table 6

Model Output: Decrease in Mean P300 Amplitude over Repetitions

Variable	Beta	SE	CI	t-value	p-value
Intercept	1.166	0.042	1.08 - 1.25	27.63	<0.001
Rep	-0.104	0.005	-0.12 - -0.09	-19.077	<0.001
Age	-0.006	0.012	-0.03 - 0.02	-0.52	0.605
Gender (F)	-0.162	0.063	-0.29 - -0.04	-2.57	0.010
SL (7)	0.209	0.087	0.04 - 0.40	2.38	0.017
meanBase	-0.223	0.044	-0.23 - -0.21	-49.98	<0.001
Rep x Age	-0.008	0.002	-0.011 - -0.005	-5.00	<0.001
Rep x SL (7)	-0.091	0.013	-0.12 - -0.07	-6.90	<0.001
Variance Components	**SD**	**Goodness of fit**			
Participant	1.10	Log Likelihood	-106636.4		
Residual	1.46				

Note. Rep = Repetition, SL7 = Sequence Length 7, SL10 = Sequence Length 10.

3.2.2 Predicting Learning Success

After finding that mean P300 amplitude decreased over repetition, the next step consisted of examining whether the knowledge index could be predicted by mean P300 amplitude of a given sequence repetition within participants. Step function selected the following model as having the best fit (with results in table 7):

$$KI \sim mP300 + Age + Gender + Age \, x \, Gender + (1|Participant). \tag{10}$$

The model had a fixed effect structure of age, gender, and their interaction and intercept as random effect. A significant main effect of mean P300 amplitude (β = -0.003, CI = [-0.006; -0.001], p = 0.025) revealed by the model suggested that increases in KI are associated with decreases in mean P300 amplitude. Furthermore, the step function

selected a model without interaction between mean P300 amplitude and age, suggesting that the predictability of mean P300 amplitude is similar across age.

Table 7

Model Output: Predicting Differences in KI with Mean P300 Amplitude

| Variable	Beta	SE	CI	t-value	p-value
Intercept	0.468	0.007	0.46 - 0.48	67.94	<0.001
meanP300	-0.003	0.001	-0.006 - -0.001	-2.25	0.025
Age	0.009	0.002	0.005 - 0.013	4.72	<0.001
Gender (F)	0.014	0.013	-0.01 - 0.04	1.12	0.264
Age x Gender (F)	0.010	0.004	0.003 - 0.017	2.84	0.004
Variance Components	**SD**	**Goodness of fit**			
Participant	0.20	Log Likelihood	-106636.4		
Residual	0.23				

3.2.3   P300 Amplitude and Learning Categories

So far it was established that on average participants increase their knowledge of sequence over the course of learning and that this was accompanied by a decrease in mean P300 amplitude over repetitions. Furthermore, the KI could be predicted by mean P300 amplitude within participants. Learning categories for individual stimuli differed in P300 amplitude regardless of repetition. The reason for this analysis was two fold. First, to disentangle the possibility that the decrease in P300 amplitude was due to habituation effects. Second, to provide a stronger link to behavioral measures of learning on a stimulus level rather than a sequence level. Plotting the average ERPs of learning categories for four randomly selected age groups (5) indicated that the P300 peaks differ between the learning categories. Furthermore, plotting the topography of the learning categories 6 indicated that the known stimuli had a lower and more widespread distribution of positive voltage compared to unknown stimuli.
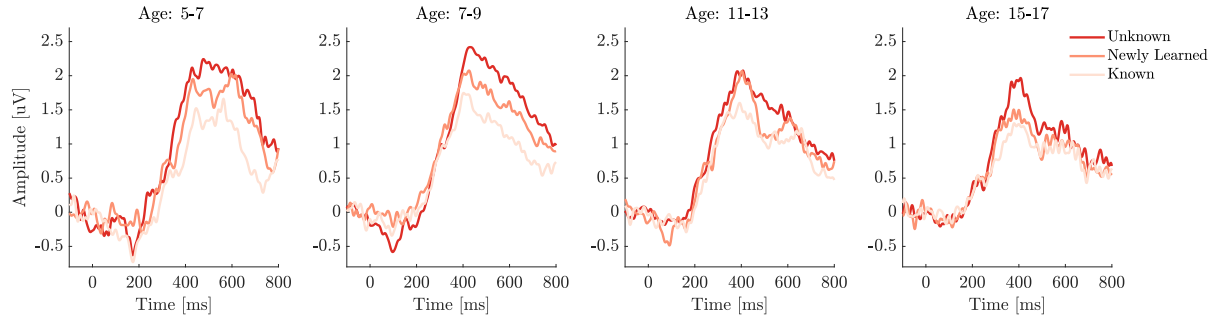
Figure 5

ERP of Learning Categories Visualized for Four Groups
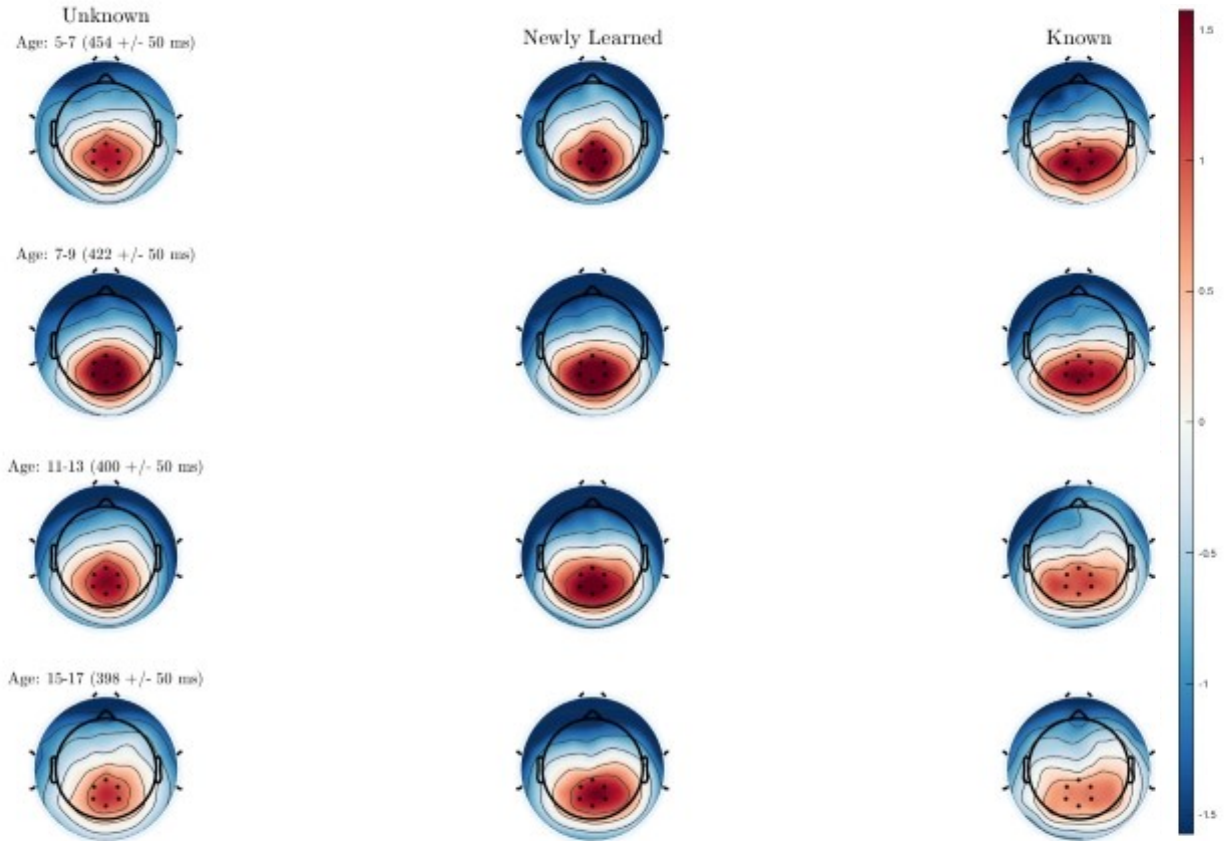
Figure 6

Topoplots of Learning Categories Visualized for Four Groups

A final mixed effects model was selected with fixed effects structure of LearningCategories (reference: NL), Age, Gender, and base, as well as a random intercept for every participant (results in table 8):

$$P300 \sim LearningCategories + Age + Gender + base + (1|Participant). \qquad (11)$$

The model revealed the following results. Learning categories forgotten (β = -0.202, CI =[-0.33; -0.07], p = 0.002) and known (β = -0.378, CI = [-0.48; -0.28], p<0.001) had significantly lower P300 amplitudes compared to newly learned, while stimuli categorized as unknown did not differ significantly (β = -0.068, CI = [-0.16; -0.02], p = 0.133). Overall, age is associated with lower P300 amplitudes (β = -0.024, CI = [-0.04; -0.01], p = 0.003) but the step function did not select an interaction between learning categories and age, suggesting similar differences between newly learned and other learning categories across age.

Table 8

Model Output: P300 Amplitude Differences in Learning Categories

Variable	Beta	SE	CI	t-value	p-value
Intercept	1.118	0.046	1.03 - 1.21	24.36	<0.001
Forgotten	-0.202	0.065	-0.33 - -0.07	-3.10	0.002
Known	-0.378	0.050	-0.48 - -0.28	-7.53	<0.001
Unknown	-0.068	0.045	-0.16 - -0.02	-1.50	0.133
Age	-0.024	0.008	-0.04 - -0.01	-2.95	0.003
Gender (F)	-0.155	0.062	-0.28 - -0.03	-2.52	0.012
mBase	-0.232	0.004	-0.24 - -0.22	-54.12	<0.001
Variance Components	**SD**	**Goodness of fit**			
Participant	0.87	Log Likelihood	-164135.2		
Residual	4.03				

Table 9

Post Hoc Contrasts of Learning Categories

Contrast	Effect Size	CI	p-value*
NL - K	0.094	0.07 - 0.12	<0.001
NL - UN	0.017	-0.005 - 0.04	0.796
NL - F	0.050	0.02 - 0.08	0.012
K - UN	-0.077	-0.10 - -0.05	<0.001
F - K	0.044	0.01 - 0.07	0.048
F - UN	-0.033	-0.06 - -0.004	0.164

*Note.* NL = Newly Learned, K = Known, UN = Unknown, F = Forgotten.
*Bonferroni corrected.

Next, a pairwise post hoc comparison of the learning categories was performed since the mixed effects model only compares each learning category with the reference category. The post hoc comparison was for average aged participants, but since the model did not include an interaction between age and learning categories the differences between learning categories were thought to be comparable across age. As shown in table 9, known stimuli had significantly lower P300 amplitudes compared to unknown stimuli or forgotten. Furthermore, forgotten stimuli did not differ significantly from unknown stimuli.

3.3 Exploratory Analysis

Using *ggmModSelect* the final model structure arrived after minimizing BIC is depicted in figure 7. The circles are nodes (i.e., selected variables), and the lines represent edges or the estimated relationship between nodes (i.e., partial correlation). All edges are undirected, with blue edges being positive partial correlations and red edges being negative partial correlations. The filled circles around each node displays the predictability (i.e., r-squared) of that node by its neighboring nodes (i.e., directly linked

nodes) (Haslbeck & Fried, 2017). See table 14 in Supplementary Material for the r-squared of each node. Node placement was determined by Fruchterman–Reingold algorithm (Fruchterman & Reingold, 1991) which creates a layout in which highly connected nodes are forced into the center of the network layout and less connected nodes are placed in outer positions of the layout (Epskamp et al., 2012).
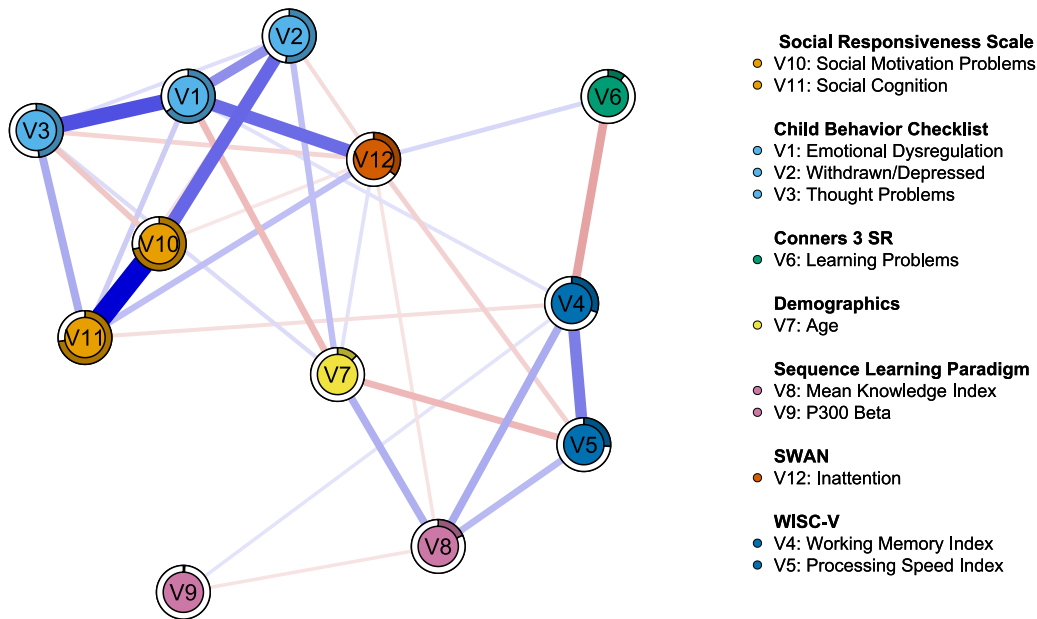
Figure 7

Gaussian Graphical Model Estimated with ggmModSelect

In the resulting network with 12 nodes, 31 non-zero edges out of 66 possible edges were included. The strongest positive relationship in the network was between Emotional Dysregulation (V1) and Thought Problems (V3) with a partial correlation of 0.48. The strongest negative relationship was between Leaning Problems (V6) and Working Memory Index (V4) with -0.25. A main interest in this analysis were the nodes P300 Beta (V9) and Mean Knowledge Index (V8) and their direct connecting nodes. After conditioning on all nodes, Mean Knowledge Index had a direct relationship to Processing Speed Index (V5, partial correlation = 0.18), Working Memory Index (V4, partial correlation = 0.23),

Age (V7, partial correlation = 0.21), P300 Beta (V9, partial correlation = -0.07) and Inattention (V12, partial correlation = -0.07). The decrease in P300 amplitude over repetition (P300 Beta) was associated negatively with the Mean Knowledge Index (V8, partial correlation = -0.07) and positively associated with Working Memory Index (V4, partial correlation = 0.07) after conditioning on all nodes. That is, a higher mean knowledge index is associated with a steeper decrease in P300 amplitude over repetitions after controlling for all nodes in the network (e.g., Age). A positive partial correlation between Working Memory Index and P300 Beta suggested that after conditioning on all nodes, higher working memory abilities are associated with a steeper decrease in P300 amplitudes over repetitions. Overall, these conditional associations explained 1% of variance in P300 Beta which may need to be considered during interpretation of the relevance of this finding (Haslbeck & Waldorp, 2018). Furthermore, bootstrapped confidence intervals (alpha = 0.05) of both edges connected to P300 Beta included zero, suggesting that these edges were not significantly different from zero (Mean Knowledge Index: $CI = [-0.11; 0]$, Working Memory Index: $CI = [0; 0.135]$).

4 Discussion

In this thesis, a previously found association between the expectancy driven P300 amplitude and successful learning (e.g., Steinemann et al., 2016; Strzelczyk et al., 2022) could be extended to a sample of children and adolescents with various psychiatric disorders and learning problems. More specifically, behavioral performance (i.e., knowledge index: KI) increased over repetitions, and age explained differences in the first repetition, as well as differences over the course of subsequent repetitions. That is, older participants had higher KI in the first repetition and increased it more strongly over subsequent repetitions. These age effects could represent differences in working memory capacity and would be consistent with previous research showing visual working memory capacity to increase with age during childhood and adolescents (Cowan et al., 2005). Furthermore, ongoing improvements in executive functions and processing efficiency during adolescents due to brain maturation and myelination of nerve fibers may play a role (Anderson et al., 2001). Successful learning can be seen as the formation of cognitive schemas to store and organize multiple items from working memory, thus WM capacity limitations can be overcome to a certain degree and performance can increase when incorporating these schemas during tasks (Bengson & Luck, 2016; Cowan, 2010; Paas & Ayres, 2014; Van Merriënboer & Sweller, 2010). P300 amplitude has been theorized to reflect processes of updating such schemas with surprise (or expectancy) to schema incongruent stimuli (Donchin, 1981). Thus, P300 amplitude changes were assessed during learning. In doing so, it was shown that the mean P300 amplitude decreases over the course of learning as the learned sequence becomes increasingly anticipated. Additionally, the decrease in mean P300 amplitude was shown to be stronger for participants with a higher age, which is, as mentioned above, accompanied by increased knowledge of the sequence (i.e., KI) over sequence repetitions. Still, the possibility existed whether this observed decrease in amplitude was due to habituation effects. Previous studies have shown that the amplitude of P300 decreases with repetition in an active oddball

discrimination task in young adults for both visual and auditory stimuli modality (Romero & Polich, 1996). The authors mentioned that habituation effects can differ depending on task characteristics and physical properties of stimuli used (Romero & Polich, 1996). For example, Courchesne (1978) showed that P300 amplitude does not habituate for target stimuli (i.e., letter A or B) which require covert response (i.e., counting) whereas stimuli that can be ignored are associated with habituation. In the context of sequence learning, P300 amplitude decreases are expected, but only meaningful if accompanied with an increase in behavioral performance. Therefore, in a subsequent analysis with the aim of strengthening the link between P300 amplitude and knowledge index (KI), it was investigated whether KI can be predicted by mean P300 amplitudes. The results showed a significant association between increased KI and reduced mean P300 amplitude within an individual. Additionally, in an attempt to disentangle P300 amplitude decreases related to successful learning from decreases occurring from habituation or simply ignoring stimuli without learning them, differences in learning categories were investigated regardless of sequence repetitions. Using the visual sequence learning paradigm and having the same sequence repeated allowed for finer differentiating of the learning status of spatial locations. Rather than classifying spatial locations during encoding based on later remembered vs. forgotten, they were categorized into learning categories based on how often they were responded to in a correct manner: from unknown over newly learned to fully committed to memory (e.g., known). Investigating the P300 amplitudes of these learning categories revealed that although unknown and newly learned did not differ significantly, once a spatial location was correctly responded to multiple times (i.e., known) the corresponding P300 amplitudes during encoding were significantly lower compared to newly learned and unknown categories. A model including an interaction between learning categories and P300 amplitude did not fit the data better suggesting that differences between learning categories were similar across age, still overall higher age was generally associated with lower P300 amplitudes. Overall, the finding that

higher age is associated with lower P300 amplitudes could result from older participants having increased knowledge of the sequence and hence having higher expectancies of upcoming spatial locations. Although it has been observed that in visual paradigms P300 amplitudes decrease with increasing age during childhood and adolescents (Riggins & Scott, 2020; Stige et al., 2007) which may need to be considered when interpreting the found age differences in amplitude. Interestingly though, age differences did not significantly explain P300 amplitude differences in the first repetition, rather it was the decrease in amplitude over repetitions that was influenced by age. Potential explanation could be that in the first repetition, for all participants, the stimuli are equally surprising and therefore no differences in P300 amplitudes are expected. Considering that P300 amplitudes are suggested to be lower in older participants, a second possible explanation could be that these differences in amplitude are mitigated because older participants update their internal model to a larger extent in the first repetition (i.e., higher KI).

Network analysis can be used as a tool to explore associations in the data and derive new hypotheses from the associations found between nodes in the network (Isvoranu et al., 2022). Exploring whether differences in mean knowledge index and P300 amplitude decrease can be explained as a network of interacting transdiagnostic constructs was the aim of the second part of this thesis. Focusing more on symptom level rather than disorder labels was motivated by the following points. Firstly, P300 amplitudes have been found to be altered in many psychiatric disorders, but in an unspecific fashion (Duncan et al., 2009) and are sometimes referred to as reflecting general vulnerability (e.g., Patrick et al., 2006). Thus, symptoms shared by disorders (e.g., transdiagnostic) may provide insight into the causes of this alteration in P300 amplitudes. Second, the heterogeneity within diagnostic labels can limit conclusions while investigating P300 alterations (e.g., Kaiser et al., 2020). The resulting network suggested that how well a participant performs overall in the VSLP (i.e., Mean Knowledge Index) provides information on P300 amplitude decrease over repetitions (i.e., P300 Beta) after controlling for all other

included variables in the network. This finding is consistent with established results from the conceptual replication part in the sense that the P300 amplitude is expectancy driven and with greater knowledge of the learned sequence, expectancy decreases. Still, while considering the bootstrapped confidence interval, this association does not seem to be significant.

4.1   Limitations

4.1.1   P300 Peak Extraction

Using group average to extract P300 peak latency, as was done in this thesis, has multiple advantages. For example, through averaging trials of many individuals, the signal-to-noise ratio decreases, and the resulting ERP waveform likely includes less random noise, making the peak better identifiable (Luck, 2014). However, averaging between individuals can be problematic, especially in the investigated age range, where the peak latencies of P300 are expected to differ (van Dinteren et al., 2014). This variation (i.e., latency jitter) can influence peak amplitude (Luck, 2014). To mitigate this problem, individuals were grouped into age bins of two years. Using a two-year bin size is not uncommon but it has been noted that bin size can affect measurements of ERP (i.e., smaller amplitudes, increased latency, diffuse topographic distribution) due to increased variability between subjects (Riggins & Scott, 2020). It is possible that even in narrow age bins as two years, latency variability may differ as a function of age. For example, in a longitudinal study that examined children from 6 to 8 years of age, DuPuis et al. (2015) showed that the intra-individual latency jitter of ERP component (i.e., ERN) decreased in this development time window. Furthermore, considering that P300 peak latency can be modulated by other factors such as gender, pubertal development, intelligence, memory capacity and psychiatric disorders, grouping individuals by age may result in a time window that doesn't capture the P300 component of each individual equally well (Brumback et al., 2012; Hansenne, 2000; Lazzaro et al., 1997; Polich, 2007; Riggins & Scott, 2020). It is worth noting that a method for calculating individual P300 peak

amplitude and latency (see., Liesefeld, 2018) failed, possibly due to the low number of trials in each individual. Single trial analysis in children has been described as challenging in previous studies (e.g., Miyazaki et al., 1994) possibly due to difficulties in discriminating the P300 signal from background noise (Riggins & Scott, 2020).

Experiments with children and adolescents are marked by multiple challenges resulting from development differences in that age range. These include not only brain maturation that should be considered when choosing a time window, but also tolerance for experimental procedures and, therefore, the paradigm should be child friendly and only include as many trials as necessary (Brooker et al., 2020). Finding a balance between child friendly number of trials and the possibility to extract individual peak amplitude should be considered in future research.

4.1.2   Task Difficulty

In the investigated age range, natural differences in cognitive performance due to brain maturation are expected to occur. Having the same task for every individual regardless of age would have been problematic, thus a shorter and longer version of the VSLP was implemented where participants under the age of nine years completed the shorter version and older participants completed the longer version. This cutoff of nine years followed from a pilot study in which floor effects were found when implementing the longer version of VSLP for all participants (Langer et al., 2017). It is possible that older participants that are close to the cutoff may experience the longer version of the task more demanding. The effects of higher task demand on P300 amplitude in VSLP are unclear. However, higher task demands have been shown to affect P300 amplitudes in dual-tasks (Isreal et al., 1980; Wickens et al., 1983).

4.1.3   Linking P300 Amplitude with Behavioral Measures

In this thesis, the P300 amplitude was linked with behavioral measures through learning categories (i.e., Unknown, Newly Learned, Known, Forgotten). This linkage is important since, as already mentioned, P300 amplitudes can be modulated by several

factors, thus confounding P300 amplitude changes over repetitions (e.g., habituation). Although the ignored and attended stimuli differ in elicited P300 amplitude, it was shown that P300 amplitude in an ignore condition decreases with repetitions in the same manner as when stimuli are attended to (Becker & Shapiro, 1980). Therefore, even ignoring the learning task could possibly lead to a decrease over repetitions. The rational for linking learning categories with P300 amplitudes was that these are relatively independent of sequence repetition. It could be assumed that this is not entirely the case since learning categories depend on sequence repetition insofar as known categories during learning should tend to increase in later repetitions (see table 10 in Supplementary Material). Additionally, stimuli are categorized as known if responded to correctly at least twice in a row. Thus, no known categories are given in the first repetition. Hence, there is the possibility that habituation contributes to the decrease in P300 amplitude found in known compared to unknown and newly learned categories. Whether this possibility occurred and how much of the effect can be ascribed to habituation wasn't answered in this thesis. Especially since no control condition was used in the paradigm. However, Steinemann et al. (2016) previously provided evidence against habituation due to time spent on task (in VSLP) by comparing sequence learning blocks with passive sequences blocks that acted as a control condition.

4.1.4 Predicting Learning Success Across Participants

By fitting an exponential to the learning curve, Steinemann et al. (2016) and Strzelczyk et al. (2022) could show that P300 amplitude can predict fast learners versus slow learners. In this thesis, the same approach to predict fast versus slow learners was attempted, but failed. Given the small numbers of sequence repetitions, fitting an exponential function (or functions of different order) was problematic. Although the aggregated learning curves showed a gradual increase over repetitions, the same must not be true for individual learning curves (Estes, 1956; Smith & Little, 2018). How excatly individuals learned over repetitions is difficult to interpret. For example, LME estimated

negative slopes for repetition in 868 participants. It is possible that many learned in the first repetition, but fail to maintain their knowledge over repetitions. Forgotten learning categories were on average 9.73 % per participant, which also indicates unstable memory or other contributing factors such as, for example, attention lapses or other problems in executive functions. Future research should explore this further, as differentiating between participants is important for the clinical utility of P300 amplitude.

4.2 Conclusion and Future Research

Previously found association between P300 amplitude and successful learning could be extended to a sample of children and adolescents with various psychiatric disorders. Therefore, highlighting its potential as a neurophysiological marker of learning. Such a marker has the possibility of identifying individuals struggling with learning, especially in cases where overt responses are difficult to measure. Second, P300 could be helpful in evaluating interventions aimed at facilitating learning through the monitoring of cognitive processes.

Future studies could consider joint modeling approach in which brain and behavior measures are modeled simultaneously (Turner, 2015). This approach provides the possibility of bridging the two different levels of analysis by formal modeling cognitive processes and relating the resulting latent parameter to neural measures (Turner et al., 2017). Lately, it was suggested to have a stronger focus on large trial numbers and analysis at the individual participant level rather than solely focusing on large N and group level analysis (Smith & Little, 2018). Especially with the challenges faced in this thesis, future studies may benefit from including more trials in a way that is still child friendly.

5 Data Citation

Functional Connectomes Project International Neuroimaging Data-Sharing Initiative http://dx.doi.org/10.15387/CMI_HBN(2017).

6 Manuscript and Code Scripts

Following GitHub repository is provided to reproduce this thesis:

- GitHub: https://github.com/oli9402/Masterarbeit_UZH

7 References

Aben, B., Stapert, S., & Blokland, A. (2012). About the Distinction between Working Memory and Short-Term Memory. *Frontiers in Psychology*, *3*. Retrieved April 3, 2023, from https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00301

Abramovitch, A., Short, T., & Schweiger, A. (2021). The C Factor: Cognitive dysfunction as a transdiagnostic dimension in psychopathology. *Clinical Psychology Review*, *86*, 102007.

Achenbach, T. M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 profile. *University of Vermont, Department of Psychiatry.*

Adam, K. C. S., Mance, I., Fukuda, K., & Vogel, E. K. (2015). The Contribution of Attentional Lapses to Individual Differences in Visual Working Memory Capacity. *Journal of Cognitive Neuroscience*, *27*(8), 1601–1616. https://doi.org/10.1162/jocn_a_00811

Alday, P. M. (2019). How much baseline correction do we need in ERP research? Extended GLM model can replace baseline correction while lifting its limits. *Psychophysiology*, *56*(12), e13451. https://doi.org/10.1111/psyp.13451

Alexander, L. M., Escalera, J., Ai, L., Andreotti, C., Febre, K., Mangone, A., Vega-Potler, N., Langer, N., Alexander, A., Kovacs, M., Litke, S., O'Hagan, B., Andersen, J., Bronstein, B., Bui, A., Bushey, M., Butler, H., Castagna, V., Camacho, N., . . . Milham, M. P. (2017). An open resource for transdiagnostic

research in pediatric mental health and learning disorders. *Scientific Data*, *4*(1), 170181. https://doi.org/10.1038/sdata.2017.181

Alexander, L. M., Salum, G. A., Swanson, J. M., & Milham, M. P. (2020). Measuring strengths and weaknesses in dimensional psychiatry. *Journal of Child Psychology and Psychiatry*, *61*(1), 40–50. https://doi.org/10.1111/jcpp.13104

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of Executive Functions Through Late Childhood and Adolescence in an Australian Sample. *Developmental Neuropsychology*, *20*(1), 385–406. https://doi.org/10.1207/S15326942DN2001_5

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive psychology*, *45*(2), 153–219.

Arvaneh, M., Robertson, I. H., & Ward, T. E. (2019). A P300-Based Brain-Computer Interface for Improving Attention. *Frontiers in Human Neuroscience*, *12*. Retrieved April 24, 2023, from https://www.frontiersin.org/articles/10.3389/fnhum.2018.00524

Astle, D. E., Holmes, J., Kievit, R., & Gathercole, S. E. (2022). Annual Research Review: The transdiagnostic revolution in neurodevelopmental disorders. *Journal of Child Psychology and Psychiatry*, *63*(4), 397–417. https://doi.org/10.1111/jcpp.13481

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (pp. 89–195). Elsevier.

Atkinson, R. C., & Shiffrin, R. M. (1971). The Control of Short-Term Memory. *Scientific American*, *225*(2), 82–91. Retrieved April 2, 2023, from https://www.jstor.org/stable/24922803

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in cognitive sciences*, *4*(11), 417–423.

Baddeley, A. (2010). Working memory. *Current biology, 20*(4), R136–R140

.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (pp. 47–89). Elsevier.

Baillet, S. (2013). Forward and Inverse Problems of MEG/EEG. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of Computational Neuroscience* (pp. 1–8). Springer. https://doi.org/10.1007/978-1-4614-7320-6_529-1

Barron, A. B., Hebets, E. A., Cleland, T. A., Fitzpatrick, C. L., Hauber, M. E., & Stevens, J. R. (2015). Embracing multiple definitions of learning. *Trends in Neurosciences, 38*(7), 405–407. https://doi.org/10.1016/j.tins.2015.04.008

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software, 67*(1). https://doi.org/10.18637/jss.v067.i01

Becker, D. E., & Shapiro, D. (1980). Directing Attention Toward Stimuli Affects the P300 But Not the Orienting Response. *Psychophysiology, 17*(4), 385–389. https://doi.org/10.1111/j.1469-8986.1980.tb00168.x

Bengson, J. J., & Luck, S. J. (2016). Effects of strategy on visual working memory capacity. *Psychonomic Bulletin & Review, 23*(1), 265–270. https://doi.org/10.3758/s13423-015-0891-7

Best, J. R., & Miller, P. H. (2010). A Developmental Perspective on Executive Function. *Child Development, 81*(6), 1641–1660. https://doi.org/10.1111/j.1467-8624.2010.01499.x

Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* Guilford Press.

Biederman, J., Petty, C. R., Day, H., Goldin, R. L., Spencer, T., Faraone, S. V., Surman, C. B., & Wozniak, J. (2012). Severity of the aggression/anxiety-depression/attention (AAA) CBCL profile discriminates

between different levels of deficits in emotional regulation in youth with ADHD. *Journal of developmental and behavioral pediatrics*, *33*(3), 236.

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Frontiers in Neuroinformatics*, *9*. Retrieved April 13, 2023, from https://www.frontiersin.org/articles/10.3389/fninf.2015.00016

Boada, R., Willcutt, E. G., & Pennington, B. F. (2012). Understanding the Comorbidity Between Dyslexia and Attention-Deficit/Hyperactivity Disorder. *Topics in Language Disorders*, *32*(3), 264–284. https://doi.org/10.1097/TLD.0b013e31826203ac

Borsboom, D. (2022). Reflections on an emerging new science of mental disorders. *Behaviour Research and Therapy*, *156*, 104127. https://doi.org/10.1016/j.brat.2022.104127

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*, 487–502. https://doi.org/10.1037/a0016797

Brites, C., Salgado-Azoni, C. A., Ferreira, T. L., Lima, R. F., & Ciasca, S. M. (2015). Development and applications of the SWAN rating scale for assessment of attention deficit hyperactivity disorder: A literature review. *Brazilian Journal of Medical and Biological Research*, *48*, 965–972. https://doi.org/10.1590/1414-431X20154528

Brooker, R. J., Bates, J. E., Buss, K. A., Canen, M. J., Dennis-Tiwary, T. A., Gatzke-Kopp, L. M., Hoyniak, C., Klein, D. N., Kujawa, A., Lahat, A., Lamm, C., Moser, J. S., Petersen, I. T., Tang, A., Woltering, S., & Schmidt, L. A. (2020). Conducting Event-Related Potential (ERP) Research with Young Children: A

Review of Components, Special Considerations and Recommendations for Research on Cognition and Emotion. *Journal of psychophysiology*, *34*(3), 137–158. https://doi.org/10.1027/0269-8803/a000243

Brumback, T. Y., Arbel, Y., Donchin, E., & Goldman, M. S. (2012). Efficiency of responding to unexpected information varies with sex, age, and pubertal development in early adolescence. *Psychophysiology*, *49*(10), 1330–1339.

Bull, R., & Scerif, G. (2001). Executive Functioning as a Predictor of Children's Mathematics Ability: Inhibition, Switching, and Working Memory. *Developmental neuropsychology*, *19*, 273–93. https://doi.org/10.1207/S15326942DN1903_3

Calvin, C. M., Batty, G. D., Der, G., Brett, C. E., Taylor, A., Pattie, A., Čukić, I., & Deary, I. J. (2017). Childhood intelligence in relation to major causes of death in 68 year follow-up: Prospective population study. *BMJ*, *357*, j2708. https://doi.org/10.1136/bmj.j2708
IQ

Chavez-Baldini, U., Nieman, D. H., Keestra, A., Lok, A., Mocking, R. J. T., de Koning, P., Krzhizhanovskaya, V. V., Bockting, C. L., van Rooijen, G., Smit, D. J. A., Sutterland, A. L., Verweij, K. J. H., van Wingen, G., Wigman, J. T., Vulink, N. C., & Denys, D. (2021). The relationship between cognitive functioning and psychopathology in patients with psychiatric disorders: A transdiagnostic network analysis. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291721001781

Chi, M. T. (1978). Knowledge structures and memory development. *Children's thinking: What develops*, *1*, 75–96.

Chiang, H.-S., Hsiao, K.-L., & Liu, L.-C. (2018). EEG-Based Detection Model for Evaluating and Improving Learning Attention. *Journal of Medical and Biological Engineering*, *38*(6), 847–856. https://doi.org/10.1007/s40846-017-0344-z

Chun, M. M. (2011). Visual working memory as visual attention sustained internally over time. *Neuropsychologia*, *49*(6), 1407–1409. https://doi.org/10.1016/j.neuropsychologia.2011.01.029

Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*(6625), 604–608. https://doi.org/10.1038/386604a0

Cohen, N. J., & Squire, L. R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: Dissociation of knowing how and knowing that. *Science*, *210*(4466), 207–210.

Conners, C. K. (2008). Conners 3.

Connolly, J. F., D'Arcy, R. C. N., Lynn Newman, R., & Kemps, R. (2000). The application of cognitive event-related brain potentials (ERPs) in language-impaired individuals: Review and case studies. *International Journal of Psychophysiology*, *38*(1), 55–70. https://doi.org/10.1016/S0167-8760(00)00130-6

Conroy, M. A., & Polich, J. (2007). Normative variation of P3a and P3b from a large sample: Gender, topography, and response time. *Journal of Psychophysiology*, *21*(1), 22–32. https://doi.org/10.1027/0269-8803.21.1.22

Courchesne, E. (1978). Changes in P3 waves with event repetition: Long-term effects on scalp distribution and amplitude. *Electroencephalography and Clinical Neurophysiology*, *45*(6), 754–766. https://doi.org/10.1016/0013-4694(78)90143-8

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current directions in psychological science*, *19*(1), 51–57.

Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, *26*(2), 197–223. https://doi.org/10.1007/s10648-013-9246-y

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N., Elliott, E. M., Saults, J. S., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive psychology*, *51*(1), 42–100.

Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, *66*, 93–99. https://doi.org/10.1016/j.jesp.2015.10.002

Crowder, R. G. (1976). *Principles of learning and memory*. Lawrence Erlbaum.

Crum, R. M., Ensminger, M. E., Ro, M. J., & McCord, J. (1998). The association of educational achievement and school dropout with risk of alcoholism: A twenty-five-year prospective study of inner-city children. *Journal of Studies on Alcohol*, *59*(3), 318–326.

Cui, T., Wang, P. P., Liu, S., & Zhang, X. (2017). P300 amplitude and latency in autism spectrum disorder: A meta-analysis. *European Child & Adolescent Psychiatry*, *26*(2), 177–190. https://doi.org/10.1007/s00787-016-0880-z

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*(4), 450–466.

De Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin & Review*, *20*(4), 631–642. https://doi.org/10.3758/s13423-013-0386-3

de Cheveigné, A. (2020). ZapLine: A simple and effective method to remove power line artifacts. *NeuroImage*, *207*, 116356. https://doi.org/10.1016/j.neuroimage.2019.116356

Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. In *Principles of frontal*

lobe function (pp. 466–503). Oxford University Press.
https://doi.org/10.1093/acprof:oso/9780195134971.003.0029

Dien, J. (1998). Issues in the application of the average reference: Review, critiques, and recommendations. *Behavior Research Methods, Instruments, & Computers*, *30*(1), 34–43. https://doi.org/10.3758/BF03209414

Donchin, E., & Cohen, L. (1967). Averaged evoked potentials and intramodality selective attention. *Electroencephalography and Clinical Neurophysiology*, *22*(6), 537–546. https://doi.org/10.1016/0013-4694(67)90061-2

Donchin, E. (1981). Surprise!. . . surprise? *Psychophysiology*, *18*(5), 493–513.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behavioral and brain sciences*, *11*(3), 357–374.

Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. *Event-related brain potentials in man*, *349*, 411
.

Donfrancesco, R., Innocenzi, M., Marano, A., & Biederman, J. (2015). Deficient Emotional Self-Regulation in ADHD Assessed Using a Unique Profile of the Child Behavior Checklist (CBCL): Replication in an Italian Study. *Journal of Attention Disorders*, *19*(10), 895–900. https://doi.org/10.1177/1087054712462884

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., Polich, J., Reinvang, I., & Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. https://doi.org/10.1016/j.clinph.2009.07.045

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467.

DuPuis, D., Ram, N., Willner, C. J., Karalunas, S., Segalowitz, S. J., & Gatzke-Kopp, L. M. (2015). Implications of ongoing neural development for the measurement of the error-related negativity in childhood. *Developmental Science*, *18*(3), 452–468. https://doi.org/10.1111/desc.12229

Ebbinghaus, H. (1885). *Über das gedächtnis: Untersuchungen zur experimentellen psychologie.* Duncker & Humblot.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of statistical software*, *48*, 1–18.

Epskamp, S., & Fried, E. I. (2015). Package 'bootnet'. *Bootstrap methods for various network estimation routines*, *5*, 0–1.

Eriksson, J., Vogel, E. K., Lansner, A., Bergström, F., & Nyberg, L. (2015). Neurocognitive Architecture of Working Memory. *Neuron*, *88*(1), 33–46. https://doi.org/10.1016/j.neuron.2015.09.020

Escera, C., Alho, K., Schröger, E., & Winkler, I. W. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neurotology*, *5*(3-4), 151–166.

Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, *53*, 134–140. https://doi.org/10.1037/h0045156

Fabiani, M., Karis, D., & Donchin, E. (1990). Effects of mnemonic strategy manipulation in a Von Restorff paradigm. *Electroencephalography and Clinical Neurophysiology*, *75*(1), 22–35. https://doi.org/10.1016/0013-4694(90)90149-E

Fernández, G., Effern, A., Grunwald, T., Pezer, N., Lehnertz, K., Dümpelmann, M., Van Roost, D., & Elger, C. (1999). Real-time tracking of memory formation in the human rhinal cortex and hippocampus. *Science*, *285*(5433), 1582–1585. https://doi.org/10.1126/science.285.5433.1582

Fiske, A., & Holmboe, K. (2019). Neural substrates of early executive function development. *Developmental Review*, *52*, 42–62. https://doi.org/10.1016/j.dr.2019.100866

Fletcher, J. M., Francis, D. J., Morris, R. D., & Lyon, G. R. (2005). Evidence-based assessment of learning disabilities in children and adolescents. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *34*(3), 506–522. https://doi.org/10.1207/s15374424jccp3403_7

Frazier, T. W., Youngstrom, E. A., Glutting, J. J., & Watkins, M. W. (2007). ADHD and Achievement: Meta-Analysis of the Child, Adolescent, and Adult Literatures and a Concomitant Study With College Students. *Journal of Learning Disabilities*, *40*(1), 49–65. https://doi.org/10.1177/00222194070400010401

Friedman, D., & Johnson Jr, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy research and technique*, *51*(1), 6–28.

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432–441.

Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and experience*, *21*(11), 1129–1164.

Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological review, 63*(3), 149.

Gathercole, S. E. (1998). The Development of Memory. *Journal of Child Psychology and Psychiatry, 39*(1), 3–27. https://doi.org/10.1111/1469-7610.00301

Gathercole, S. E., Brown, L., & Pickering, S. J. (2003). Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology, 20*(3), 109–122.

Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology, 70*(2), 177–194.

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology, 18*(1), 1–16. https://doi.org/10.1002/acp.934

Hansenne, M. (2000). [The p300 cognitive event-related potential. II. Individual variability and clinical application in psychopathology]. *Neurophysiologie clinique = Clinical neurophysiology, 30*(4), 211–231. https://doi.org/10.1016/s0987-7053(00)00224-0

Haslbeck, J. M. B., & Fried, E. I. (2017). How predictable are symptoms in psychopathological networks? A reanalysis of 18 published datasets. *Psychological Medicine, 47*(16), 2767–2776. https://doi.org/10.1017/S0033291717001258

Haslbeck, J. M. B., & Waldorp, L. J. (2018). How well do network models predict observations? On the importance of predictability in network models. *Behavior Research Methods, 50*(2), 853–861. https://doi.org/10.3758/s13428-017-0910-x

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological bulletin, 111*(1), 127.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88*(3), 297.

Hopfinger, J. B., & West, V. M. (2006). Interactions between endogenous and exogenous attention on cortical visual processing. *NeuroImage, 31*(2), 774–789. https://doi.org/10.1016/j.neuroimage.2005.12.049

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications.* Routledge.

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology, 17*(3), 259–273.

Isvoranu, A.-M., & Epskamp, S. (2021). Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. *Psychological Methods.* https://doi.org/10.1037/met0000439

Isvoranu, A.-M., Epskamp, S., Waldorp, L., & Borsboom, D. (2022). *Network psychometrics with R: A guide for behavioral and social scientists.* Routledge.

Johnson Jr., R., Pfefferbaum, A., & Kopell, B. S. (1985). P300 and Long-Term Memory: Latency Predicts Recognition Performance. *Psychophysiology, 22*(5), 497–507. https://doi.org/10.1111/j.1469-8986.1985.tb01639.x

Kahana, M. J. (2012). *Foundations of human memory.* Oxford University Press .

Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta psychologica, 86*(2-3), 199–225.

Kaiser, A., Aggensteiner, P.-M., Baumeister, S., Holz, N. E., Banaschewski, T., & Brandeis, D. (2020). Earlier versus later cognitive event-related potentials (ERPs) in attention-deficit/hyperactivity disorder (ADHD): A meta-analysis. *Neuroscience & Biobehavioral Reviews, 112*, 117–134. https://doi.org/10.1016/j.neubiorev.2020.01.019

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology: General*, *132*(1), 47.

Karis, D., Fabiani, M., & Donchin, E. (1984). "P300" and memory: Individual differences in the von Restorff effect. *Cognitive Psychology*, *16*(2), 177–216. https://doi.org/10.1016/0010-0285(84)90007-0

Kim, S., Liu, Z., Glizer, D., Tannock, R., & Woltering, S. (2014). Adult ADHD and working memory: Neural evidence of impaired encoding. *Clinical Neurophysiology*, *125*(8), 1596–1603. https://doi.org/10.1016/j.clinph.2013.12.094

Klein, S. B. (2015). What memory is. *WIREs Cognitive Science*, *6*(1), 1–38. https://doi.org/10.1002/wcs.1333

Knight, R. T., Scabini, D., Woods, D. L., & Clayworth, C. C. (1989). Contributions of temporal-parietal junction to the human auditory P3. *Brain research*, *502*(1), 109–116.

Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior Parietal Cortex Is Critical for the Manipulation of Information in Working Memory. *The Journal of Neuroscience*, *29*(47), 14980–14986. https://doi.org/10.1523/JNEUROSCI.3706-09.2009

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557–577.

Kolk, S. M., & Rakic, P. (2022). Development of prefrontal cortex. *Neuropsychopharmacology*, *47*(1), 41–57. https://doi.org/10.1038/s41386-021-01137-9

Kuncel, N. R., & Hezlett, S. A. (2010). Fact and Fiction in Cognitive Ability Testing for Admissions and Hiring Decisions. *Current Directions in Psychological Science*, *19*(6), 339–345. https://doi.org/10.1177/0963721410389459

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lacey, B. C., & Lacey, J. I. (1980). Cognitive modulation of time-dependent primary bradycardia. *Psychophysiology*, *17*(3), 209–221.

Langer, N., Ho, E. J., Alexander, L. M., Xu, H. Y., Jozanovic, R. K., Henin, S., Petroni, A., Cohen, S., Marcelle, E. T., Parra, L. C., Milham, M. P., & Kelly, S. P. (2017). A resource for assessing information processing in the developing brain using EEG and eye tracking. *Scientific Data*, *4*(1), 170040. https://doi.org/10.1038/sdata.2017.40

Lauritzen, S. L. (1996). *Graphical models* (Vol. 17). Clarendon Press.

Lazzaro, I., Anderson, J., Gordon, E., Clarke, S., Leong, J., & Meares, R. (1997). Single trial variability within the P300 (250–500 ms) processing window in adolescents with attention deficit hyperactivity disorder. *Psychiatry Research*, *73*(1), 91–101. https://doi.org/10.1016/S0165-1781(97)00107-8

Lenartowicz, A., Escobedo-Quiroz, R., & Cohen, J. D. (2010). Updating of context in working memory: An event-related potential study. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 298–315. https://doi.org/10.3758/CABN.10.2.298

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2019). Package 'emmeans'.

Liesefeld, H. R. (2018). Estimating the Timing of Cognitive Operations With MEG/EEG Latency Measures: A Primer, a Brief Tutorial, and an Implementation of Various Methods. *Frontiers in Neuroscience*, *12*. Retrieved November 1, 2022, from https://www.frontiersin.org/articles/10.3389/fnins.2018.00765

Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., Swerdlow, N. R., & Braff, D. L. (2010). Electroencephalography (EEG) and

Event-Related Potentials (ERP's) with Human Participants. *Current protocols in neuroscience / editorial board, Jacqueline N. Crawley … [et al.], CHAPTER*, Unit–6.2524. https://doi.org/10.1002/0471142301.ns0625s52

Liu, H., Lafferty, J., & Wasserman, L. (2009). The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *Journal of Machine Learning Research*, *10*(10).

Lonergan, A., Doyle, C., Cassidy, C., MacSweeney Mahon, S., Roche, R. A., Boran, L., & Bramham, J. (2019). A meta-analysis of executive functioning in dyslexia with consideration of the impact of comorbid ADHD. *Journal of Cognitive Psychology*, *31*(7), 725–749.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (Second edition). The MIT Press
.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, *54*(1), 146–157. https://doi.org/10.1111/psyp.12639

Luck, S. J., & Kappenman, E. S. (Eds.). (2012). *Oxford handbook of event-related potential components*. Oxford University Press.

Luu, P., Poulsen, C., & Tucker, D. M. (2009). Neurophysiological Measures of Brain Activity: Going from the Scalp to the Brain. In D. D. Schmorrow, I. V. Estabrooke, & M. Grootjen (Eds.), *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience* (pp. 488–494). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02812-0_57

Malmberg, K. J., Raaijmakers, J. G. W., & Shiffrin, R. M. (2019). 50 years of research sparked by Atkinson and Shiffrin (1968). *Memory & Cognition*, *47*(4), 561–574. https://doi.org/10.3758/s13421-019-00896-7

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *Journal of Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/JNEUROSCI.2925-08.2008

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American academy of child & adolescent psychiatry*, *44*(4), 377–384.

Mayes, S. D., & Calhoun, S. L. (2007). Learning, Attention, Writing, and Processing Speed in Typical Children and Children with ADHD, Autism, Anxiety, Depression, and Oppositional-Defiant Disorder. *Child Neuropsychology*, *13*(6), 469–493. https://doi.org/10.1080/09297040601112773

Medin, D. L., & Ross, B. H. (1992). *Cognitive psychology*. Harcourt Brace Jovanovich.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of verbal Learning and verbal Behavior*, *2*(1), 1–21.

Meyer, E. M., & Reynolds, M. R. (2018). Scores in Space: Multidimensional Scaling of the WISC-V. *Journal of Psychoeducational Assessment*, *36*(6), 562–575. https://doi.org/10.1177/0734282917696935

Mirsky, A. F., & Duncan, C. C. (2001). A Nosology of Disorders of Attention. *Annals of the New York Academy of Sciences*, *931*(1), 17–32. https://doi.org/10.1111/j.1749-6632.2001.tb05771.x

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*(4), 621–640. https://doi.org/10.1037/0096-3445.130.4.621

Miyazaki, M., Shibasaki, H., Suwazono, S., Honda, M., Ikeda, A., Nagamine, T., Nishida, S., Nakamura, M., Hayakawa, T., Mutoh, K., & Mikawa, H. (1994). Characteristics of auditory P300 in children: Application of single trial analysis. *Brain and Development*, *16*(5), 374–381. https://doi.org/10.1016/0387-7604(94)90124-4

Moisello, C., Meziane, H. B., Kelly, S., Perfetti, B., Kvint, S., Voutsinas, N., Blanco, D., Quartarone, A., Tononi, G., & Ghilardi, M. F. (2013). Neural Activations during Visual Sequence Learning Leave a Trace in Post-Training Spontaneous EEG. *PLOS ONE*, *8*(6), e65882. https://doi.org/10.1371/journal.pone.0065882

Nogueira, A. M., Bueno, O. F., Manzano, G., Kohn, A., & Pompéia, S. (2015). Late positive slow waves as markers of chunking during encoding. *Frontiers in Psychology*, *6*. Retrieved April 23, 2023, from https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01032

Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, *142*(7), 758–799. https://doi.org/10.1037/bul0000046

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885–958. https://doi.org/10.1037/bul0000153

Paas, F., & Ayres, P. (2014). Cognitive Load Theory: A Broader View on the Role of Memory in Learning and Education. *Educational Psychology Review, 26*(2), 191–195. https://doi.org/10.1007/s10648-014-9263-5

Paller, K. A., McCarthy, G., & Wood, C. C. (1988). ERPs predictive of subsequent recall and recognition performance. *Biological psychology, 26*(1-3), 269–276.

Patrick, C. J., Bernat, E. M., Malone, S. M., Iacono, W. G., Krueger, R. F., & McGue, M. (2006). P300 amplitude as an indicator of externalizing in adolescent males. *Psychophysiology, 43*(1), 84–92.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex.* Oxford Univ. Press.

Pedroni, A., Bahreini, A., & Langer, N. (2019). Automagic: Standardized preprocessing of big EEG data. *NeuroImage, 200*, 460–473. https://doi.org/10.1016/j.neuroimage.2019.06.046

Peng, P., & Kievit, R. A. (2020). The Development of Academic Achievement and Cognitive Abilities: A Bidirectional Perspective. *Child Development Perspectives, 14*(1), 15–20. https://doi.org/10.1111/cdep.12352

Pham, A. V., & Riviere, A. (2015). Specific Learning Disorders and ADHD: Current Issues in Diagnosis Across Clinical and Educational Settings. *Current Psychiatry Reports, 17*(6), 38. https://doi.org/10.1007/s11920-015-0584-y

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology, 9*(4), 456–479.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage, 198*, 181–197.

Polanczyk, G., & Rohde, L. A. (2007). Epidemiology of attention-deficit/hyperactivity disorder across the lifespan. *Current Opinion in Psychiatry, 20*(4), 386. https://doi.org/10.1097/YCO.0b013e3281568d7a

Polich, J. (2004). Clinical application of the P300 event-related brain potential. *Physical Medicine and Rehabilitation Clinics of North America*, *15*(1), 133–161. https://doi.org/10.1016/S1047-9651(03)00109-8

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: An integrative review. *Biological Psychology*, *41*(2), 103–146. https://doi.org/10.1016/0301-0511(95)05130-9

Polster, M. R., Nadel, L., & Schacter, D. L. (1991). Cognitive neuroscience analyses of memory: A historical perspective. *Journal of Cognitive Neuroscience*, *3*(2), 95–116 .

Potter, D. D., & Barrett, K. (1999). Assessment of mild head injury with ERPs and neuropsychological tasks. *Journal of Psychophysiology*, *13*(3), 173.

Pritchard, W. S. (1981). Psychophysiology of P300. *Psychological Bulletin*, *89*(3), 506–540. https://doi.org/10.1037/0033-2909.89.3.506

Privitera, A. J. (2022). On the utility of the p3 as a neuromarker of academic performance: A brief review. *Italian Journal of Educational Technology*.

Raaijmakers, J. G. W. (1993). The story of the two-store model of memory: Past criticisms, current status, and future directions. In *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 467–488). The MIT Press.

Ravden, D., & Polich, J. (1998). Habituation of P300 from visual stimuli. *International Journal of Psychophysiology*, *30*(3), 359–365. https://doi.org/10.1016/S0167-8760(98)00039-7

Riggins, T., & Scott, L. S. (2020). P300 development from infancy to adolescence. *Psychophysiology*, *57*(7), e13346. https://doi.org/10.1111/psyp.13346

Romero, R., & Polich, J. (1996). P3(00) habituation from auditory and visual stimuli. *Physiology & Behavior*, *59*(3), 517–522. https://doi.org/10.1016/0031-9384(95)02099-3

Rugg, M. D., & Doyle, M. C. (1994). Event-related potentials and stimulus repetition in direct and indirect tests of memory. *Cognitive electrophysiology*, 124–148.

Rushby, J. A., Barry, R. J., & Doherty, R. J. (2005). Separation of the components of the late positive complex in an ERP dishabituation paradigm. *Clinical Neurophysiology*, *116*(10), 2363–2380. https://doi.org/10.1016/j.clinph.2005.06.008

Sabol, T. J., & Pianta, R. C. (2012). Patterns of school readiness forecast achievement and socioemotional development at the end of elementary school. *Child development*, *83*(1), 282–299.

Salgari, G. C., Potts, G. F., Schmidt, J., Chan, C. C., Spencer, C. C., & Bedwell, J. S. (2021). Event-related potentials to rare visual targets and negative symptom severity in a transdiagnostic psychiatric sample. *Clinical Neurophysiology*, *132*(7), 1526–1536. https://doi.org/10.1016/j.clinph.2021.02.398

Schall, J. D. (2004). On Building a Bridge Between Brain and Behavior. *Annual Review of Psychology*, *55*(1), 23–50. https://doi.org/10.1146/annurev.psych.55.090902.141907

Schneider, W., Kron, V., Hünnerkopf, M., & Krajewski, K. (2004). The development of young children's memory strategies: First findings from the Würzburg Longitudinal Memory Study. *Journal of Experimental Child Psychology*, *88*(2), 193–209. https://doi.org/10.1016/j.jecp.2004.02.004

Semrud-Clikeman, M. (2005). Neuropsychological Aspects for Evaluating Learning

Disabilities. *Journal of learning disabilities*, *38*, 563–8.
https://doi.org/10.1177/15257401050260040601

Sloan, E., Hall, K., Moulding, R., Bryce, S., Mildred, H., & Staiger, P. K. (2017).
Emotion regulation as a transdiagnostic treatment construct across anxiety,
depression, substance, eating and borderline personality disorders: A systematic
review. *Clinical psychology review*, *57*, 141–163.

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design.
Psychonomic Bulletin & Review, *25*(6), 2083–2101.
https://doi.org/10.3758/s13423-018-1451-8

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Macmillan. Retrieved April
22, 2023, from https:
//scholar.google.com/scholar_lookup?title=Perception+and+the+conditioned+
reflex&author=Sokolov%2C+Evgenii+Nikolaevich.&publication_year=1963

Squire, L. R. (1986). Mechanisms of memory. *Science*, *232*(4758), 1612–1619.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current
perspective. *Neurobiology of Learning and Memory*, *82*(3), 171–177.
https://doi.org/10.1016/j.nlm.2004.06.005

Squire, L. R. (2009). Memory and Brain Systems: 1969–2009. *The Journal of
Neuroscience*, *29*(41), 12711–12716.
https://doi.org/10.1523/JNEUROSCI.3575-09.2009

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus
sequence on the waveform of the cortical event-related potential. *Science*,
193(4258), 1142–1146.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, *38*(4), 387–401.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The quarterly journal of experimental psychology*, *59*(4), 745–759.

Steinemann, N. A., Moisello, C., Ghilardi, M. F., & Kelly, S. P. (2016). Tracking neural correlates of successful learning over repeated sequence observations. *NeuroImage*, *137*, 152–164. https://doi.org/10.1016/j.neuroimage.2016.05.001

Stige, S., Fjell, A. M., Smith, L., Lindgren, M., & Walhovd, K. B. (2007). The Development of Visual P3a and P3b. *Developmental Neuropsychology*, *32*(1), 563–584. https://doi.org/10.1080/87565640701361096

Strzelczyk, D., Popov, T., Kelly, S. P., & Langer, N. (2022, July 19). *Neurophysiological markers of successful learning in healthy aging* (preprint). Neuroscience. https://doi.org/10.1101/2022.07.18.500426

Sur, S., & Sinha, V. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70. https://doi.org/10.4103/0972-6748.57865

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, *150*(3700), 1187–1188. https://doi.org/10.1126/science.150.3700.1187

Swanson, J. M., Sergeant, J. A., Taylor, E., Sonuga-Barke, E. J., Jensen, P. S., & Cantwell, D. P. (1998). Attention-deficit hyperactivity disorder and hyperkinetic disorder. *The Lancet*, *351*(9100), 429–433. https://doi.org/10.1016/S0140-6736(97)11450-7

Swanson, J., Deutsch, C., Cantwell, D., Posner, M., Kennedy, J. L., Barr, C. L., Moyzis, R., Schuck, S., Flodman, P., Spence, M. A., & Wasdell, M. (2001). Genes

and attention-deficit hyperactivity disorder. *Clinical Neuroscience Research*, *1*(3), 207–216. https://doi.org/10.1016/S1566-2772(01)00007-X

Szuromi, B., Czobor, P., Komlósi, S., & Bitter, I. (2011). P300 deficits in adults with attention deficit hyperactivity disorder: A meta-analysis. *Psychological Medicine*, *41*(7), 1529–1538. https://doi.org/10.1017/S0033291710001996

Teplan, M. (2002). FUNDAMENTALS OF EEG MEASUREMENT. *MEASUREMENT SCIENCE REVIEW*, *2*.

Tinga, A. M., de Back, T. T., & Louwerse, M. M. (2019). Non-invasive neurophysiological measures of learning: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *99*, 59–89. https://doi.org/10.1016/j.neubiorev.2019.02.001

Turner, B. M. (2015). Constraining cognitive abstractions through Bayesian modeling. *An introduction to model-based cognitive neuroscience*, 199–220.

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., & Van Maanen, L. (2017). Approaches to analysis in model-based cognitive neuroscience. *Journal of Mathematical Psychology*, *76*, 65–79. https://doi.org/10.1016/j.jmp.2016.01.001

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin*, *133*, 1038–1066. https://doi.org/10.1037/0033-2909.133.6.1038

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. *Medical Education*, *44*(1), 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x

van Asselen, M., Kessels, R. P. C., Neggers, S. F. W., Kappelle, L. J., Frijns, C. J. M., & Postma, A. (2006). Brain areas involved in spatial working memory. *Neuropsychologia*, *44*(7), 1185–1194. https://doi.org/10.1016/j.neuropsychologia.2005.10.005

van Dinteren, R., Arns, M., Jongsma, M. L. A., & Kessels, R. P. C. (2014). P300 Development across the Lifespan: A Systematic Review and Meta-Analysis (F. Di Russo, Ed.). *PLoS ONE*, *9*(2), e87347. https://doi.org/10.1371/journal.pone.0087347

Vaughn, S., & Fuchs, L. S. (2003). Redefining Learning Disabilities as Inadequate Response to Instruction: The Promise and Potential Problems. *Learning Disabilities Research and Practice*, *18*(3), 137–146. https://doi.org/10.1111/1540-5826.00070

Verleger, R., Heide, W., Butt, C., & Kömpf, D. (1994). Reduction of P3b in patients with temporo-parietal lesions. *Cognitive Brain Research*, *2*(2), 103–116.

Volpert-Esmond, H. I., Page-Gould, E., & Bartholow, B. D. (2021). Using multilevel models for the analysis of event-related potentials. *International Journal of Psychophysiology*, *162*, 145–156. https://doi.org/10.1016/j.ijpsycho.2021.02.006

von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of affective disorders*, *227*, 483–493.

Vossen, H., Van Breukelen, G., Hermens, H., Van Os, J., & Lousberg, R. (2011). More potential in statistical analyses of event-related potentials: A mixed regression approach. *International Journal of Methods in Psychiatric Research*, *20*(3), e56–e68. https://doi.org/10.1002/mpr.348

Walsh, M. M., Gunzelmann, G., & Anderson, J. R. (2017). Relationship of P3b single-trial latencies and response times in one, two, and three-stimulus oddball

tasks. *Biological Psychology*, *123*, 47–61.

https://doi.org/10.1016/j.biopsycho.2016.11.011

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1999). *Unobtrusive measures* (Vol. 2). Sage Publications.

Wechsler, D. J. S. A. P. C. (2014). Wechsler intelligence scale for children–Fifth Edition (WISC-V). *Bloomington, MN: Pearson.*

Wentzel, K. R., & Wigfield, A. (1998). Academic and social motivational influences on students' academic performance. *Educational Psychology Review*, 155–175.

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information-processing resources. *Science*, *221*(4615), 1080–1082.

Widmann, A., & Schröger, E. (2012). Filter effects and filter artifacts in the analysis of electrophysiological data. *Frontiers in psychology*, *3*, 233.

Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-Based Cognitive Neuroscience Approaches to Computational Psychiatry: Clustering and Classification. *Clinical Psychological Science*, *3*(3), 378–399. https://doi.org/10.1177/2167702614565359

Wigham, S., McConachie, H., Tandos, J., Le Couteur, A. S., & team, G. M. S. core. (2012). The reliability and validity of the Social Responsiveness Scale in a UK general child population. *Research in developmental disabilities*, *33*(3), 944–950.

Wilkinson, G. N., & Rogers, C. E. (1973). Symbolic description of factorial models for analysis of variance. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *22*(3), 392–399.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. *Biological psychiatry*, *57*(11), 1336–1346.

Willner, C. J., Gatzke-Kopp, L. M., Bierman, K. L., Greenberg, M. T., & Segalowitz, S. J. (2015). Relevance of a neurophysiological marker of attention

allocation for children's learning-related behaviors and academic performance. *Developmental Psychology*, *51*(8), 1148–1162. https://doi.org/10.1037/a0039311

Wrulich, M., Brunner, M., Stadler, G., Schalke, D., Keller, U., & Martin, R. (2014). Forty years on: Childhood intelligence predicts health in middle adulthood. *Health Psychology*, *33*(3), 292.

Zahedi, Y., & Abdi, M. (2012). The impact of imagery strategy on EFL learners' vocabulary learning. *Procedia-Social and Behavioral Sciences*, *69*, 2264–2272.

Zhao, T., Liu, H., Roeder, K., Lafferty, J., & Wasserman, L. (2012). The huge package for high-dimensional undirected graph estimation in R. *The Journal of Machine Learning Research*, *13*(1), 1059–1062.

8   Supplementary Material

Visual Learning Sequence Paradigm

Table 10

Total Learning Categories per Repetitions.

	Repetition 1	Repetition 2	Repetition 3	Repetition 4	Repetition 5
UN	7'759	5'378	4'560	4'005	3'598
NL	4'913	2'174	1'964	1'779	1'692
K	0	3'092	3'528	3'864	4'137
F	0	1'543	1'409	1'369	1'297

Table 11

Summary of Trials and Learning Categories

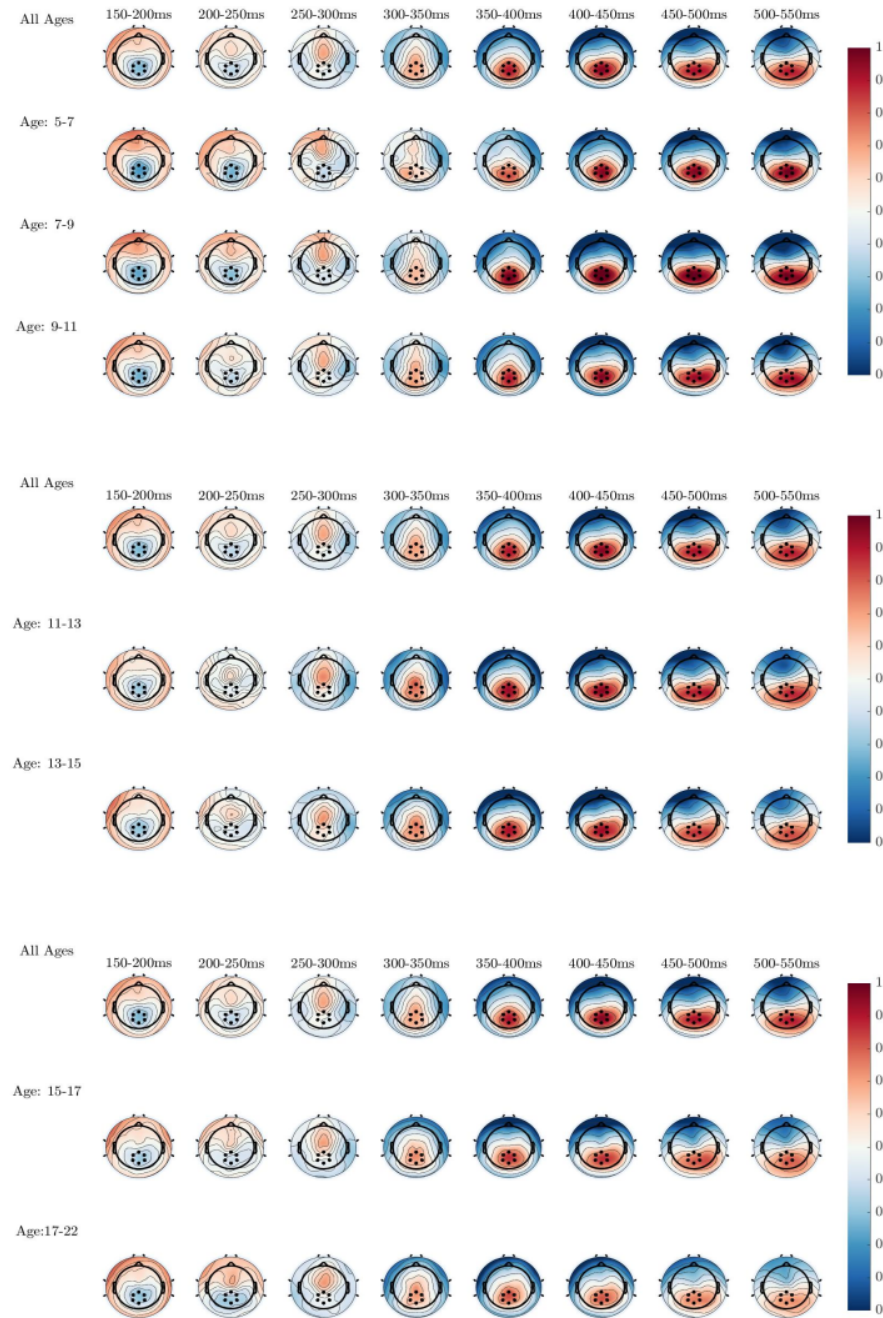|	Trials	Learning Categories				
	Mean Trials/Subject (SD)	Total	UN	NL	K	F
S7	27.42 (4.61)	14'833	6'263	3'268	3'821	1'481
S10	41.05 (6.95)	43'227	19'037	9'254	10'799	4'137
Total	36.42 (8.99)	58'060	25'300	12'522	14'620	5'618

Note. UN = Unknown, NL = Newly Learned, K = Known, F = Forgotten.

Table 12

P300 Peak Latency of Age Groups

|	Age Groups (years):						
	5-7	7-9	9-11	11-13	13-15	15-17	17-22
Peak Latency (ms)	454	422	412	400	414	398	392

89

Figure 8

Topoplots for Electrodes Selection.

Why Different Knowledge Index

Figure 9 visualizes the study design in Steinemann et al. (2016) paper. Since learning categories were created by comparing pre and posttest, the possibilty that a given stimulus location categorized as newly learned (i.e., pretest incorrect and posttest correct) could actually become known (i.e., twice correct) during the three sequence presentations led the authors to weight the newly learned learning categories with 0.5. In this thesis, behavior response is acquired after every sequence, therefore, the newly learned categories were not weighted.
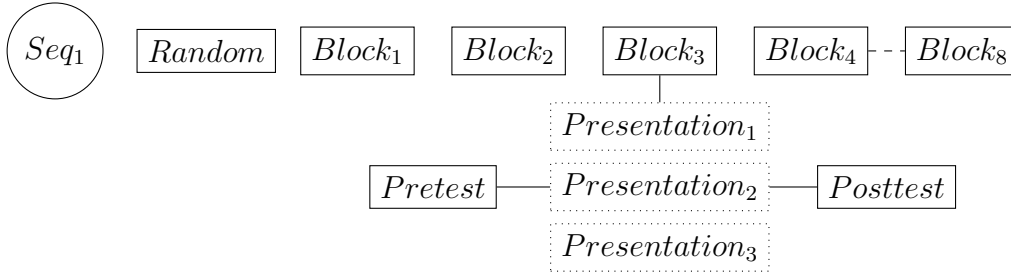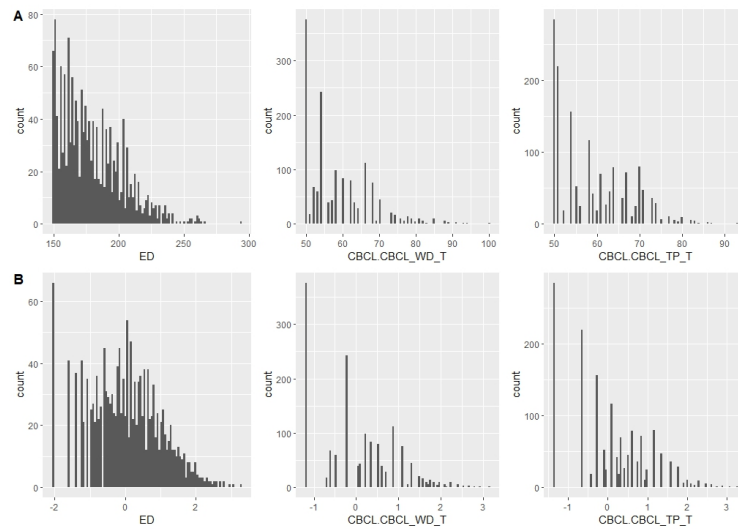

Figure 9

Study design of experiement 2 in Steinemann et al. (2016) visualized for the first sequence.

Network Analysis
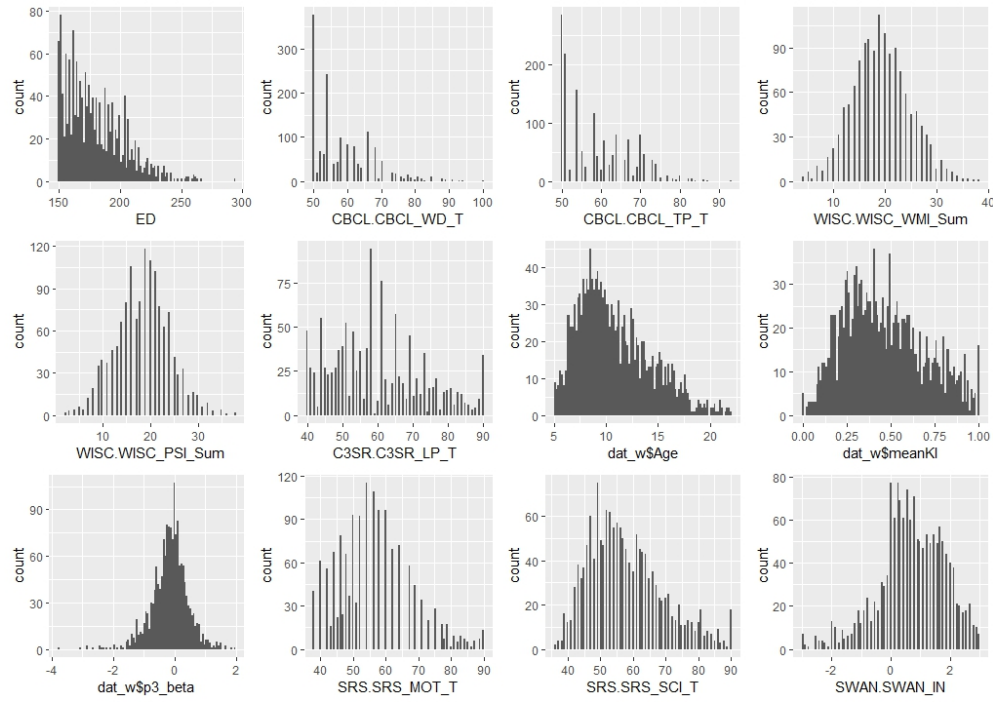
Figure 10

Distribution of Variables included in the Network Analysis.

Figure 11

Distribution of Variables transformed. A: Before transformation. B: After transformation.

Table 14

Explained Variance of each Node in Network Analysis.

Nodes	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
R^2	0.65	0.52	0.48	0.30	0.26	0.11	0.13	0.18	0.01	0.71	0.73	0.35

Table 13

Partical Correlation Matrix (estimated with ggmModSelect).

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12
V1												
V2	0.30											
V3	0.48	0.09										
V4	0.08	0.00	0.00									
V5	0.00	0.00	0.00	0.35								
V6	0.00	0.00	0.00	-0.25	0.00							
V7	-0.18	0.17	0.10	0.00	-0.19	0.00						
V8	0.00	0.00	0.00	0.23	0.18	0.00	0.21					
V9	0.00	0.00	0.00	0.07	0.00	0.00	0.00	-0.07				
V10	0.00	0.41	-0.15	0.00	0.00	0.00	0.00	0.00	0.00			
V11	0.14	-0.08	0.23	-0.09	0.00	0.00	0.00	0.00	0.00	0.00		
V12	0.40	-0.09	-0.11	0.00	-0.12	0.11	0.08	-0.07	0.00	0.17	0.00	

Philosophische Fakultät
Studiendekanat

Universität Zürich
Philosophische Fakultät
Studiendekanat
Rämistrasse 69
CH-8001 Zürich
www.phil.uzh.ch

Selbstständigkeitserklärung

Hiermit erkläre ich, dass die Qualifikations-Arbeit von mir selbst ohne unerlaubte Beihilfe verfasst worden ist und dass ich die Grundsätze wissenschaftlicher Redlichkeit eingehalten habe (vgl. dazu: https://api.swiss-academies.ch/site/assets/files/25852/kodex_layout_de_web.pdf).

CH-8400 Winterthur, 01.06.2023

....

Ort und Datum   Unterschrift