10-18-2022 1:30 PM

Using Machine Learning To Identify Neural Mechanisms Underlying the Development of Cognition in Children and Adolescents With ADHD

Brian Pho, *The University of Western Ontario*

Supervisor: Stojanoski, Bobby, *Ontario Tech University, The University of Western Ontario*
Co-Supervisor: Mohsenzadeh, Yalda, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Neuroscience

Recommended Citation
Pho, Brian, "Using Machine Learning To Identify Neural Mechanisms Underlying the Development of Cognition in Children and Adolescents With ADHD" (2022). *Electronic Thesis and Dissertation Repository*. 8917.
https://ir.lib.uwo.ca/etd/8917

Abstract

Childhood and adolescence are marked by improvements to cognition and by the emergence of neurodevelopmental disorders such as attention deficit hyperactivity disorder (ADHD). What neural mechanisms are associated with cognitive development in ADHD? In this study, I applied machine learning models to functional connectivity profiles to identify patterns of network connectivity that predict various cognitive abilities in a group of participants ages 6 to 16 with ADHD. The models successfully predicted IQ, visual spatial, verbal comprehension, and fluid reasoning in children ages 6 to 11, but not adolescents. Furthermore, the models identified connections with the default mode, memory retrieval, and dorsal attention networks as driving prediction during early and middle childhood, but connections with the somatomotor, cingulo-opercular and frontoparietal networks were more important in middle childhood. These results suggest that computational models can identify neural mechanisms associated with predicting cognitive abilities in children and adolescents with ADHD using naturalistic stimuli.

Keywords

Computational, Cognitive, Developmental, Neuroscience, Cognition, Attention Deficit Hyperactivity Disorder, ADHD, Machine Learning, Ridge Regression, Partial Least Squares, Intelligence, Children, Adolescence, Wechsler Intelligence Scale for Children, WISC, Naturalistic stimulus, Movies

Summary for Lay Audience

One of the most common disorders in children that affects the development of the brain is attention deficit hyperactivity disorder (ADHD). ADHD is defined by abnormal amounts of inattention, impulsivity, and hyperactivity—such as when a student has trouble paying attention to the teacher's school lesson. From previous scientific studies, it is believed that ADHD is linked to problems with controlling a person's behavior including suppressing inappropriate or unwanted actions, switching between different mental processes, and remembering information for a short period of time. However, it is not clear how these mental or cognitive abilities relate to children's brains with ADHD and how they change as children become adolescents. Specifically, how does the brain activity in children and adolescents diagnosed with ADHD relate to cognitive abilities such as IQ and working memory? In this study, I explored this question by using computer models to find links between brain activity—while participants watched the movie "Despicable Me"—to various cognitive abilities in children and adolescents ages 6 to 16 with ADHD. I found that the models could predict IQ, visual spatial ability, and verbal comprehension ability in early childhood. In addition to these three cognitive abilities, the models could also predict fluid reasoning and working memory ability in middle childhood. This suggests that the models can capture different cognitive abilities in children of different ages. But how do the models predict cognition from brain activity? By analyzing the models, they pointed to a set of brain networks that were more important than other networks to predict cognitive ability. These networks are believed to participate in memory, attention, and motor control, and this set of important networks changed as the models were trained on children in different developmental stages. Overall, these findings provide evidence that computer models can predict cognitive ability in children with ADHD and they contribute to our understanding of how brain activity is linked to different cognitive abilities in children with ADHD.

Acknowledgments

I would like to thank my supervisors, Dr. Bobby Stojanoski and Dr. Yalda Mohsenzadeh, for their support and feedback throughout my degree. They provided me the opportunity to study cognition using machine learning with the freedom to pursue interesting ideas, for which I am grateful for. I am also grateful for their guidance throughout the project such as when I got stuck on a programming bug or when the code generated an abnormal result. Dr. Stojanoski helped in discussing and thinking about the story that connects all of the project’s results into a narrative thread, as well as the overall project goals. Dr. Mohsenzadeh helped in interpreting the computational aspects of the project and providing ideas to overcome modeling roadblocks. I am tremendously grateful to both of them for their continued patience and persistence with me throughout this project.

I am also thankful for everyone in the DevSCN lab and CNAI lab for their input and advice on how to succeed in graduate school and academia. I am especially grateful for Kathleen Lyons’s help in explaining the Healthy Brain Network biobank, and for Brandon Hu’s time to review the cross-prediction code. Additionally, I am grateful to Haitao Yang for his IT support on the OwenLab servers, particularly to refresh the NoMachine license when it expired.

I want to acknowledge the work by the Child and Mind Institute for creating the Healthy Brain Biobank that this research depends on.

Lastly, I am thankful for my friends and family that have supported me throughout this journey.

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

Childhood and adolescence represent a period of profound change to higher-level cognition, which can have lifelong influences on school and work success, health, and even happiness. It is also a period when many psychopathologies begin to emerge (Nielsen et al., 2020). Common across different childhood psychopathologies are the detrimental effects they have on cognitive development. In particular, children with Attention Deficit Hyperactivity Disorder (ADHD) have been shown to have severe impairments in higher-level cognitive abilities, such as executive functioning, that are essential for maintaining focus and planning actions. However, identifying the neural mechanisms associated with typical cognitive development, and the neural mechanisms associated with disrupted development as seen in ADHD, is a challenge and remains unresolved. One solution is to use machine learning, which provides a valuable tool to better understand the relationship between cognition and neurodevelopmental disorders. The goal of my thesis was to apply machine learning to neuroimaging data, specifically neural activity in response to movie watching, to identify patterns of functional network connectivity that best predict cognitive ability in a large cohort of children and adolescents with ADHD.

1.1 Neural Mechanisms Supporting Cognitive Development

Cognition is the act of perceiving, thinking, and understanding. For example, thinking about what to wear to a birthday party requires, among other things, perceiving the available wardrobe, considering the weather, and understanding the social context. Thus, cognition is an umbrella term that covers various aspects of intellectual abilities—such as perception, attention, intelligence, memory, and executive function—that we rely on to make sense of our complex internal and external environments.

While some cognitive abilities develop within the first six months of life, such as face recognition (Nelson, 2001), other aspects of cognition emerge later in childhood and continue to develop throughout adolescence and into adulthood. In fact, the transition period from childhood to adolescence is marked by extraordinary developmental changes in sensory, motor, and various cognitive abilities (Casey et al., 2008), including executive function (Baum et al., 2017; Rubenstein et al., 2020). The maturation of different cognitive abilities is particularly important during childhood and adolescence because of its influence on various aspects of adult life. For instance, Calvin and colleagues found that healthy higher-level cognitive functioning in childhood is predictive of many aspects of an healthy adult life, including success in school, work, and even life expectancy (Calvin et al., 2011, 2017). Calvin suggested that intelligence, which is best thought of as the capacity to integrate various cognitive abilities to solve specific problems (Chollet, 2019), is essential for academic success and is the bridge to positive vocational and health metrics in adults.

Studies on the neurobiology of intelligence have implicated the frontoparietal network as important (Deary et al., 2010). For instance, Deary et al. (2010) argued that intelligence is not localized to any single brain area, but instead uses a distributed network consisting of the frontoparietal network, among others. The frontoparietal network—which consists of the dorsolateral prefrontal cortex and posterior parietal cortex (Gong et al., 2016; Gratton et al., 2018; Thomas Yeo et al., 2011)—has been found to increase in functional selectivity and modularity over the course of development from childhood to adulthood (Baum et al., 2017; Fair et al., 2009; Wendelken et al., 2011).

One common method for measuring and quantifying intelligence in children is the Wechsler Intelligence Scale for Children (WISC) (David Wechsler, 2014). The WISC, like many other scales of IQ, is represented by primary cognitive abilities including executive functioning, visual and spatial processing, and working memory. Therefore, the development of intelligence is built upon on the development of specific, more rudimentary cognitive abilities.

A major component of intelligence is executive function, which is associated with cognitive control and flexible behavior (Buss & Spencer, 2018). Specifically, executive function refers to three aspects of higher-order processes, namely inhibitory control, cognitive flexibility, and working memory (Diamond, 2013; Rubenstein et al., 2020). Executive function has been shown to emerge in infancy and shows clear improvements in childhood, with aspects of executive function continuing to develop in adolescence (Luna et al., 2010; Rubenstein et al., 2020). What changes in the brain underlie the development of executive function? Baum et al. (2017) found that the modular segregation of structural brain networks—most prominent was the frontoparietal network—mediates the improvement of executive function in youth. Furthermore, Baum and colleagues found that the trajectory of executive performance develops rapidly between the ages of 8 to 14 but plateaus from age 14 to 22. This is consistent with Buss & Spencer's (2018) finding that changes in the long-range cortical interaction between frontal and posterior cortical regions underlie the emergence of executive function in children. Buss & Spencer (2018) used the dimensional change card sort task to test children’s flexibility in rule-use and found that children who failed to switch rules showed weak frontal cortex activations, while children who successfully switched rules showed strong frontal cortex activations. Furthermore, older children showed contrasting posterior cortical activations on easy and hard versions of the task, reflecting continued refinement of these brain networks.

Another cognitive ability that contributes to intelligence, and is also closely linked to executive function, is working memory (Cowan, 2014). Working memory maintains brief representations of relevant information to achieve specific and immediate goals. Working memory is considered to consist of at least two subsystems in humans: one for verbal information and one for visuospatial information (Baddeley, 1992, 2010; Kandel et al., 2021). These two memory subsystems are coordinated by executive control processes in the prefrontal cortex, which are believed to allocate attentional resources to monitor, manipulate, and update stored representations between these systems (Rainer et al., 1998). The verbal subsystem maintains speech representations in conscious awareness and recruits posterior parietal cortical regions. The visuospatial subsystem maintains mental images and locations of visual objects and depends on the parietal, inferior temporal, and occipital cortical regions. Thus, in adults, working memory involves activation of specialized cortical regions based on the content of the representation, and the activation of general control mechanisms in the prefrontal cortex

(Kandel et al., 2021). However, is the separation of working memory into verbal and visuospatial components also seen in children?

The study by Alloway et al. (2006) addressed this question by measuring different memory components in children between ages 4 and 11 years old and found that a model based on distinct verbal and visuospatial components had a high correlation between the components and shared a large amount of variance. Furthermore, working memory models in adult studies (Kane et al., 2004) also had a similar amount of large shared variance, suggesting that an adult population looks similar to a child population in terms of distinguishable verbal and spatial working memory constructs. Therefore, Alloway et al. (2006) concluded that children have separable verbal and spatial working memory. Although the separation of verbal and spatial working memory is seen in both children and adults, working memory capacity has been shown to differ between children and adults. Children do not start at an adult-level of working memory capacity, rather, steady improvements are seen on digit span repetition tasks until it approaches adult-levels of performance (Cowan, 2014; Gathercole et al., 2004). What are the neural mechanisms that support this improvement in working memory capacity? Edin et al. (2007) used a computational network to model the effects of cellular maturation processes—such as myelination, synaptic strength, and synaptic pruning—and found that stronger frontoparietal synaptic connectivity for neurons encoding similar stimuli was the only developmental process that accounted for changes in brain activity associated with the development of working memory. This finding is supported by Ullman et al. (2014) who found that working memory capacity correlated with frontoparietal cortical activity. A better understanding of the development of working memory capacity in children is important because it is a strong predictor of future academic performance and is useful for the identifying children at risk for poor cognitive and academic development (Alloway & Alloway, 2010; Ullman et al., 2014).

Working memory is strongly related to another cognitive ability, fluid reasoning, which is another core component of intelligence (Chuderski & Necka, 2012; Dehn, 2017; Kyllonen & Christal, 1990). Fluid reasoning is the capacity to think logically and solve problems in novel situations, and can be assessed by Raven’s Progressive Matrices

(Cattell, 1987). It is used in insight problem solving where there is a sudden and unexpected realization of a problem’s solution (Chuderski, 2014). Fluid reasoning is strongly linked to academic success, and poor performance (i.e., lower scores) on fluid reasoning assessments significantly predicted subsequent high school dropout (Pagani et al., 2017). Neuroimaging studies have confirmed that the same regions active in working memory, particularly the frontoparietal network, are also key to fluid reasoning (Dehn, 2017; Otero, 2017). In adults, fMRI studies involving the Raven’s Progressive Matrices have shown that the rostrolateral prefrontal cortex is activated when integrating across multiple mental representations, while the parietal cortex is activated for identifying and representing visual-spatial relations used by the rostrolateral prefrontal cortex (Ferrer et al., 2009). In children, however, this cognitive ability is believed to emerge in the first two or three years of life and follows a rapid developmental trajectory in early and middle childhood, followed by a slowing developmental trajectory until early adolescence, and stabilizes in mid-adolescence (McArdle et al., 2000; Otero, 2017). The developmental changes in brain structure believed to underlie this trajectory are reductions in synaptic density first in the dorsal parietal lobes during childhood, which then spread anteriorly to dorsal frontal regions during adolescence until 20 years of age (Ferrer et al., 2009). This is also supported by the previous work done by Wright et al. (2008) who found that changes in the function of the rostrolateral prefrontal cortex over childhood and adolescence may contribute to improvements in fluid reasoning.

The literature that I have covered on the neural mechanisms supporting cognitive development in intelligence, executive function, working memory, and fluid reasoning was specific to typically developing children and adolescents. However, deficits in many of the cognitive abilities discussed above, including executive function, working memory, and fluid reasoning, have been found in atypically developing children and adolescents, specifically those diagnosed with Attention Deficit Hyperactivity Disorder (ADHD). In the next section, I focus on cognitive development in children and adolescents diagnosed ADHD. Examining the distinct developmental trajectories of cognitive abilities, and the underlying neural mechanisms, may provide a better understanding of ADHD.

1.2 Cognition and ADHD

Attention Deficit Hyperactivity Disorder (ADHD) is the most common neurodevelopment disorder among children and adolescents, affecting an estimated 4.8% of all Canadian children up to 19 years of age (Waddell et al., 2002). It is not, however, a condition restricted to childhood and adolescence as ADHD often persists into adulthood, with ramifications in the individual's personal and professional life, such as low self-esteem and impaired social function compared to non-ADHD individuals (Harpin et al., 2016; Harpin, 2005). The reason ADHD is commonly diagnosed in school-aged children is because the symptoms linked to ADHD are most salient in the classroom (Danielson et al., 2018). For instance, ADHD is best characterized by a persistent pattern of inattention (inability to maintain focus), impulsivity (acting on instinct without thinking), and/or hyperactivity (excessive restlessness and movement) that can interfere with not only completing school-based tasks, but also extends to daily functioning (American Psychiatric Association, 2013). Multiple studies have suggested that the symptoms characterizing ADHD are not only associated with a specific set of behavioral consequences, but are also associated with poor sleep (Kirov & Brand, 2014; Konofal et al., 2010; Owens, 2005) and worse academic performance leading to lower average income (Arnold et al., 2020; Biederman & Faraone, 2006; Daley & Birchwood, 2010).

The symptomology of ADHD is also associated with cognitive functioning, which is likely related to academic and career success. Indeed, one of the most common aspects of ADHD is deficits in executive functioning (Rucklidge & Tannock, 2002). As described above, executive functioning is comprised of three components: inhibitory control, cognitive flexibility, and working memory. An early model of ADHD by Barkley (1997) argued that ADHD mainly results from a core deficit in only one of these three components—inhibitory control. The model was updated a few years later following work by Willcutt et al. (2005), who conducted a meta-analysis of 83 studies where executive function was measured in groups with and without ADHD, and found that the groups with ADHD showed significant impairment on all three components of executive function. However, the moderate effect sizes and lack of consistency of executive function deficits for the ADHD groups lead the researchers to conclude that executive function deficiencies were neither necessary nor sufficient to cause all cases of ADHD. The link between ADHD and executive functioning was expanded by Castellanos et al. (2006) who divided executive functioning into two parts: “cold” and “hot.” Castellanos et al. (2006) found ADHD to be associated with the ‘cool’ executive function but that the expected essential role of inhibitory and executive deficits was not supported for a substantial proportion of children with ADHD. The ‘cool’ executive function is associated with the orbital and medial prefrontal cortex and is used in suppressing automatic responses and maintaining task instructions in working memory, while the ‘hot’ executive function is associated with the dorsolateral prefrontal cortex and is used in high affective or flexible appraisal problems (Zelazo & Mller, 2007). This aligns with evidence from the biological bases of ADHD that have shown reductions in the prefrontal cortex, basal ganglia, cerebellum, and parieto-temporal regions, implicating multiple cognitive processes such as cognitive control, working memory, and attention (Durston, 2003; Friedman & Rapoport, 2015; Krain & Castellanos, 2006).

Two other cognitive abilities that have been found to be impaired in children with ADHD are working memory and processing speed (Mayes & Calhoun, 2006; Moura et al., 2019; Thaler et al., 2013). Mayes & Calhoun (2006) used scores on the WISC assessment to determine if they could be helpful in diagnosing children with ADHD and found that all children diagnosed with ADHD scored lowest on the working memory and processing speed measures. Coming to a similar conclusion, Moura et al. (2019) studied the specific cognitive impairments in children between the ages of 6 and 12 with ADHD using the WISC and found that the children had marked deficits in subtests requiring working memory and processing speed. Other measures of working memory and processing speed, often in the context of reading, have produced similar results in associating poor performance on both of these cognitive abilities with ADHD symptomology (Fosco et al., 2020; Jacobson et al., 2011; Katz et al., 2011; Rucklidge & Tannock, 2002). Recently, there has been a lot of interest in improving working memory in children diagnosed with ADHD through training, often demonstrating significant improvements on working memory tasks (Klingberg et al., 2002, 2005). These results suggest that there is a strong link between ADHD and deficits in working memory and processing speed.

Although there is a vast literature on cognition and ADHD, many studies focus on adults or single-age cohorts of children. These studies neglect the importance of cognitive development from childhood to adolescence and eventually to adulthood in individuals with ADHD. Relatedly, many of the findings emerging from these studies focus on uncovering a snapshot of the neural mechanisms associated with cognition in ADHD in specific age cohorts, without taking development into consideration. For example, experiments have uncovered a link between ADHD and deficits in executive function, working memory, and processing speed, but a deeper understanding of how these cognitive abilities relate to the developing brain remains poorly understood. This gap in the literature—how cognition develops in children and adolescents with ADHD and the neural mechanisms associated with the specific cognitive deficits in ADHD—forms the focus of my thesis.

1.3 Modeling and Predicting Cognition

One powerful approach that can capture the neural mechanisms associated with cognitive development are computational models. Computational models are powerful because they can exploit patterns in large datasets that may otherwise go unnoticed. Computational approaches, such as machine learning, have become popular in recent years in addressing cognitive neuroscience questions (Hassabis et al., 2017; Nielsen et al., 2020). Machine learning models are a set of methods that can ‘learn’, that is, leverage data to improve performance on classification and prediction tasks (Lecun et al., 2015). Two of the main approaches in machine learning are supervised learning and unsupervised learning. Supervised learning involves building a model using data that contains both inputs and desired outputs. Unsupervised learning, on the other hand, requires building a model using only input data. Examples of supervised learning models include linear regression and support vector machines, and examples of unsupervised learning models include k-means clustering and principal component analysis. Two recent and exciting advances in the field have been the use of deep neural networks for natural language generation (Brown et al., 2020) and to solve the protein folding problem (Jumper et al., 2021). But can computational models be applied to study the link between brains, cognition, and development?

The answer is yes. In fact, machine learning has discovered many insights about the relationship between brains and cognition (Bertolero & Bassett, 2020; Chen et al., 2022; Cui et al., 2022; Marek et al., 2022; Rosenberg et al., 2018; Shen et al., 2017). For example, machine learning has been successfully applied by Tian et al. (2021)—who used resting-state functional connectivity networks with Ridge regression in healthy young adults (n=400)—to achieve modest prediction accuracies (r=0.2-0.4) for fluid and crystalized intelligence. However, a similar approach can also be used to predict links between neural activity and cognition in younger participants. The study by Sripada et al. (2019) applied machine learning models to a resting-state dataset (n=2013) collected from 9 to 10 year old children to predict individual differences in three higher-order cognitive functions: general ability, speed/flexibility, and learning/memory. The model successfully predicted individual differences in these three higher-order cognitive functions with correlation scores of r=0.29-0.33, r=0.06-0.09 and r=0.10-0.15 for general ability, speed/flexibility, and learning/memory respectively. When the model was analyzed, connections with the default mode network and three task control networks—frontoparietal, salience, and dorsal attention—were implicated in predicting differences in General Ability in early adolescence. A more recent study conducted by Finn & Bandettini (2021) used movie-watching functional connectivity in healthy adults (n=176) to predict cognition and emotion scores, such as processing speed and life satisfaction, and achieved similar results (Spearman r=0.2-0.4) to Tian & Zalesky (2021) and Sripada et al. (2020). This result shows how naturalistic stimuli—in this case movies—can improve the prediction of cognition and emotion by amplifying individual differences in relevant brain networks. Taken together, these three studies converge on a correlation score range between 0.2 to 0.4 for predicting cognition using various datasets and models. I expect that my study, which incorporates aspects inspired from these studies—the Ridge model used in Tian & Zalesky (2021), the data from youth in Sripada et al. (2020), and the movie-watching stimuli in Finn & Bandettini (2021)—to also achieve prediction scores between 0.2 to 0.4; although a key difference between the study by Finn & Bandettini (2021) and my study is they used naturalistic stimuli to predict cognition in adults, while I use neural activity in response to movie watching to predict cognition in children and adolescents.

1.4 Naturalistic Stimuli in fMRI

Naturalistic stimuli like movies are a unique and useful tool for exploring the link between various social cognitive processes and the neural mechanisms that support them. For instance, movies can be used during fMRI acquisition to examine the relationship between social cognition and neural synchrony in children with and without autism (Lyons et al., 2020). But they have also been shown to be useful for predicting cognition as Cantlon & Li (2013) used neural activity in the intraparietal sulcus, while children and adults watched the television show “Sesame Street”, to predict performance on tests assessing mathematics ability. In a more recent study, Caldinelli & Cusack (2022) used the “Forrest Gump” movie to examine dynamic functional connectivity during movie-watching and found that the frontoparietal network was not as flexible as other networks across the brain, concluding that the frontoparietal network did not have features of a flexible hub during movie watching.

Various studies have also found that the use of movie-watching stimuli during fMRI improves functional connectivity-based prediction of behavior compared to resting-state fMRI (Caldinelli & Cusack, 2022; Finn & Bandettini, 2021; Gruskin & Patel, 2022). One reason movie-fMRI performs better than rest-fMRI is because movies are associated with less body and head movement in participants, especially in children. The reduced movement during movie-watching results in less noise in the fMRI data, leading to improved modeling of the data compared to rest-fMRI. Another reason naturalistic stimuli typically performs better than task- and rest-fMRI is because it more closely resembles daily life. For instance, in movies, we encounter characters in a setting that perform various actions to achieve their goals, similar to the interactions people have in their own lives. This leads to improvements over resting-state fMRI for predicting cognition because movie-watching requires the integration of various cognitive systems to follow the complexities of the plot and because individuals often have a unique interpretation of the movie, resulting in enhanced individual signals and thus richer brain dynamics that can be captured by predictive models (Meer et al., 2020; Vanderwal et al., 2017). Every person has had unique experiences in their life and movies are one method to draw out these unique experiences by immersing the viewer in a similar natural sensory environment. For my study, the enhanced individual signals captured while children and adolescents watch the movie “Despicable Me” will be used by computational models to predict cognition.

1.5 Objectives and Hypotheses

In the current study, I aim to explore three questions: (1) can cognition be predicted in a large group of children and adolescents diagnosed with ADHD using functional neural activity during movie watching? (2) does prediction accuracy remain constant across development? and (3) if not, what are the neural mechanisms associated with developmental differences in predicting cognition? I integrate various aspects of cognitive development, ADHD, computational modeling, and naturalistic stimuli to explore these questions. Previous research has covered some of these aspects, such as predicting cognition in adults (Tian & Zalesky, 2021) and analyzing the link between cognition and ADHD (Claesdotter et al., 2018), but no study to date has combined all of these aspects.

To explore the first question, I use movie-watching fMRI—in ADHD and typically developing (TD) children and adolescents—to predict scores on six cognitive abilities as measured by the WISC. I hypothesize, based on the literature, that there is a strong relationship between brain activity induced by movie-watching and higher-level cognition (Finn & Bandettini, 2021; Vanderwal et al., 2017), and that I can generate models to predict this relationship in children and adolescents with ADHD. Specifically, I predict the model will achieve Pearson correlation scores between 0.2 and 0.4 for all the cognitive abilities except for working memory and processing speed in the ADHD group. The predicted score range is based on similar computational studies using functional connectivity to predict cognition in typically developing children and adults (Finn & Bandettini, 2021; Sripada et al., 2020; Tian & Zalesky, 2021). The hypothesized inability to predict working memory and processing speed in the ADHD group is because of various studies that have uncovered a link between deficits in these two cognitive abilities and ADHD (Becker et al., 2021; Jacobson et al., 2011; Mayes & Calhoun, 2006; Thaler et al., 2013).

Next, to explore the second question and uncover any developmental effects, I split the ADHD group by age into three age bins and modeled the neural activity (in response to the movie) to predict the same six cognitive abilities for each age bin. By splitting the ADHD group into three age bins, the models will either 1) predict the same set of cognitive abilities for all three age bins, or 2) predict a different set of cognitive abilities for each age bin. If the same set of abilities is predicted for some age bins—such as predicting scores in visual spatial ability for participants between the ages of 6-8 and 9-11, but not for ages 12-16—one possibility is that there is a shared functional connectivity profile for participants ages 6-8 and 9-11 that is used to predict visual spatial ability. In contrast, if a different set of abilities is predicted between age bins—such as only predicting scores in working memory ability for participants ages 9-11—this suggests that the model captures a functional connectivity profile specific to this age bin that is not shared by other age bins.

To address the third question, I focus on a set of differential neural mechanisms underlying cognitive development and I test if the functional connectivity profile is shared between age bins by using the model trained on one age bin to predict the same cognitive ability in a different age bin. For example, I use a model trained on visual spatial ability in participants ages 6-8 to predict visual spatial ability in participants ages 9-11. If the prediction is successful, this would suggest that both age bins share a functional connectivity profile for that specific cognitive ability—in the previous example, the functional connectivity profile used to predict visual spatial ability is shared between participants ages 6-8 and 9-11. If the prediction is unsuccessful (meaning the model could not generalize), this would suggest that the age bins have unique functional connectivity profiles. The shared functional connectivity profile is important because I can analyze the shared profile for network connections that are equally important for predicting cognitive ability in participants across age bins. To analyze the shared profile, I calculate the difference between the models' feature weights trained on each age bin, revealing connections that changed between age bins and connections that did not change between age bins. I hypothesize that frontoparietal and memory network connections change between age bins based on studies of the neural mechanisms associated with the development of intelligence, executive function, and working memory capacity (Baum et al., 2017; Deary et al., 2010; Hitch & Halliday, 1983). Conversely, I expect visual networks connections to not change between age bins due to the work by Alloway et al. (2006) who found that children have verbal and spatial working memory at an adult level. In addition to these hypotheses, I take an exploratory approach in search of unique and unexpected network patterns associated with cognition that might generate new hypotheses for future studies.

Chapter 2

2 Methods

The period from childhood to adolescence is notable for its extraordinary physical, behavioral, and cognitive development. For instance, structural and functional changes in the frontal and parietal lobes during this period are associated with improvements in higher-level cognitive abilities such as executive function and working memory (Baum et al., 2017; Fair et al., 2009; Kandel et al., 2021). However, the literature on cognitive development focuses on typically developing individuals and neglects atypically developing individuals, including those with neurodevelopmental disorders. The neural mechanisms underlying cognitive development in one of these neurodevelopment disorders, ADHD, are incompletely understood. I seek to address this gap in the literature by exploring three questions: (1) can cognition be predicted in a large group of children and adolescents diagnosed with ADHD using functional neural activity during movie watching? (2) does prediction accuracy remain constant across development? and (3) if not, what are the neural mechanisms associated with developmental differences in predicting cognition? To predict cognitive abilities in children and adolescents, I first acquired and preprocessed neuroimaging and phenotypic data in a cohort of ADHD and typically developing youth. The preprocessed data (i.e., individual functional connectivity profiles) were then fed into machine learning models to predict six cognitive abilities. After the models successfully predicted five of the cognitive abilities in the ADHD dataset, I divided the dataset into three age bins and repeated the modeling to test if the prediction accuracy would remain constant. And lastly, to explore the neural mechanisms underlying development, I used the models trained on one age bin and tested them on a different age bin.

2.1 Data and Preprocessing

2.1.1 Participants

The participant data was obtained from the Healthy Brain Network (HBN) biobank collected by the Child Mind Institute (Alexander et al., 2017). The HBN is an ongoing initiative to create and share multimodal data from thousands of New York City children and adolescents between the ages of 5 to 21. The biobank uses a community-referred recruitment model with advertisements to parents, community members, educations, and local care providers. Participants were screened and excluded if there were impairments that would interfere with the study procedure, safety concerns, or medical concerns. The study was approved by the Chesapeake Institutional Review Board and details on the HBN biobank can be found at: http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/

I acquired 1,116 participants aged 6 to 16 from HBN releases 1 to 8 that included a T1-weighted anatomical MRI scan, a BOLD functional MRI scan, and their phenotypic data. From the 1,116 participants, 880 of them passed standard MRI preprocessing (see section 2.1.2) without errors and were used to generate individual functional connectivity matrices. Subsequently, all structural and functional MRI data—and functional connectivity matrices—were visually inspected for artifacts and 154 participants were excluded for not meeting the standard level of quality, leaving 726 cleaned functional connectivity matrices. I then filtered for participants with an FSIQ greater than 70, excluding 18 more participants. The last step was to filter the 708 participants for either an ADHD diagnosis or no diagnosis. This left 229 participants with a diagnosis that was not ADHD, 373 participants with a diagnosis that was ADHD, and 106 participants without a diagnosis (typically developing (TD) group). The participants diagnosed with ADHD included comorbidities.

I used the following phenotypic data provided by the HBN: age, sex, clinical diagnosis, and six measures of cognition (see section 2.1.3). The demographic information for the dataset is provided in Table 1. For age, I only included participants between the ages of 6 to 16. In addition to running models across the entire ADHD group (n=373), participants were divided into discrete age bins to examine developmental changes in the neural mechanisms associated with cognition with ages 6-8 representing early childhood (Bin 1; n=114), ages 9-11 representing middle childhood (Bin 2; n=147), and ages 12-16 representing adolescence (Bin 3; n=112). The age 9 cutoff between Bins 1 and 2 was chosen because around this age is when children enter school and when symptoms of neurodevelopmental disorders not previously noticed begin to appear (Berger, 2017). The age 12 cutoff between Bins 2 and 3 is important because the diagnosis of ADHD in the DSM-V states that symptoms must start before the age of 12 and this age is typically when puberty begins, marking the start of adolescence (American Psychiatric Association, 2013; Berger, 2017). Group assignment was based on rounding down the age provided by the HBN phenotypic profile. For example, a child with an age of 8.9 years old was put into Bin 1. The TD group (n=106) was not divided into age bins due to its low sample size. The phenotypic data also included up to ten diagnoses per participant, which were used to group the participants into either the ADHD or TD group. Clinical diagnoses were provided by licensed clinicians after interviews with the parents and child (Alexander et al., 2017). To classify a participant into the ADHD group, each participant's diagnosis list was checked for containing the keyword "ADHD". To classify a participant into the TD group, no diagnosis must be present for the participant.

Table 1: Participant demographics.

For each group and measure, the mean, standard deviation (in brackets), and range are provided.

	Group				
	ADHD				**TD**
Measure	**All**	**Bin 1**	**Bin 2**	**Bin 3**	**All**
N	373	114	147	112	106
Age	10.57 (2.53) 6.03 to 15.98	7.73 (0.76) 6.03 to 8.98	10.34 (0.88) 9.04 to 11.96	13.75 (1.13) 12.03 to 15.98	10.12 (2.78) 6.05 to 16.50
Sex (M/F)	274/99	77/37	118/29	79/33	62/44
WISC FSIQ	100.13 (16.31) 70.00 to 147.00	104.13 (15.47) 73.00 to 138.00	98.56 (15.90) 70.00 to 141.00	98.11 (16.94) 71.00 to 147.00	108.58 (14.20) 76.00 to 145.00

WISC VSI	102.32 (17.12) 57.00 to 155.00	106.38 (16.64) 67.00 to 147.00	100.73 (16.69) 64.00 to 144.00	100.28 (17.46) 57.00 to 155.00	105.21 (15.01) 64.00 to 141.00
WISC VCI	105.04 (15.87) 70.00 to 155.00	108.53 (16.06) 70.00 to 146.00	104.44 (15.34) 70.00 to 155.00	102.29 (15.71) 70.00 to 142.00	110.58 (13.93) 78.00 to 155.00
WISC FRI	101.33 (16.19) 67.00 to 144.00	104.58 (15.40) 67.00 to 140.00	99.14 (15.75) 67.00 to 134.00	100.90 (16.98) 67.00 to 144.00	107.42 (15.11) 69.00 to 155.00
WISC WMI	98.54 (15.18) 62.00 to 138.00	99.53 (14.83) 67.00 to 138.00	97.21 (14.04) 65.00 to 138.00	99.29 (16.77) 62.00 to 135.00	103.92 (14.29) 72.00 to 135.00
WISC PSI	93.65 (15.35) 53.00 to 148.00	97.33 (15.29) 56.00 to 148.00	93.12 (13.94) 56.00 to 123.00	90.59 (16.38) 53.00 to 132.00	106.55 (15.80) 66.00 to 155.00

2.1.2 MRI acquisition and preprocessing

The HBN neuroimaging data consisted of eight releases at the time of this study totaling 1,116 MRI participants. The MRI data consists of a T1-weighted anatomical MRI scan and a functional MRI scan acquired while participants watched a ten-minute movie clip from ‘Despicable Me’ (from 1:02:09 to 1:12:09). All MRI data was collected on a 3 T Siemens scanner using a Siemens 32-channel head coil. The structural MRI scans were acquired in 224 sagittal (TR=2500 ms, TE=3.15 ms, resolution=0.8 x 0.8 mm^2). The functional MRI scans were acquired with a gradient-echo planar imaging pulse sequence (TR=800 ms, TE=30 ms, Flip Angle=31 degrees, whole brain coverage 60 slices, resolution=2.4 x 2.4 mm^2).

The MRI data were preprocessed and analyzed using the Automatic Analysis (AA) toolbox (Cusack et al., 2015), SPM 8, in-house MATLAB scripts, and visual inspection. Preprocessing of functional data included motion correction (using six motion parameters: left/right, anterior/posterior, superior/inferior, chin up/down, top of head left/right, nose left/right), and functional and structural scans were co-registered and normalized to the Montreal Neurological Institute (MNI) template. Functional data were then spatially smoothed using a Gaussian filter (8 mm kernel) and low-frequency noise, such as drift, was removed by high-pass filtering with a threshold of 1/128 Hz. The data was denoised using Bandpass filter regressors with cerebrospinal fluid, white matter signals, motion parameters, their lag-3 2nd-order volterra expansion (Friston et al., 2000), and spikes (based on mean signal variance across volumes) as nuisance regressors.

2.1.3 Cognitive ability

Cognitive ability was measured using scores on the Wechsler Intelligence Scale for Children Fifth Edition (WISC-V; WISC), which uses 16 subtests to measure intelligence in children (David Wechsler, 2014). The WISC-V measures a child’s intellectual ability on five primary indices: Visual Spatial Index (VSI), Verbal Comprehension Index (VCI), Fluid Reasoning Index (FRI), Working Memory Index (WMI), and Processing Speed Index (PSI). VSI measures the ability to evaluate visual details and understand visual spatial relationships to construct geometric designs from a model. VCI measures the ability to access and apply acquired word knowledge. FRI measures the ability to detect the underlying conceptual relationships among visual objects and requires reasoning to identify and apply rules. WMI measures the ability to register, maintain, and manipulate visual and auditory information in conscious awareness. PSI measures the speed and accuracy of visual identification, decision making, and decision implementation.

In addition to the five primary indices, WISC-V also provides a Full Scale IQ (FSIQ) score, which is derived from seven subtests drawn from the five primary indices. Thus, the FSIQ represents general intellectual ability across a diverse set of cognitive abilities. Both the five primary indices and FSIQ are on a standard score metric with a mean of 100 and a standard deviation of 15. The primary index scores range from 45 to

155, while the FSIQ score ranges from 40 to 160. For both the primary indices and FSIQ, scores between 90 to 109 are considered average for a typically developing child.

2.1.4 Functional connectivity

Functional connectivity matrices were generated using 264 regions-of-interest (ROI) as defined in the Power et al. (2011) atlas. Individual ROIs comprised of spheres of 5 mm in radius with spatial smoothing full-width half maximum of 6 mm and z-score standardization. The activity in each sphere was then correlated to every other sphere in the brain, resulting in a 264 x 264 functional connectivity matrix. This process was repeated for every participant. To vectorize the matrix, the top-right triangle was extracted while discarding the diagonal and bottom-left triangle values, resulting in a 34,716 x 1 vector.

2.2 Computational Models

I used two computational models to capture the relationship between functional brain activity and measures of cognitive ability: partial least squares (PLS) regression and Ridge regression. Ridge and PLS models are ideal for high-dimensional input data with multicollinearity among dependent variables and have built-in anti-overfitting (regularization) parameters.

2.2.1 Standard scaler

Before fitting each model, the training dataset was used to fit a standard scaler model. The standard scaler is used to rescale features such that they have the properties of a standard normal distribution with a mean of zero and a variance of one. To avoid data leakage between the training and testing set, the standard scaler was only fit on the training set and was applied to both the training and testing set. The features are standardized because regularized linear models assume that all features are centered around zero and have variance in the same scale. If a feature has variance on a different scale compared to other features, this feature might dominate the objective function used to train the model. For example, if one feature ranged from 0-1 and another feature ranged from 100-900, then the feature with the 100-900 range may trigger large feature weight updates, which can prevent the model from converging.

2.2.2 Partial least squares

Partial least squares (PLS) is a statistical method of finding a linear regression model by projecting the samples (X) and the targets (y) to a new latent space. A PLS model will search the multidimensional sample space that explains the maximum multidimensional variance in the target space. That is, PLS searches for the latent variables most strongly associated between the sample and target data. For this study, I used the univariate version of PLS, which is a form of regularized linear regression. The univariate version is in the same class as Ridge regression and principal component regression where the number of components controls the strength of regularization (Wegelin, 2000).

2.2.3 Ridge regression

Ridge is another method in the class of regularized linear regression but instead of using partial least squares, it uses complete least squares with L2 regularization. Regularization is a technique to impose constraints on the feature weights of a linear model to reduce overfitting on the data, thereby improving generalizability. L2 regularization, also known as Ridge, implements regularization by adding the square of the absolute value of the feature weights in the least-squares penalty. This penalty on the size of feature weights enables the model to be more robust to the collinearity of features by sparsifying the feature weights.

2.3 Model Training and Validation

2.3.1 Hyperparameter search

To determine the optimal number of components for the PLS model, a hyperparameter search was performed by varying the number of components and selecting the number of components that resulted in the highest score. The explored range for the number of components was between two and six (inclusive) with four components resulting in the highest test score on all cognitive measures. To determine the optimal alpha value for the Ridge model, a hyperparameter search was done by varying alpha and selecting the value that resulted in the highest test score. The explored range for alpha was between 1 to 10,000 (exclusive) with a step size of 100. The hyperparameter search was conducted using the 10-fold 10-repeat cross validation scheme described below.

2.3.2 Cross validation

After optimizing for the component parameter for PLS and the alpha parameter for Ridge, both models were cross-validated using a 10-fold 10-repeat scheme. The scheme starts by splitting the dataset into 10 equal chunks where 9 of the chunks are used to train the model and the remaining 1 chunk is used to test the model. This process is repeated 9 times where each chunk is used for testing and the rest are used for training, resulting in 10 folds and 10 Pearson r scores. After these 10 folds, the dataset is shuffled and a new set of 10 folds is generated. This process is repeated 9 more times, resulting in 100 variations in the training and testing set and 100 Pearson r scores. The final cross-validation score is the mean of these 100 Pearson r scores. Cross-validation was performed to avoid overfitting as the model's performance on a random sampling of the data may not represent the model's predictive performance on unseen data. Thus, by splitting the dataset into 100 variations for training and testing, I achieved a more robust and accurate estimate of the model's performance on the data.

2.3.3 Evaluation

To evaluate the model's performance, I used the Pearson correlation (r) between the predicted and actual cognitive ability scores. For example, a set of participants will have a cognitive ability score associated with each participant; these are the true scores. I can also input the set of participants' functional connectivity matrices into a computational model and have the model predict each participant's cognitive ability score; these are the predicted score. To compare how close the predicted cognitive scores are to the true cognitive scores, one measure is the Pearson correlation. This allow me to compare my results to Finn & Bandettini (2021), Sripada et al. (2020), and Tian & Zalesky (2021) and captures whether the model is predicting the relationship between functional connectivity and cognitive ability. To interpret the Pearson correlation, an absolute value of one implies that a linear equation perfectly describes the relationship between the input data (functional connectivity matrix) and the output data (cognitive ability score). A positive correlation implies that as the input increases, the output increases. A negative correlation implies that as the input increases, the output decreases. A correlation of zero implies that there is no linear relationship between the variables.

2.3.4 Permutation testing

Permutation statistics were used to evaluate the significance of the cross-validation score, which is the correlation between the predicted and actual cognitive scores. A permutation-based p-value represents how likely the performance of a model—the cross-validation score—would be obtained by chance. The null hypothesis is that the model fails to find any statistical dependency between features (functional connectivity) and targets (cognitive ability). This manifests as the model predicts random cognitive scores, which is reflected in random correlations between the predicted and actual cognitive scores. I can generate a null distribution of scores reflecting the null hypothesis by randomly permuting or shuffling the targets, and then training and testing the model on this shuffled dataset. By randomly shuffling the targets, I break any dependency between features and targets. I shuffled the targets 500 times, generating 500 permutation scores to form a null distribution. The p-value is calculated as the fraction of permutation scores greater than the cross-validation score. For example, if 5 of the 500 permutation scores were greater than the cross-validation score, then the calculation outputs a p-value of 0.01. A low p-value implies a low probability that the performance of the model was obtained by chance. A high p-value implies either a lack of dependency between features and targets, or that the model was not able to learn the dependency in the data. In this case, using a different model that can capture the dependency in the data may result in a lower p-value.

To correct for multiple hypotheses, I used the max-statistic method (Nichols & Hayasaka, 2003). This method combines the null distribution among a group of tests into one null distribution for the entire group. For this study, I combined the null distributions for all cognitive measures within each age group for each model. The 95th percentile value was used as the threshold for significance representing an alpha of 0.05. Cross-validation scores above the threshold were considered significant with max-statistic correction, while scores below the threshold were not considered significant.

2.4 Model Feature Weight Analysis

After the models were trained, I analyzed the model’s feature weight using two methods. The first method was to measure the feature weight reliability using the intraclass correlation coefficient (ICC). This method was used by Tian et al. (2021) to measure the reliability of feature weights between different computational models. The second method was to use the feature weights trained on one subset of the dataset and test them on a different subset. This out-of-sample cross-validation—referred to as cross-prediction—was only performed on the ADHD group.

The ICC is a measure of how correlated the same variables are in different observations and can range from zero to one with an ICC of one indicating identical values for variables in different observations. For my purposes, ICC was used to measure how similar the PLS feature weights were to the Ridge feature weights. Thus, I could test whether both models were using similar or different feature weights to predict cognitive ability. This use-case of ICC is identical to a one-way ANOVA fixed effects model.

I performed cross-prediction by training a model on one age bin and testing it on another age bin. For instance, a model trained on Bin 1 with FSIQ was used to predict FSIQ in Bin 2. This allowed me to explore the generalizability of the models, and the shared and distinct feature weights across age bins. To perform the cross prediction, I started by splitting all age bins into training and testing sets using the 10-fold 10-repeat cross validation scheme. I then trained a model on the in-group age bin and tested it on the out-group age bin. This process was repeated for all 100 folds and the final score was the average score from these folds. I also applied permutation testing to this final score by randomly permuting the cognitive abilities within their respective groups and rerunning the modeling.

I extracted and visualized the model’s feature weights to better understand which network connections it found important to predict cognition. A model’s feature weights are a vector where each value represents the weight (or importance) associated with a feature (functional connection). This vector was multiplied by a participant’s functional connectivity matrix and the result was added with the linear intercept (i.e., a single value) to calculate the predicted cognitive score. This score was compared against the participant’s actual cognitive score to update the feature weights. Thus, the feature weights represent a heat map of important functional connections for predicting a cognitive ability. To obtain the final feature weights from the 10-fold 10-repeat cross-validation scheme, I averaged the 100 feature weights generated by each model in each fold. This resulted in a 34,716 x 1 feature weight vector where each value is the weight assigned to a feature. This node-level feature weight vector was condensed to a network-level feature weight vector by taking the mean of all connections between two networks to get one weight per network-network pair instead of one weight per node-node pair. By repeating this procedure for all 13 Power atlas networks, I ended up with a 13 x 13 network-level feature weight matrix. This reduction was performed to make the feature weights more interpretable and to focus on a network-level analysis. While these feature weights have information for understanding how the model predicts cognition, they do not have information on how cognition develops.

To explore how cognition develops, I subtracted the feature weights of models trained on early childhood (Bin 1) from models trained on middle childhood (Bin 2) with respect to cognitive ability. I then took the absolute value of the feature weight differences. The absolute value was performed to highlight the magnitude of change between Bin 1 and Bin 2 feature weights, and to deemphasize the direction of change. The absolute-value-feature-weight-differences matrix represents the network connections that change in importance between Bins 1 and 2. Connections with large changes in importance are considered as “distinct” network profiles representing the largest differences, while connections with small changes in importance are considered as ‘shared’ network profiles representing the smallest differences. Both the distinct and shared network profiles are important when considering cognitive development as connections that change are equally important to connections that do not change between early and middle childhood.

2.5 Data Pipeline

An overview of the pipeline for the neuroimaging data is shown in Figure 1. There were three overall stages in the pipeline: preprocessing, modeling, and analysis. The preprocessing stage was used to clean the data of artifacts and noise, and to generate functional connectivity matrices. The modeling stage was when the data was used to train machine learning models to predict cognitive ability from functional connectivity. The last stage, analysis, extracted the model's feature weights to gain insights into the neural mechanisms underlying cognition and cognitive development. For the phenotypic data, I only excluded participants with an incomplete WISC assessment or with an FSIQ below 70.

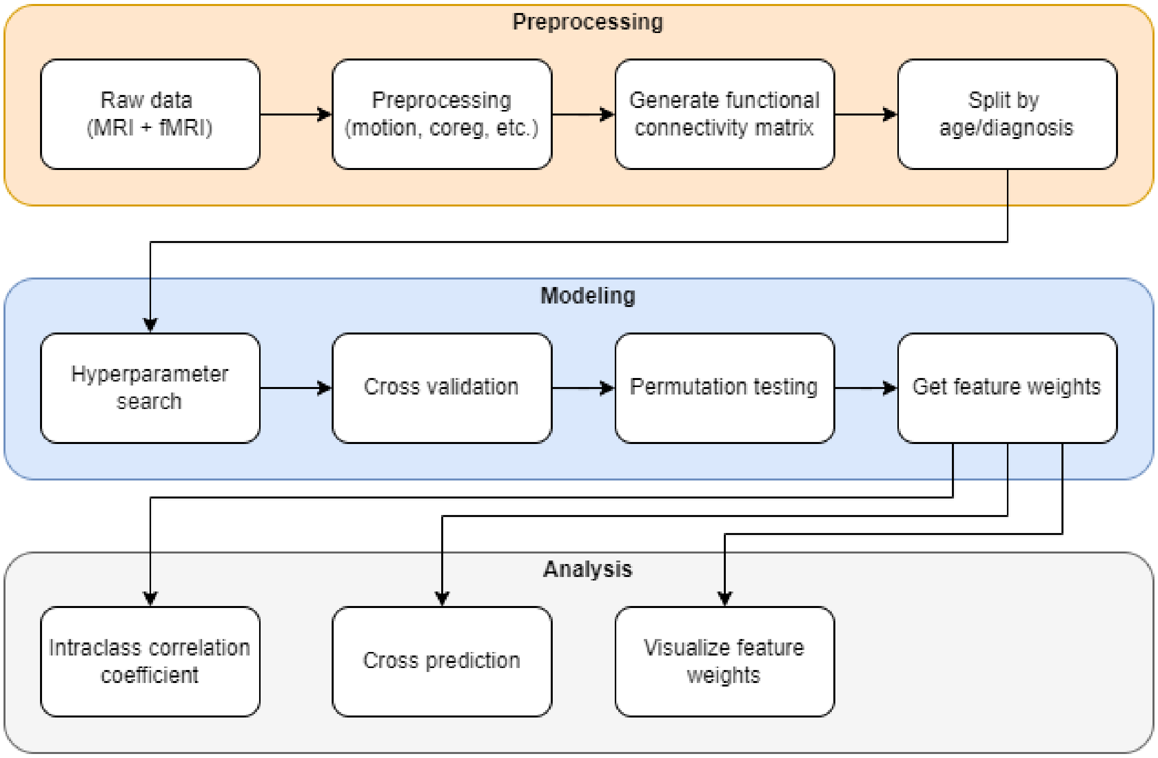


Figure 1: Processing stages for the neuroimaging data.

There are three overall stages to the data pipeline: preprocessing, modeling, and analysis. Preprocessing involves correcting the raw MRI and fMRI data for motion, coregistering the structural and functional images, normalizing to a standard template, generating a functional connectivity matrix, and splitting the participants by age or diagnosis. Next is modelling and it starts with searching for the optimal parameters for the model, then training and validating the model using the functional connectivity matrices, randomly permutating the data, and ends with extracting the model’s feature weights. Lastly, the feature weights were analyzed by calculating the intraclass correlation coefficient, using the weights to cross-predict cognition in a different age bin, and visualizing the feature weights.

Chapter 3

3 Results

3.1 Predicting age and sex in ADHD and TD

I started with predicting age and sex because the results can be used to validate my data against the age and sex prediction results from the literature. Using Ridge regression, I successfully predicted the age (r^2=0.45, p=.01) of individuals in the ADHD group (n=373) and predicted their sex—using a Ridge classifier—with an accuracy of 74% (p<.001). Similarly, I successfully predicted the age (r^2=0.13, p=.01) and sex (60% accuracy, p=.01) of individuals in the TD group (n=106). The sex prediction was class-balanced to ensure that the model was not constantly predicting the most prevalent sex. These findings are consistent with prior work using computational models to predict both age and sex based on resting-state functional connectivity (Dosenbach et al., 2010; Fair et al., 2013; Rudolph et al., 2017). It is important to note that less of the variance was explained when predicting age and sex in the TD group. This may suggest that the neural connectivity data for individuals in the TD group were more heterogeneous than the ADHD group.

3.2 Predicting cognitive ability in ADHD and TD

For the entire ADHD dataset (n=373) and entire TD dataset (n=106), Table 2 shows the results for predicting cognitive performance—using neural connectivity profiles in response to movie-watching—in children and adolescents. When a Ridge regression model was applied to each cognitive ability, I found that the model could predict FSIQ (r=0.38, p=.002), VSI (r=0.31, p=.002), VCI (r=0.39, p=.002), FRI (r=0.30, p=.002), and WMI (r=0.21, p=.004), but failed to predict PSI (r=0.05, p=.26) in the group of participants diagnosed with ADHD. Conversely, in TD, I could not predict FSIQ (r=0.04, p=.42), VSI (r=0.16, p=.11), VCI (r=0.20, p=.05), FRI (r=-0.07, p=.73), WMI (r=0.12, p=.21), and PSI (r=-0.06, p=.70). These p-values were corrected for multiple comparisons using the max-statistic method. To determine whether these results were driven by model choice, I replicated the analysis using a partial least squares (PLS) model. I found no difference between the performance of the two models, except that VCI (r=0.23, p=.04) could be predicted in TD using PLS. Based on these results, I excluded the TD group from further analyses due to the inability to predict any of the six cognitive abilities at statistical significance and with consistency. Moreover, given the similar results profile between Ridge and PLS, I focused on Ridge in subsequent analyses.

Table 2: Scores for predicting six cognitive abilities in ADHD and TD using partial least squares and Ridge regression.

The Pearson r correlation test score represents the linear correlation between the model's predicted values of the cognitive ability and the true values. The p-value was calculated by comparing the observed Pearson r score to a null distribution of Pearson r scores generated from 500 random permutations of the dataset. Both the PLS and Ridge models predicted FSIQ, VSI, VCI, FRI, and WMI in the ADHD group at significance (p<.011) but failed to predict PSI (p=.23). For the TD group, only VCI was predicted at significance (p=.04) using PLS. PLS and Ridge achieved similar Pearson r correlation scores on both the ADHD and TD groups. *values indicate statistically significant at p<.05 max-statistic corrected.

	ADHD (n=373)				**TD (n=106)**			
	PLS		**Ridge**		**PLS**		**Ridge**	
WISC Primary Index	Pearson r	P-value	Pearson r	P-value	Pearson r	P-value	Pearson r	P-value
Intelligence Quotient (FSIQ)	0.37	.002*	0.38	.002*	0.04	.388	0.04	.417
Visual Spatial (VSI)	0.28	.002*	0.31	.002*	0.14	.129	0.16	.107
Verbal Comprehension (VCI)	0.37	.002*	0.39	.002*	0.23*	.035*	0.20	.051
Fluid Reasoning (FRI)	0.30	.002*	0.30	.002*	-0.07	.737	-0.07	.734
Working Memory (WMI)	0.17	.011*	0.21	.004*	0.08	.259	0.12	.213
Processing Speed (PSI)	0.06	.227	0.05	.257	-0.10	.791	-0.06	.698

To gain a better understanding of the network connections that the models extracted for predicting cognitive ability, I visualized the Ridge feature weights trained on the ADHD group for FSIQ, VSI, VCI, FRI, and WMI (Figure 2; left column). For FSIQ, the model assigned large positive weights to internetwork connections between memory retrieval and dorsal attention, and to intranetwork connections within memory retrieval. In contrast, the same model assigned large negative weights to connections between dorsal and ventral attention, and between subcortical with sensory/somatomotor (mouth and hand) and cingulo-opercular task control. For VSI, strong positive weights were again assigned between memory retrieval and dorsal attention, and strong negative weights were assigned between subcortical and sensory/somatomotor (mouth). For VCI, networks connected with the memory retrieval network, such as memory retrieval and cerebellar, positively predicted VCI performance, while connections between dorsal and ventral attention, ventral attention with subcortical, and visual with sensory/somatomotor (mouth) negatively predicted VCI performance. For FRI, I found a similar pattern of large positive weights as seen in FSIQ, namely connections between memory retrieval and dorsal attention, and connections within memory retrieval. A strong negative weight was found between frontoparietal task control and visual. And lastly, for WMI, I noticed a large positive weight for intranetwork connections in sensory/somatomotor (mouth) and in memory retrieval. To summarize across all cognitive measures, I found that the models consistently assigned the largest positive weights to intranetwork connections within two networks: memory retrieval and sensory/somatomotor (mouth), while internetwork connections with the largest positive weights were between memory retrieval and dorsal attention, and between memory retrieval and cerebellar. The largest negative weights were commonly assigned to internetwork connections between frontoparietal task control and visual, between dorsal and ventral attention, and to networks connected with subcortical areas.

3.3 Developmental changes in links between functional connectivity and cognitive abilities in ADHD

To examine developmental changes in the relationship between neural connectivity profiles and cognitive ability, I divided the ADHD data into three age bins and independently modeled the six WISC-V measures for each age bin (Table 3). The model successfully predicted FSIQ (r=0.27, p=.02), VSI (r=0.24, p=.02), and VCI (r=0.22, p=.03) for Bin 1 (ages 6-8), and FSIQ (r=0.35, p=.002), VSI (r=0.21, p=.02), VCI (r=0.35, p=.002), FRI (r=0.31, p=.004), and WMI (r=0.29, p=.004) for Bin 2 (ages 9-11). Interestingly, the model did not predict (p>.17) any WISC-V measure for individuals in Bin 3 (ages 12-16). All p-values were corrected for multiple comparisons using the max-statistic method. To check if the larger sample size of Bin 2 (n=147) affected the results, I repeated the modeling on a reduced-sample-size Bin 2 (n=113) that closely matched the sample sizes of Bins 1 and 3, and found that I could still predict FSIQ (r=0.37, p=.002), VSI (r=0.27, p=.01), VCI (r=0.37, p=.002), FRI (r=0.30, p=.006), WMI (r=0.35, p=.002), but not PSI (r=0.04, p=.38). Thus, it is unlikely that the sample size of Bin 2 affected the ability to successfully predict cognitive abilities. Based on these results, subsequent analyses—such as feature weight visualization—focused on Bins 1 and 2. The Ridge feature weights for predicting FSIQ, VSI, VCI, FRI, and WMI in Bin 1 and Bin 2 are shown in Figure 2 (center and right columns respectively).

Table 3: Scores for predicting six cognitive abilities in ADHD across three age bins using Ridge.

Bin 1 represents early childhood (ages 6-8), Bin 2 represents middle childhood (ages 9-11), and Bin 3 represents adolescence (ages 12-16). The Ridge model successfully predicted FSIQ, VSI, and VCI in Bin 1 (p<.03); FSIQ, VSI, VCI, FRI, and WMI in Bin 2 (p<.02); and no cognitive ability in Bin 3 (p>.17). *values indicate statistically significant at p<.05 max-statistic corrected.

	Age Bins					
	Bin 1 (n=114)		**Bin 2 (n=147)**		**Bin 3 (n=112)**	
WISC Primary Index	Pearson r	P-value	Pearson r	P-value	Pearson r	P-value
Intelligence Quotient (FSIQ)	0.27	.019*	0.35	.002*	0.11	.177
Visual Spatial (VSI)	0.24	.017*	0.21	.021*	0.09	.229
Verbal Comprehension (VCI)	0.22	.027*	0.35	.002*	0.04	.403

Fluid Reasoning (FRI)	0.05	.347	0.31	.004*	-0.01	.518
Working Memory (WMI)	0.05	.357	0.29	.004*	0.10	.183
Processing Speed (PSI)	-0.09	.792	0.06	.263	0.09	.257

When predicting FSIQ for Bin 1 participants, the model found that the strength of internetwork connections between memory retrieval and dorsal attention, and intranetwork connections within the memory retrieval network, were important positive predictors of FSIQ, while connections between dorsal and ventral attention were important negative predictors. In comparison, predicting VSI shared the same positive importance of the memory retrieval and dorsal attention connection as FSIQ, but the model differed in the strong negative weight for the connection between memory retrieval and sensory/somatomotor (mouth). For VCI, the model learned strong positive weights for connections between cingulo-opercular task control and sensory/somatomotor (mouth), and for internal sensory/somatomotor (mouth) connections. Strong negative weights were learned for cerebellar and sensory/somatomotor (mouth), and for dorsal and ventral attention. As the models trained on Bin 1 were not able to predict FRI and WMI at significance, I will not interpret those feature weights. When taken together, the feature weights for FSIQ, VSI, and VCI had positive weights for network connections between memory retrieval and dorsal attention, within memory retrieval, within sensory/somatomotor (mouth), and between cingulo-opercular task control and sensory/somatomotor (mouth). Negative weights were learned for connections between dorsal and ventral attention, between memory retrieval and sensory/somatomotor (mouth), and between cerebellar and sensory/somatomotor (mouth). Next I turn to models trained on Bin 2.

For the Bin 2 feature weights, I noticed a general pattern of less-extreme feature weights across all cognitive measures. This is seen in Figure 2 (right column) as the Bin 2 feature weights had fewer darker-colored cells representing fewer extreme feature weight values in contrast to the entire ADHD group and Bin 1 feature weights. This suggests that the model is not relying on any specific network connection(s) across the brain, but instead is using a distributed approach to predict cognitive ability. One connection did stand out, however, for predicting VCI: the connections within the sensory/somatomotor (mouth) network. This intranetwork connection was given a large negative weight, implying that although it is useful to predict VCI performance, it does so by decreasing the predicted VCI score. Interestingly, this connection was assigned a large positive in Bin 1, which shows that the models switched from a positive to a negative weight from Bin 1 to Bin 2 when predicting VCI.


Feature weights to predict FSIQ in All
Feature weights to predict FSIQ in Bin 1
Feature weights to predict FSIQ in Bin 2
0.008
0.006
0.004
0.002
0.000
−0.002
−0.004
−0.006
Feature weights to predict VSI in All
Feature weights to predict VSI in Bin 1
Feature weights to predict VSI in Bin 2
0.008
0.006
0.004
0.002
0.000
−0.002
−0.004
Feature weights to predict VCI in All
Feature weights to predict VCI in Bin 1
Feature weights to predict VCI in Bin 2
0.006
0.004
0.002
0.000
−0.002
−0.004
−0.006
Feature weights to predict FRI in All
Feature weights to predict FRI in Bin 1
Feature weights to predict FRI in Bin 2
0.008
0.006
0.004
0.002
0.000
−0.002
−0.004
Feature weights to predict WMI in All
Feature weights to predict WMI in Bin 1
Feature weights to predict WMI in Bin 2
0.010
0.008
0.006
0.004
0.002
0.000
−0.002
−0.004
VIS
FPN
DMN
SMH
SMM
CON
AUD
SAL
MEM
VAN
CER
SUB
DAN

Figure 2: Feature weights used to predict five cognitive abilities in the entire ADHD group, Bin 1, and Bin 2.

Each row represents one of five WISC measures: FSIQ, VSI, VCI, FRI, and WMI. Each column represents one of three ADHD groups: All (ages 6-16), Bin 1 (ages 6-8), and Bin 2 (ages 9-11). Each scale applies to the feature weight matrices in the row. A feature weight matrix represents the average feature weight for all connections between two networks shown for all networks. Darker cells in the feature weight matrix represent more extreme values, while lighter cells represent values closer to zero. Red cells represent positive values (increases in value for that network connection increased the predicted cognitive score), while blue cells represent negative values (increases in value for that network connection decreased the predicted cognitive score). Diagonal cells represent intranetwork connections, while off-diagonal cells represent internetwork connections. The networks are visual (VIS), frontoparietal task control (FPN), default mode (DMN), sensory/somatomotor (hand; SMH), sensory/somatomotor (mouth; SMM), cingulo-opercular task control (CON), auditory (AUD), salience (SAL), memory retrieval (MEM), ventral attention (VAN), cerebellar (CER), subcortical (SUB), and dorsal attention (DAN).

3.4 Intraclass correlation coefficient in ADHD

To explore the similarity between PLS and Ridge, I calculated the intraclass correlation coefficient (ICC) between each model’s feature weights trained on the ADHD group for each cognitive ability (Table 4). The ICC revealed strong correlations (> 0.90) in the weights produced by the Ridge and PLS models, which suggest convergence between the features that PLS and Ridge find important. Specifically, the lowest ICC (0.92, p<.001) was found for Bin 1 in PSI, while the second-highest ICC (0.98, p<.001) was found for Bin 2 in FSIQ, WMI, PSI, and for Bin 3 in PSI. The highest ICC (0.99, p<.001) was found for Bin 1 in FRI. Thus, the consistency between model output (Table 2) and feature weights (Table 4) strongly suggests that both PLS and Ridge used near-identical network patterns to predict cognition.

Table 4: Intraclass correlation coefficients between partial least squares (PLS) and Ridge feature weights.

The ICC describes how strongly values for the same variables resemble each other. A high ICC implies similar feature weights between PLS and Ridge, while a low ICC implies dissimilar feature weights. The lowest ICC was 0.92 (p<.001) when predicting PSI in Bin 1 and the highest ICC was 0.99 (p<.001) when predicting FRI in Bin 1. Overall, the high ICC values for all cognitive measures across all groups suggest that PLS and Ridge used very similar feature weights when predicting cognition. All values have a p<.001 and were corrected for multiple comparisons using FDR.

WISC Primary Index	All Ages	Bin 1	Bin 2	Bin 3
Intelligence Quotient (FSIQ)	0.96	0.97	0.98	0.97
Visual Spatial (VSI)	0.96	0.95	0.94	0.93
Verbal Comprehension (VCI)	0.96	0.97	0.96	0.96
Fluid Reasoning (FRI)	0.96	0.99	0.96	0.97
Working Memory (WMI)	0.95	0.96	0.98	0.93
Processing Speed (PSI)	0.96	0.92	0.98	0.98

3.5 Cross-prediction across age bins in ADHD

Using cross-prediction (out-of-sample cross-validation), one of the most robust tests of generalizability, I found that models trained on Bin 1 and tested on Bin 2 can successfully predict FSIQ (r=0.33, p=.002), VSI (r=0.36, p=.002), VCI (r=0.32, p=.002), and FRI (r=0.15, p=.02) (Figure 3). I also found the reverse pattern—a model trained on Bin 2 and tested on Bin 1—can successfully predict FSIQ (r=0.36, p=.002), VSI (r=0.40, p=.002), VCI (r=0.30, p=.002), and FRI (r=0.20, p=.01). This was surprising because I had hypothesized that the models would extract unique network patterns for each age bin, and that those unique patterns would not generalize to other developmental stages. Instead, the ability to cross-predict FSIQ, VSI, VCI, and FRI from Bin 1 to Bin 2, and from Bin 2 to Bin 1, suggests that the models extracted similar, generalizable network patterns to predict these cognitive abilities in both Bins 1 and 2. Although the models successfully cross-predicted FSIQ, VSI, VCI, and FRI, it failed to cross-predict WMI (r=0.03, p=.35) and PSI (r=0.07, p=.18) when trained on Bin 1 and tested on Bin 2. It also failed to cross-predict for the reverse pattern—trained on Bin 2 and tested on Bin 1—in WMI (r=0.03, p=.37) and PSI (r=0.03, p=.40). One explanation is that the neural mechanisms that support WMI and PSI could be distinct between early childhood (Bin 1) and middle childhood (Bin 2). The hypothesis that WMI and PSI are supported by distinct neural mechanisms between Bins 1 and 2, however, is not directly supported by my findings because there could be shared neural mechanisms between WMI and PSI. But since the models did not predict—and thus did not capture any neural mechanisms to predict—WMI within Bin 1 and PSI within Bins 1 and 2, the hypothesis remains speculative. This hypothesis does not apply to the shared neural mechanisms associated with FSIQ, VSI, VCI, and FRI because the models did capture a neural mechanism for predicting these cognitive abilities within Bins 1 and 2. This neural mechanism was then used to successfully cross-predict FSIQ, VSI, VCI, and FRI in participants in a different age bin, so the neural mechanism predicted cognition in both age bins. Having found that the models trained on one age group can be used to predict cognition in another age group, is it possible to identify which connections were shared and thus drove the cross-prediction, and which connections were distinct between age bins?

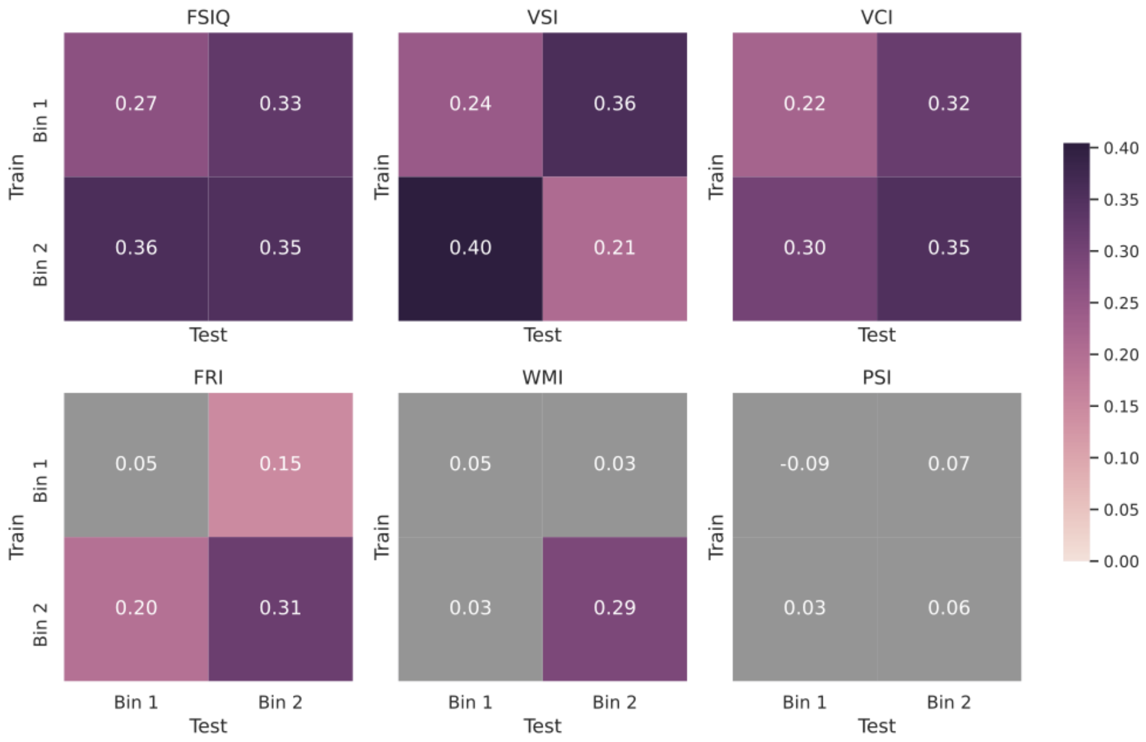


Figure 3: Scores for cross-predicting six cognitive ability between Bin 1 and Bin 2. For each matrix, rows represent the age bin (Bin 1 or Bin 2) the model was trained on, while columns represent the age bin (Bin 1 or Bin 2) the model was tested on. The top-left to bottom-right diagonal represents training and testing the model within the same age bin (same scores as in Table 3), while the bottom-left to top-right diagonal represents the training the model on Bin 2 and testing on Bin 1, and training the model on Bin 1 and testing on Bin 2 respectively. Values within each cell are the Pearson r correlation test score and represent the linear correlation between the model's predicted values of the cognitive ability and the true values. Purple cells indicate statistically significant at p<.05 after being corrected for multiple comparisons using the max-statistic method, while grey cells indicate not statistically significant.

To explore which network patterns were shared and to what extent, I identified the most similar and dissimilar feature weights that were trained on Bin 1 and Bin 2. I did this by subtracting the Bin 2 feature weights from the Bin 1 feature weights with respect to each cognitive measure (e.g., for FSIQ, VSI, VCI, FRI, and WMI). This highlighted both the shared and distinct network connections between Bins 1 and 2. The focus of this analysis was to identify which networks were most similar and most dissimilar across age bins by taking the absolute value of the differences in feature weights. The results of this analysis are presented in Figure 4; the top ten most dissimilar (largest differences) represent distinct network connections, and the top ten most similar (smallest differences) represent shared network connections. For FSIQ, I found the most distinct intranetwork connections included the memory retrieval, cingulo-opercular, and sensory/somatomotor (mouth) networks. Most distinct internetwork connections included those between the memory retrieval network with frontoparietal, sensory/somatomotor (mouth), salience, and dorsal attention networks. Connections between the dorsal attention network with the frontoparietal network were also most distinct between the two age groups. Intranetwork connections were most similar between the age groups within the frontoparietal network. However, a number of internetworks connections had similar feature weight profiles including connections between the ventral attention and default mode networks, between dorsal attention and salience networks, as well as connections between the visual network with the frontoparietal, default mode, and cingulo-opercular networks. I also found connections between subcortical areas and auditory, ventral attention, and cerebellum remained the same between Bin 1 and 2.

The pattern of connections for VSI that were most dissimilar between the two age groups included connections within sensory/somatomotor (mouth) but also between sensory/somatomotor (mouth) and auditory networks, and memory retrieval, in addition to connections between dorsal attention to frontoparietal networks, and cerebellar connections to sensory/somatomotor (hand), cingulo-opercular, and salience networks. The feature weights that remained stable between the two age groups included connections with the default mode, cingulo-opercular task control, and frontoparietal networks.

When predicting VCI, the largest change (i.e., most dissimilar) in feature weights between early childhood (Bin 1) and middle childhood (Bin 2) were found for internal connections within the sensory/somatomotor (mouth) and subcortical networks, and large differences in the feature weights were associated with connections between the memory retrieval network with the default mode and sensory/somatomotor (hand) networks, and connections between sensory/somatomotor (hand) and the dorsal attention network.

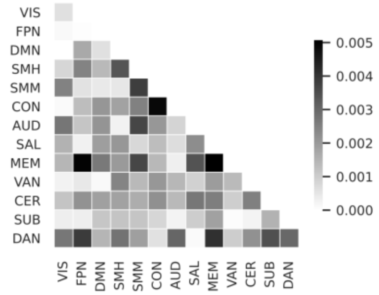
Conversely, the smallest change in feature weights (i.e., most similar) were found with salience and dorsal attention internetwork connections. Similar feature weights were also found for several internetwork connections, which include between memory retrieval and auditory connections, as well as ventral attention network connections to the default mode network, and dorsal attention network connections to the memory retrieval network.

In contrast, dissimilar feature weights between the two age groups associated with predicting FRI included within cingulo-opercular, memory retrieval, and cerebellar network connections, while between-network connections included memory retrieval to frontoparietal, salience, and dorsal attention networks, in addition to connections between the dorsal attention network with the frontoparietal network. Interestingly, there were also large differences in feature weights with cerebellar connections, specifically to sensory/somatomotor (mouth) and cingulo-opercular networks. Key similar network configurations based on feature weights between the two age groups include within default mode and subcortical networks, and similar between networks include default mode network connections with the dorsal attention network, frontoparietal network connections with the salience network, and sensory/somatomotor networks (hand and mouth) connections to auditory, salience, memory, ventral attention, and visual attention networks.

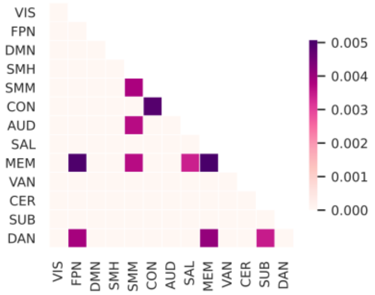
Finally, the most dissimilar feature weights associated with predicting WMI include within cingulo-opercular, memory retrieval, and cerebellar connections, which are the same cortical networks that were most dissimilar for predicting FRI. In addition to dissimilar intranetwork connections, I also found dissimilar internetwork connections, notably between the dorsal attention network and the sensory/somatomotor (mouth), auditory, and memory retrieval networks, as well as between the frontoparietal network and the memory retrieval network. Interestingly, I found no within-network connections that were shared between the two age groups. Feature weight profiles that were the most similar between the two age groups included connections between the salience network with the frontoparietal, sensory/somatomotor (mouth), auditory, and dorsal attention networks; connections between ventral attention network with the default mode network; and connections between the frontoparietal network with both auditory and visual networks. Unlike for other cognitive measures, feature weights associated with the memory retrieval network were not shared between early and middle childhood.

To summarize, the feature weight profiles of the most similar networks between the two age groups across all cognitive measures comprised of four intranetwork connections: the frontoparietal, default mode, subcortical, and dorsal attention networks. The feature weights associated with internetwork connections that were most similar between the two age groups primarily included the frontoparietal, default mode, subcortical, and salience structures, although other networks were also found (but to a lesser degree) to be shared between age groups. Conversely, there were relatively more intranetwork connections that were dissimilar between age groups across the cognitive measures, which predominately included the sensory/somatomotor (mouth), cingulo-opercular, and memory retrieval networks, but also included cerebellar and ventral attention networks. The internetwork connections that were most dissimilar included the memory retrieval, dorsal attention, sensory/somatomotor networks (mouth and hand) networks. Moreover, connections between cingulo-opercular network and other parts of the brain were more often shared than not between the age groups across the different cognitive measures. It is important to note that the model was not able to predict FRI and WMI in Bin 1, thus those feature weights and their differences with Bin 2 do not reflect direct comparisons. That said, my goal was to identify distinct and shared internetwork and intranetwork connections between the two age groups, and not between the different cognitive measures, which is the focus of future work.

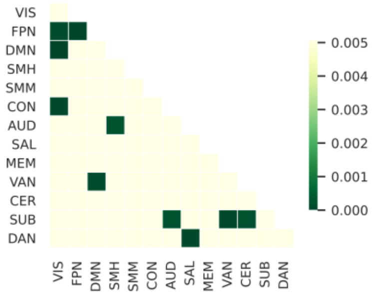
All Bin 1 and 2 differences for FSIQ



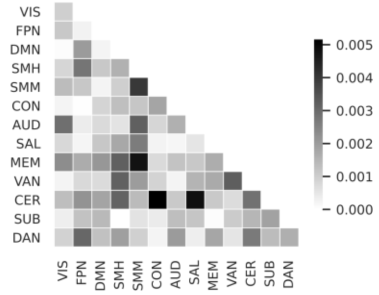
Distinct Bin 1 and 2 connections for FSIQ



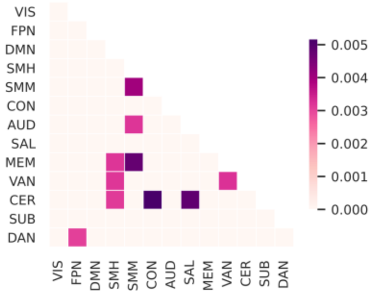
Shared Bin 1 and 2 connections for FSIQ



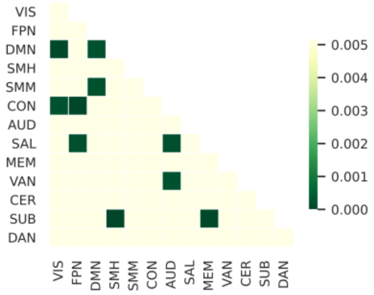
All Bin 1 and 2 differences for VSI



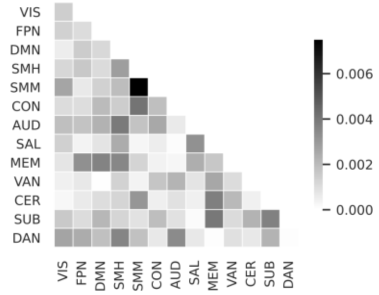
Distinct Bin 1 and 2 connections for VSI



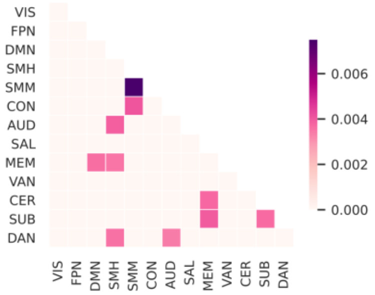
Shared Bin 1 and 2 connections for VSI



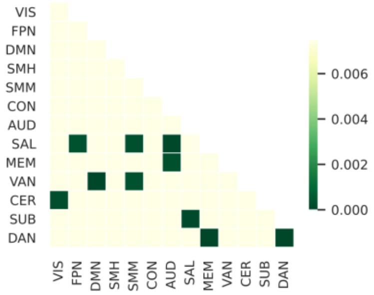
All Bin 1 and 2 differences for VCI



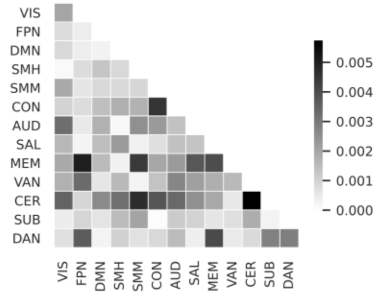
Distinct Bin 1 and 2 connections for VCI



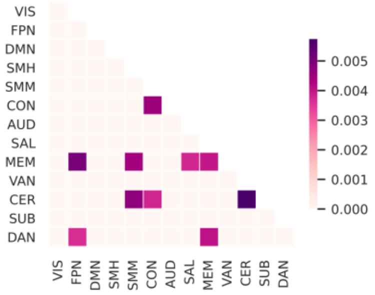
Shared Bin 1 and 2 connections for VCI



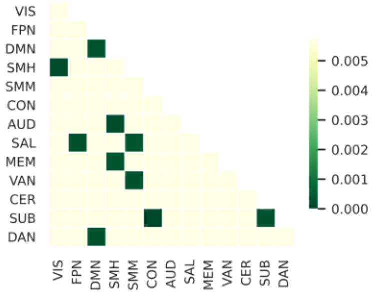
All Bin 1 and 2 differences for FRI



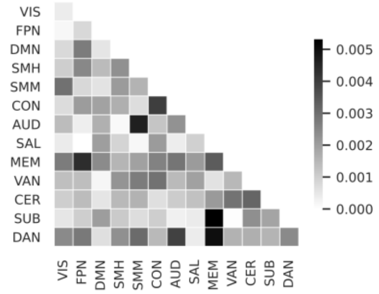
Distinct Bin 1 and 2 connections for FRI



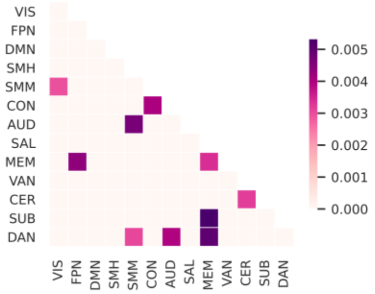
Shared Bin 1 and 2 connections for FRI



All Bin 1 and 2 differences for WMI



Distinct Bin 1 and 2 connections for WMI



Shared Bin 1 and 2 connections for WMI

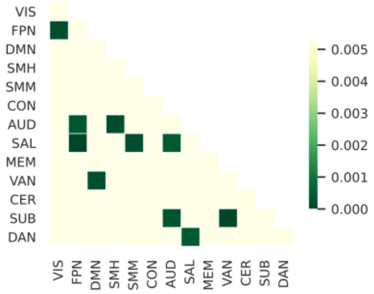


Figure 4: Difference in feature weights between Bin 1 and Bin 2 for five cognitive abilities.

Each row represents one of five WISC measures: FSIQ, VSI, VCI, FRI, and WMI. The left column (grey) represents all feature weight differences between Bin 1 and 2, the center column (pink) represents network connections given dissimilar values for Bin 1 and 2, and the right column (green) represents network connections given similar values between Bin 1 and 2. The dissimilar/distinct network profiles were obtained by thresholding all feature weight differences between Bins 1 and 2 by the ten largest differences. The shared network profiles were obtained by thresholding all feature weight differences between Bins 1 and 2 by the ten smallest differences. For the left and center columns, darker cells represent a larger difference between the feature weights assigned to Bin 1 and 2 when predicting cognition, while lighter cells represent a smaller difference. Diagonal cells represent intranetwork connections, while off-diagonal cells represent internetwork connections.

Chapter 4

4 Discussion

The aim of the current study was to examine three questions: (1) can cognition be predicted in a large group of children and adolescents diagnosed with ADHD using functional neural activity during movie watching? (2) does prediction accuracy remain constant across development? and (3) if not, what are the neural mechanisms associated with developmental differences in predicting cognition? While previous studies have focused on predicting cognition in children (Sripada et al., 2020) and in adults (Tian & Zalesky, 2021) using resting-state stimuli, or in response to naturalistic stimuli, such as movies (Finn & Bandettini, 2021), no study to date has taken a computational modeling approach to predict age, sex, and various aspects of higher-level cognition using neural activity in response to a movie in a group of children and adolescents with ADHD.

Overall, my results indicate that applying machine learning to fMRI data—when recording brain activity during movie watching—is a viable tool for predicting demographic and higher-level cognitive abilities in children and adolescents diagnosed with ADHD. Specifically, in a large cohort of early childhood, middle childhood, and adolescent participants, I was able to predict the age and sex of an individual, and identify neural mechanisms associated with different aspects of cognitive functioning including intelligence, visual spatial, verbal comprehension, fluid reasoning, and working memory abilities, but not processing speed. Moreover, when the ADHD group was divided into distinct age bins, I found that different aspects of higher-level cognition were better predicted than others, which were supported by distinct and overlapping neural mechanisms.

Consistent with my predictions, I also found a specific trajectory associated with cognitive development in my cohort of children and adolescents diagnosed with ADHD. By dividing the ADHD group into three age bins, and I found the model’s ability to predict certain cognitive abilities was dependent on the age group to which the participant belonged. Specifically, I found that I was able to predict IQ, visual spatial, and verbal comprehension abilities for participants in early childhood (ages 6-8); IQ, visual spatial, verbal comprehension, fluid reasoning, and working memory abilities for participants in middle childhood (ages 9-11); but could not predict any cognitive ability for participants in the adolescent age group (ages 12-16). It is important to note that cognitive scores were normalized by age, thus disentangling them from age. One pattern that emerged from my results is an “inverted-U” pattern, which suggests that the link between intra brain functional connectivity in response to movie watching and cognitive ability is strongest and most consistent during middle childhood, weaker and more variable for early childhood and—according to my results—not detectable for adolescence. This is surprising because I could predict five of the six cognitive abilities when all subjects were grouped together and expected that subgroups from this dataset would lead to predicting subgroups of the five cognitive abilities. This expectation holds for early and middle childhood as I could predict three and five cognitive abilities respectively. However, adolescence violated this expectation, suggesting that the link between functional connectivity and cognitive ability in adolescence may either be too weakly represented in the data, or too variable to be captured (cognitive ability is underdeveloped or not brought out by movies for everyone to the same degree) or that linear models are insufficient to capture the link.

The strongest association between functional connectivity and cognitive ability was found for participants in middle childhood. This result is in line with the study by Sripada et al. (2020) who found that higher-order cognitive functions could be predicted in children ages 9-10 from resting state fMRI. One note, however, is that Sripada et al. (2020) did not test their model on older or younger children thus, the inverted-U pattern may be unique to my study.

While I did generate a model to successfully predict some aspects of cognition in early and middle childhood, the cognitive abilities that were best predicted for these two developmental stages were different. One important difference was that the model could predict working memory in middle childhood but not in early childhood. This is interesting because the work by Alloway et al. (2006) in children ages 4 to 11 suggests that the separation of verbal and visuospatial working memory is stable across this development period, which seems contrary to my findings. If working memory is stable during this period, then there should be no changes in verbal and visuospatial working memory from early to middle childhood, and yet the models could predict working memory in middle but not early childhood. One reason for the apparent discrepancy in my results with that of Alloway and colleagues is that while verbal and visuospatial working memory abilities may be stable during this period, the associated neural mechanisms supporting these abilities may undergo differential development and are not stable over time. Thus, the link between brain activity and working memory ability might not be stable in children ages 4 to 11, which may explain why the model could not predict working memory in early childhood but could in middle childhood. In addition, it is known that children ages 6 to 12 diagnosed with ADHD have deficits in working memory, so the development of working memory in children with ADHD may differ from the typically developing children studied by Alloway et al. (2006), or that the link between brain activity and working memory may be weak in children with ADHD (Mayes & Calhoun, 2006; Moura et al., 2019).

Another important difference in what the model could predict was found for fluid reasoning. While the model could predict fluid reasoning in middle childhood, it could not in early childhood. The explanation for the inability to predict working memory may also apply for fluid reasoning as both cognitive abilities rely on the development of the frontoparietal network (Otero, 2017; Ullman et al., 2014), which undergoes considerable change during this developmental period and may differ for those with ADHD (A. Breukelaar et al., 2018; Sun et al., 2020). For instance, Dehn (2017) argues that fluid reasoning depends on working memory to hold information for a period of awareness until the reasoning task is complete. Both cognitive abilities appear to rely on the prefrontal cortex (PFC) as this region is active during processing speed and working memory functioning. Furthermore, ADHD has been associated with deficits in fluid reasoning and fMRI studies implicate a role for the parietal and cerebellar regions (Mano et al., 2019; Tamm & Juranek, 2012). Therefore, developmental changes in the activity of the frontoparietal network may affect the prediction of working memory and fluid reasoning scores.

Although there were differences in predicting cognitive abilities in the three age bins, one similar result across the entire ADHD group was the inability to predict processing speed. This is consistent with the literature as children with ADHD had one of the most pronounced deficits in subtests measuring processing speed (Mayes & Calhoun, 2006; Moura et al., 2019). With deficits in processing speed, there may be a weak relationship between neural activity associated with movie watching and processing speed ability, and/or the representation of processing speed in brain activity may have greater variability, in part due to distinct developmental progression across individuals. In related work, Sripada et al. (2020) had the lowest prediction scores on measures of speed/flexibility out of three higher-order cognitive functions (General Ability, Speed/Flexibility, and Learning/Memory) for 9-10 year old children. They concluded that either the speed/flexibility cognitive function may not be strongly represented in resting state connectivity, the computational models used may not be sufficiently sensitive to their representation, or a combination of both. My study used movie-watching rather than rest, and a different model (i.e., Ridge regression versus principal component analysis), but surprisingly, I obtained a similar result as Sripada et al. (2020). Although not identical to what Sripada and colleagues found—I used different measures of cognitive speed (i.e., WISC-V versus NIH Toolbox)—tracing the origin of the term "Speed/Flexibility" reveals that it was actually renamed from the "Executive Function" component in Thompson et al. (2019). Thus, the low prediction score for Speed/Flexibility in Sripada et al. (2020), and the inability to predict processing speed in this study, may suggest that neither rest nor movie-watching functional connectivity capture the neural mechanisms related to processing speed, which is a component of executive function (Diamond, 2013).

To test whether the models for predicting cognition are dependent on the developmental stage or generalize across age bins, I used a model generated for one age group to predict cognitive ability in another age group. This procedure, referred to as cross-prediction, was only performed on participants in early and middle childhood; the adolescent group was not included because the models could not predict cognition within this stage. The results of the cross-prediction revealed robust out-of-sample cross-validation for IQ, visual spatial, verbal comprehension, and fluid reasoning abilities between early and middle childhood. This implies that the important brain network connections (feature weights) for predicting these four cognitive abilities from functional connectivity—generated from one developmental stage—are also important for predicting the same four abilities in a different developmental stage. I called these important connections for predicting cognition in both developmental stages “shared network connections”. The shared networks for predicting cognition between early and middle childhood were divided into two groups: intranetwork (within) connections and internetwork (between) connections. Shared intranetwork connections were predominately made up of four networks: frontoparietal, default mode, subcortical, and dorsal attention, but also included sensory (e.g., visual and auditory) networks. The shared internetwork connections were comprised primarily of the five networks: frontoparietal, default mode, memory retrieval, dorsal attention, and salience. Interestingly, Sripada et al. (2020) found four out of these five networks when assessing the importance of specific networks to predict general cognitive ability—the frontoparietal, default mode, dorsal attention, and salience networks. To review, these shared intra- and internetwork connections were most important for cross-predicting IQ, visual spatial, verbal comprehension, and fluid reasoning abilities in early and middle childhood. This does not imply that these network connections remain stable during this developmental period—perhaps the connections do change but not in a way that is relevant to cognition. What the shared network connections do imply is that the models did not change their importance for these network connections when predicting cognition for early and middle childhood. It also suggests that these connections are what enable the successful cross-prediction of the four listed cognitive abilities. But why would the models highlight these networks? One possibility is that many of the shared networks, which have been linked to higher-level cognitive processing—such as the frontoparietal, memory retrieval, dorsal attention, and salience networks—bridge cognitive maturity and the degree to which they are recruited during movie watching is similar between the two age groups. That is, young children with greater scores on cognitive abilities are recruiting (or not recruiting) these networks during movie watching to the same degree as older children, while the same relationship is true for early and middle childhood participants with lower scores on cognitive abilities.

Although I found many important shared network connections between the two age groups, the proportion of explained variance remained relatively low. This suggests that the shared networks do not capture all, or even most, of the developmental neural mechanisms supporting higher-level cognition in early and middle childhood. That is, not all network connections were shared for predicting cognition between developmental stages; these other connections are called “distinct network connections”. In contrast to the shared network connections, the distinct network connections for predicting cognition between early and middle childhood were mainly found for intranetwork connections within the sensory/somatomotor (mouth), cingulo-opercular, and memory retrieval networks, but also included subcortical, cerebellar, and ventral attention networks. I found that internetwork connections to and from the sensory/somatomotor networks (mouth and hand), dorsal attention, and memory retrieval networks were also part of the distinct network connections. These network connections represent developmental changes in the neural mechanisms associated with cognition between early and middle childhood. Again, I ask: why did the model highlight these networks? The sensory/somatomotor (hand) and memory retrieval networks connections were also highlighted when predicting general cognitive ability in a cohort of middle childhood participants in Sripada et al. (2020). This may suggest that in the network configuration for sensory/somatomotor (hand) and memory retrieval networks in early childhood changes in ways that are important for predicting cognition in middle childhood. This would explain why both my results and Sripada and colleague’s results find the sensory/somatomotor (hand) and memory retrieval networks as important for predicting cognitive ability in middle childhood. Contrary to what I had predicted, I did not find the frontoparietal network as one of the main distinct networks between early and middle childhood. This is surprising because studies such as Baum et al. (2017) and Deary et al. (2010) implicate changes of the frontoparietal network in the development of executive function and intelligence. That is, I expected that changes in the frontoparietal network connections would be captured by the model to predict cognitive abilities related to executive function (e.g., working memory) and intelligence, but no changes in the model’s importance of the network were found between early and middle childhood. However, a recent study by Caldinelli & Cusack (2022) argues that the frontoparietal network is not a flexible hub during naturalistic cognition in adults. Thus, perhaps in a naturalistic stimuli context— such as the use of movie-watching in my study and Caldinelli and Cusack's study—the frontoparietal network is not flexible and its activity does not strongly change. This hypothesis is further supported by the shared network connections as the frontoparietal network is one of the shared network connections and was found to not strongly change in importance between early and middle childhood.

In the context of computational modeling approaches for findings overlapping feature weights across groups, discovering similar network connections between within- and cross-prediction was expected. There are two possible outcomes for this type of approach: 1) within-prediction without cross-prediction, which implies that there are no shared network connections between the developmental stages, and 2) within-prediction with cross-prediction, which implies that there are shared network connections between the developmental stages. Given this, it was surprising to achieve cross-prediction without within-prediction for fluid reasoning. This may be explained as some neural connectivity profiles in early childhood are more mature-looking and thus can predict fluid reasoning for middle childhood connectivity profiles. But these mature-looking profiles are poor for predicting fluid reasoning in the immature profiles present in early childhood. This is not to say that fluid reasoning has greater variance in early childhood compared to middle childhood as shown in Table 1, but that the link between brain connectivity and fluid reasoning may have greater variance in early childhood compared to middle childhood. Another potential reason the cross-prediction score is greater than the within-prediction score is because of my cross-validation scheme (10-fold 10-repeat). This scheme used 90% of the data to generate the model and the remaining 10% to test the model, which can lead to an inadequate number of test participants, resulting in higher cross-prediction scores than within-prediction scores. However, this is unlikely to account for my results because the within-prediction scores were also obtained from the same 10-fold 10-repeat cross-validation scheme and the within-prediction test scores were not greater than the within-prediction train scores. This suggests that the model had an adequate number of test participants for the within-prediction and, by extension, the cross-prediction. Another question related to the cross-prediction test score is why it is not higher or perfect (i.e., Pearson r=1). In context of the cross-prediction results, the distinct network connections represent part of the reason the cross-prediction score is not higher. This is because these connections differ between early and middle childhood, among many other differences unaccounted for by the model, and thus cannot generalize to other development stages, leading to an imperfect cross-prediction score.

While I found that various aspects of cognition could be predicted and cross-predicted for participants in early and middle childhood diagnosed with ADHD, I could not predict any cognitive ability in the TD group. This is intriguing because other studies have predicted cognition in a typically-developing group of participants (Finn & Bandettini, 2021; Sripada et al., 2020; Tian & Zalesky, 2021). One possible reason is because I had a much smaller sample size (n=106) compared to Finn & Bandettini (2021) (n=176), Sripada et al. (2020) (n=2016), and Tian & Zalesky (2021) (n=958), which leads to weak power in predicting cognition. However, I was able to predict some cognitive measures with a comparable ADHD sample such as the early childhood participants (n=114). Thus, I seem to have enough data but sample size is only one indirect measure of data quality. Another measure is the amount of signal and noise in the data; the signal being the component of brain activity associated to cognition—the information the model uses to predict cognition—and the noise being the component of brain activity unrelated to cognition. This leads to the hypothesis that the TD data has more noise or is more heterogeneous leading to poor predictive performance. It is difficult, however, to distinguish between signal and noise in data because variations in the data might reflect true differences between individuals' brain activity, or it might also reflect artifacts from the process of collecting the data. Although preprocessing does help reduce the noise in the data using assumptions and known noise patterns, it cannot account for all sources of noise and therefore cannot remove all noise from the data. One test of how much noise is in the data is to use it to predict measures such as age and sex. When the TD data was used to predict age and sex, it had lower explained variance and lower accuracy compared to the ADHD group. From this, I argue that sex is represented in neural connectivity more strongly than age followed by cognition. This is supported by the finding that sex is predicted with an explained variance of 53% (Tian & Zalesky, 2021), age is predicted with 42% (Rudolph et al., 2017), and cognition is predicted with 16% (Finn & Bandettini, 2021). Thus, since age was predicted with 13% explained variance for the TD in this study versus the expected 42% from the literature, I believe the TD data to be noisier compared to the ADHD data, leading to poor performance when predicting cognition.

I also generated models to predict the age and sex of individuals in either the ADHD or TD group. I found that I could successfully (above chance) predict age and sex for participants in each group. This is an important result because it demonstrates that a dimensional data-driven approach, such as machine learning, can be used to extract information from the neural connectivity profile to predict both age and sex. The distinct feature-weight profile used by the model likely reflect that male and female (sex was limited to these two categories) children and adolescents have distinct functional patterns of brain activity and are relying on different neural mechanisms to process the movie. This result replicates and expands on previous work that have generated models to predict age and sex. For example, Rudolph et al. (2017) used functional connectivity data from 212 participants aged 10 to 25 years old while they completed a rapid event-related emotional go/no go impulse control task. In the neutral emotional context (no anticipation of a punishment or reward), Rudolph and colleague predicted age with an explained variance of 42% using partial least squares regression. My result (45% explained variance) for predicting age in ADHD are in line with what Rudolph and colleagues found (42% explained variance) for predicting age in typically developing participants. However, unlike Rudolph and colleagues' findings, my found much lower (13%) explained variance for predicting sex in a group of typically developing children and adolescents. Moreover, Tian & Zalesky (2021) were interested in the reliability of feature weights between computational models and used resting-state functional connectivity matrices from 400 participants (aged 22 to 37 years old) to predict sex. Using Ridge regression, they found that a participant's sex could be predicted with a high degree of accuracy (Pearson correlation of 0.73), which is considerably higher than what they found for predicting other phenotypic measures, such as cognition. This lead Tian and Zalesky to conclude that an overt biological attribute, such as sex, could be reliably localized to specific important connections of a predictive model unlike cognitive performance, which could not be reliably localized. I found similar results for my dataset and model as sex could be predicted with a greater accuracy than cognition in both the

ADHD and TD groups. Replicating results from the literature is important because it validates the ADHD and TD datasets used in my study. That is, I confirmed that the data I used contain important information that can be used to make predictions about individuals and specific groups. However, it is important to highlight that both the explained variance for age prediction, and the accuracy for sex prediction, is lower for the TD group than for the ADHD group, suggesting that the TD data may have more noise compared to the ADHD data. If the TD data has more noise, then this might explain why I could not predict any cognitive ability in the TD group, a finding that I return to below.

Importantly, I was able to replicate all findings using a different model: partial least squares. In fact, I found a very high correspondence between the feature weights generated by Ridge and partial least squares as measured by the high intraclass correlations. This suggests that the models' ability to predict cognition is likely not driven by model choice as both the model output (its correlation score) and the model internals (its feature weights) are extremely similar between Ridge and partial least squares. This is similar to the results by Tian & Zalesky (2021) who found little difference between a Lasso model and Ridge's correlation score when predicting fluid and crystalized intelligence. Thus, perhaps in the space of regularized linear models, the choice of model does not lead to significant performance differences.

4.1 Limitations and Future Directions

One limitation of my study is that I was unable to predict cognition in typically developing children, and therefore it is unclear whether the pattern of neural connections associated with the development of cognition I found is specific to children and adolescents with ADHD or generalizes to other groups. That is, without a comparison group, my results may be specific to cognitive development in participants diagnosed with ADHD. Therefore, future studies should focus on predicting various cognitive abilities in a cohort of children and adolescents who are typically developing. This would help uncover which patterns of neural activity are most associated with cognitive development in typically developing children, provide insight as to why I was unable to predict cognition in this group, and clarify whether my findings are specific to ADHD or if they apply to children in general. In addition to TD, it would also be interesting to compare the cognitive development of other neurodevelopmental disorders to ADHD's cognitive development, such as Autism Spectrum Disorder, leading to a richer understanding of cognitive development in typically and atypically developing children.

While I was able to identify specific patterns of neural connectivity that differentially predicted cognition in early and middle childhood, there are many additional avenues to explore the relationship between neural activity and cognition. The goal of the current analysis was to identify whether there were differences between the neural mechanisms associated with cognitive ability in early and middle childhood, however, future studies should examine the direction of the distinct network connections between early and middle childhood, and determine whether those differences are associated with a specific developmental period. Future analyses could probe the specific feature weight differences associated with each developmental stage—early childhood and middle childhood—by not performing the absolute value on the feature weight differences. This would result in a map of feature weight differences that shows increases and decreases in the importance of different network connections between early and middle childhood, and not only changes in importance as in this study. For example, maybe the connection between the memory retrieval and frontoparietal network has a positive weight in early childhood but a negative weight in middle childhood. My current analysis only reveals that this connection changed, but future analyses could explore the direction of change (i.e., from positive to negative).

Another avenue for future work in this area is to explore the use of different computational models and parameters to better capture the relationship between brain activity and cognitive ability. For instance, future studies can use a dataset incorporating functional brain activity captured during different movies, use different computational models, or use different measures of cognitive ability. Replicating the current findings using an ever larger dataset would also be worthwhile; in a recent study, it was found that reproducing brain-wide association studies requires at least thousands of participants, thus a larger dataset is desirable (Marek et al., 2022). In addition to more data, exploring how nonlinear models—such as deep neural networks—capture the relationship between brain activity and cognitive ability may lead to better predictive performance by leveraging more data and using more parameters. One caveat, however, is that these models are more complex and understanding how the model achieves its predictions is more difficult (Voss et al., 2021). And lastly, cognition is not limited to the tasks performed during a WISC-V assessment, thus different tasks that measure different cognitive abilities would lead to a richer understanding of cognition overall.

One final future direction to explore is to associated specific brain networks with specific cognitive abilities using a causal approach. My study used a correlational approach as I did not manipulate the network connections used by the model to predict cognition. Instead, manipulation of the input data could be implemented by removing connections (analogous to a white matter lesion) or by removing networks (analogous to a grey matter lesion) and testing the impact of this change on the model’s ability to predict cognition. The study by Hebling Vieira et al. (2021) implemented the latter approach but found that the removal of no network was sufficient to disrupt model performance. However, the removal of networks could be extended by removing different permutations of pairs, triplets, etc. of networks and testing model performance.

4.2 Conclusion

To conclude, I found that I could model and predict various cognitive abilities in a large group of children and adolescents diagnosed with ADHD using functional neural activity during movie watching, that prediction accuracy does not remain constant across development but instead follows an inverted-U developmental trajectory from early childhood to adolescence, and that certain neural mechanisms linked to higher-level cognition were shared for predicting cognition between early and middle childhood, while a different set of neural mechanisms were distinct for predicting cognition between early and middle childhood. It is important to note the exploratory nature of this study when interpreting these results, and that further investigation into the relationship between ADHD, cognition, cognitive development, and naturalistic stimuli is needed. This could include adding a comparison group, attempting different models, or removing functional connections.

References

A. Breukelaar, I., Williams, L. M., Antees, C., Grieve, S. M., Foster, S. L., Gomes, L., & Korgaonkar, M. S. (2018). Cognitive ability is associated with changes in the functional organization of the cognitive control brain network. *Human Brain Mapping*, *39*(12), 5028–5038. https://doi.org/10.1002/hbm.24342

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). Data Descriptor: An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Scientific Data*, *4*(1), 1–26. https://doi.org/10.1038/sdata.2017.181

Alloway, T. P., & Alloway, R. G. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698–1716. https://doi.org/10.1111/j.1467-8624.2006.00968.x

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

Arnold, L. E., Hodgkins, P., Kahle, J., Madhoo, M., & Kewley, G. (2020). Long-Term Outcomes of ADHD: Academic Achievement and Performance. *Journal of Attention Disorders*, *24*(1), 73–85. https://doi.org/10.1177/1087054714566076

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), R136–R140. https://doi.org/10.1016/j.cub.2009.12.014

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Baum, G. L., Ciric, R., Roalf, D. R., Betzel, R. F., Moore, T. M., Shinohara, R. T., Kahn, A. E., Vandekar, S. N., Rupert, P. E., Quarmley, M., Cook, P. A., Elliott, M. A., Ruparel, K., Gur, R. E., Gur, R. C., Bassett, D. S., & Satterthwaite, T. D. (2017). Modular Segregation of Structural Brain Networks Supports the Development of Executive Function in Youth. *Current Biology*, *27*(11), 1561-1572.e8. https://doi.org/10.1016/j.cub.2017.04.051

Becker, A., Daseking, M., & Koerner, J. K. A. (2021). Cognitive profiles in the wisc-v of children with adhd and specific learning disorders. *Sustainability (Switzerland)*, *13*(17), 9948. https://doi.org/10.3390/su13179948

Berger, K. S. (2017). *The Developing Person Through the Life Span* (Tenth). Worth

Publishers.

Bertolero, M. A., & Bassett, D. S. (2020). Deep neural networks carve the brain at its joints. *BioRxiv*, 36–38. http://arxiv.org/abs/2002.08891

Biederman, J., & Faraone, S. V. (2006). The effects of attention-deficit/hyperactivity disorder on employment and household income. *MedGenMed Medscape General Medicine*, *8*(3), 12. /pmc/articles/PMC1781280/

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *2020-Decem*. https://doi.org/10.48550/arxiv.2005.14165

Buss, A. T., & Spencer, J. P. (2018). Changes in frontal and posterior cortical activity underlie the early emergence of executive function. *Developmental Science*, *21*(4), e12602. https://doi.org/10.1111/desc.12602

Caldinelli, C., & Cusack, R. (2022). The fronto-parietal network is not a flexible hub during naturalistic cognition. *Human Brain Mapping*, *43*(2), 750–759. https://doi.org/10.1002/hbm.25684

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

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2011). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology*, *40*(3), 626–644. https://doi.org/10.1093/ije/dyq190

Cantlon, J. F., & Li, R. (2013). Neural Activity during Natural Viewing of Sesame Street Statistically Predicts Test Scores in Early Childhood. *PLoS Biology*, *11*(1), e1001462. https://doi.org/10.1371/journal.pbio.1001462

Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. In *Annals of the New York Academy of Sciences* (Vol. 1124, pp. 111–126). Blackwell Publishing Inc. https://doi.org/10.1196/annals.1440.010

Castellanos, F. X., Sonuga-Barke, E. J. S., Milham, M. P., & Tannock, R. (2006). Characterizing cognition in ADHD: Beyond executive dysfunction. *Trends in Cognitive Sciences*, *10*(3), 117–123. https://doi.org/10.1016/j.tics.2006.01.011

Cattell, R. B. (1987). *Intelligence: Its Structure, Growth and Action*. North-Holland.

Chen, J., Tam, A., Kebets, V., Orban, C., Ooi, L. Q. R., Asplund, C. L., Marek, S., Dosenbach, N. U. F., Eickhoff, S. B., Bzdok, D., Holmes, A. J., & Yeo, B. T. T. (2022). Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nature Communications*, *13*(1), 2020.06.24.168724. https://doi.org/10.1038/s41467-022-29766-8

Chollet, F. (2019). *On the Measure of Intelligence*.

https://doi.org/10.48550/arxiv.1911.01547

Chuderski, A. (2014). How well can storage capacity, executive control, and fluid reasoning explain insight problem solving. *Intelligence, 46*(1), 258–270. https://doi.org/10.1016/j.intell.2014.07.010

Chuderski, A., & Necka, E. (2012). The contribution of working memory to fluid reasoning: Capacity, control, or both? *Journal of Experimental Psychology: Learning Memory and Cognition, 38*(6), 1689–1710. https://doi.org/10.1037/a0028465

Claesdotter, E., Cervin, M., Åkerlund, S., Råstam, M., & Lindvall, M. (2018). The effects of ADHD on cognitive performance. *Nordic Journal of Psychiatry*, *72*(3), 158–163. https://doi.org/10.1080/08039488.2017.1402951

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

Cui, Z., Pines, A. R., Larsen, B., Sydnor, V. J., Li, H., Adebimpe, A., Alexander-Bloch, A. F., Bassett, D. S., Bertolero, M., Calkins, M. E., Davatzikos, C., Fair, D. A., Gur, R. C., Gur, R. E., Moore, T. M., Shanmugan, S., Shinohara, R. T., Vogel, J. W., Xia, C. H., … Satterthwaite, T. D. (2022). Linking Individual Differences in Personalized Functional Network Topography to Psychopathology in Youth. *Biological Psychiatry*, 2021.08.02.454763. https://doi.org/10.1016/j.biopsych.2022.05.014

Cusack, R., Vicente-Grabovetsky, A., Mitchell, D. J., Wild, C. J., Auer, T., Linke, A. C., & Peelle, J. E. (2015). Automatic analysis (aa): Efficient neuroimaging workflows and parallel processing using Matlab and XML. *Frontiers in Neuroinformatics, 8*(JAN), 90. https://doi.org/10.3389/fninf.2014.00090

Daley, D., & Birchwood, J. (2010). ADHD and academic performance: Why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? *Child: Care, Health and Development, 36*(4), 455–464. https://doi.org/10.1111/j.1365-2214.2009.01046.x

Danielson, M. L., Bitsko, R. H., Ghandour, R. M., Holbrook, J. R., Kogan, M. D., & Blumberg, S. J. (2018). Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment Among U.S. Children and Adolescents, 2016. *Journal of Clinical Child and Adolescent Psychology*, *47*(2), 199–212. https://doi.org/10.1080/15374416.2017.1417860

David Wechsler. (2014). *Wechsler intelligence scale for children-fifth edition*. Pearson.

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews Neuroscience*, *11*(3), 201–211. https://doi.org/10.1038/nrn2793

Dehn, M. J. (2017). How working memory enables fluid reasoning. *Applied Neuropsychology: Child*, *6*(3), 245–247. https://doi.org/10.1080/21622965.2017.1317490

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64,* 135–168.

https://doi.org/10.1146/annurev-psych-113011-143750

Dosenbach, N. U. F., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., Nelson, S. M., Wig, G. S., Vogel, A. C., Lessov-Schlaggar, C. N., Barnes, K. A., Dubis, J. W., Feczko, E., Coalson, R. S., Pruett, J. R., Barch, D. M., Petersen, S. E., & Schlaggar, B. L. (2010). Prediction of individual brain maturity using fMRI. *Science*, *329*(5997), 1358–1361. https://doi.org/10.1126/science.1194144

Durston, S. (2003). A review of the biological bases of ADHD: What have we learned from imaging studies? *Mental Retardation and Developmental Disabilities Research Reviews*, *9*(3), 184–195. https://doi.org/10.1002/mrdd.10079

Edin, F., Macoveanu, J., Olesen, P., Tegnér, J., & Klingberg, T. (2007). Stronger synaptic connectivity as a mechanism behind development of working memory-related brain activity during childhood. *Journal of Cognitive Neuroscience*, *19*(5), 750–760. https://doi.org/10.1162/jocn.2007.19.5.750

Fair, D. A., Cohen, A. L., Power, J. D., Dosenbach, N. U. F., Church, J. A., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Computational Biology*, *5*(5), e1000381. https://doi.org/10.1371/journal.pcbi.1000381

Fair, D. A., Nigg, J. T., Iyer, S., Bathula, D., Mills, K. L., Dosenbach, N. U. F., Schlaggar, B. L., Mennes, M., Gutman, D., Bangaru, S., Buitelaar, J. K., Dickstein, D. P., Martino, A. Di, Kennedy, D. N., Kelly, C., Luna, B., Schweitzer, J. B., Velanova, K., Wang, Y. F., … Milham, M. P. (2013). Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Frontiers in Systems Neuroscience*, *6*(FEB), 1–31. https://doi.org/10.3389/fnsys.2012.00080

Ferrer, E., O’Hare, E. D., & Bunge, S. A. (2009). Fluid reasoning and the developing brain. *Frontiers in Neuroscience*, *3*(MAY), 3. https://doi.org/10.3389/neuro.01.003.2009

Finn, E. S., & Bandettini, P. A. (2021). Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage*, *235*, 117963. https://doi.org/10.1016/j.neuroimage.2021.117963

Fosco, W. D., Kofler, M. J., Groves, N. B., Chan, E. S. M., & Raiker, J. S. (2020). Which ‘Working’ Components of Working Memory aren’t Working in Youth with ADHD? *Journal of Abnormal Child Psychology*, *48*(5), 647–660. https://doi.org/10.1007/s10802-020-00621-y

Friedman, L. A., & Rapoport, J. L. (2015). Brain development in ADHD. *Current Opinion in Neurobiology*, *30*, 106–111. https://doi.org/10.1016/j.conb.2014.11.007

Friston, K. J., Mechelli, A., Turner, R., & Price, C. J. (2000). Nonlinear responses in fMRI: The balloon model, Volterra kernels, and other hemodynamics. *NeuroImage*, *12*(4), 466–477. https://doi.org/10.1006/nimg.2000.0630

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The Structure of Working Memory from 4 to 15 Years of Age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Gong, D., He, H., Ma, W., Liu, D., Huang, M., Dong, L., Gong, J., Li, J., Luo, C., & Yao, D. (2016). Functional Integration between Salience and Central Executive Networks: A Role for Action Video Game Experience. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/9803165

Gratton, C., Sun, H., & Petersen, S. E. (2018). Control networks and hubs. *Psychophysiology*, *55*(3), e13032. https://doi.org/10.1111/psyp.13032

Gruskin, D. C., & Patel, G. H. (2022). Brain connectivity at rest predicts individual differences in normative activity during movie watching. *NeuroImage*, *253*, 119100. https://doi.org/10.1016/j.neuroimage.2022.119100

Harpin, V. A. (2005). The effect of ADHD on the life of an individual, their family, and community from preschool to adult life. *Archives of Disease in Childhood*, *90*(SUPPL. 1). https://doi.org/10.1136/adc.2004.059006

Harpin, V., Mazzone, L., Raynaud, J. P., Kahle, J., & Hodgkins, P. (2016). Long-Term Outcomes of ADHD: A Systematic Review of Self-Esteem and Social Function. *Journal of Attention Disorders*, *20*(4), 295–305. https://doi.org/10.1177/1087054713486516

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-Inspired Artificial Intelligence. In *Neuron* (Vol. 95, Issue 2, pp. 245–258). Cell Press. https://doi.org/10.1016/j.neuron.2017.06.011

Hebling Vieira, B., Dubois, J., Calhoun, V. D., & Garrido Salmon, C. E. (2021). A deep learning based approach identifies regions more relevant than resting-state networks to the prediction of general intelligence from resting-state fMRI. *Human Brain Mapping*, *42*(18), 5873–5887. https://doi.org/10.1002/hbm.25656

Hitch, G. J., & Halliday, M. S. (1983). Working memory in children. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *302*(1110), 325–340. https://doi.org/10.1098/rstb.1983.0058

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mark Mahone, E. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology*, *17*(3), 209–224. https://doi.org/10.1080/09297049.2010.532204

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kandel, E. R., Koester, J. D., Mack, S. H., & Siegelbaum, S. A. (2021). *Principles of Neural Science* (6th ed.). McGraw-hill.

Kane, M. J., Tuholski, S. W., Hambrick, D. Z., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-

3445.133.2.189

Katz, L. J., Brown, F. C., Roth, R. M., & Beers, S. R. (2011). Processing speed and working memory performance in those with both ADHD and a reading disorder compared with those with ADHD alone. *Archives of Clinical Neuropsychology*, *26*(5), 425–433. https://doi.org/10.1093/arclin/acr026

Kirov, R., & Brand, S. (2014). Sleep problems and their effect in ADHD. *Expert Review of Neurotherapeutics*, *14*(3), 287–299. https://doi.org/10.1586/14737175.2014.885382

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlström, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD - A randomized, controlled trial. *Journal of the American Academy of Child and Adolescent Psychiatry*, *44*(2), 177–186. https://doi.org/10.1097/00004583-200502000-00010

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology*, *24*(6), 781–791. https://doi.org/10.1076/jcen.24.6.781.8395

Konofal, E., Lecendreux, M., & Cortese, S. (2010). Sleep and ADHD. *Sleep Medicine*, *11*(7), 652–658. https://doi.org/10.1016/j.sleep.2010.02.012

Krain, A. L., & Castellanos, F. X. (2006). Brain development and ADHD. *Clinical Psychology Review*, *26*(4), 433–444. https://doi.org/10.1016/j.cpr.2006.01.005

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389–433. https://doi.org/10.1016/S0160-2896(05)80012-1

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the Development of Cognitive Control through Adolescence? In *Brain and Cognition* (Vol. 72, Issue 1, pp. 101–113). Brain Cogn. https://doi.org/10.1016/j.bandc.2009.08.005

Lyons, K. M., Stevenson, R. A., Owen, A. M., & Stojanoski, B. (2020). Examining the relationship between measures of autistic traits and neural synchrony during movies in children with and without autism. *NeuroImage: Clinical, 28*, 102477. https://doi.org/10.1016/j.nicl.2020.102477

Mano, Q. R., Jastrowski Mano, K. E., Guerin, J. M., Gibler, R. C., Becker, S. P., Denton, C. A., Epstein, J. N., & Tamm, L. (2019). Fluid reasoning and reading difficulties among children with ADHD. *Applied Neuropsychology: Child*, *8*(4), 307–318. https://doi.org/10.1080/21622965.2018.1466706

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022).

Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654–660. https://doi.org/10.1038/s41586-022-04492-9

Mayes, S. D., & Calhoun, S. L. (2006). WISC-IV and WISC-III profiles in children with ADHD. *Journal of Attention Disorders*, *9*(3), 486–493. https://doi.org/10.1177/1087054705283616

McArdle, J. J., Hamagami, F., Meredith, W., & Bradway, K. P. (2000). Modeling the dynamic hypotheses of Gf-Gc theory using longitudinal life-span data. *Learning and Individual Differences*, *12*(1), 53–79. https://doi.org/10.1016/S1041-6080(00)00036-4

Meer, J. N. va. der, Breakspear, M., Chang, L. J., Sonkusare, S., & Cocchi, L. (2020). Movie viewing elicits rich and reliable brain state dynamics. *Nature Communications*, *11*(1), 1–14. https://doi.org/10.1038/s41467-020-18717-w

Moura, O., Costa, P., & Simões, M. R. (2019). WISC-III Cognitive Profiles in Children with ADHD: Specific Cognitive Impairments and Diagnostic Utility. *Journal of General Psychology*, *146*(3), 258–282. https://doi.org/10.1080/00221309.2018.1561410

Nelson, C. A. (2001). The Development and Neural Bases of Face Recognition. *Infant and Child Development*, *10*(1–2), 3–18. https://doi.org/10.1002/icd.239

Nichols, T., & Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research*, *12*(5), 419–446. https://doi.org/10.1191/0962280203sm341ra

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., & Greene, D. J. (2020). Machine Learning With Neuroimaging: Evaluating Its Applications in Psychiatry. In *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* (Vol. 5, Issue 8, pp. 791–798). Elsevier Inc. https://doi.org/10.1016/j.bpsc.2019.11.007

Otero, T. M. (2017). Brief review of fluid reasoning: Conceptualization, neurobasis, and applications. *Applied Neuropsychology: Child*, *6*(3), 204–211. https://doi.org/10.1080/21622965.2017.1317484

Owens, J. A. (2005). The ADHD and sleep conundrum: A review. *Journal of Developmental and Behavioral Pediatrics*, *26*(4), 312–322. https://doi.org/10.1097/00004703-200508000-00011

Pagani, L. S., Brière, F. N., & Janosz, M. (2017). Fluid reasoning skills at the high school transition predict subsequent dropout. *Intelligence*, *62*, 48–53. https://doi.org/10.1016/j.intell.2017.02.006

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., Vogel, A. C., Laumann, T. O., Miezin, F. M., Schlaggar, B. L., & Petersen, S. E. (2011). Functional Network Organization of the Human Brain. *Neuron*, *72*(4), 665–678. https://doi.org/10.1016/j.neuron.2011.09.006

Rainer, G., Asaad, W. F., & Miller, E. K. (1998). Memory fields of neurons in the primate prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(25), 15008–15013.

https://doi.org/10.1073/pnas.95.25.15008

Rosenberg, M. D., Casey, B. J., & Holmes, A. J. (2018). Prediction complements explanation in understanding the developing brain. *Nature Communications*, *9*(1), 1–13. https://doi.org/10.1038/s41467-018-02887-9

Rubenstein, J., Rakic, P., Chen, B., Kwan, K. Y., & Claire, H. (2020). Neural Circuit and Cognitive Development. In *Neural Circuit and Cognitive Development* (2nd ed.). Elsevier Inc. https://doi.org/10.1016/c2017-0-00830-3

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescent with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(8), 988–1003. https://doi.org/10.1111/1469-7610.00227

Rudolph, M. D., Miranda-Domínguez, O., Cohen, A. O., Breiner, K., Steinberg, L., Bonnie, R. J., Scott, E. S., Taylor-Thompson, K., Chein, J., Fettich, K. C., Richeson, J. A., Dellarco, D. V., Galván, A., Casey, B. J., & Fair, D. A. (2017). At risk of being risky: The relationship between "brain age" under emotional states and risk preference. *Developmental Cognitive Neuroscience*, *24*, 93–106. https://doi.org/10.1016/j.dcn.2017.01.010

Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols*, *12*(3), 506–518. https://doi.org/10.1038/nprot.2016.178

Sripada, C., Rutherford, S., Angstadt, M., Thompson, W. K., Luciana, M., Weigard, A., Hyde, L. H., & Heitzeg, M. (2020). Prediction of neurocognition in youth from resting state fMRI. *Molecular Psychiatry*, *25*(12), 3413–3421. https://doi.org/10.1038/s41380-019-0481-6

Sun, Y., Zhao, L., Lan, Z., Jia, X. Z., & Xue, S. W. (2020). Differentiating boys with ADHD from those with typical development based on whole-brain functional connections using a machine learning approach. *Neuropsychiatric Disease and Treatment*, *16*, 691–702. https://doi.org/10.2147/NDT.S239013

Tamm, L., & Juranek, J. (2012). Fluid reasoning deficits in children with ADHD: Evidence from fMRI. *Brain Research*, *1465*, 48–56. https://doi.org/10.1016/j.brainres.2012.05.021

Thaler, N. S., Bello, D. T., & Etcoff, L. M. (2013). WISC-IV Profiles Are Associated With Differences in Symptomatology and Outcome in Children With ADHD. *Journal of Attention Disorders*, *17*(4), 291–301. https://doi.org/10.1177/1087054711428806

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fisch, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Thompson, W. K., Barch, D. M., Bjork, J. M., Gonzalez, R., Nagel, B. J., Nixon, S. J., &

Luciana, M. (2019). The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: Findings from the ABCD study's baseline neurocognitive battery. *Developmental Cognitive Neuroscience*, *36*, 100606. https://doi.org/10.1016/j.dcn.2018.12.004

Tian, Y., & Zalesky, A. (2021). Machine learning prediction of cognition from functional connectivity: Are feature weights reliable? *NeuroImage*, *245*, 2021.05.27.446059. https://doi.org/10.1016/j.neuroimage.2021.118648

Ullman, H., Almeida, R., & Klingberg, T. (2014). Structural maturation and brain activity predict future working memory capacity during childhood development. *Journal of Neuroscience*, *34*(5), 1592–1598. https://doi.org/10.1523/JNEUROSCI.0842-13.2014

Vanderwal, T., Eilbott, J., Finn, E. S., Craddock, R. C., Turnbull, A., & Castellanos, F. X. (2017). Individual differences in functional connectivity during naturalistic viewing conditions. *NeuroImage*, *157*, 521–530. https://doi.org/10.1016/j.neuroimage.2017.06.027

Voss, C., Cammarata, N., Goh, G., Petrov, M., Schubert, L., Egan, B., Lim, S., & Olah, C. (2021). Visualizing Weights. *Distill*, *6*(2). https://doi.org/10.23915/distill.00024.007

Waddell, C., Offord, D. R., Shepherd, C. A., Hua, J. M., & McEwan, K. (2002). Child psychiatric epidemiology and Canadian public policy-making: The state of the science and the art of the possible. *Canadian Journal of Psychiatry*, *47*(9), 825–832. https://doi.org/10.1177/070674370204700903

Wegelin, J. (2000). A survey of Partial Least Squares (PLS) methods, with emphasis on the two-block case. *Technical Report 371, Department of Statistics, University of Washington, Seattle, 2000.*, *371*, 371. https://www.stat.washington.edu/www/research/reports/2000/tr371.pdf

Wendelken, C., O'Hare, E. D., Whitaker, K. J., Ferrer, E., & Bunge, S. A. (2011). Increased functional selectivity over development in rostrolateral prefrontal cortex. *Journal of Neuroscience*, *31*(47), 17260–17268. https://doi.org/10.1523/JNEUROSCI.1193-10.2011

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. https://doi.org/10.1016/j.biopsych.2005.02.006

Wright, S. B., Matlen, B. J., Baym, C. L., Ferrer, E., & Bunge, S. A. (2008). Neural correlates of fluid reasoning in children and adults. *Frontiers in Human Neuroscience*, *1*(MAR), 8. https://doi.org/10.3389/neuro.09.008.2007

Zelazo, P. D., & Mller, U. (2007). Executive Function in Typical and Atypical Development. *Blackwell Handbook of Childhood Cognitive Development*, 445–469. https://doi.org/10.1002/9780470996652.ch20

Curriculum Vitae

Name: Brian Pho

Post-secondary Education and Degrees:

University of Calgary
Calgary, Alberta, Canada
2015-2020 B.S. Software Engineering

Western University
London, Ontario, Canada
2020-2022 M.S. Neuroscience

Honours and Awards:

Program for Undergraduate Research Experience (PURE) Award
2019

Del Henderson Memorial Bursary
2017

Jason Lang Scholarships
2016-2018

Horatio Alger Scholarship
2015

Related Work Experience

Teaching Assistant
Western University
2020-2022

Learning Assistant
University of Calgary
2019-2020

Undergraduate Researcher Summer Student
University of Calgary
2019-2019

Software Developer Intern
Pason Systems Corp.
2018-2019

Instructional Assistant
Kumon Inc.
2013-2018