ADHD-200 Patient Characterization and Classification using Resting State Networks: A Dissertation

ADHD-200 PATIENT CHARACTERIZATION AND CLASSIFICATION USING RESTING STATE NETWORKS

A Dissertation Presented

By

Suzanne Marie Czerniak

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

March 28th, 2014

Clinical and Population Health Research

ADHD-200 PATIENT CHARACTERIZATION AND CLASSIFICATION USING RESTING STATE NETWORKS

A Dissertation Presented
By

Suzanne Marie Czerniak

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

Constance Moore, Ph.D., Thesis Advisor

Eric Mick, Sc.D., Member of Committee

Matthew Gounis, Ph.D., Member of Committee

Blaise Frederick, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Jean King, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

Anthony Carruthers, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

Clinical and Population Health Research

March 28th, 2014

The work herein is dedicated to those who have travelled a different path:

"The grey rain-curtain turn[s] all to silver glass and [is] rolled back ... [behold] white shores and beyond them a far green country under a swift sunrise."

J.R.R. Tolkien

ACKNOWLEDGEMENTS

I would first like to thank my thesis advisor, Dr. Constance Moore, and my committee members, Dr. Jean King, Dr. Eric Mick, Dr. Matthew Gounis, and Dr. Blaise Frederick, for their time, advice, and support in conceptualizing, writing, modifying, and defending this dissertation. I would especially like to thank Dr. Constance Moore for her mentorship and guidance, as well as Dr. Jean King for her leadership. Dr. David Kennedy, who has given feedback on this project from the very start, also deserves hearty acknowledgment for his contributions.

I would also like to extend my sincerest gratitude to Dr. Elif Sikoglu, without whom this work would not exist. Her patience, encouragement, and unwavering support not only made me believe I could succeed, but provided the fundamental substrate on which that success was built. From the bottom of my heart, Elif: thank you.

I feel especially fortunate that my time as a graduate student was spent at the CCNI: the people here have become not only my colleagues, but my friends. Dr. Wei Huang: thank you for the million tiny conversations about resting state, choosing seeds, managing husbands, and of course: ICA. Dr. Ana Liso Navarro: thank you for bringing a smile with you into every room, and for showing me how awesome child psychiatrists are. Laurellee Payne-Roberts: thank you for being a natural mouse-whisperer, and for never failing to make me laugh. Yvette Gonzalez: without you, no paperwork would be signed, no meetings scheduled, and no rooms booked. Thank you for helping me in all the little ways that turn out to not be so little after all. Ryan Rogan: thank you for always listening to me, and for not resenting my repeated requests for the same thing you'd

already sent me three times, and I just couldn't remember where I saved it. Dr. Zhifeng Liang: thank you for all the discussions we had when I was first beginning to understand this whole imaging thing, and for continuing to have those discussions with me despite my initial confusion. Kelly Tam: thank you for making things go – for planning lab meetings, presentations, and seminars, and generally running the lab. Meghan Heffernan: thank you for cheerfully welcoming me to the CCNI and helping me get properly situated here. And to Janaque Fernando, Matthew Thompson, John Gannon, SarahRose Gabor, and all the other volunteers: thank you for taking work off our shoulders!

My deepest and most long-standing thanks go to my father, Dr. Richard Czerniak, and my stepmom, Cecelia Czerniak. You have both been incredibly supportive of my life decisions. Dad, thank you for always encouraging me to ask "why?" and for teaching me that until you can explain it to somebody else, you don't really understand it yourself. Lia, thank you for talking me down off of many, MANY ledges and for giving me the best advice in the myriad of new situations I found myself in – you have taught me how to act like an adult, even when I felt like an inexperienced child.

And finally, last but not least, I want to thank Sasha Hayes-Rusnov, who has put up with the cranky, stressed-out, terrified-of-failure me. You must love me something fierce to take such good care of me, even when I'm not taking care of myself.

Thank you all.

ABSTRACT

Attention Deficit/Hyperactivity Disorder (ADHD) is a common psychiatric disorder of childhood that is characterized by symptoms of inattention, impulsivity/hyperactivity, or a combination of both. Intrinsic brain dysfunction in ADHD can be examined through various methods including resting state functional Magnetic Resonance Imaging (rs-fMRI), which investigates patients' functional brain connections in the absence of an explicit task. To date, studies of group differences in resting brain connectivity between patients with ADHD and typically developing controls (TDCs) have revealed reduced connectivity within the Default Mode Network (DMN), a resting state network implicated in introspection, mind-wandering, and day-dreaming. However, few studies have addressed the use of resting state connectivity measures as a diagnostic aide for ADHD on the individual patient level. In the current work, we attempted first to characterize the differences in resting state networks, including the DMN and three attention networks (the salience network, the left executive network, and the right executive network), between a group of youth with ADHD and a group of TDCs matched for age, IQ, gender, and handedness. Significant over- and under-connections were found in the ADHD group in all of these networks compared with TDCs. We then attempted to use a support vector machine (SVM) based on the information extracted from resting state network connectivity to classify participants as "ADHD" or "TDC." The IFG-middle temporal network (66.8% accuracy), the parietal association network (86.6% specificity and 48.5% PPV), and a physiological noise component (sensitivity 39.7% and NPV 69.6%) performed the best classifications. Finally, we attempted to combine and

utilize information from all the resting state networks that we identified to improve classification accuracy. Contrary to our hypothesis, classification accuracy decreased to 54-55% when this information was combined. Overall, the work presented here supports the theory that the ADHD brain is differently connected at rest than that of TDCs, and that this information may be useful for developing a diagnostic aid. However, because ADHD is such a heterogeneous disorder, each ADHD patient's underlying brain deficits may be unique making it difficult to determine what connectivity information is diagnostically useful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter II

Chapter III

PREFACE

Introduction Figure 1was not directly generated by the author, and was taken from the Centers for Disease Control and Prevention website (http://www.cdc.gov/ncbddd/adhd/prevalence.html#current). In a personal communication to the author, it was confirmed that this figure is in the public domain, is free to use, and has no license number attached to it.

INTRODUCTION

ADHD definition, prevalence, and diagnostic process

Attention Deficit/Hyperactivity Disorder (ADHD) is psychiatric disorder usually diagnosed in childhood. Children with ADHD exhibit multiple inattentive behaviors, hyperactive/impulsive behaviors, or some combination of the two, that impair their ability to function or develop normally (American Psychiatric Association, 2013). Whether or not a child meets diagnostic criteria for ADHD is determined using the DSM-V, which lists symptoms and behaviors that fall under the two major domains of Inattention or Hyperactivity and Impulsivity (Table 1) (American Psychiatric Association, 2013). A child must have consistently exhibited 6 or more behaviors from one or both domains for at least 6 months, in two or more settings (such as both at home and at school) in order to receive the diagnosis (American Psychiatric Association, 2013). It is also necessary that at least several of the symptoms from either domain were present in the patient before the age of 12, and that the patient's symptoms are not caused exclusively by another psychiatric disorder (such as a mood or anxiety disorder) (American Psychiatric Association, 2013). Based on which domain(s) a child's symptoms fall under, he or she may be classified as one of the 3 subtypes of ADHD: inattentive subtype (inattentive, but not hyperactive/impulsive, criteria are met), hyperactive/impulsive subtype (hyperactive/impulsive, but not inattentive, criteria are met), or combination subtype (both inattentive and hyperactive/impulsive criteria are met) (American Psychiatric Association, 2013).

While it is largely accepted that ADHD is the most common pediatric psychiatric disorder, its prevalence varies from country to country, by gender, and by age. Recent

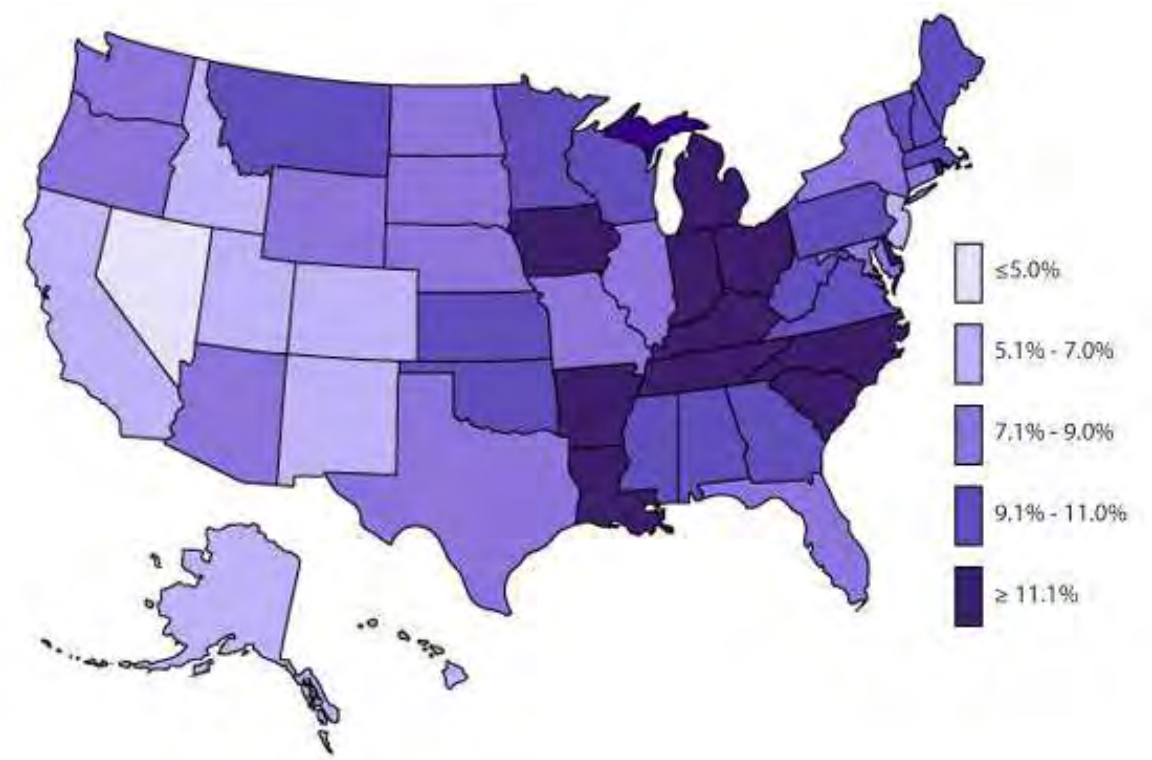
Table 1. DSM-V criteria for diagnosing ADHD. This table shows inattentive (column 1) and hyperactive or impulsive (column 2) symptoms that may lead to a diagnosis of ADHD. A child must have consistently exhibited 6 or more behaviors from one or both columns to be considered for the diagnosis.

Table 1. DSM-V Criteria for Diagnosing ADHD

Inattentive Diagnostic Criteria (must have 6 or more of the following)	**Hyperactive or Impulsive Diagnostic Criteria** (must have 6 or more of the following)
*Often fails to give close attention to details or makes careless mistakes in schoolwork, at work, or during other activities (e.g., overlooks or misses details, work is inaccurate).	*Often fidgets with or taps hands or feet or squirms in seat.
*Often has difficulty sustaining attention in tasks or play activities (e.g., has difficulty remaining focused during lectures, conversations, or lengthy reading).	*Often leaves seat in situations when remaining seated is expected (e.g., leaves his or her place in the classroom, in the office or other workplace, or in other situations that require remaining in place).
*Often does not seem to listen when spoken to directly (e.g., mind seems elsewhere, even in the absence of any obvious distraction).	*Often runs about or climbs in situations where it is inappropriate. (Note: In adolescents or adults, may be limited to feeling restless.)
*Often does not follow through on instructions and fails to finish schoolwork, chores, or duties in the workplace (e.g., starts tasks but quickly loses focus and is easily sidetracked).	*Often unable to play or engage in leisure activities quietly.
*Often has difficulty organizing tasks and activities (e.g., difficulty managing sequential tasks; difficulty keeping materials and belongings in order; messy, disorganized work; has poor time management; fails to meet deadlines).	*Is often "on the go," acting as if "driven by a motor" (e.g., is unable to be or uncomfortable being still for extended time, as in restaurants, meetings; may be experienced by others as being restless or difficult to keep up with).
*Often avoids, dislikes, or is reluctant to engage in tasks that require sustained mental effort (e.g., schoolwork or homework; for older adolescents and adults, preparing reports, completing forms, reviewing lengthy papers).	*Often talks excessively.
*Often loses things necessary for tasks or activities (e.g., school materials, pencils, books, tools, wallets, keys, paperwork, eyeglasses, mobile telephones).	*Often blurts out an answer before a question has been completed (e.g., completes people's sentences; cannot wait for turn in conversation).
*Is often easily distracted by extraneous stimuli (for older adolescents and adults, may include unrelated thoughts).	*Often has difficulty waiting his or her turn (e.g., while waiting in line).
*Is often forgetful in daily activities (e.g., doing chores, running errands; for older adolescents and adults, returning calls, paying bills, keeping appointments).	*Often interrupts or intrudes on others (e.g., butts into conversations, games, or activities; may start using other people's things without asking or receiving permission; for adolescents and adults, may intrude into or take over what others are doing).

estimates place the US prevalence of ADHD at about 11% of school-aged youth (Visser et al., 2014), while the world-wide prevalence of ADHD in children and adolescents is thought to be closer to 5% (Polanczyk et al., 2007). In the US, southern and mid-western states (such as Arkansas and Kentucky) have the highest prevalence of ADHD (14.6% and 14.8%, respectively), whereas southwestern states (such as Nevada) have the lowest prevalence (4.2%) (Figure 1) (http://www.cdc.gov/ncbddd/adhd/data.html). In general, boys are more likely than girls to receive an ADHD diagnosis. In a 2011-2012 Centers for Disease Control (CDC) report, it was estimated that 1 in 5 high school-aged boys had received an ADHD diagnosis at some point in their lives, whereas the same could be said of only 1 in 11 high school girls (Visser et al., 2014). These findings are in line with earlier reports that in population-based studies, ADHD was 3 times more prevalent among boys than girls (Gaub and Caryn L. Carlson, 1997). This gender asymmetry (3:1 or more recently, ~ 2:1, boys to girls) may be reflective of an underlying referral bias. As previous studies have shown, boys with ADHD tend to express more hyperactive/disruptive symptoms than girls, which may alert parents and teachers to the presence of the underlying disorder; on the other hand girls (who express more inattentive symptoms) may not be as readily identified (Biederman et al., 2002; Biederman et al., 2005; Gaub and Caryn L. Carlson, 1997; Hinshaw et al., 2006; Quinn, 2008). Finally, the prevalence of ADHD changes with patient age. For many children, ADHD symptoms may spontaneously resolve as they grow up: a recent study found that less than a third (~29%) of childhood ADHD cases persisted into adulthood (Barbaresi et al., 2013). This finding is line with previous works demonstrating the persistence of

Figure 1. The current prevalence of ADHD in the US by state, 2011-2012 (http://www.cdc.gov/ncbddd/adhd/prevalence.html#current).


≤5.0%
5.1% - 7.0%
7.1% - 9.0%
9.1% - 11.0%
≥ 11.1%

ADHD into adulthood was 15% - 65%, depending on the criteria used (Biederman et al., 2006; Faraone et al., 2006). In general, the average prevalence of adult ADHD is estimated to be 2.5% (American Psychiatric Association, 2013; Simon et al., 2009).

ADHD is usually diagnosed by a licensed health care professional; for example, a child's pediatrician or a psychiatrist specializing in pediatric disorders. Generally during an evaluation for ADHD, the clinician will take a detailed history of the type, onset, duration, frequency, and severity of ADHD-related symptoms by interviewing the child's parent (AACAP Official Action, 2007). In addition to the interview, the parent may be asked to complete one or more of several behavior rating questionnaires that have been shown to be useful in eliciting and assessing ADHD-related symptoms (for example, the Conners Parent Rating Scale) (AACAP Official Action, 2007; Pelham et al., 2005). If the parent consents, the clinician may also contact the child's school and request that his or her teacher also complete corresponding questionnaires, such as the Conners Teacher Rating Scale (AACAP Official Action, 2007; Pelham et al., 2005). While these scales are useful to the clinician as one element of the child's evaluation, it has been shown that when used alone, their specificity in diagnosing ADHD is low (~36%) (Parker and P. Corkum, 2013). It is therefore important that a full clinical evaluation be conducted before the ADHD diagnosis is given. The fact remains, however, that currently ADHD diagnosis relies heavily on parent and/or teacher report(s), which may be unintentionally subjective (Parens and J. Johnston, 2009). In this context, a quantifiable, intrinsic marker of ADHD may prove valuable as a diagnostic adjunct.

Neuroimaging using magnetic resonance techniques

Neuroimaging techniques may provide a way to more objectively assess underlying brain deficits in patients with ADHD, as well as those with other psychiatric or neurologic disorders. There are many viable options for imaging the brain, including positron emission tomography (PET), electroencephalography (EEG), and magnetic resonance (MR) techniques, to name a few. Ideally, a technique used for imaging the human brain would have low invasiveness (not require surgery or contrast to implement), high spatial resolution (ability to see changes occurring at the millimeter level or smaller), and high temporal resolution (ability to see changes occurring in milliseconds to seconds) (Huettel et al., 2009). While functional magnetic resonance imaging (fMRI) usually has a temporal resolution of seconds to minutes, it has excellent spatial resolution (millimeters or below) and is non-invasive in that it does not require surgery, contrast, or exposure to ionizing radiation (Huettel et al., 2009). It is therefore optimally suited for mapping functional changes throughout the brain, and will be the method of choice for this work.

The basis of fMRI is the blood oxygen level dependent (BOLD) signal. MRI techniques take advantage of the magnetic properties of certain atomic nuclei (hydrogen nuclei, or protons, are the most commonly studied), and manipulate them using a combination of a large magnetic field, the superposition of sets of smaller magnetic fields (gradients), and a series of radiofrequency pulses (Huettel et al., 2009). The signal generated by these manipulations depends on how the nuclei under study return to their usual (equilibrium) states, which is highly dependent on the environment (Huettel et al.,

2009). In the case of fMRI imaging, blood (specifically the balance between oxygenated and deoxygenated blood) is the most important environmental factor. As the polarizability of deoxygenated blood creates its own magnetic fields, its presence will cause the MR signal from surrounding nuclei to lose coherence more quickly, resulting in a faster decrease in MR signal (Thulborn et al., 1982). Areas of the brain that have a higher concentration of deoxygenated blood show less MR signal (Ogawa et al., 1990; Thulborn et al., 1982). Therefore, BOLD signal can be exploited to indirectly measure brain activity. As neurons fire, their oxygen consumption increases, causing a concomitant increase in deoxygenated blood in the surrounding brain capillaries (Malonek and A. Grinvald, 1996; Ogawa et al., 1990). In response, the body increases blood flow to the active area to replace the deoxygenated blood with oxygenated blood, decreasing the amount of deoxygenated blood in the area (Malonek and A. Grinvald, 1996; Ogawa et al., 1990). This decrease in deoxygenated blood in the active area of the brain is read by the MR signal receiver as an increase in BOLD signal (Huettel et al., 2009). Thus, increased brain activity ultimately yet indirectly results in increased BOLD signal.

The first fMRI studies were designed to measure brain activation by presenting a research participant with a task (i.e., visual or motor task), and then measuring the BOLD signal change in response to task performance (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1992). Hence, this technique is sometimes referred to simply as "fMRI" and sometimes as "task-based fMRI"; task-based fMRI will be the preferred term in this work. Task-based fMRI allows for the investigation of which brain regions "light

up" during the task: usually the BOLD signal is acquired while a participant is lying quietly ("at rest") or while performing a control task ("control task"), and then again while they are performing the task of interest ("task"). The difference between these two states ("task" – "at rest" or "task" – "control task") is then interpreted as how involved that region was in the underlying mental processes that supported task performance (Huettel et al., 2009). Task-based fMRI has since been used extensively to investigate the neural underpinnings of many different types of cognitive tasks, as well as examine how the degree of activation may differ between a group of patients and a group of healthy controls (Huettel et al., 2009; Samanez-Larkin and M. D'Esposito, 2008). However, it is also possible to study brain BOLD signal in the absence of an explicit task: this is known as resting state fMRI (rs-fMRI). rs-fMRI grew out of the observation that even during the "at rest" portions of a motor task, the areas that activated during the task showed fluctuations in BOLD signal (Biswal et al., 1995). Furthermore, these "at rest" fluctuations were synchronous among the areas identified by the task (Biswal et al., 1995). This finding led to the concept of "resting state functional connectivity," where areas of the brain that show highly correlated BOLD signal changes are thought to work together (be functionally connected), even in the absence of an explicit task (Biswal et al., 1995).

Resting state functional connectivity (rs-FC), or areas of the brain that are connected at rest, can be investigated in many different ways (although all of these somehow rely on measuring the similarity of BOLD signal fluctuations in two or more brain areas). The two methods that will be discussed here are (1) hypothesis-driven seed-

based and (2) Independent Component Analysis (ICA) techniques (Cole et al., 2010; Greicius, 2008). While both of these methods have been used extensively to investigate rs-FC, they differ in a few important ways. Seed-based methods require an *a priori* hypothesis, in that the investigator specifies an area of the brain as the "seed," and then examines the strength of the connections between the seed area and the rest of the brain (Lee et al., 2013). This is usually done by extracting the average time course of BOLD signal in the seed area, and then calculating the correlation between the seed area's time course and the time course of other areas in the brain (Lee et al., 2013). These other areas are usually referred to as regions of interest (ROIs) and can be any size, from the individual volumes of the brain that BOLD signal was originally acquired from (voxels) to larger brain areas defined by their structure (for example, gyri) or function (for example, the dorsolateral prefrontal cortex, DLPFC) (Lee et al., 2013; Vincent et al., 2008). The strength of the connection between the seed and each ROI can then be compared across two groups of participants (for example, patients with ADHD and typically developing controls, TDCs) to see if there are areas of the brain that are under- or over-connected to the seed in the patient group (Greicius, 2008). By contrast, ICA is a data-driven approach to revealing spatial patterns of brain connections (Beckmann et al., 2005; Calhoun et al., 2001). Instead of specifying a seed area, ICA uses an algorithm to simultaneously decompose whole-brain BOLD signal into a set of time courses (components) and brain maps that describe how well each voxel's time course represents the component time course (Beckmann et al., 2005; Calhoun et al., 2001; Hyvärinen and Erkki Oja, 2000; McKeown and Terrence J. Sejnowski, 1998). In this way, sets of areas

sharing a similar BOLD time course (the component time course) can be identified (Greicius, 2008). The number of component time courses (and corresponding maps) the ICA algorithm finds is a variable controlled by the investigator; currently no "best" number of components is agreed-upon (Cole et al., 2010). The sets of functionally connected areas shown by the ICA-generated spatial maps are usually referred to as resting state networks (RSNs); the integrity of these networks can then be compared across patient and TDC groups (Greicius, 2008). Similar patterns of brain connectivity have been found using both seed-based and ICA techniques.

Between 5 and 12 resting state networks (RSNs) have been consistently found across multiple participants (Allen et al., 2011; Damoiseaux et al., 2006; De Luca et al., 2006; Thomason et al., 2011; van den Heuvel and H. E. Hulshoff Pol, 2010). Ostensibly, every component resolved through ICA can be thought of as a RSN; however, in reality some components represent noise (participant head motion, cardiac and respiratory cycles, etc), while others have not been universally agreed on as "true" RSNs (Cole et al., 2010). Several RSNs consist of areas that have been shown to support different brain functions in task-based studies, such as the visual network (comprised of areas of the occipital lobe), motor network (comprised of bilateral pre-central gyri), and left and right fronto-parietal (executive) networks (comprised of ipsilateral DLPFC and parietal cortices) (Biswal et al., 1995; Seeley et al., 2007; van den Heuvel and H. E. Hulshoff Pol, 2010). By contrast, the default mode network (DMN) has been identified as specifically a task-independent brain network (Buckner et al., 2008; Raichle et al., 2001). It is comprised of the medial prefrontal cortex/anterior cingulate cortex, the posterior

cingulate cortex/precuneus, and the inferior parietal lobules and is associated with daydreaming and introspection (Allen et al., 2011; Biswal et al., 1995; Damoiseaux et al., 2006; Raichle et al., 2001; van den Heuvel and H. E. Hulshoff Pol, 2010).

fMRI findings in ADHD research

Considerable research has been done investigating differences between patients with ADHD and TDCs using task-based fMRI techniques. From this perspective, task-based fMRI studies have sought to uncover aberrations in brain function using cognitive tasks that are difficult for ADHD patients (such as inhibitory control- and attention-related tasks). However, it is difficult to draw over-arching conclusions from this literature, as each study may use a different type or version of a task. Despite this drawback, a 2006 meta-analysis attempted to integrate findings from 16 task-based studies of ADHD (Dickstein et al., 2006). The authors found that there were several brain areas that consistently under-activated in patients with ADHD compared to TDCs regardless of task (Dickstein et al., 2006). These areas included portions of the frontal lobe (including anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), and inferior prefrontal cortex), along with portions of the basal ganglia, thalamus, and parietal cortices (Dickstein et al., 2006). The authors also performed a sub-analysis examining only studies that used a task designed to elicit an inhibitory-related brain response; here, ADHD patients showed under-activation in portions of the frontal lobe (including the inferior frontal gyrus (IFG) and bilateral pre-central gyri), along the midline of the brain (cingulate gyrus), in the parietal lobe (superior parietal lobule), and in the basal ganglia (caudate body) (Dickstein et al., 2006). Taken together, these

observations were interpreted as supporting the theory that fronto-striatal connections (connections between the frontal lobe, particularly the DLPFC, and the caudate and/or putamen), are compromised in ADHD (Dickstein et al., 2006). More recently, a 2013 meta-analysis examining a total of 34 task-based studies of ADHD (21 addressing inhibition tasks and 13 addressing attention tasks) found evidence similar to the 2006 meta-analysis (Hart et al., 2013). Here, ADHD patients showed under-activation in portions of the frontal lobe (the inferior frontal cortex, ACC, and supplemental motor area) as well as the basal ganglia and thalamus in response to inhibition tasks (Hart et al., 2013). Furthermore, patients with ADHD showed under-activation in the DLPFC, parietal areas, and the basal ganglia and thalamus in response to attention tasks (Hart et al., 2013). These findings support the theory that deficits in activation of different brain areas may underlie specific behavioral differences (i.e., performance on inhibition versus attention tasks) in patients with ADHD (Hart et al., 2013).

Compared to task-based fMRI experiments, results from rs-fMRI studies can be more easily synthesized as they investigate the differences in the way ADHD patients' brains are connected in the absence of task. However, research into differences in RSN integrity between ADHD patients and TDCs has been less comprehensive than the task-based literature. Here, investigations have largely focused on differences in DMN-related connections across the two groups, finding decreased resting state integrity within the DMN in ADHD patients (Castellanos et al., 2008; Fair et al., 2010; Posner et al., 2014; Qiu et al., 2011; Uddin et al., 2008). However, other RSN differences have also been found in patients with ADHD, including differences in the left and right executive

networks, salience network, and motor network (Castellanos and Erika Proal, 2012). This small but significant body of research only highlights the need for further investigation of ADHD-related differences in rs-FC, not only within the DMN but across other RSNs as well.

Using rs-fMRI results to diagnose ADHD

Until recently, the goal of many fMRI studies was to compare differences in brain activation or brain connections across two groups of research participants (for example, a patient group and a TDC group). As this work will investigate resting state measures, task-based approaches for group comparisons will not be elucidated here (for a good review of group comparisons using task-based fMRI, see (Samanez-Larkin and M. D'Esposito, 2008). For group comparisons of rs-fMRI measures, each connection between two brain areas is typically evaluated independently using a univariate test, such as a 2-sample t-test (Fox and M. Greicius, 2010; Greicius, 2008). The result of the t-test is then used to infer if the average strength of a connection between two areas or within an RSN is significantly larger or smaller in the patient group compared to the TDC group (Greicius, 2008). This approach results in the performance of a multitude of t-tests (one for each connection under study). In this way, a voxel-by-voxel map of significant over- or under-connections can be built to show which areas of the brain, on average, differ in the patient group. While this type of analysis is useful for understanding which brain areas may underlie the deficits associated with the condition of interest in the patient group, it has two major drawbacks. The first is that the results from this method can only be used to describe patients as a group, and cannot be used to make inferences about

individuals. That is, an increase in the average connectivity between two areas in the patient group may be found significant by t-test, but any single patient in that group may have increased, no change in, or decreased connectivity. It is therefore difficult to clinically apply the results of this approach on an individual level. Second, the use of multiple serial t-tests means that each connection is evaluated independently, in isolation from whatever is happening in all the other voxels in the brain. Simultaneous configurations of voxel connection that may underlie a disorder therefore cannot be investigated with this approach.

More recently, investigators have begun using techniques that take into account patterns of voxel activity or connectivity in individual participants' brains. These are known as multi-voxel pattern analysis (MVPA) methods, and include machine learning algorithms such as neural networks and support vector machines (SVMs) (Norman et al., 2006). Machine learning is a branch of computer science that focuses on teaching algorithms to identify patterns within data, with the goal of learning a general rule from that set of data that can then be used to make predictions in novel situations (Mitchell, 1997). SVMs are usually employed for binary classification; that is, they learn to discriminate between outcome A (for example, an ADHD diagnosis) and outcome B (a TDC) (Boser et al., 1992). An SVM typically accomplishes this by first taking a set of features (here, RSN integrity in each voxel) describing each individual in a group and mapping those individuals to a multi-dimensional space corresponding to the number of features used (Boser et al., 1992; Noble, 2006). That is, each participant is placed into the multi-dimensional space at a location that corresponds to their set of features. The SVM

then finds a hyperplane that optimally separates individuals with outcome A from individuals with outcome B based on their location (Noble, 2006). The SVM is thus "trained" to see the difference between an ADHD participant and a TDC participant based on their input features. From the point of view of the SVM, if a new participant has a set of features that locates them on the "ADHD" side of the hyperplane, they are predicted to have an ADHD diagnosis. Conversely, if the new participant's features cause them to fall on the "TDC" side of the hyperplane, they are predicted to NOT have an ADHD diagnosis. In this way, participant-level predictions about ADHD status can be made using patterns of brain connectivity.

In order to determine how accurate these SVM-generated predictions are, the investigator usually takes a dataset where the outcome is known for all participants and splits it up into training and test sets (Kohavi, 1995). The test set is held in reserve while the training set is used to teach the SVM to discriminate between the two classes (here, ADHD and TDC) (Ambroise and Geoffrey J. McLachlan, 2002; Kohavi, 1995). After the SVM is trained it is applied to the test set. How well the SVM performs in classifying participants in the test set can be ascertained by calculating metrics such as the accuracy (total number of correct predictions as ADHD or TDC over the total number of participants), sensitivity (total number of correct predictions as ADHD over the number of actual ADHD diagnoses), specificity (total number of correct predictions as TDC over the number of actual TDCs), positive predictive value (total number of correct predictions as ADHD over total predictions of ADHD), and negative predictive value (total number of correct predictions as TDC over total predictions of TDC). If the SVM

performs well on these metrics (for example, shows both high sensitivity and specificity), it may be considered for use as a diagnostic aid for children being evaluated for ADHD.

Several attempts at using machine learning techniques, including SVMs, to diagnose ADHD have already been described (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012; Zhu et al., 2008). Many of these investigations used the publicly-available ADHD-200 database, which is part of the 1000 Functional Connectomes Project and constitutes a total of 973 rs-fMRI acquisitions from participants diagnosed with ADHD as well as TDCs (http://fcon_1000.projects.nitrc.org/indi/adhd200/) (Milham et al., 2012). These scans were gathered across 8 different sites (Peking University, Bradley Hospital/Brown University, Kennedy Krieger Institute, NeuroIMAGE group, New York University Child Study Center, Oregon Health and Science University, University of Pittsburgh, and Washington University in St. Louis) and were originally released as training set of 776 scans, with 197 scans withheld as the test set. Teams of researchers could then compete in the ADHD-200 Global Competition (closed September 2011), where they attempted to build the best diagnostic classifier using the 776 released scans to train their algorithms. The trained algorithms were then submitted to the ADHD-200 consortium for testing on the withheld 197 scans. Out of 21 competing teams, the winners were able to classify ADHD versus TDC with a sensitivity of 21% and a specificity of 94% (Eloyan et al., 2012). Overall, the average prediction accuracy across competing teams was 49.52%, almost 11 percentage points greater than the 38.75% that would be expected by chance (http://fcon_1000.projects.nitrc.org/indi/adhd200/results.html). While these results

clearly show that there is important and usable information contained within an ADHD patient's neuroimaging data, it is also evident that there is considerable room for diagnostic improvement.

Focus of this dissertation

In the current work, three separate but related research objectives were pursued. First, we asked the question, "How does resting state function connectivity differ in a group of children who have ADHD compared to a group of TDCs?" To answer this, we sought to resolve RSNs through ICA in a subset of ADHD-200 participants. We further investigated how these RSNs differed in integrity in the subtypes of ADHD by examining differences between the combination subtype of ADHD versus TDC and the inattentive subtype of ADHD versus TDC. Based on the previous studies detailed above, we predicted that connectivity within the DMN would be decreased in patients with ADHD, regardless of subtype. As other networks have been less studied than the DMN in this disorder, we also specifically looked for differences in attention networks (the left and right executive networks and the salience network) between ADHD and TDC groups.

The second question we asked was, "Can one use RSN integrities to accurately predict ADHD diagnosis on the individual level?" We again used a subset of the ADHD-200 participants to address this question: first we divided participants into training and test sets, and then performed ICA separately on each. The integrities of the resulting RSNs (20 total) were then used as features for an SVM analysis. Each RSN was examined separately, to see which one yielded the highest accuracy for predicting ADHD

diagnosis. We again predicted that the DMN would be the most accurate RSN for discriminating ADHD participants from TDCS, based on previous studies' findings.

Finally, we asked "Can we integrate information from all the RSNs to raise diagnostic accuracy?" To address this question, we examined 3 different ways to combine the information provided by the 20 networks we identified (including 12 RSNs previously found in the literature and 8 components which were unclassified). The three methods we used were individual prediction scores, which summed a participant's binary diagnosis of ADHD or TDC across networks; decision values, which were used as a proxy for the confidence in the diagnosis of ADHD or TDC; and finally probability estimates, which determined the likelihood of a given participant having ADHD between 0% and 100%. We predicted that by using resting state connectivity information from every network, we would be able to diagnose ADHD more accurately than by using any single network.

In summary, ADHD is a disorder that interferes with a child's ability to function and develop normally. It is the most prevalent psychiatric disorder that affects children and adolescents, and can last into adulthood. The way it is currently diagnosed can be unintentionally subjective, so finding an intrinsic measure that could serve as a diagnostic aid is a worthwhile goal. To this end, both task-based fMRI and rs-fMRI have been applied to ADHD patient groups to investigate brain areas and networks that may underlie this disorder. Results from resting state studies have clearly implicated the DMN as abnormally connected in ADHD, and recent work using machine learning techniques has shown resting state information can be used to predict the presence of ADHD above

what would be expected by chance. The overarching goal of this work is therefore to first characterize ADHD patients (taken from a publically-available database of rs-fMRI acquisitions) in terms of their resting state connectivity, and then to see if an accurate diagnostic classifier (SVM) can be built using this information.

CHAPTER I

Introduction

Attention Deficit/Hyperactivity Disorder (ADHD) is a common psychiatric disorder that begins in childhood and may persist through adolescence and into adulthood (Biederman et al., 2006; Faraone et al., 2006). It is currently thought to affect up to 11% of school-age youth in the US (Visser et al., 2014) and 5.3% of children and adolescents world-wide (Polanczyk et al., 2007). ADHD is identified by the presence of hyperactive, impulsive, or inattentive behaviors that are beyond what would be expected as average for child of that age (American Psychiatric Association, 2013). Based on what type of symptoms a child exhibits, they may be classified as one of the three subtypes of ADHD: Hyperactive/Impulsive, Inattentive, or Combination (exhibiting both hyperactive and inattentive behaviors) (American Psychiatric Association, 2013). The behaviors that drive the ADHD diagnosis often lead to functional difficulty in a variety of areas. Academically, children with ADHD perform worse on measures of reading and mathematics than their typically developing peers (Biederman et al., 1996; Loe and Heidi M. Feldman, 2007). Socially, children with ADHD exhibit an impaired ability to identify causal relationships, and thus have difficulty understanding and responding appropriately to interpersonal problems (Pugzles Lorch et al., 2004; Sibley et al., 2010; Storer et al., 2014). At home, children and adolescents with ADHD exhibit greater parent-child conflict than their typically developing peers (Barkley et al., 1991; Edwards et al., 2001; Storer et al., 2014). Attempts to understand the cognitive underpinnings of these functional deficits have led to the investigation of executive function in youth with ADHD. Executive function, which includes such cognitive processes as inhibitory

control, selective attention, and working memory, has consistently been found to be impaired in children with ADHD compared to typically developing controls (TDCs) (Homack and Cynthia A. Riccio, 2004; Van Mourik et al., 2005; Willcutt et al., 2005). It has therefore been posited that ADHD can be thought of primarily as a disorder of executive function (Barkley, 1997; Willcutt et al., 2005).

One way to non-invasively investigate areas of the brain that may underlie ADHD-related deficits in executive function utilizes Magnetic Resonance Imaging (MRI) techniques. Specifically, task-based functional MRI (fMRI) is a method that quantifies relative changes in Blood Oxygen Level-Dependent (BOLD) signal while a participant is performing a specific cognitive task (Huettel et al., 2009). These BOLD signal changes are an indirect measure of brain activity: areas of the brain that have a higher concentration of oxygenated blood (due to increased neuronal firing in response to task performance) will generate a larger BOLD signal (Ogawa et al., 1990). Typically, BOLD signal is measured throughout the brain before the participant performs a cognitive task ("at rest"), and then again during task performance ("task") (Bandettini et al., 1992; Kwong et al., 1992; Ogawa et al., 1992). The change in BOLD signal in each brain area between these two states ("task" minus "at rest") is interpreted as a reflection of how involved, or active, that brain area was in the performance of the task (Huettel et al., 2009). Differences between youth with ADHD and TDCs in task-related brain responses have already been thoroughly investigated using this technique (Booth et al., 2005; Durston et al., 2003; Durston et al., 2007; Pliszka et al., 2006; Rubia et al., 1999; Rubia et al., 2005; Smith et al., 2006; Suskauer et al., 2008; Tamm et al., 2004; Vaidya et al.,

2005). While it is difficult to find consensus among the results of these studies, as many utilized different versions of executive function tasks in their methods, a 2006 meta-analysis attempted to synthesize ADHD-related differences in brain activation. It included 16 studies of children and adults with ADHD and found that portions of the frontal lobe, including anterior cingulate cortex (ACC), dorsolateral prefrontal cortex (DLPFC), and inferior prefrontal cortex, along with portions of the basal ganglia, thalamus, and parietal cortices were consistently under-activated when compared to TDCs (Dickstein et al., 2006). These findings support the theory that connections between the frontal lobe, particularly the DLPFC, and the striatum (caudate and putamen), or fronto-striatal connections, are compromised in ADHD (Dickstein et al., 2006). A more recent review considered the evidence not only from tasks designed to probe executive function, but also reward processing and timing, and concluded that 3 related but separate circuits are implicated in ADHD: the dorsofronto-striatal, orbitofrontal-striatal, and fronto-cerebellar circuits (Durston et al., 2011).

Task-based fMRI is useful for understanding which brain areas are involved in specific cognitive functions and how activation in these areas differs between patient and TDC groups; however, it is also possible to solely examine the brain "at rest." In resting state fMRI (rs-fMRI), the connections between different areas of the brain are studied in the absence of a task. In an rs-fMRI analysis, BOLD signal fluctuations from different brain regions are compared, and areas with similar changes in BOLD signal over time are thought to be working together, or functionally connected (Biswal et al., 1995). This is known as resting state functional connectivity (rs-FC). Sets of areas that are consistently

functionally connected can form resting state networks (RSNs) (Allen et al., 2011; Damoiseaux et al., 2006; De Luca et al., 2006; van den Heuvel and H. E. Hulshoff Pol, 2010). Between 5 (De Luca et al., 2006) and 12 (Thomason et al., 2011) RSNs have been identified; several of these map to areas of the brain associated with different functions identified through task-based studies, such as the visual network, motor network, and left and right fronto-parietal (executive) networks (Biswal et al., 1995; van den Heuvel and H. E. Hulshoff Pol, 2010). Perhaps the most famous RSN (and the one most often implicated in ADHD) is the default mode network (DMN), which has been associated with daydreaming and introspection (Buckner et al., 2008; Raichle et al., 2001). The DMN is made up of the medial prefrontal cortex/anterior cingulate cortex, the posterior cingulate cortex/precuneus, and the inferior parietal lobules (Allen et al., 2011; Biswal et al., 1995; Damoiseaux et al., 2006; Raichle et al., 2001; van den Heuvel and H. E. Hulshoff Pol, 2010). Several studies have found decreased resting state integrity of the DMN in children with ADHD (Castellanos et al., 2008; Fair et al., 2010; Qiu et al., 2011; Uddin et al., 2008). However, differences have also been found within many other RSNs in children with ADHD, including the left and right executive networks, salience network, and motor network (Castellanos and Erika Proal, 2012). Differences in fronto-striatal connections have also been found between ADHD and typically developing youth, corroborating task-based findings (Liston et al., 2011).

There are several ways to identify RSNs, both through hypothesis-driven and data-driven approaches. Here, we utilized a data-driven approach known as independent component analysis (ICA) (Beckmann et al., 2005; Calhoun et al., 2001; Hyvärinen and

Erkki Oja, 2000; McKeown and Terrence J. Sejnowski, 1998). ICA simultaneously resolves rs-fMRI signal from a group of individuals into a set of independent time courses (components) and a corresponding set of brain maps showing each voxel's contribution to that particular time course (Calhoun et al., 2009). Since areas of the brain with similar BOLD time courses are thought to be functionally connected, these maps can be interpreted as depicting putative RSNs. However, not all component maps resolved through ICA will represent RSNs; some will represent noise. Separating components that represent RSNs from components that represent noise is an important step after ICA: here, we employed a template-matching strategy where the results of our ICA analysis were spatially correlated to previously-published templates of pediatric RSNs (Thomason et al., 2011).

In the present study, we investigated group-level differences in RSN integrity between children with ADHD and TDCs using the publicly-available ADHD-200 database. Three different analyses were conducted: 1- Any ADHD versus TDC, in which the ADHD group included all 3 subtypes of ADHD; 2- Combination ADHD versus TDC, in which the ADHD group included only children with the Combination subtype of ADHD; and 3- Inattentive ADHD versus TDC, in which the ADHD group included only children with the Inattentive subtype of ADHD. Based on the behavioral and cognitive deficits associated with ADHD along with the results of previous imaging studies detailed above, we predicted that the RSNs that would show significant differences between ADHD and TDCs were the DMN and attention networks, such as the left and right executive networks and the salience networks, which all include portions of the

frontal lobe. We further expected that integrity in both the DMN and attention networks would be decreased in children with ADHD.

Materials and Methods

Data Sets: Three subsets of participants' rs-fMRI data were taken from the ADHD-200 database (http://fcon_1000.projects.nitrc.org/indi/adhd200/). Eight different sites contributed to the ADHD-200 database for a total of 973 rs-fMRI data acquisitions. These sites were: Peking University (Peking), Bradley Hospital/Brown University (Brown), Kennedy Krieger Institute (KKI), NeuroIMAGE, New York University Child Study Center (NYU), Oregon Health and Science University (OHSU), University of Pittsburgh (Pittsburgh), and Washington University in St. Louis (WashU) (Milham et al., 2012). At the time the publically-available ADHD-200 database was accessed, no information was available from the Brown Site, resulting in a maximum of 776 scans from the 7 remaining sites for the analyses presented here. The acquisition parameters for each site, including the number of participants used from each site for each of the 3 analyses described below, are presented in Table 1.1. Of the 776 available scans, 210 were deemed unacceptable for use: 115 failed the pre-processing procedure described below. 95 scans did not pass the quality control measure provided on the ADHD-200 website and an additional 59 scans were excluded due to participant head motion that exceeded 2.0 mm translational motion or 2.0° angular motion. Furthermore, any TDCs that had a history of medication use were excluded (n=5). This left a total of 502 participants suitable for the following analyses: 164 participants had a diagnosis of ADHD and 338 were TDCs. All scans were accompanied by phenotypic data describing

Table 1.1. Image acquisition parameters. This table contains information about how each rs-fMRI scan was performed at each of the ADHD-200 sites. How many participants from each site were included in this analysis can be seen in columns 2-4 (column 2 corresponds to the Any ADHD versus TDC analysis, column 3 corresponds to the Combination ADHD versus TDC analysis, and column 4 corresponds to the Inattentive ADHD versus TDC analysis). The total number of participants (n =) used for each analysis can be seen in the column header.

Table 1.1. Image Acquisition Parameters

Site	Any ADHD vs TDC (n = 324)	Comb. ADHD vs TDC (n = 220)	Inatt. ADHD vs TDC (n = 205)	Magnet Field Strength (Tesla)	MRI System	TR (msec)	TE (msec)	Flip Angle	Field of View (mm)	No. of Slices	Slice Thickness (mm)	Eyes Open or Closed
Peking	34 32	16 12	31 23	3.0 T	Siemens Magnetom Trio, A Tim	2000	30	90°	200	33 30	3.5 4.5	Open or Closed
KKI	45	40	28	3.0 T	Philips	2500	30	75°	256	47	3.0	Open
Neuro-IMAGE	0	0	0	1.5 T	Siemens Magnetom Avanto	1960	40	80°	224	35	3.0	Closed
NYU	140	101	72	3.0 T	Siemens Magnetom Allegra	2000	15	90°	240	33	4.0	Closed
OHSU	37	26	22	3.0 T	Siemens Magnetom Trio, A Tim	2500	30	90°	240	36	3.8	Open
Pittsburgh	18	11	16	3.0 T	Siemens Magnetom Trio, A Tim	1500	29	70°	200	29	4.0	Open
WashU	18	14	13	3.0 T	Siemens Magnetom Trio, A Tim	2500	27	90°	256	32	4.0	Open

ADHD: Attention Deficit/Hyperactivity Disorder; Peking: Peking University; KKI: Kennedy Krieger Institute; MRI: magnetic resonance imaging; NYU: New York University Child Study Center; OHSU: Oregon Health and Science University; Pittsburgh: University of Pittsburgh; TDC: Typically Developing Control; TE: echo time; TR: repetition time; WashU: Washington University in St. Louis

participants' subtype of ADHD, if appropriate (combination, hyperactive, or inattentive), and their age, IQ scores, gender, and handedness (left, right, or ambidextrous).

Data Preprocessing and Preparation: Pre-processing of all 502 rs-fMRI scans was performed using the Data Processing Assistant for Resting-State fMRI (DPARSFA, http://www.restfmri.net; (Song et al., 2011). DPARSFA was developed as a user-written extension for SPM8 (Statistical Parameter Mapping– Welcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) and works in conjunction with the previously-developed Resting-State fMRI Data Analysis toolkit (REST). Preprocessing steps for each scan included removing the first 10 time points, correction for slice acquisition time differences, realignment for motion correction, normalization to the MNI EPI template (voxel size 3X3X3), spatial smoothening with a 4 mm FWHM kernel, detrending, and temporal band-pass filtering to 0.01-0.08 Hz (Deligiannidis et al., 2013).

Participant Matching: Out of the 502 scans, 3 analyses were conducted: 1- Any ADHD versus TDC, which used all 164 participants diagnosed with ADHD, including the combination subtype (n=91), the hyperactive subtype (n=2), and the inattentive subtype (n = 71). 160 TDCs were then matched to the 164 ADHD participants based on age, gender, handedness, and IQ score similarities (Table 1.2) leading to a total of 324 participants. 2- Combination ADHD versus TDC, which used only ADHD participants diagnosed with the combination subtype (n=91); these were matched to 126 TDCs based on age, gender, handedness, and IQ score (Table 1.3) leading to a total of 220

Table 1.2. Cohort matching for Any ADHD versus TDC. This table compares age, IQ, proportion of male participants, and handedness of the Any ADHD and the TDC group. Values are mean (Standard Deviation) for the first two rows and percent in the last two rows.

Table 1.3. Cohort matching for Combination ADHD versus TDC. This table compares age, IQ, proportion of male participants, and handedness of the Combination ADHD and the TDC group. Values are mean (Standard Deviation) for the first two rows and percent in the last two rows.

Table 1.4. Cohort matching for Inattentive ADHD versus TDC. This table compares age, IQ, proportion of male participants, and handedness of the Inattentive ADHD and the TDC group. Values are mean (Standard Deviation) for the first two rows and percent in the last two rows.

Table 1.2. Cohort Matching for Any ADHD versus TDC			
	ADHD (n = 164)	**TDC (n= 160)**	**p value**
Mean Age in Years (SD)	11.2 (2.5)	11.6 (2.4)	0.12
Mean IQ (SD)	105.3 (13.8)	108.1 (12.3)	0.06
Percent Male Participants	74%	79%	0.27
Percent Right-Handed Participants	83%	88%	0.20
ADHD: Attention Deficit/Hyperactivity Disorder; SD: Standard Deviation; TDC: Typically Developing Control			

Table 1.3. Cohort Matching for Combination Subtype of ADHD versus TDC			
	ADHD (n = 91)	**TDC (n= 129)**	**p value**
Mean Age in Years (SD)	10.6 (2.3)	11.2 (2.5)	0.09
Mean IQ (SD)	106.8 (13.4)	106.4 (11.9)	0.85
Percent Male Participants	80.0%	74.4%	0.34
Percent Right-Handed Participants	78.9%	84.3%	0.22
ADHD: Attention Deficit/Hyperactivity Disorder; SD: Standard Deviation; TDC: Typically Developing Control			

Table 1.4. Cohort Matching for Inattentive Subtype of ADHD versus TDC			
	ADHD (n = 71)	**TDC (n= 134)**	**p value**
Mean Age in Years (SD)	12.0 (2.6)	11.8 (2.5)	0.48
Mean IQ (SD)	103.3 (14.4)	106.2 (12.9)	0.16
Percent Male Participants	67.6%	56.7%	0.13
Percent Right-Handed Participants	90.1%	89.5%	0.92
ADHD: Attention Deficit/Hyperactivity Disorder; SD: Standard Deviation; TDC: Typically Developing Control			

participants. 3- Inattentive ADHD versus TDC, which examined only ADHD participants diagnosed with the inattentive subtype (n=71); these were matched to 134 TDCs based on age, gender, handedness, and IQ score (Table 1.4) leading to a total of 205 participants. The 2 participants with the hyperactive subtype of ADHD were not studied further.

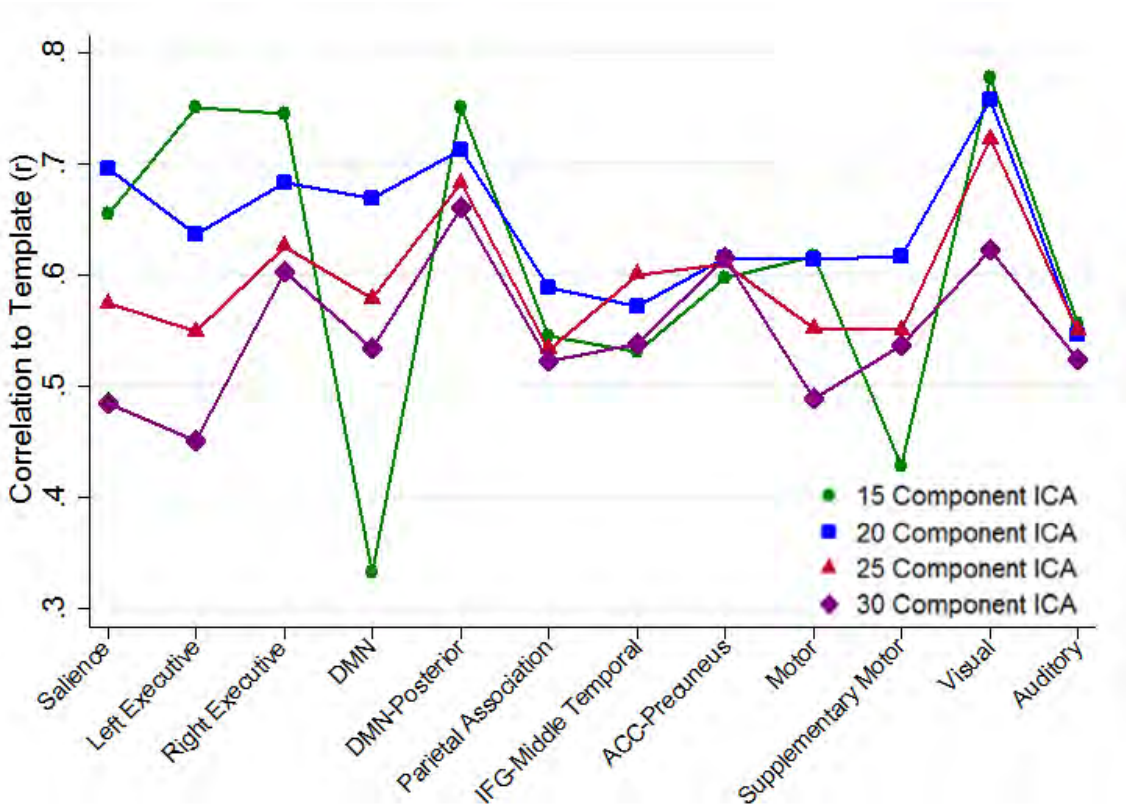
Independent Component Analysis (ICA): For each of the 3 analyses, ICA was performed using the Group ICA of fMRI (GIFT) toolbox (http://mialab.mrn.org/software/) following the method outlined in Calhoun et al (Calhoun et al., 2001; Calhoun et al., 2009). The general procedure for implementing ICA was as follows: First, standard subject-specific Principle Component Analysis (PCA) was done to reduce the dimensionality of each individual's data (Calhoun et al., 2009). For each analysis, individual participants' PCA results were then concatenated into one group (including both ADHD and TDC participants; n= 324 for Any ADHD versus TDC, n=220 for Combination ADHD versus TDC, and n=205 for Inattentive ADHD versus TDC) and a 2nd, group-level PCA was performed to further reduce the dimensionality of the rs-fMRI data. Next, the INFOmax algorithm was employed to obtain maximally spatially independent group-level components (Bell and Terrence J. Sejnowski, 1995). Lastly, the dual-regression option in the GIFT toolbox was used to predict each individual participant's component time courses and maps (Filippini et al., 2009; Zuo et al., 2010).

Initially, ICA was repeated 4 times for Any ADHD versus TDC: each time the number of components used to resolve the independent, underlying networks obtained from the rs-fMRI data was varied. 15, 20, 25, or 30 components were used to determine the number of components that most successfully resolved pediatric resting state

networks. The metric used to determine the best number of components was the spatial correlation between the resulting independent components and 12 pre-defined brain network templates published by Thomason et al (Thomason et al., 2011). These templates were generated using 65 healthy children and adolescents aged 9-15 years old. Of the 4 trials (15, 20, 25, or 30 components), the 20-component ICA yielded components that most consistently matched (had the highest spatial correlation to) the pediatric resting state network templates (Figure 1.1). Thus, 20 components was taken as the optimal number to use in our ICA analyses, and the ICAs for the Combination ADHD versus TDC and the Inattentive ADHD versus TDC analyses were each carried out using only 20 components. Template matching also allowed for the identification of components that represented "true," previously-established resting state networks. Therefore, template matching was also performed for the Combination ADHD versus TDC and Inattentive ADHD versus TDC analyses, to determine which of the 20 components resolved through ICA represented previously-established resting state networks, and which represented noise.

Statistics: Voxel-wise 2-sample t-tests for each resting state network were performed using SPM8 across ADHD and TDC groups for each of the three analyses. The site of scan was controlled for in each t-test. Areas where ADHD component coherency was greater than TDC were considered over-connected to the network in question. Areas where ADHD component coherency was less than that of TDC were considered under-connected to the network in question. Resulting maps of t-scores describing the

Figure 1.1. Correlation of networks resolved using 15-, 20-, 25-, and 30-component analyses to templates of previously-derived pediatric resting state networks (Thomason et al., 2011). ACC: Anterior Cingulate Cortex; DMN: Default Mode Network; IFG: Inferior Frontal Gyrus.


Correlation to Template (r)
.3
.4
.5
.6
.7
.8
15 Component ICA
20 Component ICA
25 Component ICA
30 Component ICA
Salience
Left Executive
Right Executive
DMN
DMN-Posterior
Parietal Association
IFG-Middle Temporal
ACC-Precuneus
Motor
Supplementary Motor
Visual
Auditory

differences between ADHD and TDC groups were thresholded at $p < 0.01$ and a cluster size of 25 or more contiguous voxels.

Results

Any ADHD versus TDC: There were no significant differences in age, IQ scores, gender, or handedness between the ADHD and TDC groups (Table 1.2). ICA performed using 20 components consistently yielded the best matches to pediatric RSN templates, with the visual network template having the highest correlation ($R = 0.76$) and the auditory network template having the lowest correlation ($R = 0.55$) (Figure 1.1). Examples of components resolved through ICA and found to match pediatric RSN templates are shown in Figure 1.2. The salience network, left executive network, and right executive network (attention networks) as well as the DMN and DMN-posterior (default mode networks) were chosen for display as they were the networks expected to differ most between ADHD and TDC groups, based on the findings of previous studies (Castellanos et al., 2008; Dickstein et al., 2006; Durston et al., 2011; Fair et al., 2010; Liston et al., 2011; Qiu et al., 2011; Uddin et al., 2008). Areas of the brain that showed differences in RSN integrity between ADHD and TDC groups are shown in Figure 1.3 (attention networks) & Figure 1.4 (default mode networks). Areas found to be over-connected in the ADHD group are show in red (positive T-scores) whereas areas that were under-connected in the ADHD group are shown in blue (negative T-scores). The list of the differences in attention networks (the salience network, Figure 1.3A; the left executive network, Figure 1.3D, and the right executive network, Figure 1.3G) can be found in Table 1.5. The salience network showed significant over-connections in medial and

Figure 1.2. Examples of networks resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. A. salience network; B. left executive network; C. right executive network; D. default mode network; E. default mode network, posterior portion. Colorbar represents t-scores.

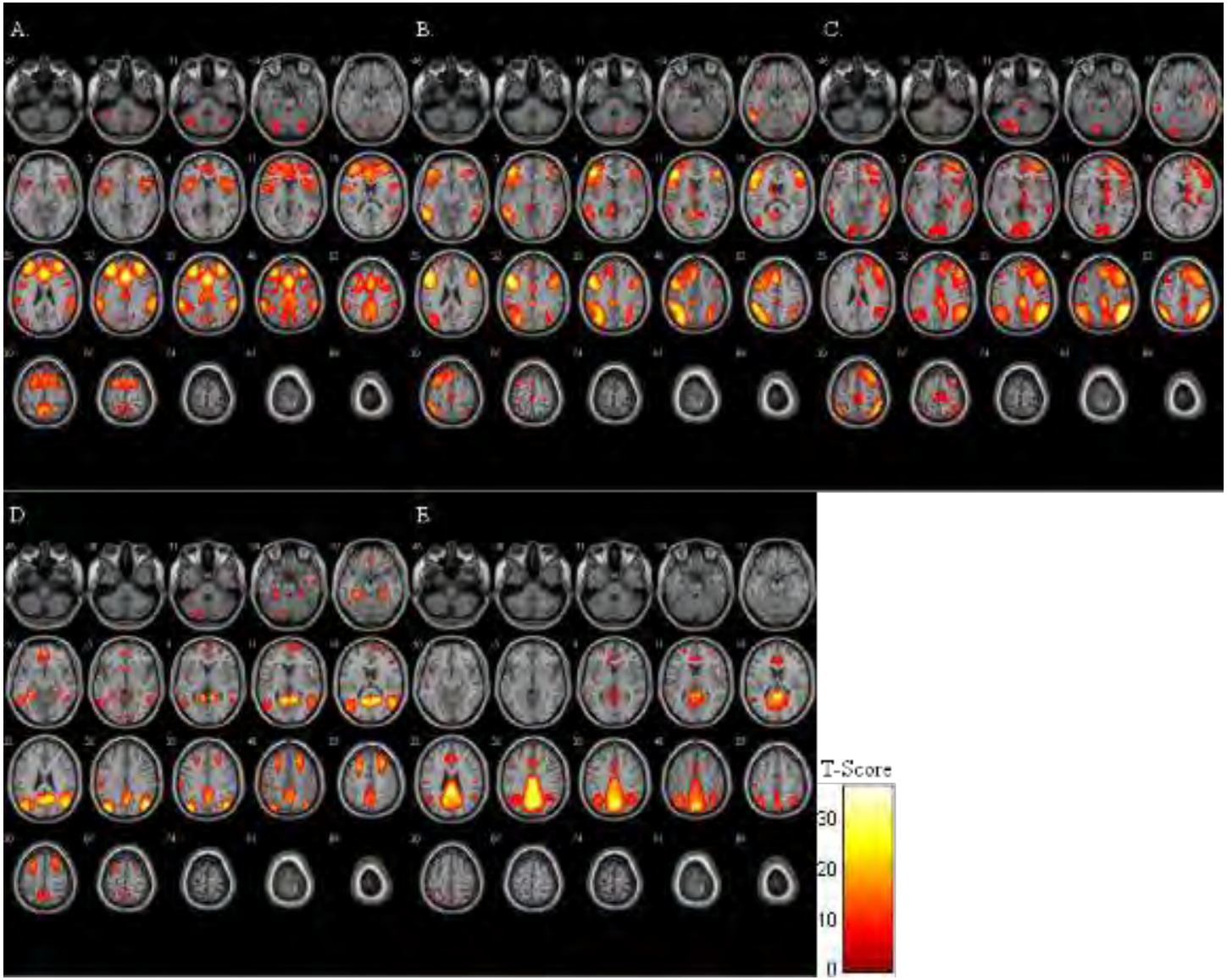

A.
B.
C.
D.
E.
T-Score
30
20
10
0

Figure 1.3. Areas of the brain that show significant differences in connectivity in attention networks between children with ADHD and TDCs. **A**. any ADHD diagnosis versus TDC connectivity in the salience network; **B**. combination subtype ADHD diagnosis versus TDC connectivity in the salience network; **C**. inattentive subtype ADHD diagnosis versus TDC connectivity in the salience network; **D**. any ADHD diagnosis versus TDC connectivity in the left executive network; **E**. combination subtype ADHD diagnosis versus TDC connectivity in the left executive network; **F**. inattentive subtype ADHD diagnosis versus TDC connectivity in the left executive network; **G**. any ADHD diagnosis versus TDC connectivity in the right executive network; **H**. combination subtype ADHD diagnosis versus TDC connectivity in the right executive network; **I**. inattentive subtype ADHD diagnosis versus TDC connectivity in the right executive network. Colorbar represents T-Scores: red areas correspond to significant over-connections, blue areas correspond to significant under-connections when compared to TDC. ACC: anterior cingulate cortex; IFG: inferior frontal gyrus; IPL: inferior parietal lobule; MedFG: medial frontal gyrus; MFG: middle frontal gyrus; SFG: superior frontal gyrus.

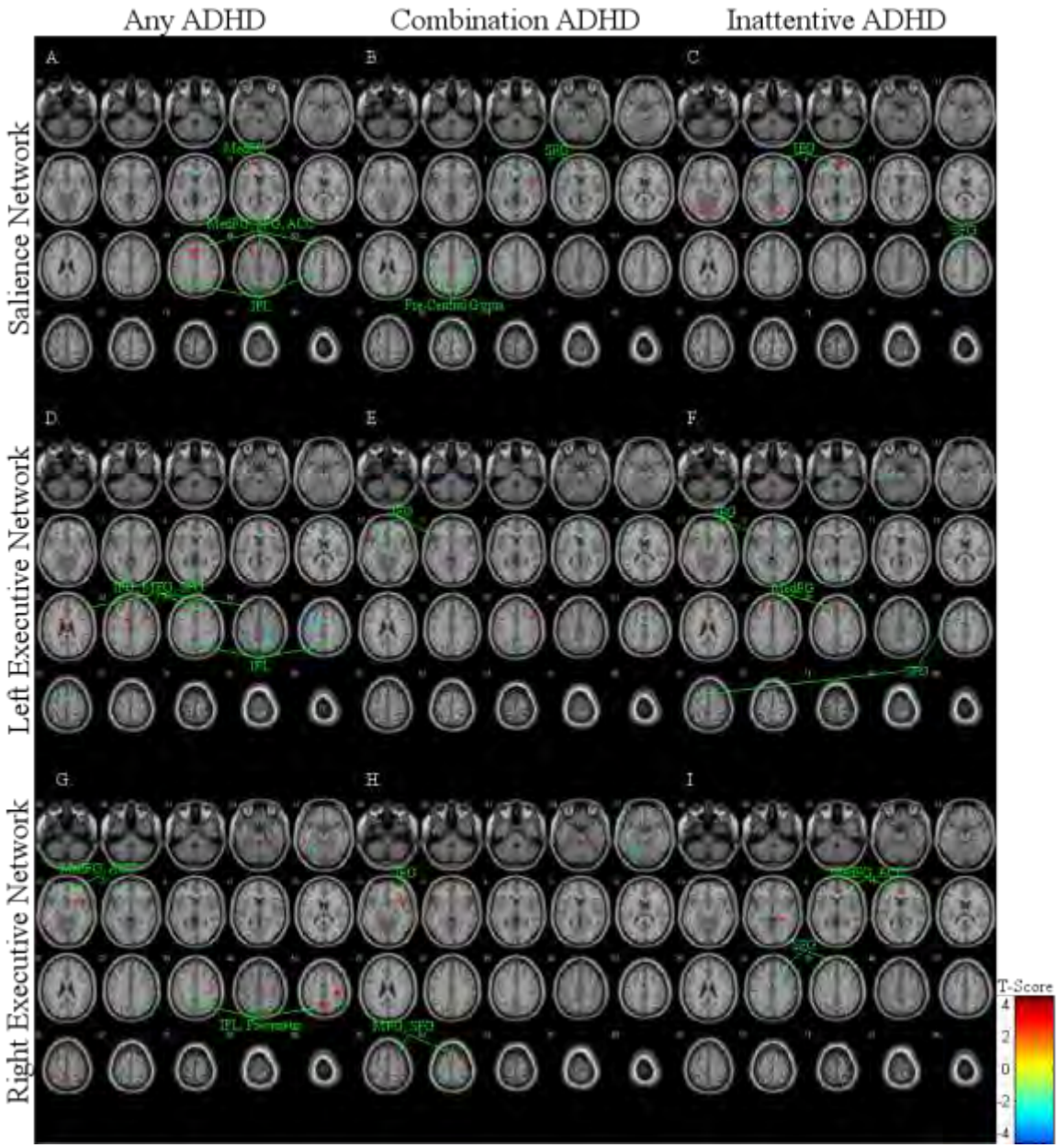

Any ADHD
Combination ADHD
Inattentive ADHD
Salience Network
Left Executive Network
Right Executive Network
A
B
C
D
E
F
G
H
I
T-Score
4
2
0
-2
-4

Figure 1.4. Areas of the brain that show significant differences in connectivity in the default mode network (DMN) between children with ADHD and TDCs. **A**. any ADHD diagnosis versus TDC connectivity in the DMN; **B**. combination subtype ADHD diagnosis versus TDC connectivity in the DMN; **C**. inattentive subtype ADHD diagnosis versus TDC connectivity in the DMN; **D**. any ADHD diagnosis versus TDC connectivity in the DMN-posterior; **E**. combination subtype ADHD diagnosis versus TDC connectivity in the DMN-posterior; **F**. inattentive subtype ADHD diagnosis versus TDC connectivity in the DMN-posterior. Colorbar represents T-Scores: red areas correspond to significant over-connections, blue areas correspond to significant under-connections. IFG: inferior frontal gyrus; IPL: inferior parietal lobule; MedFG: medial frontal gyrus; MFG: middle frontal gyrus; SFG: superior frontal gyrus.

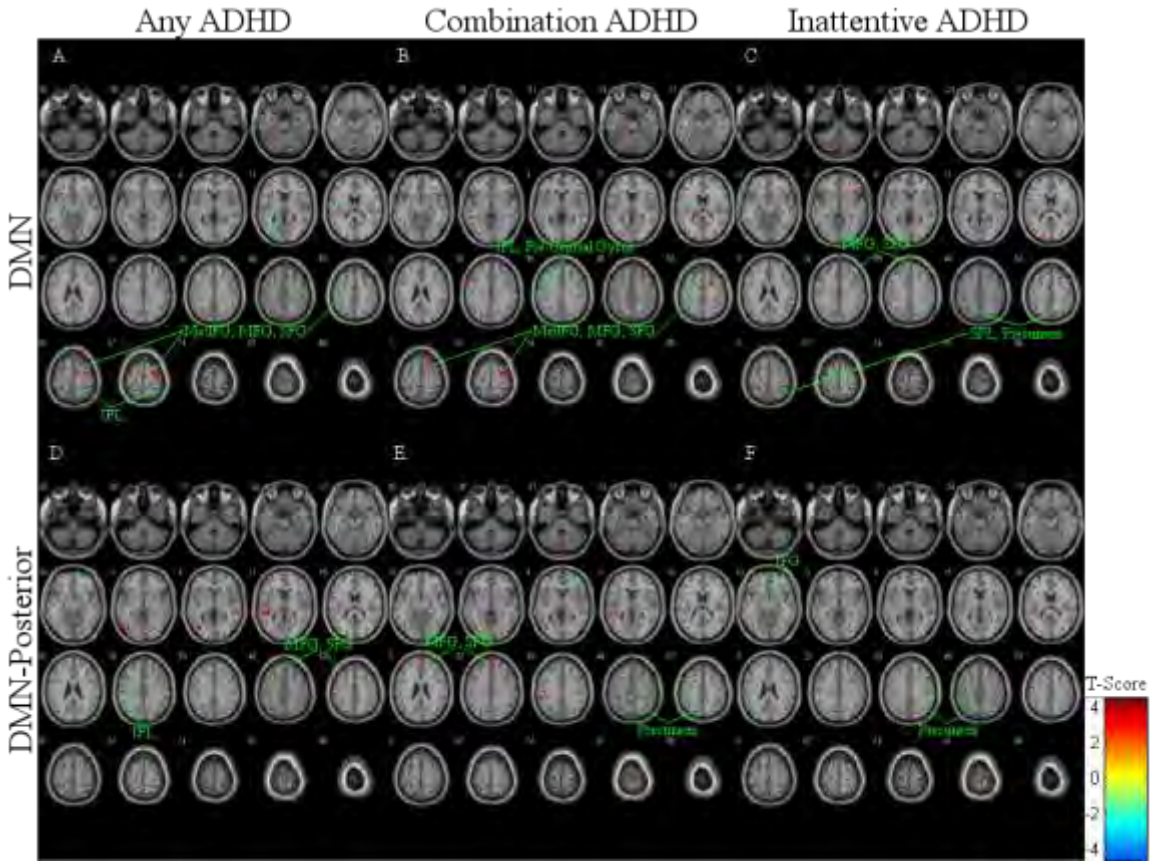

Any ADHD
Combination ADHD
Inattentive ADHD
DMN
DMN-Posterior
A
B
C
D
E
F
T-Score
4
2
0
-2
-4

Table 1.5. Significantly over- or under-connected brain areas in attention networks for Any ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.5. Significantly Over- or Under-Connected Brain Areas in Attention Networks for Any ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
Salience	Over-Connections	-57	24	3	28	Inferior Frontal Gyrus	45, 47	26
						Superior Temp. Gyrus	38	1
		-12	60	12	44	Medial Frontal Gyrus	10	15
		-9	48	51	135	Superior Frontal Gyrus	8, 9	17
						Medial Frontal Gyrus	6, 8	59
						Cingulate Gyrus	32	21
	Under-Connections	-57	-51	42	34	Inferior Parietal Lobule	40	19
		-39	-39	57	36	Inferior Parietal Lobule	40	13
						Postcentral Gyrus	22	22
Left Executive	Over-Connections	45	21	36	91	Middle Frontal Gyrus	9, 10	14
						Inferior Frontal Gyrus	46	4
		0	-24	24	46	Cingulate Gyrus	23	5
						R Caudate Body		3
		-18	36	48	39	Superior Frontal Gyrus	8	14
						Middle Frontal Gyrus	9	2
	Under-Connections	-6	-69	0	28	Lingual Gyrus	18, 30	8
						Cuneus	23	2
		-45	21	15	33	Inferior Frontal Gyrus	45	1
		-27	-54	39	59	Superior Parietal Lobule	7	13
						Inferior Parietal Lobule	40	1
						Precuneus	19, 39	13
		6	-63	42	33	Precuneus	7	21
		27	-54	45	59	Inferior Parietal Lobule	40	8
						Precuneus	7	14
		-21	-57	57	27	Superior Parietal Lobule	7	11
						Precuneus	7	9
		-36	-15	57	43	Precentral Gyrus	4	16
		-6	0	54	57	Medial Frontal Gyrus	6	29
						Superior Frontal Gyrus	6	16
		-36	18	54	66	Middle Frontal Gyrus	6	19
						Superior Frontal Gyrus	8	7
Right Executive	Over-Connections	30	-39	-21	39	R Cerebellum		19
						Fusiform Gyrus	20, 37	8
						Parahippocampal Gyrus	36	1
		30	15	-6	50	Inferior Frontal Gyrus	47	15
						Insula	13	8
		-36	-45	42	30	Inferior Parietal Lobule	40	7
		45	-21	48	119	Postcentral Gyrus	2,3	46
						Precentral Gyrus	4	14
		-3	-60	54	79	Precuneus	7	40
	Under-Connections	-18	-60	-15	47	L Cerebellum		36
						Fusiform Gyrus	19	5
		-6	42	-3	29	Anterior Cingulate	24, 32	9
						Medial Frontal Gyrus	10	2
		48	-60	33	34	Angular Gyrus	39	5
						Supramarginal Gyrus	40	1
		42	-57	45	47	Inferior Parietal Lobule	40	18
						Superior Parietal Lobule	7	5
		-24	-42	66	31	Postcentral Gyrus	2, 5	20

superior portions of the frontal lobe, including the ACC, compared to TDC. The salience network also showed significant under-connection to the inferior parietal lobule in the ADHD group. Similarly, the left executive network also showed over-connection to portions of the frontal lobe (including the superior, middle, and inferior frontal gyri) in the ADHD group, while also showing under-connection to the inferior parietal lobule. By contrast, the right executive network showed over-connections to areas including the precuneus and inferior parietal lobule as well as under-connections to the ACC and medial frontal gyrus.

All differences in the DMN (Figure 1.4A) and the DMN-posterior (Figure 1.4D) are listed in Table 1.6. Most notably, both default networks showed significant over-connections in the ADHD group to portions of the frontal lobe (particularly the middle and superior frontal gyri) and well as significant under-connections to the inferior parietal lobule.

Combination ADHD versus TDC: There were no significant differences in age, IQ scores, gender, or handedness between the ADHD and TDC groups (Table 1.3). ICA was performed only once using 20 components to resolve the resting state networks. Resulting networks were comparable to those found in the Any ADHD versus TDC analysis (examples shown in Figure 1.2). Significant over- and under-connections between Combination ADHD and TDC in attention networks can be seen in Figure 1.3 (the salience network, Figure 1.3B; the left executive network, Figure 1.3E, and the right executive network, Figure 1.3H) and are listed in Table 1.7. Most notably, there were

Table 1.6. Significantly over- or under-connected brain areas in the DMN for Any ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.6. Significantly Over- or Under-Connected Brain Areas in the DMN for Any ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
DMN	Over-Connections	-42	0	60	25	Precentral Gyrus	6	6
						Middle Frontal Gyrus	6	10
		39	-9	63	139	Precentral Gyrus	6 & 4	50
						Middle Frontal Gyrus	6	40
						Medial Frontal Gyrus	6	23
						Superior Frontal Gyrus	6	14
		24	30	57	42	Superior Frontal Gyrus	6 & 8	31
						Middle Frontal Gyrus	6	4
		9	-45	21	25	Posterior Cingulate	23 & 31	10
		21	-45	12	33	Lingual Gyrus	19	1
						Posterior Cingulate	30	3
		3	15	-9	26	Anterior Cingulate	25	11
						Left Caudate Head		5
						Right Caudate Head		5
		39	0	-12	26	Insula	13	3
		-21	-54	-24	27	Cerebellum - Culmen		14
						Cerebellum - Declive		11
						Cerebellum - Dentate		2
	Under-Connections	39	-48	57	30	Inferior Parietal Lobule	40	6
						Superior Parietal Lobule	5 & 7	2
		39	54	24	49	Middle Frontal Gyrus	10 & 46	19
						Superior Frontal Gyrus	10	13
		-27	-66	12	47	Cuneus	18	1
		-48	-33	9	27	Transverse Temp. Gyrus	41	6
						Insula	13	7
		-21	-24	12	40	Left Thalamus		25
		36	15	-30	28	Superior Temp. Gyrus	38	8
Posterior DMN	Over-Connections	-39	-72	0	54	Middle Occipital Gyrus	19 & 37	40
		-42	-36	12	56	Superior Temp. Gyrus	22 & 41	43
						Transverse Temp. Gyrus	41	11
		-39	24	51	25	Middle Frontal Gyrus	8	20
						Superior Frontal Gyrus	8	2
	Under-Connections	33	51	6	43	Middle Frontal Gyrus	10 & 11	19
						Superior Frontal Gyrus	10	11
		-33	-57	30	36	Inferior Parietal Lobule	40	2

Table 1.7. Significantly over- or under-connected brain areas in attention networks for Combination ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.7. Significantly Over- or Under-Connected Brain Areas in Attention Networks for Combination ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
Salience	Over-Connections	48	3	3	30	Insula	13	20
		-12	57	12	34	Superior Frontal Gyrus	10	16
	Under-Connections	-27	0	-3	33	L Putamen		17
						L Globus Pallidus		3
		-27	45	21	61	Superior Frontal Gyrus	10	26
		-30	-81	24	27	Superior Occip. Gyrus	19	11
		57	-9	30	54	Precentral Gyrus	4, 6	34
						Postcentral Gyrus	1-3, 40, 43	20
		-60	-6	30	27	Precentral Gyrus	4, 6	23
						Postcentral Gyrus	3	4
		-57	-54	42	28	Inferior Parietal Lobule	40	13
Left Executive	Over-Connections	-51	30	-9	26	Inferior Frontal Gyrus	47	1
	Under-Connections	-18	15	-9	33	L Putamen		12
						Caudate Head		9
						Inferior Frontal Gyrus	47	2
Right Executive	Over-Connections	27	-33	27	25	R Cerebellum		17
						Fusiform Gyrus	20	2
		36	15	-18	55	Inferior Frontal Gyrus	47	22
						Insula	13	5
						Superior Temp. Gyrus	38	1
		42	-21	45	46	Postcentral Gyrus	3	20
						Precentral Gyrus	4	5
		24	9	63	30	Middle Frontal Gyrus	6	14
						Superior Frontal Gyrus	6	12
	Under-Connections	-24	-60	-18	66	L Cerebellum		47
						Fusiform Gyrus	19, 37	6
		-39	-15	3	35	Insula	13	9
						Claustrum		4
		-24	-45	66	55	Postcentral Gyrus	2, 5	20
						Superior Parietal Lobule	40	2

significant over-connections in the ADHD group within all 3 attention networks to portions of the frontal lobe, including the inferior frontal gyrus, middle frontal gyrus, and superior frontal gyrus. There was also significant under-connection between the salience network and the pre-central gyri bilaterally in the ADHD group.

Significant over- and under-connections within default mode networks are shown in Figure 1.4 (DMN, Figure 1.4B; DMN-posterior, Figure 1.4E). An exhaustive list of significantly difference areas between Combination ADHD and TDC groups can be found in Table 1.8. Most notably, there were several significant over-connections to clusters in the frontal lobe (middle frontal gyri and superior frontal gyri for both DMN and DMN-posterior) in the ADHD group. Both the DMN and the DMN-posterior also showed significant under-connection to the precuneus or inferior parietal lobule.

Inattentive ADHD versus TDC: There were no significant differences in age, IQ scores, gender, or handedness between the ADHD and TDC groups (Table 1.5). ICA was performed only once using 20 components to resolve the resting state networks. Resulting networks were comparable to those found in the Any ADHD versus TDC analysis (examples shown in Figure 1.2). Significant over- and under-connections between Inattentive ADHD and TDC in attention networks can be seen in Figure 1.3 (the salience network, Figure 1.3C; the left executive network, Figure 1.3F, and the right executive network, Figure 1.3I) and are listed in Table 1.9. Most notably, there were significant over-connections in the ADHD group within all 3 attention networks to medial portions of the frontal lobe, including the medial frontal gyrus, inferior frontal

Table 1.8. Significantly over- or under-connected brain areas in the DMN for Combination ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.8. Significantly Over- or Under-Connected Brain Areas in the DMN for Combination ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
DMN	Over-Connections	-21	-54	-24	57	Cerebellum - Culmen		26
						Cerebellum - Declive		23
						Cerebellum - Tuber		5
						Cerebellar Tonsil		3
		-54	12	-12	31	Superior Temporal Gyrus	38	18
						Middle Temporal Gyrus	21	13
		6	30	-12	25	Anterior Cingulate	32, 24	21
						Medial Frontal Gyrus	11	4
		-24	-45	6	26	L Caudate Tail		2
		30	-78	18	44	Cuneus	19, 7	26
		-3	-45	21	60	Posterior Cingulate	23, 30	28
		-57	15	30	29	Inferior Frontal Gyrus	9, 45	23
						Middle Frontal Gyrus	9	6
		39	-9	63	204	Superior Frontal Gyrus	6, 8	73
						Middle Frontal Gyrus	6	47
						Precentral Gyrus	4, 6	29
						Medial Frontal Gyrus	6,32	12
	Under-Connections	-51	-33	9	26	Transverse Temporal Gyrus	41	4
						Insula	13	3
		-63	-21	39	48	Postcentral Gyrus	3, 1, 2	25
						Precentral Gyrus	4, 6	14
						Inferior Parietal Lobule	40	2
		-30	42	42	33	Middle Frontal Gyrus	9	18
						Superior Frontal Gyrus	8	14

…Continued on next page…

…Table 1.8 continued from previous page…

Posterior DMN	Over-Connections	-3	-69	-30	36	Cerebellum -Declive	20
						Cerebellum - Uvula	8
						Cerebellum - Declive of Vermis	4
						Cerebellum - Pyramis	4
		-39	-75	0	39	Middle Occipital Gyrus	19	32
						Inferior Temporal Gyrus	37	5
		-42	-39	12	48	Transverse Temporal Gyrus	41	18
		-51	-39	39	36	Inferior Parietal Lobule	40	1
						Postcentral Gyrus	1, 2	5
		3	48	27	44	Medial Frontal Gyrus	9	22
						Superior Frontal Gyrus	9	6
	Under-Connections	33	51	3	28	Middle Frontal Gyrus	10	8
						Superior Frontal Gyrus	10	2
		-21	21	42	32	Middle Frontal Gyrus	8	18
						Superior Frontal Gyrus	8	10
		0	-36	45	29	Paracentral Lobule	5 & 31	5
						Precuneus	7	23

Table 1.9. Significantly over- or under-connected brain areas in attention networks for Inattentive ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.9. Significantly Over- or Under-Connected Brain Areas in Attention Networks for Inattentive ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
Salience	Over-Connections	15	-12	-18	29	Parahippocampal Gyrus	28, 34	22
		-27	-66	-12	89	Fusiform Gyrus	19	15
						Lingual Gyrus	18	56
		15	-81	-12	88	Lingual Gyrus	18	59
		-54	21	6	41	Inferior Frontal Gyrus	45, 47	36
						Precentral Gyrus	44	2
		9	57	-3	43	Superior Frontal Gyrus	10	11
		-18	-78	39	25	Precuneus	7	24
	Under-Connections	21	-93	3	54	Cuneus	18, 19	23
		24	15	0	34	Insula	13	9
		3	-69	51	39	Precuneus	7	31
		-24	27	54	36	Superior Frontal Gyrus	8	27
						Middle Frontal Gyrus	8	9
Left Executive	Over-Connections	9	-42	-42	42	R Cerebellum - Dentate		8
						R Cerebellum - Nodule		8
						Pons		8
						Medulla		1
		-33	24	-3	26	Inferior Frontal Gyrus	47	4
						Insula	13	5
		0	45	33	57	Medial Frontal Gyrus	6, 9	16
						Superior Frontal Gyrus	8	2
	Under-Connections	-30	-84	-33	25	Cerebellum - Declive		15
						Cerebellum - Uvula		10
		-24	-21	-24	30	Parahippocampal Gyrus	28, 35	8
		-9	-72	0	43	Lingual Gyrus	17, 18	7
		18	-39	0	39	Parahippocampal Gyrus	30	3
						Thalamus		4
		45	15	-3	26	Insula	13	7
						Inferior Frontal Gyrus	47	7
		-18	-78	33	45	Precuneus	7	14
						Cuneus	19	7
		-36	18	54	36	Middle Frontal Gyrus	6, 8	21
						Superior Frontal Gyrus	8	11

…Continued on next page…

…Table 1.9 continued from previous page…

Right Executive	Over-Connections	-3	42	12	36	Anterior Cingulate	32	6
						Medial Frontal Gyrus	10	1
	Under-Connections	54	-21	30	52	Inferior Parietal Lobule	40	15
						Postcentral Gyrus	2	3
		39	42	36	41	Superior Frontal Gyrus	9, 10	19
						Middle Frontal Gyrus	9, 10	19

gyrus, and the ACC. By contrast, there were also significant under-connections in the ADHD group within all 3 attention networks to lateral portions of the frontal lobe, specifically to clusters including the superior frontal gyrus.

Significant over- and under-connections within default mode networks are shown in Figure 1.4 (DMN, Figure 1.4C; DMN-posterior, Figure 1.4F). An exhaustive list of significantly different areas between Inattentive ADHD and TDC groups can be found in Table 1.10. Most notably, there were significant over-connections to the precuneus/superior parietal lobule in both the default networks in the ADHD group. By contrast, there were significant under-connections to clusters in the frontal lobe (including the inferior frontal gyrus, middle frontal gyrus, and superior frontal gyrus) in both default networks in the ADHD group.

Discussion

The aim of the present study was to determine how RSN integrity differed between children diagnosed with ADHD and their typically developing peers. A total of 3 analyses were conducted, the first including children with any of the 3 subtypes of ADHD, the second including only children with the combination subtype of ADHD, and the third including only children with the inattentive subtype of ADHD. We hypothesized that the DMN and attention networks would show significant differences in coherence between ADHD and TDC groups in all 3 analyses. We found evidence that supported our hypothesis, although the brain areas where RSN integrity was abnormal in the ADHD group differed between the analyses, based on what type of ADHD was under scrutiny. We further predicted that the ADHD groups would show decreased, rather than

Table 1.10. Significantly over- or under-connected brain areas in the DMN for Inattentive ADHD versus TDC. This table lists areas of the brain that were significantly different between ADHD and TDC groups. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Table 1.10. Significantly Over- or Under-Connected Brain Areas in the DMN for Inattentive ADHD versus TDC

RSN	Type of Connection	Peak Coordinates X	Y	Z	Cluster Size	Anatomical Area	Brod-mann Area	Voxels in Area
DMN	Over-Connections	-24	-69	-30	59	Cerebellum - Uvula		26
						Cerebellum - Pyramis		18
						Cerebellum - Declive		9
						Cerebellum - Tuber		6
		0	18	-6	29	Anterior Cingulate	25	14
						L Caudate Head		5
		-3	-99	15	36	Cuneus	18 & 19	30
		9	-72	57	47	Precuneus	7	18
						Superior Parietal Lobule	7	13
	Under-Connections	27	63	27	33	Middle Frontal Gyrus	10	11
						Superior Frontal Gyrus	9 & 10	9
		45	39	36	31	Middle Frontal Gyrus	46	18
						Superior Frontal Gyrus	9	5
		36	-48	39	26	Inferior Parietal Lobule	40	11
Posterior DMN	Over-Connections	0	-69	6	25	Cuneus	30	7
						Posterior Cingulate	23	12
		-45	-33	12	26	Superior Temporal Gyrus	41	11
						Transverse Temporal Gyrus	41	5
						Insula	13	7
		24	-48	45	26	Precuneus	7 & 31	8
	Under-Connections	15	24	-9	36	Inferior Frontal Gyrus	11 & 47	2
		-9	-21	9	26	L Thalamus		25
		48	-69	30	27	Precuneus	19 & 39	12
		-42	-57	27	49	Angular Gyrus	39	5

increased, RSN integrity compared to TDCs. While there were areas of decreased coherence observed for both attention networks and the DMN, we found multiple areas of increased coherence in each of these networks as well. A brief summary of our interpretations of these findings is presented in Table 1.11.

For the Any ADHD versus TDC analysis, the 3 attention networks investigated (the salience network and the left and right executive networks) displayed different patterns of over- and under- connection to different brain areas in the Any ADHD group. The salience network, which has been found to underpin executive function by identifying decision-relevant stimuli and at the core is comprised of the dorsal anterior cingulate (dACC) and orbitofrontal/insular cortices (Figure 1.2A), showed increased connections within itself (Seeley et al., 2007). That is, the medial frontal gyrus (MedFG), the superior frontal gyrus (SFG), and the anterior cingulate cortex (ACC), which fell within the areas defined as the salience network in Figure 1.2A, all showed increased integrity in the any ADHD group (Figure 1.3A). While it is difficult to interpret what over-connection indicates in an absolute sense, here we speculate that this could be a potential compensatory mechanism for children with ADHD: as identifying decision-relevant stimuli (paying attention) is difficult for this group, perhaps over-connections within the RSN that supports this function develop. Conversely, it is equally possible that this over-connection is a hallmark of the disorder and not a compensatory response; with cross-sectional data it is impossible to know. Still, the pattern of under-connections to the

Table 1.11. Summary of Chapter 1 findings. This table condenses the information given in the discussion section for a quick over-view of our results. Attention network findings are given in the top half of the table, while DMN findings are shown in the bottom half. The 2nd column describes over-connections in the ADHD group compared to the TDC group; the 3rd column describes under-connections in the ADHD group compared to the TDC group.

Table 1.11. Summary of Chapter 1 Findings		
Attention Networks	**Over-Connected**	**Under-Connected**
Any ADHD vs TDC	**Salience Network**: ↑ connections within itself (MedFG, SFG, and ACC ↑ integrity) **L. Exec. Network**: ↑ connections to the R. Exec. Network (right IFG, MFG, SFG ↑ integrity) **R. Exec. Network**: ↑ connection to DMN (↑ integrity in IPL)	**Salience Network**: ↓ connection to the DMN (IPL ↓ integrity) **L. Exec. Network**: ↓ connection to the DMN (IPL ↓ integrity) **R. Exec. Network**: ↓ connections within R. Exec. Network (MedFG, ACC ↓ integrity)
Combination ADHD vs TDC	**Salience Network**: similar to findings to Any ADHD analysis **L. Exec. Network**: ↑ connections within itself (left IFG↑ integrity) **R. Exec. Network**: ↑ connections within itself (right IFG, MFG, SFG ↑ integrity)	**Salience Network**: ↓ connection to bilateral motor areas **L. Exec. Network**: no change in connections to DMN **R. Exec. Network**: no change in connections to DMN
Inattentive ADHD vs TDC	**Salience Network**: ↑ connection to midline areas of frontal lobe (IFG ↑ integrity) **L. Exec. Network**: ↑ connection to midline areas of frontal lobe (IFG, MedFG ↑ integrity) **R. Exec. Network**: ↑ connection to midline areas of frontal lobe (MedFG, ACC ↑ integrity)	**Salience Network**: ↓ connection to lateral areas of frontal lobe (SFG ↓ integrity) **L. Exec. Network**: ↓ connection to lateral areas of frontal lobe (SFG ↓ integrity) **R. Exec. Network**: ↓ connection to lateral areas of frontal lobe (SFG ↓ integrity)
Default Networks	**Over-Connected**	**Under-Connected**
Any ADHD vs TDC	**DMN & DMN-Posterior**: ↑ connections to attention networks (MedFG, MFG, and SFG ↑ integrity)	**DMN & DMN-Posterior**: ↓ connections within itself (IPL, precuneus ↓ integrity)
Combination ADHD vs TDC	**DMN & DMN-Posterior**: ↑ connections to attention networks (MedFG, MFG, and SFG ↑ integrity)	**DMN & DMN-Posterior**: ↓ connections within itself (IPL, precuneus ↓ integrity)
Inattentive ADHD vs TDC	**DMN & DMN-Posterior**: ↑ connections within itself (precuneus ↑ integrity)	**DMN & DMN-Posterior**: ↓ connections to areas in frontal lobe (IFG, MFG, SFG ↓ integrity)
ACC: anterior cingulate cortex; DMN: default mode network; IFG: inferior frontal gyrus; IPL: inferior parietal lobule; MedFG: medial frontal gyrus; MFG: middle frontal gyrus; R. Exec.: Right Executive; SFG: superior frontal gyrus; TDC: typically developing control		

salience network in the Any ADHD group can be seen as supporting the compensatory theory: here, the inferior parietal lobule (IPL, an area associated with the DMN) was less strongly connected to the salience network in the Any ADHD group. This may be interpreted as hyper-segregation in the Any ADHD group: the salience network has become less strongly connected to, and therefore more thoroughly separated from, areas that underlie the DMN (which is associated with daydreaming and acts in opposition to attention networks) (Buckner et al., 2008; Raichle et al., 2001). Again, it is possible that for children with ADHD, hyper-segregating the salience network from the DMN may be the result of attempts to pay better attention.

While the pattern of under-connection to (and so potential hyper-segregation from) the IPL also held true for the left executive network in the Any ADHD group, the pattern of over-connection for this RSN was slightly different (Figure 1.3D). Here, instead of over-connection within the areas defined as the left executive network in Figure 1.2B, the left executive network showed multiple over-connections within the areas defined as the right executive network in Figure 1.2C, specifically the IFG, middle frontal gyrus (MFG), and superior frontal gyrus (SFG, over-connections shown in Figure 1.3D). The left and right executive networks have been found to underpin executive function by preparing a reaction to decision-relevant stimuli; the dorsolateral prefrontal cortices (DLPFCs) and portions of parietal cortex form their core (Seeley et al., 2007). The over-connections between the left and right executive networks may indicate a closer relationship between these two RSNs in the Any ADHD group. On the other hand, under-connections within the right executive network (MedFG, ACC) and over-connections

between the right executive network and the IPL (Figure 1.3G) may indicate that the right executive network is compromised in the Any ADHD group. The over-connections between these two executive RSNs described above could therefore be interpreted as compensating for deficits in the right executive network.

In the Combination ADHD versus TDC analysis, there were over-connections for the Combination ADHD group within the areas defining all 3 attention networks (Figure 1.3B, 3E, and 3H). This is similar to the findings of the Any ADHD analysis for the salience and left executive networks, and is in line with the idea that children with ADHD may show over-connections within their attention networks as a compensatory response. By contrast, there were not consistent under-connections to areas of the DMN (IPL or precuneus) in the Combination ADHD group, arguing against hyper-segregation of these networks. There was, however, significant under-connection within the salience network to the bilateral pre-central gyri (Figure 1.3E). This is of particular interest, as these areas have been shown to under-activate in response to a motor task in participants with ADHD (Mostofsky et al., 2006; Valera et al., 2010). The areas that comprise the resting state motor network (Supplemental Figure 1.4A) include those areas activated by motor tasks (Biswal et al., 1995), and so it appears that here the under-connection within the salience network occurred in areas usually ascribed to the motor network. One possible interpretation of this finding is that the salience network may have impaired influence in these motor areas in children with ADHD. As this under-connection was present specifically in the Combination ADHD group, which includes children with hyperactive symptoms, it is possible that this impaired influence may contribute to the hyperactive

symptoms of children with ADHD. Further study may be warranted to elucidate this relationship: it may be worthwhile to investigate whether there is under connection between the salience network and motor areas in a group of children diagnosed with the hyperactive subtype of ADHD, as they would experience purely hyperactive symptoms.

The results of the Inattentive ADHD versus TDC analysis showed a different pattern of over- and under- connection within the attention networks than the Any ADHD or Combination ADHD analyses. Here, the over-connections within the frontal lobes for all 3 attention networks seemed to occur along the midline (IFG, MedFG, ACC), whereas under-connections within the frontal lobes tended to occur laterally (SFG). These over-connected medial areas cover one of the main hubs of the salience network (the dACC), as well as the medial prefrontal cortex (mPFC; ACC and MedFG) and the ventromedial prefrontal cortex (vmPFC; IFG) (Seeley et al., 2007). While the salience network is known to subserve the identification of relevant stimuli, a wider swath of the mPFC has been shown to play a role in performance monitoring (such as detecting errors and decision uncertainty) (Ridderinkhof et al., 2004; Seeley et al., 2007). Impairments in the vmPFC have been associated with difficulty making decisions with long-term consequences (Bechara et al., 2000). The combination of these areas (the dACC hub of the salience network, the mPFC, and the vmPFC) can therefore be tied to focusing on relevant cues and making good decisions (Bechara et al., 2000; Ridderinkhof et al., 2004; Seeley et al., 2007). Over-connections within attention networks to these areas in the Inattentive ADHD group may be interpreted as a potential compensatory mechanism, similar to the over-connections seen in the salience network for the Any ADHD analysis:

as identifying decision-relevant stimuli and making good decisions based on those stimuli is difficult for this group, perhaps over-connections in areas that support these functions develop. Conversely, the lateral under-connections to the SFG seem to fall in areas belonging to the left and right executive networks. As these networks are responsible for preparing reactions to stimuli/decisions (Seeley et al., 2007), perhaps the under-connection to the salience network and within the left and right executive networks themselves is specifically related to the inattentive symptoms experienced by this group.

The pattern of over- and under- connection within the networks found to represent the DMN was similar for the Any ADHD and Combination ADHD analyses (Figure 1.4). In both cases, the DMN/DMN-posterior showed significant under-connection in the IPL and precuneus in the ADHD group, consistent with previous findings of reduced integrity within the DMN in ADHD (Castellanos et al., 2008; Fair et al., 2010; Qiu et al., 2011; Uddin et al., 2008). Both the Any ADHD and Combination ADHD analyses also revealed over-connection within the DMN/DMN-posterior and areas of the frontal lobe, including the MedFG, MFG, and SFG. These areas fall within one or more of the attention networks detailed above; the largest over-connections appeared to fall within areas usually ascribed to the right executive network (Figure 1.4A and 4B). Taken together, we interpret the findings that areas of the DMN are less coherently connected to each other and that areas not normally associated with the DMN (MedFG, MFG, SFG) become more coherent with the DMN as evidence that the DMN is less segregated from attention networks in children with ADHD. On the contrary, the Inattentive ADHD analysis revealed an opposite pattern of over- and under-connection in default networks. In this

analysis, the DMN/DMN-posterior showed over-connections to the precuneus and under-connections to portions of the frontal lobe, including the IFG, MFG, and SFG. This reversal of findings (over-connection within the DMN and under-connection to areas associated with attention networks) implies that in children with the inattentive subtype of ADHD, the DMN may be hyper-segregated from attention networks. This is in line with the know function of the DMN (task-free introspection) and the behavioral deficits of the inattentive subtype (in ability to focus and pay attention) (American Psychiatric Association, 2013; Buckner et al., 2008; Raichle et al., 2001).

While the analyses conducted in this study revealed multiple differences in rs-FC between children with ADHD and TDCs, it is important to note that only a subset of the results are presented and discussed here. In addition to the 5 networks (salience, left executive, right executive, DMN, and DMN-posterior) examined above, an additional 7 networks were also investigated (parietal, IFG-middle temporal, ACC-precuneus, motor, supplementary motor, visual, and auditory) for differences between Any ADHD versus TDC, Combination ADHD versus TDC, and Inattentive ADHD versus TDC groups. These results are shown in Supplemental Figures 1 – 7. For each figure, the component that best represented the RSN in question is shown in part A, and then for that RSN the results of the Any ADHD versus TDC comparison are shown in part B, the results of the Combination ADHD versus TDC comparison are shown in part C, and finally the results of the Inattentive ADHD versus TDC comparison are shown in part D. Furthermore, for the 5 networks discussed above, only differences in areas relevant to either the DMN or

the attention networks were discussed; for example, differences in RSN integrity within the occipital lobe were not talked about here.

Overall, there were a few trends in the patterns of over- and under-connections in our analyses that could be used to summarize our data (Table 1.11). First, there tended to be over-connections within areas related to attention networks (salience or executive networks) across all 3 analyses (Any ADHD, Combination ADHD, and Inattentive ADHD). The notable exception to this was the right executive network in the Any ADHD analysis, which showed over-connection to the IPL. Second, the pattern of under-connections within each attention network seemed to be different for the Any ADHD, Combination ADHD, and Inattentive ADHD analyses: the Any ADHD analysis revealed under-connection to the IPL for the salience and left executive networks, but over-connection to areas of the frontal lobe in the right executive network. The salience network in the Combination ADHD analysis, on the other hand, showed under-connection to bilateral pre-central cortices, and all 3 attention networks showed under-connections to lateral portions of the frontal lobe in the Inattentive analysis. Lastly, the patterns of over- and under- connection in the default networks tended towards over-connection in areas of the frontal lobe and under-connection within the DMN itself, particularly in the IPL and precuneus. The notable exception here was the Inattentive ADHD analysis, which showed the reverse pattern: over-connection within the DMN and under-connection to areas within the frontal lobe. Taken as a whole, we believe these findings support the conclusion that (1) attention networks are generally over-connected

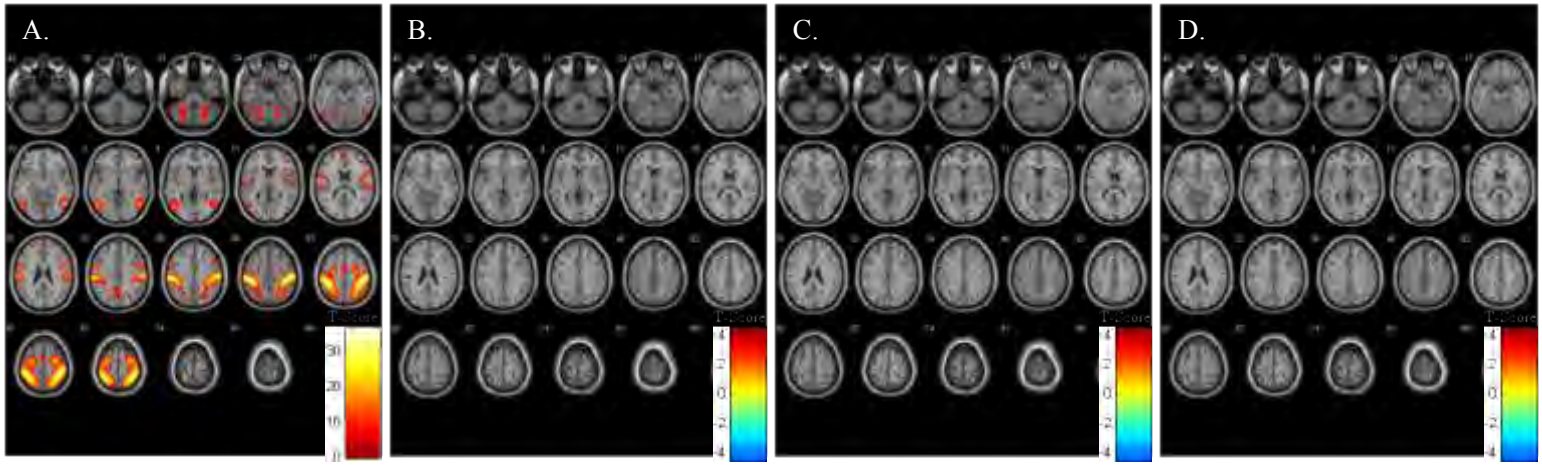
in ADHD and (2) the DMN is under-connected in combination subtype ADHD, but is over-connected in inattentive subtype ADHD.

Our findings have several implications for future studies using RSNs to further understand ADHD. The fact that significant differences in the ADHD group were found both in the connections (1) between brain areas within the attention and default networks themselves and (2) across these networks means that studying any single network in isolation may not reveal the full picture of atypical connectivity for ADHD patients. While the DMN has been the network of choice for investigating ADHD-related connectivity differences, the findings of the current study support investigation of attention networks as well. Going forward, even a cursory examination of the differences across ADHD and TDC groups in **all** RSNs may be advisable, as opposed to focusing on any single network or circuit. It is also important to bear in mind that the differences found here describe group changes in connectivity; therefore these RSN abnormalities cannot be used to make inferences about individual patient deficits. These group-level differences are useful for pointing out candidate brain areas (or here, RSNs) for future studies, but cannot be used diagnostically. To this end, the next two chapters will address methods for using information about individual patients' RSN integrities to build a classifier capable of diagnosing ADHD on an individual level.

Supplemental Information

SI Figure 1.1. The parietal association network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the parietal association network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the parietal association network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

Parietal Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC

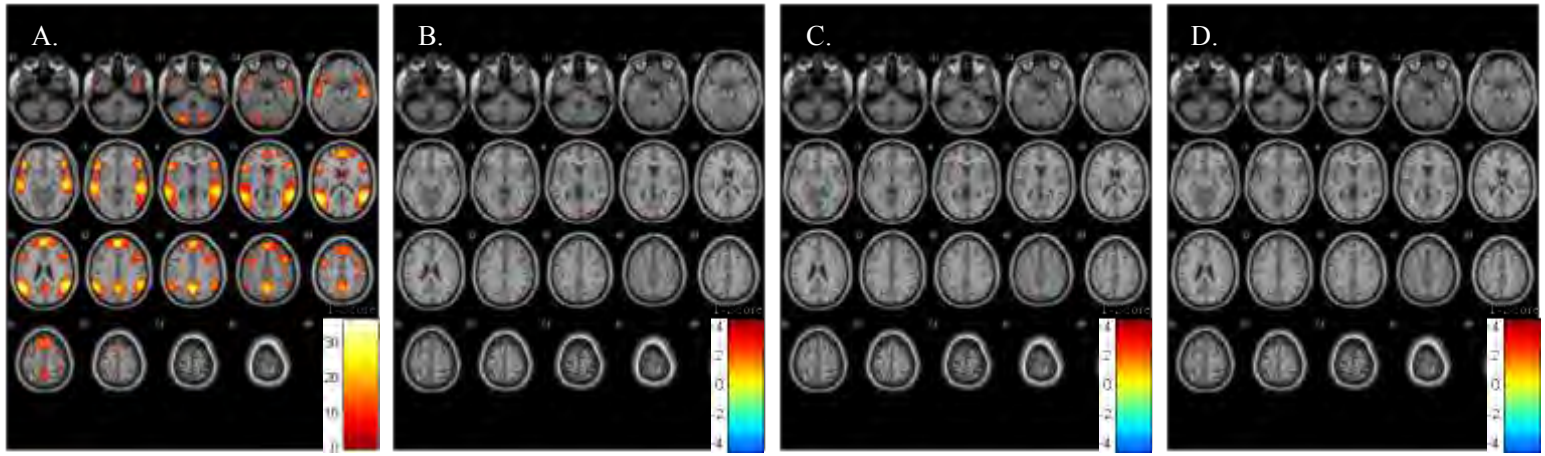


SI Table 1.1. Significantly over- or under-connected brain areas in the parietal association network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Parietal Association Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	24	-45	-30	30	Cerebellum - Culmen		26
						Cerebellum - Declive		1
		-15	42	-15	29	Middle Frontal Gyrus	11, 47	6
		-45	-24	24	40	Insula	13	4
						Postcentral Gyrus	43	1
		6	-51	72	25	Precuneus	7	6
	Under-Connections	57	-39	15	33	Inferior Parietal Lobule	40	2
						Insula	13	1
						Superior Temporal Gyrus	22, 42	5
		-6	-66	24	43	Precuneus	7, 31	19
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	33	-45	-27	31	Cerebellum - Culmen		24
		30	-24	-12	25	Hippocampus		3
		15	45	36	31	Superior Frontal Gyrus	9	19
						Medial Frontal Gyrus	9	12
	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-33	-87	-12	26	Middle Occipital Gyrus	18	9
		51	-72	12	40	Middle Temporal Gyrus	39	10
		45	36	36	27	Middle Frontal Gyrus	9, 46	19
						Superior Frontal Gyrus	9	5
	Under-Connections	15	-42	-42	32	Cerebellum - Tonsil		19
						Cerebellum - Dentate		1
						Pons		6
						Right Brainstem		6
						Superior Temporal Gyrus	22, 37	7
						Middle Occipital Gyrus	19	1
		60	-36	12	32	Superior Temporal Gyrus	41, 22	5
		12	30	21	25	Medial Frontal Gyrus	9	3
						Anterior Cingulate	32	2
		9	51	30	56	Superior Frontal Gyrus	9	37
						Medial Frontal Gyrus	9	13
		-3	-39	51	25	Paracentral Lobule	5	9

SI Figure 1.2. The inferior frontal gyrus-middle temporal network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the inferior frontal gyrus-middle temporal network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the inferior frontal gyrus-middle temporal network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

IFG-Mid Temporal Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC



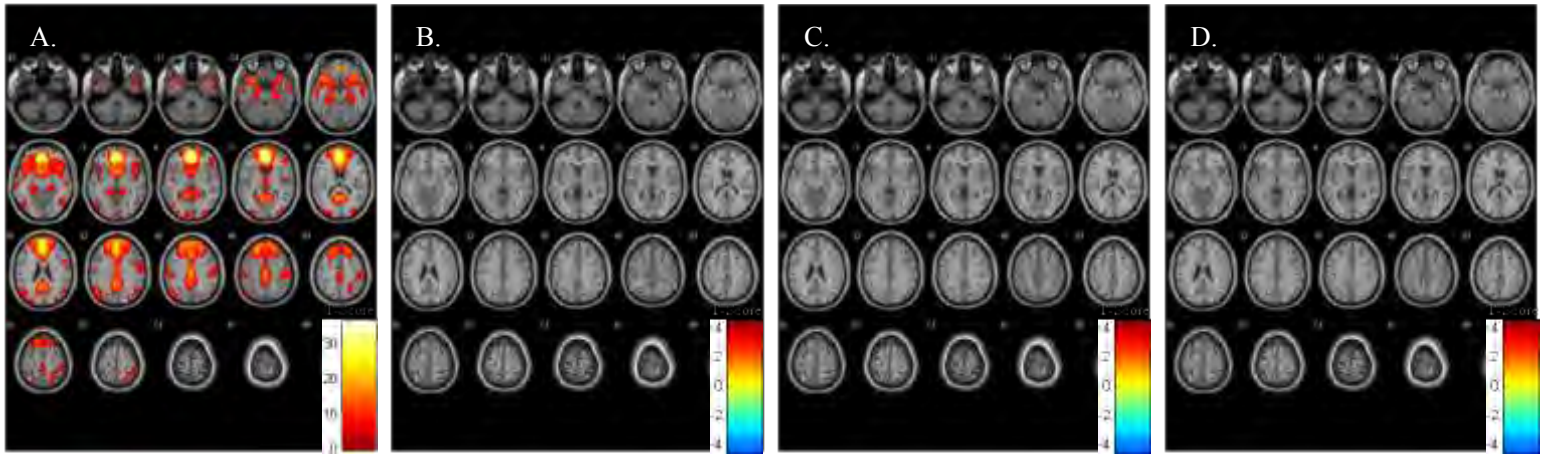
SI Table 1.2. Significantly over- or under-connected brain areas in the inferior frontal gyrus-middle temporal network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

IFG-Middle Temporal Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	33	-87	0	103	Middle Occipital Gyrus	18	14
						Cuneus	17, 19	6
						Lingual Gyrus	18	1
		-18	-12	3	30	L. Thalamus		25
		54	42	6	25	Inferior Frontal Gyrus	45, 46	6
		36	0	36	30	Middle Frontal Gyrus	6, 9	5
						Precentral Gyrus	6	2
	Under-Connections	42	-69	-33	47	Cerebellum - Tuber		29
						Cerebellum - Declive		11
						Cerebellum - Uvula		7
		-6	36	-12	27	Medial Frontal Gyrus	11	4
						Anterior Cingulate	32	5
		-39	-63	12	48	Middle Temporal Gyrus	19, 37, 39	11
		9	-36	12	37	R Thalamus		2
		45	-48	33	46	Inferior Parietal Lobule	40	5
		0	-30	30	36	Post. Cingulate Gyrus	23, 31	10
		36	-57	51	31	Inferior Parietal Lobule	40	12
						Superior Parietal Lobule	7	10
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	27	-93	6	25	Middle Occipital Gyrus	3	4
		24	9	18	31	R Putamen		3
		36	0	36	26	Inferior Frontal Gyrus	6	4
						Middle Frontal Gyrus	6	3
	Under-Connections	39	-72	-33	57	Cerebellum - Tuber		26
						Cerebellum - Declive		18
						Cerebellum - Uvula		11
						Cerebellum - Pyramis		1
						Cerebellum - Culmen		1
		21	-66	-9	57	Lingual Gyrus	19	8
						Fusiform Gyrus	18	3
						Cerebellum		25
		9	-36	15	44	R Thalamus		8
		36	-60	45	74	Inferior Parietal Lobule	40	22
						Superior Parietal Lobule	7	14 of 16
						Supramarginal Gyrus	39	1 of 9
		3	21	63	30	Superior Frontal Gyrus	6	9 of 21

	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	15	-57	-36	28	Cerebellum - Dentate		9
						Cerebellum - Declive		5
						Cerebellum - Pyramis		1
						Cerebellum - Culmen		1
						Cerebellum - Nodule		1
		-15	-48	-27	26	Cerebellum - Culmen		14
						Cerebellum - Declive		6
						Cerebellum - Dentate		4
		33	-33	-18	29	Parahippocampal Gyrus	36	6
						Hippocampus		5
						Caudate Tail		1
		-6	-66	15	42	Precuneus	30, 31	16
		-27	-33	18	32	Insula	13	3
						Caudate Tail		1
		45	6	39	32	Inferior Frontal Gyrus	6, 9	8
	Under-Connections							
		-42	-63	18	36	Middle Temporal Gyrus	21, 22, 39	13
		24	48	3	32	Middle Frontal Gyrus	10	12
						Superior Frontal Gyrus	10	3
						Inferior Frontal Gyrus	10	3
						Medial Frontal Gyrus	10	1
		21	-84	18	52	Cuneus	18, 19	12
		12	33	39	71	Medial Frontal Gyrus	6, 8, 9	44
						Superior Frontal Gyrus	8	22
		24	-18	54	41	Precentral Gyrus	4, 6	11

SI Figure 1.3. The anterior cingulate cortex-precuneus network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the anterior cingulate cortex-precuneus network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the anterior cingulate cortex-precuneus network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

ACC-Precuneus Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC



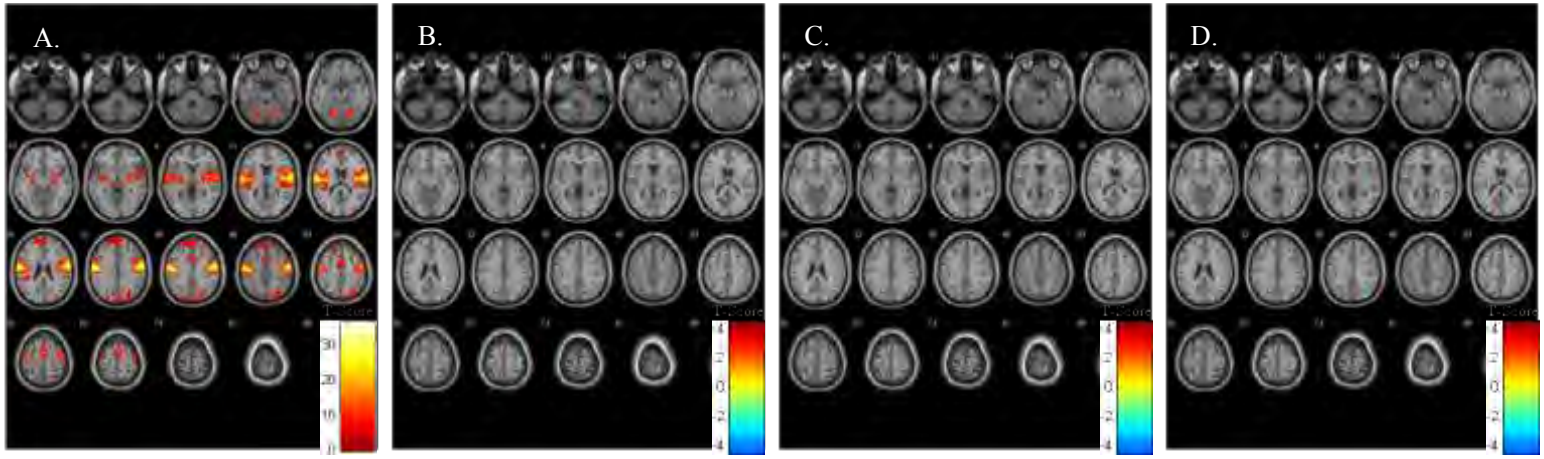
SI Table 1.3. Significantly over- or under-connected brain areas in the anterior cingulate cortex-precuneus network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

ACC-Precuneus Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-12	39	21	30	Medial Frontal Gyrus	9	9
						Anterior Cingulate	32	8
		51	-48	24	38	Inferior Parietal Lobule	40	12
						Supramarginal Gyrus	22	2
		-45	-45	24	36	Inferior Parietal Lobule	40	17
		48	15	30	28	Middle Frontal Gyrus	9	6
						Inferior Frontal Gyrus	46	2
		-39	-30	69	30	Postcentral Gyrus	3	7
						Precentral Gyrus	4 & 6	3
		-36	-51	69	27	Postcentral Gyrus	7 & 5	6
	Under-Connections	24	-48	-6	32	Parahippocampal Gyrus	19	4
		21	6	30	31	Cingulate Gyrus	24	2
		-27	-36	45	41	Inferior Parietal Lobule	40	5
		57	-54	42	27	Inferior Parietal Lobule	40	14
		-51	-57	48	31	Inferior Parietal Lobule	40	22
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	48	-48	18	27	Inferior Parietal Lobule	40	3
						Supramarginal Gyrus	22	2
						Insula	13	2
		-60	-36	33	30	Inferior Parietal Lobule	40	19
		48	15	30	26	Middle Frontal Gyrus	9, 46	6
						Inferior Frontal Gyrus	46	2
		-6	-78	42	26	Precuneus	7	13
		-39	-27	63	62	Precentral Gyrus	4, 6	9
						Postcentral Gyrus	3	12
	Under-Connections	-36	-69	0	27	Middle Occipital Gyrus	19, 37	6
		-27	-36	45	51	Cingulate Gyrus	31	2
	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-12	39	-9	27	Medial Frontal Gyrus	10	6
						Anterior Cingulate	32	2
		18	-39	72	38	Postcentral Gyrus	2, 3, 5, 40	12
						Precentral Gyrus	4	3
	Under-Connections	-27	-27	-24	25	Cerebellum -Culmen		9
						Parahippocampal Gyrus	20, 35, 36	7
		18	-69	18	29	Cuneus	18	4

						Precuneus	31	3
		-45	-27	36	50	Postcentral Gyrus	2	10
						Inferior Parietal Lobule	40	12
						Precentral Gyrus	3	1
		21	-69	39	29	Precuneus	7	1
		0	-42	39	26	Cingulate Gyrus	31	8
						Precuneus	7	2

SI Figure 1.4. The motor network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the motor network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the motor network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

Motor Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC

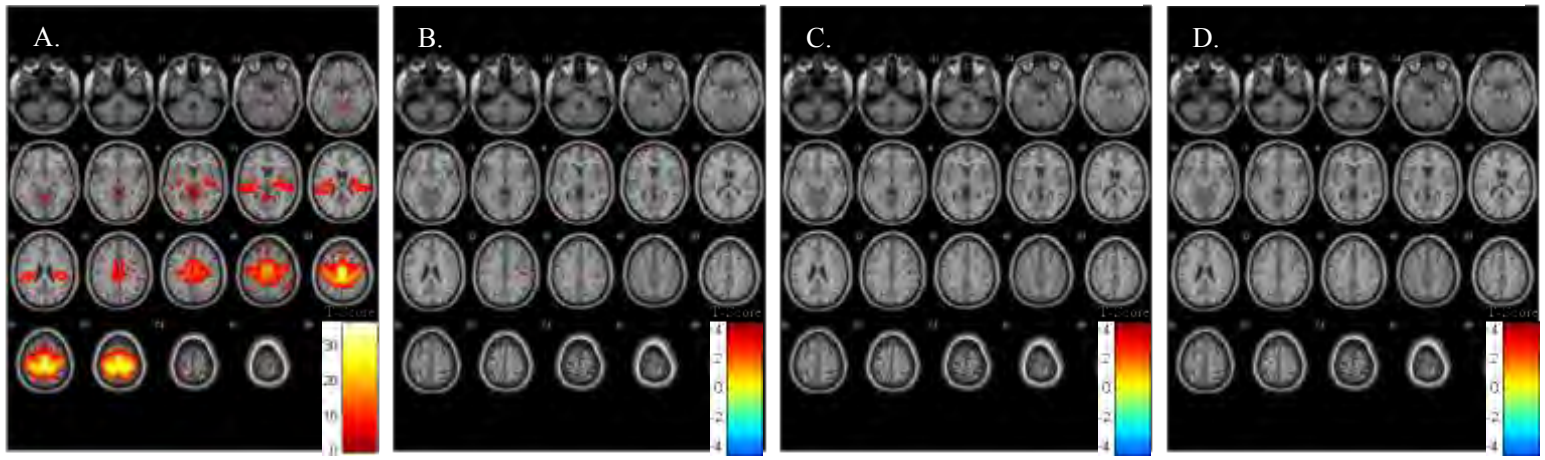


SI Table 1.4. Significantly over- or under-connected brain areas in the motor network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Motor Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	3	-39	-33	28	R Cerebellum		15
						L Cerebellum		2
						Pons		11
		-18	-63	18	84	Precuneus	7, 31	15
						Cuneus	18	11
	Under-Connections	-30	-54	-27	77	L Cerebellum		72
						Parahippocampal Gyrus	36	1
						Fusiform Gyrus	30	2
		48	-18	60	54	Precentral Gyrus	4, 6	31
						Middle Frontal Gyrus	6	1
						Postcentral Gyrus	1, 3, 40	7
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-9	-72	18	35	Precuneus	31	5 of 13
						Cuneus	18	5 of 7
	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-6	-78	21	78	Precuneus	31	9
						Cuneus	18	7
						Posterior Cingulate	23	1
		-39	-15	51	25	Precentral Gyrus	4	15
	Under-Connections	36	-12	0	33	Insula	13	6
						Claustrum		7
		18	-21	42	40	Cingulate Gyrus	24, 31	8
		42	-72	39	40	Precuneus, Angular Gyrus	19, 39	25
		39	-21	66	52	Precentral Gyrus	4, 6	12
						Postcentral Gyrus	1, 3	10

SI Figure 1.5. The supplementary motor network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the supplementary motor network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the supplementary motor network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

Supplementary Motor | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC

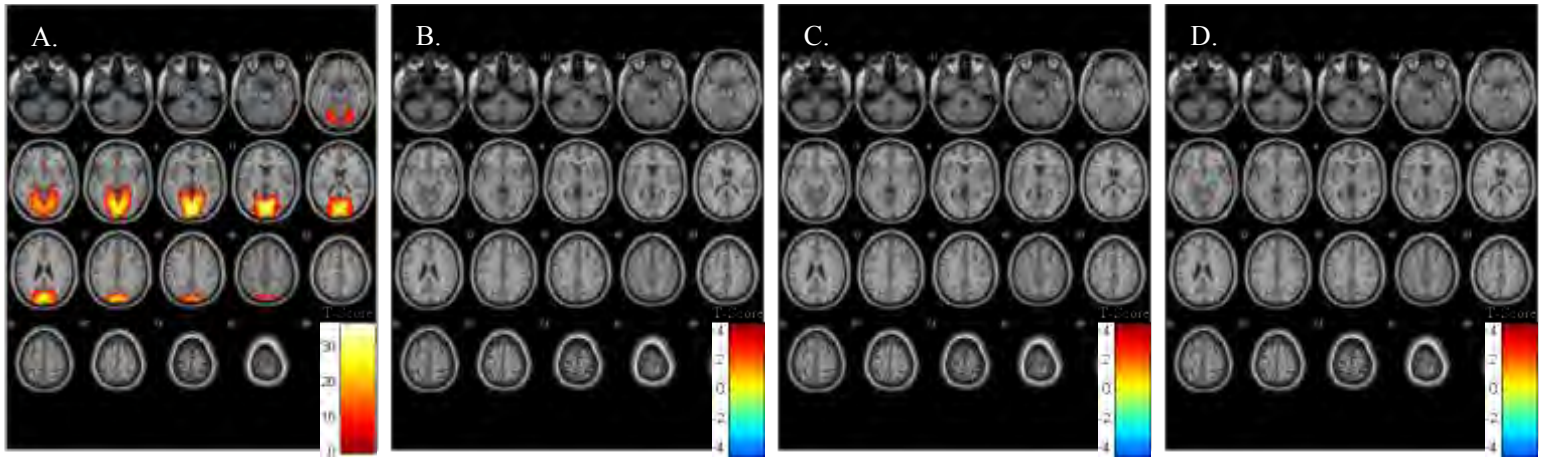


SI Table 1.5. Significantly over- or under-connected brain areas in the supplementary motor network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Supple-mentary Motor Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-21	63	9	34	Superior Frontal Gyrus	10	22
						Middle Frontal Gyrus	10	11
		45	-18	33	102	Postcentral Gyrus	2	6 of 44
						Precentral Gyrus	6	4 of 34
		-42	12	24	29	Inferior Frontal Gyrus	9	3 of 5
	Under-Connections	18	21	12	80	R Putamen		19
						Caudate Head		1
		-21	6	-3	30	L Putamen		27
		-60	-48	9	44	Superior Temporal Gyrus	21, 22	24
						Middle Temporal Gyrus	21,22	20
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-21	63	9	62	Middle Frontal Gyrus	10	43
						Superior Frontal Gyrus	10	15
						Medial Frontal Gyrus	10	1
		45	-18	33	78	Postcentral Gyrus	2,3	3
						Precentral Gyrus	6	5
	Under-Connections	30	-54	-33	34	Cerebellum - Culmen		19
						Cerebellum - Tonsil		3
						Cerebellum - Tuber		1
						Cerebellum - Pyramis		1
		3	54	-3	34	Medial Frontal Gyrus	10	9
		-60	-57	0	60	Middle Temp. Gyrus	21,22,37,39	47
						Superior Temp. Gyrus	21,22,37,39	13
		18	21	12	52	R Putamen		1
	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	51	-3	18	51	Inferior Frontal Gyrus	6, 44	2
		0	27	39	29	Anterior Cingulate	32	11
						Medial Frontal Gyrus	6,9	6
	Under-Connections	-24	6	-9	45	L Putamen		28
		69	-24	33	49	Precentral Gyrus	4, 6	21
						Postcentral Gyrus	1,2,3	5
						Inferior Parietal Lobule	40	2
		-24	-33	69	76	Postcentral Gyrus	2,3,5,40	16
						Precentral Gyrus	4	11

SI Figure 1.6. The visual network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the visual network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the visual network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

Visual Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC



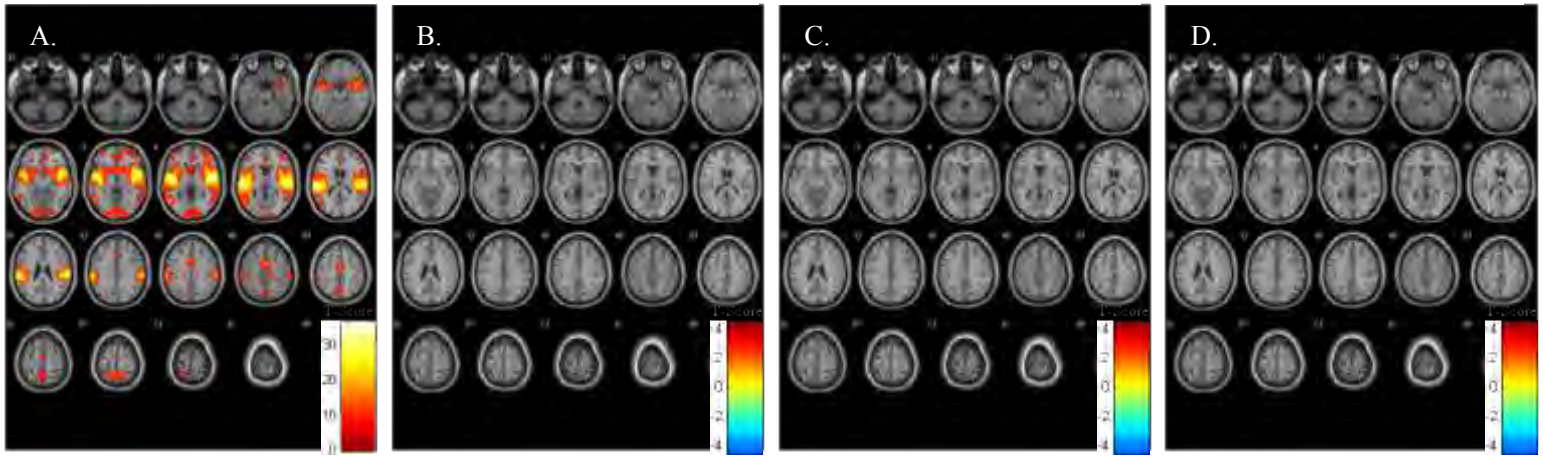
SI Table 1.6. Significantly over- or under-connected brain areas in the visual network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Visual Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	27	3	6	77	R Putamen		28
						Insula	13	1
						Claustrum		4
		21	-81	0	74	Lingual Gyrus	17, 18	45
						Middle Occipital Gyrus	17, 18	10
						Cuneus	17, 18	4
		-42	-6	15	35	Insula	13	5
		27	18	36	99	Middle Frontal Gyrus	6,8,9	29
						Superior Frontal Gyrus	6,8,9	9
						Medial Frontal Gyrus	6,8,9	3
		-45	-57	51	43	Inferior Parietal Lobule	40	17
						Superior Parietal Lobule	7	3
	Under-Connections	3	-75	-12	28	R Cerebellum		24
						Lingual Gyrus		3
		6	-57	-12	29	Cerebellum - Culmen		28
						Cerebellum - Declive		1
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	27	-48	-21	33	R Cerebellum		17
						Fusiform Gyrus	19	2
		18	-63	-9	46	Lingual Gyrus	17,18,19	27
						Cuneus	17,18,19	17
		-39	-3	15	48	Insula	13	7
						Superior Temporal Gyrus	22	1
						Precentral Gyrus	44	1
						Claustrum		1
		36	0	12	26	Insula	13	3
						R Putamen		6
						Claustrum		4
		27	12	45	118	Middle Frontal Gyrus	6,8,9	35
						Superior Frontal Gyrus	6,8,9	7
						Medial Frontal Gyrus	6,8,9	1
		-45	-57	51	33	Inferior Parietal Lobule	40	16
	Under-Connections	-9	-78	-15	30	Lingual Gyrus	18, 19	13
						Fusiform Gyrus	18, 19	2
						L Cerebellum		1
	Inattentive ADHD versus TDC							

		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	0	-45	6	33	Posterior Cingulate	29	6 of 12
						L Cerebellum		1
		3	39	0	32	Anterior Cingulate	32	11 of 26
						Medial Frontal Gyrus	10	1
		51	3	33	25	Inferior Frontal Gyrus	9	21
						Middle Frontal Gyrus	9	4
	Under-Connections	-6	-75	-15	25	L Cerebellum		12
						R Cerebellum		8
		6	-54	-9	47	R Cerebellum		38
						L Cerebellum		9
		30	-87	36	31	Precuneus	19	13
						Cuneus	19	5
		15	-42	69	34	Postcentral Gyrus	3,5	5 of 30

SI Figure 1.7. The auditory network was 1 of 12 RSNs resolved through 20-component Independent Component Analysis for the Any ADHD versus TDC analysis. **A.** Resulting component of ICA analysis identified as the auditory network. Colorbar represents t-scores from a 1-sample t-test. Areas of the brain that had significant differences in connectivity in the auditory network between children with ADHD and TDCs are shown for **B**. any ADHD diagnosis versus TDC connectivity **C**. combination subtype ADHD diagnosis versus TDC connectivity **D**. inattentive subtype ADHD diagnosis versus TDC connectivity. Colorbar represents t-scores from 2-sample t-tests.

Auditory Network | Any ADHD vs TDC | Combination ADHD vs TDC | Inattentive ADHD vs TDC



SI Table 1.7. Significantly over- or under-connected brain areas in the auditory network. This table lists areas of the brain that were significantly different between ADHD and TDC groups in the Any ADHD versus TDC analysis, the Combination ADHD versus TDC analysis, and the Inattentive ADHD versus TDC analysis. The peak MNI coordinates of each cluster are given in columns 3-5, and the size of the cluster is given in column 6.

Auditory Network	**Any ADHD versus TDC**							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	-6	-93	-3	35	Cuneus	17	13
	Under-Connections	-42	39	12	28	Inferior Frontal Gyrus	46	4
						Middle Frontal Gyrus	10	1
		-33	-36	9	25	Superior Temp. Gyrus	41	1
		42	-30	12	29	Transverse Temp. Gyrus	40, 41	20
						Superior Temp. Gyrus	40, 41	5
		57	-51	15	33	Superior Temp. Gyrus	22, 40	29
						Supramarginal Gyrus	22, 40	4
		45	-9	51	27	Precentral Gyrus	4, 6	9
						Postcentral Gyrus	3	1
	Combination ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Under-Connections	-48	-36	12	26	Superior Temp. Gyrus	41	21
						Transverse Temp. Gyrus	41	5
		57	-51	15	48	Superior Temp. Gyrus	22, 42	12
		54	-9	57	27	Precentral Gyrus	4, 6	6
						Postcentral Gyrus	3	2
	Inattentive ADHD versus TDC							
		X	Y	Z	Cluster Size	Anatomical Area	Brodmann Area	Voxels in Area
	Over-Connections	6	-66	-30	33	R Cerebellum		18
						L Cerebellum		15
		-36	-90	18	50	Middle Occipital Gyrus	18, 19	40
						Superior Occip. Gyrus	18, 19	3
		33	21	12	27	Insula	13	7
						Inferior Frontal Gyrus	45	1
		-39	6	9	34	Insula	13	12
						Precentral Gyrus	44, 45	5
						Inferior Frontal Gyrus	44, 45	4
		45	-75	33	78	Precuneus	19, 39	29
						Angular Gyrus	19, 39	23
						Middle Temporal Gyrus	19, 39	12
						Superior Occip. Gyrus	19, 39	5
		-33	-81	39	32	Precuneus	19	17
						Cuneus	19	9
		66	-18	36	26	Postcentral Gyrus	1,3	7
						Precentral Gyrus	4	6

CHAPTER II

Introduction

Attention Deficit/Hyperactivity Disorder (ADHD) is currently the most prevalent pediatric psychiatric disorder, estimated to affect up to 11% of American children and 5.3% of children and adolescents world-wide (Polanczyk et al., 2007; Visser et al., 2014). It is diagnosed based on the presence of age-inappropriate symptoms of inattention, hyperactivity, or a combination of both (American Psychiatric Association, 2013). These symptoms may persist into young adulthood; hence children with ADHD are likely to suffer worse academic, social, and psychological outcomes than their typically developing peers (Biederman et al., 2006; Faraone et al., 2006; Huh et al., 2011; Mannuzza and R. G. Klein, 2000). Although a large body of research has focused on finding the underlying basis of ADHD (Cortese, 2012), to date the precise causative factor(s) of ADHD remain elusive.

Neuroimaging techniques hold the potential to elucidate the neural underpinnings of many psychiatric disorders and may represent a possible future diagnostic aid for those disorders, including ADHD (Freilich and W. D. Gaillard, 2010; Sava and D. A. Yurgelun-Todd, 2008). Specifically, in recent years there has been an interest in utilizing resting-state functional connectivity (rs-FC) to better understand the functional deficits that underlie ADHD. rs-FC measures are based on blood oxygenation level-dependent (BOLD) signal, acquired during a resting state functional magnetic resonance imaging (rs-fMRI) scan: different regions of the brain with synchronous BOLD signal are thought to be functionally connected, reflecting the Hebbian tenant that “cells that fire together, wire together” (Hebb, 1949). Sets of different brain areas that consistently co-activate

during ‘rest’ are known as resting state networks (RSNs) (Allen et al., 2011; Damoiseaux et al., 2006; De Luca et al., 2006; van den Heuvel and H. E. Hulshoff Pol, 2010). One example of an RSN is the Default Mode Network (DMN), which is comprised of the medial prefrontal cortex/anterior cingulate cortex, the posterior cingulate cortex/precuneus, and the inferior parietal lobules (Allen et al., 2011; Biswal et al., 1995; Damoiseaux et al., 2006; Raichle et al., 2001; van den Heuvel and H. E. Hulshoff Pol, 2010). The DMN has been consistently found across multiple studies done with participants at ‘rest’, and is speculated not to be related to any externally-driven mental processes but rather to be involved in daydreaming and introspection. Previous studies using rs-FC measures have shown differences between ADHD and healthy participants, most notably within DMN (Castellanos et al., 2008; Liston et al., 2011; Qiu et al., 2011; Uddin et al., 2008; Wang et al., 2009).

RSNs can be identified several different ways; one of these is through the data-driven approach of independent component analysis (ICA) (Beckmann et al., 2005; Calhoun et al., 2001; Hyvärinen and Erkki Oja, 2000; McKeown and Terrence J. Sejnowski, 1998). In performing ICA, resting state BOLD signal data is parceled into multiple spatially independent components, each of which can be described by both a set of brain regions and a waveform characterizing the BOLD signal fluctuation in these areas (Calhoun et al., 2009). In other words, if **x** is acquired rs-FC data (BOLD signal at each time point within each brain area, or voxel, signal was acquired from), **c** represents spatially independent, underlying signals, namely components (i.e., spatial maps of temporally coherent areas), and **A** is the mixing matrix, then **x** can be represented as:

$$x = Ac \quad (2.1)$$

Since all brain regions within a given component spatial map share a similar BOLD time course, they are thought to be connected; thus components can be seen as representing putative RSNs. This methodology can be applied to sets of data (Group ICA) such that independent components within resting state data acquired from a group of participants can be identified (Calhoun et al., 2001; Calhoun et al., 2009). The result of Group ICA is a set of spatial maps that describe each component's location for the entire group; subject-level component maps can then be found through back-reconstruction, where the information from the group map and components is used to predict individual spatial maps and time courses.

Univariate testing has traditionally been used to compare resting state measures (including those that result from ICA analysis) across two groups of subjects, for example, ADHD versus typically developing controls (TDC). Specifically, voxel-wise comparisons are made throughout the brain to understand how measures of connectivity at each point in the brain differ between the ADHD versus TDC samples. While useful in understanding the neurobiological underpinnings of ADHD, these methods do not take into account patterns of activity across voxels over time, and cannot make subject-level predictions. More recently, machine learning (Mitchell, 1997) techniques have been applied to rs-FC data in the pursuit of finding a subject-level diagnostic aid for various psychiatric disorders, including ADHD (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012). One implementation of machine learning is through Support Vector Machine (SVM) classification: an SVM algorithm is 'trained'

using labeled examples from each class (e.g., ADHD or TDC) to determine a hyperplane that linearly separates the two classes while maintaining the maximum distance from the closest, i.e. most similar, training examples (Cortes and Vladimir Vapnik, 1995). In other words, the SVM is trained to classify sets of observations with known outcomes, with the ultimate goal of predicting unknown outcomes using a separate set of input observations.

The goal of the present study is to determine the utility of RSNs, determined through ICA, in accurately discriminating between ADHD and TDC. To investigate this, we applied a SVM algorithm to the landmark ADHD-200 dataset, a publicly-available database of almost 1000 rs-fMRI and structural data acquisitions from both participants with and without ADHD. We hypothesized that the RSNs that would be most useful in discriminating ADHD from TDC participants would be the DMN, as well as networks that included areas of the frontal lobe (such as executive control and salience networks).

Materials and Methods

Data Sets: The data used for this analysis consists of a subset of the 973 rs-fMRI data acquisitions that comprise the ADHD-200 dataset. These 973 scans were pooled from 8 different sites, namely: Peking University (Peking), Bradley Hospital/Brown University (Brown), Kennedy Krieger Institute (KKI), NeuroIMAGE, New York University Child Study Center (NYU), Oregon Health and Science University (OHSU), University of Pittsburgh (Pittsburgh), and Washington University in St. Louis (WashU) (http://fcon_1000.projects.nitrc.org/indi/adhd200/; (Milham et al., 2012). The ADHD-200 dataset was originally split into a globally-released training set (776 scans) and a withheld test set (197 scans); here only data from the original training set is used, which

contained information about participants' age, gender, handedness, and IQ scores in addition to neuroimaging data. Of the 776 rs-fMRI scans available in the original training set, not all were deemed suitable for use in this analysis. All rs-fMRI scans that did not pass the quality control measure provided on the ADHD-200 website were excluded. An additional 8 scans were excluded as they were determined to have low signal quality, as screened by eye. We also excluded any scans that failed the pre-processing procedure. After these exclusions, only 12 left-handed participants remained; they were not included in the analysis. Overall, out of the original 776 participants' data, a total of 139 participants were excluded, leaving 637 rs-fMRI data acquisitions for the analysis.

In preparation for the 10-fold cross-validation procedure described below, these 637 scans were randomly divided into 10 groups of 63 scans each, leaving 7 extra scans that were not further included. Thus a total of 630 scans were used in this analysis (220 from participants diagnosed with ADHD and 410 from TDC). The number of participants from each site along with the acquisition parameters used at those sites is listed in Table 2.1. Briefly, sites used Siemens and Philips systems. TR values ranged from 1500-2500 msec and TE values from 15-40 msec, and most sites used a flip angle of 90°. The field of view ranged from 200-256 mm with the number of slices acquired ranging from 29 to 47; slice thickness varied from 3.0 to 4.5 mm. Most sites instructed participants to keep their eyes open and to fixate on a center cross during resting state acquisition.

Table 2.1. Image acquisition parameters. This table contains information about how each rs-fMRI scan was performed at each of the ADHD-200 sites. How many participants from each site were included in this analysis can be seen in column 2.

Table 2.1. Image Acquisition Parameters

Site	Number of Participants	Magnet Field Strength (Tesla)	MRI System	TR (msec)	TE (msec)	Flip Angle	Field of View (mm)	Number of Slices	Slice Thickness (mm)	Eyes Open or Closed
Peking	84 42	3.0 T	Siemens Magnetom Trio, A Tim	2000	30	90°	200	33 30	3.5 4.5	Open or Closed
KKI	74	3.0 T	Philips	2500	30	75°	256	47	3.0	Open
NeuroIMAGE	4	1.5 T	Siemens Magnetom Avanto	1960	40	80°	224	35	3.0	Closed
NYU	215	3.0 T	Siemens Magnetom Allegra	2000	15	90°	240	33	4.0	Closed
OHSU	78	3.0 T	Siemens Magnetom Trio, A Tim	2500	30	90°	240	36	3.8	Open
Pittsburgh	79	3.0 T	Siemens Magnetom Trio, A Tim	1500	29	70°	200	29	4.0	Open
WashU	54	3.0 T	Siemens Magnetom Trio, A Tim	2500	27	90°	256	32	4.0	Open

Peking: Peking University; KKI: Kennedy Krieger Institute; MRI: magnetic resonance imaging; NYU: New York University Child Study Center; OHSU: Oregon Health and Science University; Pittsburgh: University of Pittsburgh; TE: echo time; TR: repetition time; WashU: Washington University in St. Louis

Data Preprocessing and Preparation: All 630 rs-fMRI scans were preprocessed using the Data Processing Assistant for Resting-State fMRI (DPARSFA, http://www.restfmri.net) (Song et al., 2011). DPARSFA is plug-in software that works with SPM8 (Statistical Parameter Mapping– Welcome Department of Imaging Neuroscience, London, UK; (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) and the Resting-State fMRI Data Analysis toolkit (REST). As a first step of preprocessing, first 10 time points were removed from all scans. Then images were corrected for slice acquisition time differences, realigned for motion correction, normalized to the MNI EPI template (voxel size 3X3X3), spatially smoothened with a 4 mm FWHM kernel, detrended and temporally band-pass filtered to 0.01-0.08 Hz (Deligiannidis et al., 2013).

The resulting pre-processed 630 scans were randomly divided into 10 groups of 63, following a 10-fold cross-validation (CV) strategy (Ambroise and Geoffrey J. McLachlan, 2002; Kohavi, 1995). For 10-fold CV, 9 folds of the data are used to select features (i.e., rs-fMRI data from different brain networks) and then train a classifier (here, a SVM) to separate examples (i.e., participants) into different classes (i.e., ADHD or TDC). After training, the classifier is tested on the remaining, withheld 1 fold of the data so that an unbiased estimate of the classifier's accuracy can be obtained. The 9 folds of the data used to train the classifier are collectively referred to as the training set and the withheld 1 fold is referred to as the test set. After this first iteration of training and testing, this process is repeated 9 more times, such that each fold serves as the test set for 1 iteration, and so that 10 un-biased estimates of the classifier's accuracy are attained. Therefore, the following steps (ICA, Feature Selection, and Classifier Implementation)

were repeated a total of 10 times (10 iterations); in the following sections the method is described for the first iteration only.

Independent Component Analysis (ICA): Group ICA was performed using the Group ICA of fMRI (GIFT) toolbox (http://mialab.mrn.org/software/). Group ICA was done separately on the test set and training set using 20 components for each set. The number of components used for ICA implementation must be specified by the investigator; here ICA was repeated with 15, 20, 25, and 30 components to determine which results produced the best match to RSNs previously found in children ages 9-15 years old (Thomason et al., 2011). Spatial correlations between the results of the 15-, 20-, 25-, or 30- component analyses and templates of these previously determined networks were used to conclude that 20 components yielded the best match. This also allowed us to identify which components represented RSNs, and which did not.

ICA was performed 20 times in total (for the training and the test sets separately within each of 10 iterations) generally following the method of Calhoun et al (Calhoun et al., 2001; Calhoun et al., 2009). For each ICA, standard subject-specific Principle Component Analysis (PCA) was done to reduce the dimensionality of each individual's data (Calhoun et al., 2009). The individual PCA results were then concatenated across participants (for a total of 567 participants in each training set and 63 participants in each test set) and a 2nd, group-level PCA was performed. These within-participant and across-participants data reduction steps are done primarily to ease the computational burden of working with high-dimensional rs-fMRI data. The resulting compressed data was then used as input for the algorithm that simultaneously determined the mixing matrix (**A** in

equation (2.1) above) and the independent components (**c** in equation (2.1) above). Here, the INFOmax algorithm was employed to obtain maximally spatially independent group-level components (Bell and Terrence J. Sejnowski, 1995). The dual-regression method (available as part of the GIFT toolbox) was then used to predict each individual's component time courses and maps (Filippini et al., 2009; Zuo et al., 2010).

Component Matching: As group ICA was performed separately on the training and test sets for each of 10 iterations, and the order of the resulting components is not consistent across ICA implementations, it was necessary to "match" the components of the test group to the components of the training group within an iteration, as well as across iterations. Here, the resulting group-level components from each iteration's test set were matched to the corresponding group-level components of that iteration's training set using spatial correlation measures. Group-level components from the training sets were also matched between iterations by spatial correlation to allow for comparison across all 10 iterations.

Feature Selection: After group ICA implementation and component matching, each participant's rs-fMRI data was represented as a set of 20 components within each of the 10 iterations; each component could be visualized as a spatial map detailing component coherence at each of about 54,000 brain areas. While all of these areas could be used to train and then to test a classifier, the relatively high dimensionality of the component data (about 54,000 data points per component) compared to the number of participants (n = 630) meant that this would most likely lead to over-fitting. That is, instead of learning the broad patterns within components that could differentiate between participants with and

without ADHD even outside the training set, the classifier would fit to differences specific to the given training set and would be unable to generalize to outside examples. It was therefore necessary to reduce the number of data points, or features, which would be used to train the classifier. To this end, a feature selection algorithm was implemented to determine what subset of data points (brain areas) should be used to train the classifier. Here, feature selection for each component was done using the Fisher Score algorithm available through the Arizona State University Feature Selection website (http://featureselection.asu.edu/software.php):

$$\text{Fisher Score} = \frac{n_{TDC}*(\mu_{TDC}-\mu)^2 + n_{ADHD}*(\mu_{ADHD}-\mu)^2}{n_{TDC}*\sigma_{TDC}^2 + n_{ADHD}*\sigma_{ADHD}^2} \quad (2.2)$$

where n_{class}, μ_{class}, and σ_{class}^2 represent the number of participants, the mean of the feature in question's values, and the variance of that feature's values for the given class, i.e. TDC or ADHD. μ is the mean of all values for the feature in question, i.e. including both TDC and ADHD participants. Note that feature selection was done using **only** data from the training set, and not the test set, so that features used to train the classifier would not bias its ultimate application to the test set (Ambroise and Geoffrey J. McLachlan, 2002). The feature selection algorithm provided a ranking of how relevant each feature was in determining class. In order to decide how many of the highest-ranked features would be needed to achieve high classification accuracy, the classifier was trained multiple times using an increasing number of features. The top 10 features were used initially; the number of features was then increased in increments of 10 up to a maximum

of 900. Classifier performance on the test set using each number of features was used to decide what feature number was best; this was done separately for each of the 20 components found through ICA and throughout each of 10 iterations.

Classifier Implementation: LIBSVM (version 3.16, http://www.csie.ntu.edu.tw/~cjlin/libsvm/) was used in conjunction with Matlab to implement the SVMs. Both a linear and a radial basis function (RBF) kernel were employed for 3 out of the 10 iterations; the RBF kernel consistently yielded higher classification accuracy in the test set. Therefore, the analysis was performed using the RBF kernel, which is useful when the distinction between two classes is nonlinear, as it maps input feature vectors nonlinearly into higher dimensional space. Given feature vectors x_1 and x_2, the RBF kernel is defined as:

$$K(x_1,x_2) = \exp(-\gamma||x_1 - x_2||^2) \tag{2.3}$$

where $\gamma > 0$ and is inversely proportional to the width of the kernel. Both the γ -parameter and C-parameter (a parameter used during SVM implementation which allows for less than perfect separation of the two classes when training the SVM) were optimized through automatic grid searching and standard fivefold cross validation available in LIBSVM. Once the optimal γ- and C-parameters, along with the best number of features to use (described above), were determined, the SVM was trained with 9 folds of the data (training set) and tested on the remaining 1 fold (test set).

The entire process (ICA, Feature Selection, and Classifier Implementation) was then repeated 9 more times using a different fold for the test set each time, for each of 20 components, for a total of 10 training and testing iterations.

After all 10 iterations were completed the average accuracy, sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for each of the 20 components across all iterations were calculated for the SVM's performance on the test group. Accuracy was calculated as simply the number of correct diagnoses made by the SVM (either ADHD or TDC) over the total number of subjects in the test group, multiplied by 100. The sensitivity was calculated as a ratio of how many participants the SVM found that truly had ADHD (true positives, TP) out of the total number of participants who truly had ADHD (the ones the SVM found - TPs - and the ones the SVM didn't find - false negatives - FNs) in the test group, multiplied by 100, or:

$$\text{Sensitivity} = \frac{TP}{TP+FN} \times 100 \tag{2.4}$$

The specificity was calculated as a ratio of how many participants the SVM found that truly did NOT have ADHD (true negatives - TNs) out of the total number of participants who truly did NOT have ADHD (the ones the SVM found - TNs - and the ones the SVM didn't find, i.e. false positives - FPs) in the test group, multiplied by 100, or:

$$\text{Specificity} = \frac{TN}{TN+FP} \times 100 \tag{2.5}$$

The PPV was calculated as a ratio of how many participants the SVM found that truly had ADHD (true positives, TPs) out of the total number of participants the SVM found to have ADHD (the ones that truly had ADHD - TPs - and the ones that did NOT truly have ADHD - FPs) in the test group, multiplied by 100, or:

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \times 100 \tag{2.6}$$

The NPV was calculated as a ratio of how many participants the SVM found that truly did NOT have ADHD (true negatives, TNs) out of the total number of participants the SVM found to NOT have ADHD (the ones that truly did NOT have ADHD - TNs - and the ones that actually had ADHD – false negatives, FNs) in the test group, multiplied by 100, or:

$$\mathrm{NPV} = \frac{\mathrm{TN}}{\mathrm{TN+FN}} \times 100 \tag{2.7}$$

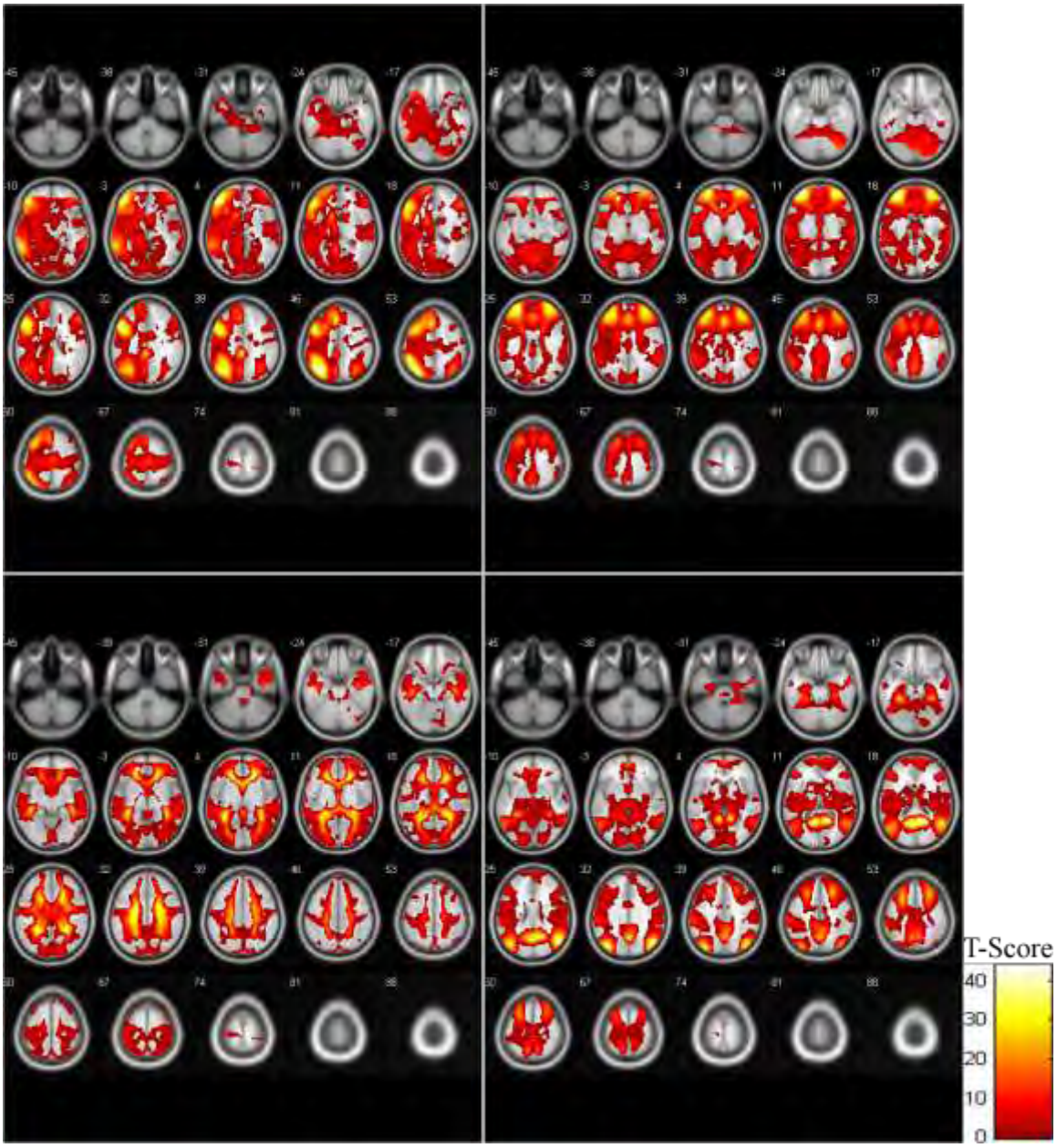
Results

Participants: The average age of participants who were reported to have ADHD and TDC participants were 11 ± 2.7 years and 12 ± 3.2 years, respectively. Participant IQ also differed across the two groups: ADHD patients and TDC had average IQ of 106 ± 13.8 and 113 ± 13.2, respectively. There were a significantly higher percentage of males among ADHD patients: 76% males in comparison to 50% males in the TDC group.

However, age, IQ, and percent male subjects were similar across all 10 subdivisions of the data used for the 10-fold cross validation. No left-handed participants were included in this analysis.

Independent Component Analysis (ICA): Group ICA was performed a total of 20 times (for the training and the test sets in each of 10 iterations). For each ICA, 20 components were used to resolve the RSNs, but not all components ultimately matched to pediatric resting state network templates (Thomason et al., 2011). Figure 2.1 shows four examples out of the 20 components generated for the training set in the first iteration. These examples were found to either match the template of an RSN or to represent noise; namely the left executive network (component 1, Figure 2.1A), the salience network (component 5, Figure 2.1B), a noise component (component 7, Figure 2.1C) and the DMN (component 13, Figure 2.1D).The left executive network, salience network, and DMN are shown because we predicted they would yield the highest classification accuracy; the noise component is shown because it was ultimately the most sensitive component for the discrimination of participants with ADHD. Other components that were found to match the previously-published RSN templates (Thomason et al., 2011) were the right executive network (component 2), the supplementary motor network (component 3), the parietal association network (component 6), the visual network (component 10), a posterior default mode network (component 14), the motor network (component 15),an inferior frontal gyrus/middle temporal network (component 18), an anterior cingulate/precuneus network (component 19), and the auditory network (component 20). The remaining 8 components did not match the RSN templates used.

Figure 2.1. Four examples of the group components found through ICA using the training set from iteration 1. Colorbar represents T-scores. The component that best matched **A**. the left executive network, **B**. the salience network, and **C**. a noise component not matched to the RSN templates. **D**. The component that best matched the DMN.


T-Score
40
30
20
10
0

Component Matching: Since group ICA was done on the training and test sets separately, it was necessary to match components across the two groups. Similarly, component matching was required across iterations in order to determine the average accuracy of the classifier for each RSN. For matching across iterations, only the training components from the first iteration were used as templates. The training components within each iteration were then used as templates to match test components. In both these cases, component matching was done using spatial correlation, and the pair that had the highest correlation was considered to be a match. Table 2.2 reports the average for all 10 iterations' spatial correlation value for matching test-to-training set components and the average training-to-training set correlation for each component across iteration. The highest test-to-training set correlation was 0.926 for the visual network (component 10), and the highest training-to-training set correlation across iterations was 0.995 for the posterior default mode network (component 14).

Feature Selection: Feature number selection for four sample components can be seen in Figure 2.2. To determine the optimal feature number for each component, the SVM was trained 90 times: the first with 10 features, then with 10-feature increment increases up to 900 features. This was done for each training set of the 10 iterations and the corresponding test set accuracy for each feature number increment across 10 iterations was then determined. The number of features that yielded the highest average test set accuracy over all 10 iterations was then chosen as the best fit for number of features.

Classifier Findings: The accuracy, sensitivity, specificity, PPV and NPV for each

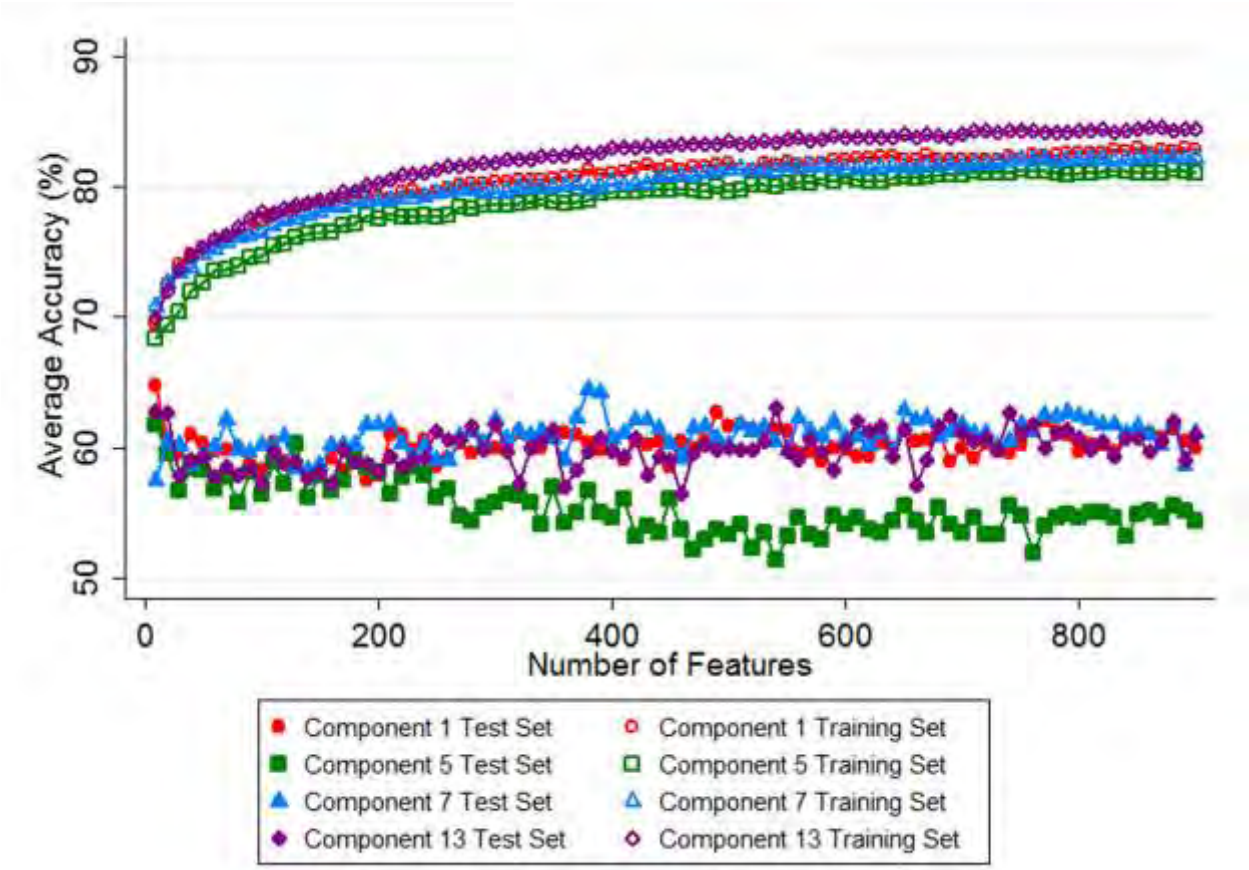
Table 2.2. Average spatial correlations for matching components within and across 10 iterations. This table shows the average correlations for comparing the training set resting state networks (components) to the test set resting state networks (components) (column 2) and correlations for comparing training set resting state networks (components) across iterations (column 3).

Table 2.2. Average Spatial Correlations for Matching Components Within and Across 10 Iterations

RSN (Component Number)	Test-to-Training Sets Correlation ± Standard Deviation (n=10)	Training-to-Training Sets Correlation ± Standard Deviation (n=9)
Left executive (1)	0.770 ± 0.082	0.930 ± 0.088
Right executive (2)	0.829 ± 0.062	0.976 ± 0.013
Supplementary motor (3)	0.795 ± 0.037	0.982 ± 0.013
(4)	0.534 ± 0.041	0.955 ± 0.022
Salience (5)	0.459 ± 0.166	0.898 ± 0.082
Parietal association (6)	0.872 ± 0.019	0.989 ± 0.006
(7)	0.465 ± 0.244	0.543 ± 0.398
(8)	0.815 ± 0.060	0.985 ± 0.009
(9)	0.402 ± 0.247	0.689 ± 0.444
Visual (10)	0.926 ± 0.026	0.983 ± 0.008
(11)	0.692 ± 0.060	0.911 ± 0.065
(12)	0.468 ± 0.251	0.530 ± 0.354
DMN (13)	0.414 ± 0.217	0.983 ± 0.007
Posterior DMN (14)	0.916 ± 0.009	0.995 ± 0.002
Motor (15)	0.872 ± 0.020	0.986 ± 0.006
(16)	0.718 ± 0.112	0.910 ± 0.067
(17)	0.813 ± 0.057	0.974 ± 0.021
IFG/middle temporal (18)	0.689 ± 0.094	0.934 ± 0.064
AC/precuneus (19)	0.760 ± 0.122	0.896 ± 0.124
Auditory (20)	0.609 ± 0.221	0.784 ± 0.275

AC: anterior cingulate; DMN: default mode network; IFG: inferior frontal gyrus; RSN: resting state network

Figure 2.2. Four examples of the average training and test sets' accuracy across all 10 iterations for each number of features (10 to 900 in increments of 10) used in the SVM. Component 1 is the left executive network, component 5 is the salience network, component 7 is noise, and component 13 is the DMN. The number of features that yielded the highest average accuracy in the test group was taken as the best number of features to use for SVM implementation.


Average Accuracy (%)
50
60
70
80
90
0
200
400
600
800
Number of Features
Component 1 Test Set
Component 5 Test Set
Component 7 Test Set
Component 13 Test Set
Component 1 Training Set
Component 5 Training Set
Component 7 Training Set
Component 13 Training Set

component averaged across all 10 iterations can be seen in Table 2.3. The IFG-middle temporal network (component 18) was the most accurate at predicting ADHD status (66.8% accuracy) while the auditory network (component 20) had the lowest accuracy at 59.5%. A noise component (component 7) had both the highest sensitivity and NPV at 39.7% and 69.6%, respectively. By contrast, the left executive network (component 1) had the lowest sensitivity at 15.4% and the auditory network (component 20) had the lowest NPV at 63.0%. Finally, the parietal association network (component 6) had both the highest specificity and PPV at 86.6% and 48.5%, respectively, while the visual network (component 10) had the lowest specificity at 65.3% and the salience network (component 5) had the lowest PPV at 20.6%.

Discussion

This study focused on the usefulness of different resting state networks (RSNs), identified using Independent Component Analysis (ICA), in classifying participants as ADHD or TDC. A large number of the publicly-available rs-fMRI acquisitions from ADHD-200 dataset were employed in this work, which allowed for the implementation of methods that could prove challenging or yield unreliable results when applied to smaller datasets. Here, a machine learning technique known as a SVM was used to classify participants as ADHD or TDC based on the integrity of their RSNs. Two of the RSNs that were expected to be most predictive of diagnostic status –the left executive network and the salience network – were found to be among the least helpful during classification. In addition, the independent components identified as representing the DMN were not particularly accurate in identifying participants in ADHD; this was

Table 2.3. Average accuracy, sensitivity, specificity, PPV, and NPV across all 10 iterations' test sets for each component. This table shows several metrics used to evaluate how well the SVM performed in classifying ADHD versus TDC participants. In all columns, a higher percentage equates to better classifier performance.

Table 2.3. **Average Accuracy, Sensitivity, Specificity, PPV, and NPV across All 10 Iterations' Test Sets for Each Component.**

RSN (Component Number)	Test Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)
Left executive (1)	64.8	15.4	86.5	31.8	65.7
Right executive (2)	63.3	26.7	77.7	40.1	66.4
Supplementary motor (3)	61.6	27.0	80.3	44.0	67.3
(4)	64.1	28.7	75.6	39.1	66.3
Salience (5)	61.7	28.6	71.8	20.6	65.5
Parietal association (6)	64.4	24.9	86.6	48.5	68.4
(7)	64.4	39.7	73.0	45.4	69.6
(8)	61.4	23.8	78.6	34.8	65.8
(9)	64.3	22.0	79.8	42.5	65.7
Visual (10)	62.9	32.2	65.3	26.6	64.4
(11)	65.2	29.5	74.2	39.2	66.3
(12)	64.8	27.7	73.8	31.4	65.5
DMN (13)	63.0	29.6	74.6	38.9	66.1
Posterior DMN (14)	61.6	30.1	74.9	40.0	66.8
Motor (15)	65.2	37.5	68.5	35.9	67.3
(16)	61.6	28.6	74.0	38.0	66.0
(17)	63.3	23.8	80.7	30.8	66.8
IFG/middle temporal (18)	66.8	32.8	76.8	43.5	68.1
AC/precuneus (19)	64.6	26.4	75.5	38.9	65.5
Auditory (20)	59.5	26.8	67.1	23.7	63.0

AC: anterior cingulate; DMN: default mode network; IFG: inferior frontal gyrus; NPV: negative predictive value; PPV: positive predictive value; RSN: resting state network

inconsistent with our hypothesis that this network would be among the most accurate. By contrast, the components identified as the IFG-middle-temporal network and the parietal association network were among the most useful RSNs for participant classification. Finally, component 7 (which was identified by eye as physiological noise) yielded the highest sensitivity in classifying ADHD participants.

The IFG-middle-temporal network, as the name implies, includes portions of the inferior frontal gyri (IFG) and middle temporal gyri. As the IFG have been implicated in ADHD using both structural (DTI) and functional (task-based and rs-fMRI) measures, it is reasonable that a network including these areas could be useful in predicting ADHD diagnosis (Liston et al., 2011). Similarly, the parietal association network is comprised of large portions of the parietal cortices. The relationship between parietal cortex function and ADHD has been less explored; however, at least one report has found aberrant parietal cortex activation in adolescents with ADHD (Tamm et al., 2006). It is therefore understandable that while neither the IFG-middle-temporal network nor the parietal association network has been implicated in ADHD to the same degree as the DMN, they are both of use in predicting the presence of ADHD.

The most unexpected result of our investigation was that component 7, which we identified as physiological noise, yielded the highest sensitivity when used to classify ADHD participants. We speculate that this component primarily reflects signal(s) originating from the cardiac cycle, as the spatial location of this component (shown in Figure 2.1C) includes areas that have previously been shown to display significant cardiac-induced signal changes (Dagli et al., 1999). Moreover, a large body of literature

exists linking heart rate variability (HRV, the variation in the amount of time between consecutive heart beats) to mental effort allocated to a given task (Jorna, 1992), and several studies have explored HRV in youth with ADHD (Börger et al., 1999; Börger and Jaap van der Meere, 2000). While increased mental effort on a task normally leads to a decrease in HRV, these studies found that ADHD children showed increased HRV compared to TDCs (Börger et al., 1999; Börger and Jaap van der Meere, 2000; Jorna, 1992). Differences in heart rate (tachycardia) and measures of HRV have also been found in children with ADHD compared to control children at rest (Buchhorn et al., 2012; Tonhajzerova et al., 2009). It is therefore possible that these reported heart rate-related differences between ADHD and TDC youth underlie the ability of component 7 (physiological noise potentially driven by the cardiac cycle) to classify ADHD participants in the present study.

Previous work with the ADHD-200 dataset has involved a range of different classification approaches, including SVMs with different kernels (linear, quadratic, cubic, and RBF), multi-kernel learning, gradient boosting, adaptive boosting (AdaBoost), random forests, and C4.5 decision trees (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012). These different implementations of machine learning were used in combination with diverse feature selection methods, ultimately yielding classification accuracies ranging from 55% to 78% (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012). In the present work, classification accuracy varied from 59.5% to 66.8%, depending on the component used. While our results are therefore comparable to previous findings, our

method did not result in substantial improvements to current classification accuracy. With respect to methodology, our approach was similar to that of Brown et al (Brown et al., 2012). However, in our approach we examined each RSN/component individually, whereas they examined only the DMN individually, and then used a combination of information from all 20 components as another measure.

While we found that certain RSNs yielded better classification metrics (accuracy, sensitivity, specificity, PPV, and NPV) than others, even the highest sensitivity (39.7%, component 7/physiological noise) and specificity (86.6%, parietal association network) we achieved are too low for this method to be advanced as a stand-alone diagnostic test. This is perhaps unsurprising, as the validity of the participant classification was based on ADHD status determined using clinical evaluation and/or behavioral measures, which can be unintentionally subjective (Parens and J. Johnston, 2009). We also did not consider the effect of gender in our analysis, which may have affected the SVM's ability to accurately distinguish ADHD children from TDC. Several studies have shown that boys diagnosed with ADHD are more likely to exhibit hyperactive/impulsive symptoms than girls, while girls are more likely to experience inattentive symptoms (Biederman et al., 2002; Gershon and Jonathan Gershon, 2002; Hinshaw et al., 2006; Quinn, 2008). This gender-driven dissimilarity in ADHD expression may imply deficits in different RSNs between boys and girls, such that when they are considered as a single group (as done in this analysis) the difference between ADHD and TDC groups is attenuated.

The fundamental challenge in accurately and objectively diagnosing ADHD stems from the inherent heterogeneity of the disorder: the label "ADHD" refers to a set of

symptoms that reflect multiple independent underlying deficits that may vary from patient to patient (Fair et al., 2012). While neuroimaging markers clearly hold some value in aiding ADHD diagnoses, our findings suggest that resting state fMRI data alone is not sufficient to distinguish "ADHD" (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012). Future research studies may focus on investigating the predictive power of RSNs in detecting subgroups of ADHD, following the Research Domain Criteria (RDoC) approach laid out by the NIMH (Fair et al., 2012; Insel et al., 2010; Oldehinkel et al., 2013). In this regard, the RDoC construct of cognitive control (which includes assessments of impulsivity and distractibility) may prove particularly useful in subdividing "ADHD," providing a basis for the comparison of RSN integrity within and across subgroups.

CHAPTER III

Introduction

Attention Deficit/Hyperactivity Disorder (ADHD) is a childhood psychiatric disorder reported to affect up to 11% of children aged 4-17 years old in the US (Visser et al., 2014) and 5.3% of children and adolescents world-wide (Polanczyk et al., 2007). It is characterized by symptoms of inattention, such as difficulty attending to details, and/or hyperactivity and impulsivity, such as the inability to sit still when the situation demands it or to stop fidgeting (American Psychiatric Association, 2013). Based on the type of symptoms experienced, patients may be categorized into one of the three clinically-recognized subtypes of ADHD: inattentive subtype (primarily inattentive symptoms), hyperactive subtype (primarily hyperactive/impulsive symptoms), or combination subtype (both inattentive and hyperactive/impulsive symptoms) (American Psychiatric Association, 2013). Of these subtypes, combination has been found to be most common (~60% of diagnoses), followed by inattentive (~26% of diagnoses) (Biederman et al., 2005). The hyperactive subtype is least common (Biederman et al., 2005). Studies have shown that ADHD is more prevalent among boys than girls (3:1 in population-based studies); however, this may be due to referral bias underpinned by gender differences in expression of the disorder (Biederman et al., 2005; Gaub and Caryn L. Carlson, 1997). Boys with ADHD are more likely to exhibit hyperactive and disruptive symptoms than girls, leading to their referral and eventual diagnosis; girls, on the other hand, experience more inattentive symptoms than boys, and may be overlooked (Biederman et al., 2002; Biederman et al., 2005; Gaub and Caryn L. Carlson, 1997; Hinshaw et al., 2006; Quinn, 2008).

While considerable effort has gone into investigating the underlying cause(s) of ADHD, to date none of the resulting research has been translated into the clinical arena (Cortese, 2012). Currently, ADHD is diagnosed by health care professionals based largely on parent and/or teacher reports of a child's symptoms (Parens and J. Johnston, 2009). Attempts to find a universal, intrinsic marker of ADHD to supplement behavioral measures have met with limited success due to the inherent heterogeneity of ADHD (Dias et al., 2013). That is, two patients diagnosed with ADHD may not have the same subtype of ADHD, and even patients with the same subtype may have experienced different symptoms leading to the same diagnosis (Dias et al., 2013). This inhomogeneity may give the impression of diagnostic subjectivity, reflected in the popular opinion that many children receiving a diagnosis of ADHD do not actually have the disorder: 82% of American participants in a 2005 study agreed that "ADHD is overdiagnosed today," while only 48% agreed that "ADHD is biologically based" (Norvilitis and Ping Fang, 2005). These findings stand in stark contrast to several more recent studies concluding that ADHD may in fact be underdiagnosed (Fabiano et al., 2013; Sciutto and Miriam Eisenberg, 2007). The confusion over what truly constitutes an ADHD diagnosis, and so the true prevalence of the disorder, only highlights the need for quantifiable, objective diagnostic aids.

One approach to quantifying the behavioral deficits associated with ADHD (inattention, impulsivity, etc) is through the use of neuropsychological testing. Neuropsychological tests assess brain function by determining how quickly and accurately a person can perform certain mental tasks (Lezak, 2004). These tasks may be

designed to specifically assess one domain, such as executive function, or they may be combined into a battery to assess multiple cognitive domains, such as executive function, intelligence, and/or memory (Lezak, 2004). As executive function, or "top-down" control of decision making, includes impulse control and the ability to maintain attention to relevant stimuli, performance on tasks of executive function have been thoroughly investigated in children with ADHD (Homack and Cynthia A. Riccio, 2004; Van Mourik et al., 2005; Willcutt et al., 2005). A majority of these studies have found that children with ADHD exhibit poorer performance on measures of executive function than their typically developing peers (Homack and Cynthia A. Riccio, 2004; Van Mourik et al., 2005; Willcutt et al., 2005). However, the average difference between the two groups can only be described as medium, indicating a considerable overlap in individual performance between children with and without ADHD (Homack and Cynthia A. Riccio, 2004; Van Mourik et al., 2005; Willcutt et al., 2005). Therefore, while assessment of executive function using neuropsychological testing may be diagnostically helpful in some cases, in general it is neither necessary nor sufficient to diagnose ADHD (Willcutt et al., 2005).

Another way to non-invasively assess brain function is possible through utilization of magnetic resonance imaging (MRI) techniques; specifically resting state functional MRI (rs-fMRI). While neuropsychological testing quantifies brain function through the administration of cognitive tasks, rs-fMRI measures fluctuations in blood oxygen level-dependent (BOLD) signal in the absence of a task. BOLD signal is considered to be indirect but objective measure of brain activity. Hence, brain areas with

synchronous BOLD fluctuations are thought to work together, or be functionally connected (Huettel et al., 2009). Different sets of functionally connected brain areas can form resting state networks (RSNs), which describe how a person's brain is connected at rest (Allen et al., 2011; Damoiseaux et al., 2006; De Luca et al., 2006; van den Heuvel and H. E. Hulshoff Pol, 2010). The RSN most often implicated in ADHD is the default mode network (DMN), a task-negative network associated with introspection and day-dreaming (Buckner et al., 2008; Castellanos et al., 2008; Liston et al., 2011; Qiu et al., 2011; Raichle and Abraham Z. Snyder, 2007; Raichle et al., 2001; Uddin et al., 2008; Wang et al., 2009). It has been theorized that children with ADHD have difficulty disengaging their DMNs and transitioning to goal-oriented brain activity, and that this may be a key deficit in forming the neurobiological basis of ADHD (Fassbender et al., 2009). Other studies have found decreased resting state integrity of the DMN in children with ADHD, clearly implicating the DMN as a target for further study (Castellanos et al., 2008; Fair et al., 2010; Qiu et al., 2011; Uddin et al., 2008). However, many other differences in rs-fMRI measures have been found between ADHD and typically developing youth, especially within fronto-striatal circuits (Liston et al., 2011). Previous attempts to use rs-fMRI measures as a diagnostic aid for ADHD have yielded mixed results, with accuracies in identifying ADHD patients versus typically developing controls (TDCs) ranging from 55 % to 78% (Brown et al., 2012; Cheng et al., 2012; Colby et al., 2012; Dai et al., 2012; Eloyan et al., 2012). Therefore, similar to neuropsychiatric testing, it appears that these methods cannot currently be used as stand-alone diagnostic tools.

In the present study, we attempted to build on our previous work in Chapter 2 investigating which individual RSN(s) would be most helpful in predicting ADHD status. In that work we found that, similar to other studies, the most predictive network (the inferior frontal gyrus-middle temporal network) only achieved a diagnostic accuracy of 66.7% (for details, please see Chapter 2). Our goal in this study was to combine information from all 20 resting state networks that we identified to see if, used together, this information could improve diagnostic accuracy above what we found for individual networks. The 20 networks we identified included 12 RSNs previously found in the literature (Thomason et al., 2011) and 8 components which were unclassified. Three separate but related approaches were used to combine the information provided by RSNs: (1) Individual Prediction Scores, which summed a participant's binary diagnosis of ADHD or TDC across networks; (2) Decision Values, which were used as a proxy for the confidence in the diagnosis of ADHD or TDC; and (3) Probability Estimates, which determined the likelihood of a given participant having ADHD between 0% and 100%. We hypothesized that by combining information provided by resting state networks, a diagnostic accuracy higher than 66.7% could be achieved.

Materials and Methods

Data Preparation: 630 out of the 973 rs-fMRI data acquisitions that comprise the ADHD-200 dataset were used for this analysis. Preprocessing was done using the Data Processing Assistant for Resting-State fMRI (DPARSFA, http://www.restfmri.net) (Song et al., 2011), Statistical Parametric Mapping software (SPM8, Welcome Department of Imaging Neuroscience, London, UK; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/)

and the Resting-State fMRI Data Analysis toolkit (REST). Group ICA was performed using the Group ICA of fMRI (GIFT) toolbox (http://mialab.mrn.org/software/) to resolve resting state networks (RSNs). Feature selection for each RSN/component was done using the Fisher Score algorithm available through the Arizona State University website (http://featureselection.asu.edu/software.php). LIBSVM (version 3.16, http://www.csie.ntu.edu.tw/~cjlin/libsvm/) was used in conjunction with Matlab to implement support vector machine (SVM) classification. A 10-fold cross-validation strategy was used, such that 90% of the participants were used to train the SVM and the remaining 10% served as the test set. This allowed the diagnostic SVM classifier to be implemented a total of 10 times for each of 20 RSNs/components, using a different 10% of the participants as a test set each time. In this way, each participant was included once in the test group, and at this time was assigned a diagnosis of ADHD or TDC by the SVM for each of the 20 components. The output of this methodology was therefore a set of 20 diagnoses per participant, 1 for each of the 20 RSNs/components resolved through ICA. For more detail, please refer to the methods section of Chapter 3.

(1) Individual Prediction (IP) Scores: Using the results of the SVM, four individual prediction (IP) scores were generated: 1- All Score; 2- RSN Score; 3- Attention Score; 4- Noise Score. For each of 20 components, each participant was assigned a diagnosis of ADHD (1) or TDC (0) by the SVM. These values were summed for each individual yielding a score (All Score) between 0 and 20, where a score of 0 indicated that all 20 components classified that individual as TDC and a score of 20 indicated that all 20 components classified that individual as ADHD (Table 3.1). An RSN Score was also

Table 3.1. Example generation of Individual Predictions (IP) scores. Subjects are numbered 1-630 in column 1. The true status of each participant (0 = TDC, 1 = ADHD) is shown in column 2 (labeled diagnosis). Columns 2-22, labeled C1-C20, are the statuses given by the support vector machine for each component/resting state network, respectively. Scores for each subject were generated as follows: All score: sum of status given by all components, outlined in red; RSN score: sum of status given by 12 components that represent resting state networks (RSNs), outlined in orange; Attention score: sum of status given by 5 components that represent attention networks, outlined in green; Noise score: sum of status given by 8 unclassified components, outlined in purple.

Table 3.1. Example Generation of Individual Predictions (IP) Scores

Subject	Diagnosis	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20
1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
3	0	0	1	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
630	0	0	0	0	0	1	0	0	1	1	1	0	1	0	0	0	1	0	0	0	0

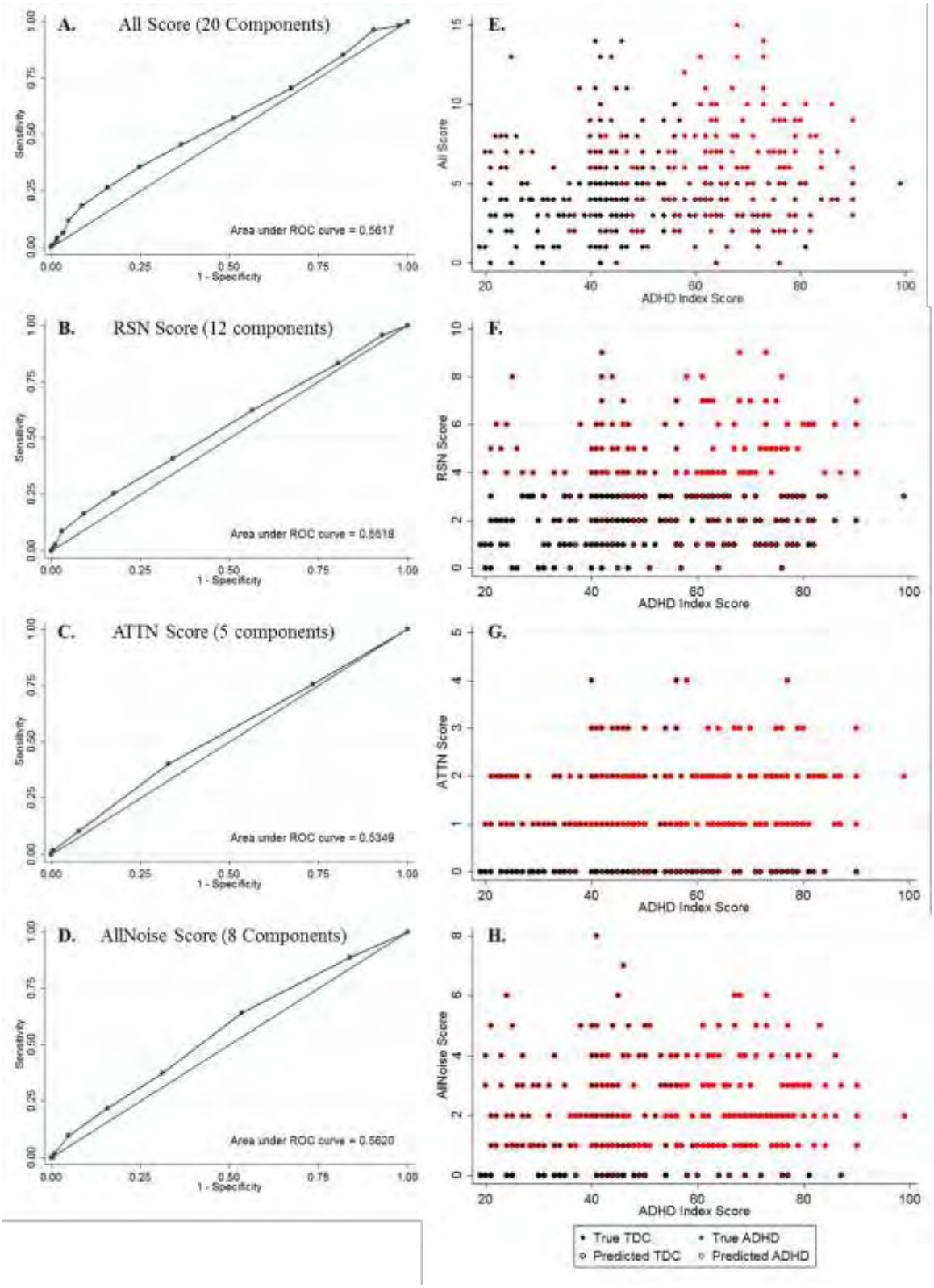
computed using only the 12 components previously found to represent resting state networks (Thomason et al., 2011); here a score of 12 indicated complete diagnosis as ADHD. Another score (Attention Score) using the 5 components found to represent the DMN and attention networks (the DMN, the posterior DMN, the salience network, and the left and right executive networks) was calculated; here a maximum score of 5, indicating complete diagnosis as ADHD, was possible. Finally, an additional score using the remaining 8 unclassified components (Noise score) was calculated (see Table 3.1).

STATA 12 software (Stata Statistical Software: Release 12. (2011) College Station, TX: StataCorp LP) was used to generate receiver operating characteristic (ROC) curves for each of these 4 IP scores (see Figure 3.1). These curves plot the sensitivity versus 1-specificity for each possible cut point of a given score, allowing for selection of the optimal cut point, that is, the one that maximizes sensitivity and minimizes 1-specificity. Youden's J-Statistic, given below in equation (3.1), was calculated for each ROC curve and used to determine the optimal cut point for each score (Youden, 1950).

$$J = \max\{\text{Sensitivty} + \text{Specificity} - 1\} \tag{3.1}$$

Visually, J determines the cut point by finding the maximum distance between the line of equality (extending from the origin to (1, 1) and describing accuracy no better than chance) and each point on the ROC curve. The point at which J is maximized is taken as the cutoff value for the given score; participants with scores above this point are considered "positive" for the test (here, diagnosed with ADHD), while scores below this

Figure 3.1. Receiver Operating Characteristic (ROC) curves for **A.** all 20 components, **B.** 12 components that are previously defined as resting state networks (RSNs), **C.** 5 components that attention/default mode networks, and **D.** 8 unclassified/noise components. Corresponding plots of **E.** All Score, **F.** RSN Score, **G.** Attention Score, and **H.** Noise Score versus ADHD index score measured by either the ADHD Rating Scale IV or Conner's Parent Rating Scale (Revised, Long Version). Filled black or red circles represent typically developing controls (TDC) or participants diagnosed with ADHD, respectively; open black or red circles represent the status of the participants as predicted by the cut point of the IP score.


A. All Score (20 Components)
Area under ROC curve = 0.5617
B. RSN Score (12 components)
Area under ROC curve = 0.5618
C. ATTN Score (5 components)
Area under ROC curve = 0.5349
D. AllNoise Score (8 Components)
Area under ROC curve = 0.5620
Sensitivity
1 - Specificity
E.
All Score
F.
RSN Score
G.
ATTN Score
H.
AllNoise Score
ADHD Index Score
True TDC
True ADHD
Predicted TDC
Predicted ADHD

point indicate absence of illness (here, TDC). The sensitivity, specificity, and accuracy for diagnosing ADHD based on the cut off score for each of the 4 IP measures was also determined.

In addition to ROC curves, plots of each IP score versus ADHD Index Score were also generated. Either the ADHD Rating Scale IV (DuPaul et al., 1997) or Conner's Parent Rating Scale (Revised, Long Version) (Conners et al., 1998) was used to quantify participants ADHD behaviors and symptoms. For both of these measures there is no single cut-point that determines ADHD diagnosis; instead scores are compared to normative data and converted to a percentile based on age and sex. For the ADHD Rating Scale IV, scores that rank above the 80th, 85th, 90th, 93rd percentiles may be considered abnormal, while for the Conner's Parent Rating Scale (Revised, Long Version), placement in the 93rd percentile or above is considered abnormal(Collett et al., 2003; Conners et al., 1998; DuPaul et al., 1997). Nevertheless, increased ADHD Index Score is associated with increased severity of symptoms/likelihood of ADHD diagnosis. Here, ADHD Index Score is used to demonstrate how intrinsic measures of brain function (i.e. resting state functional connectivity, rs-FC), relate to clinical measures, i.e. ADHD Index.

(2) Decision Values: Decision values for each component were calculated as part of SVM implementation with the LIBSVM 3.16 package (Chang and Chih-Jen Lin, 2011). Each time the SVM was implemented, it determined the hyperplane that best separated the RSN integrities, or features, of ADHD participants from those of TDC participants by finding the greatest distance between the hyperplane and the closest participants used to train the SVM. The decision value for each participant in the test set was proportional to

how far that participant ultimately fell from the separating hyperplane when the SVM was used to diagnose ADHD vs TDC. This distance can therefore be thought of as a measure of how confident the SVM was in its diagnosis: test participants whose features (RSN integrities) fell very close to the hyperplane were seen by the SVM as very similar, while those far away from the hyperplane, very different. The test participants with the shortest distance to the separating hyperplane (those with the smallest decision values) would be expected to be most frequently misdiagnosed, while those far away from the hyperplane (with the greatest decision values) would be expected to be most often correct. Decision values for each test participant resulting from SVM implementation for the left executive network (component 1), salience network (component 5), physiological noise (component 7), and the DMN (component 13) are shown here as examples plotted versus ADHD Index score (see Figure 3.2).

(3) Probability Estimates: Probability estimates were also calculated as part of the SVM implementation with the LIBSVM 3.16 package (Chang and Chih-Jen Lin, 2011). Here, each participant was assigned a probability between 0 and 1 of having a diagnosis of ADHD based on their input features (component integrities). Probability estimates for each test participant for the left executive network (component 1), salience network (component 5), physiological noise (component 7), and the DMN (component 13) are shown here as examples plotted versus ADHD Index score (see Figure 3.3).

Results

(1) Individual Prediction Scores: ROC curves for each of the 4 IP scores (All, RSN, Attention, and Noise Scores) (Figure 3.1A-D) and Youden's J-Statistic (equation 3.1)

Figure 3.2. Plots of decision values for **A.** the left executive network (component 1), **B.** the salience network (component 5), **C.** a noise component (component 7), and **D.** the default mode network (DMN, component 13) versus ADHD index score measured by either the ADHD Rating Scale IV or Conner's Parent Rating Scale (Revised, Long Version). Filled black or red circles represent typically developing controls (TDC) or participants diagnosed with ADHD, respectively; open black or red circles represent the status of the participants as predicted by the network/component in question.

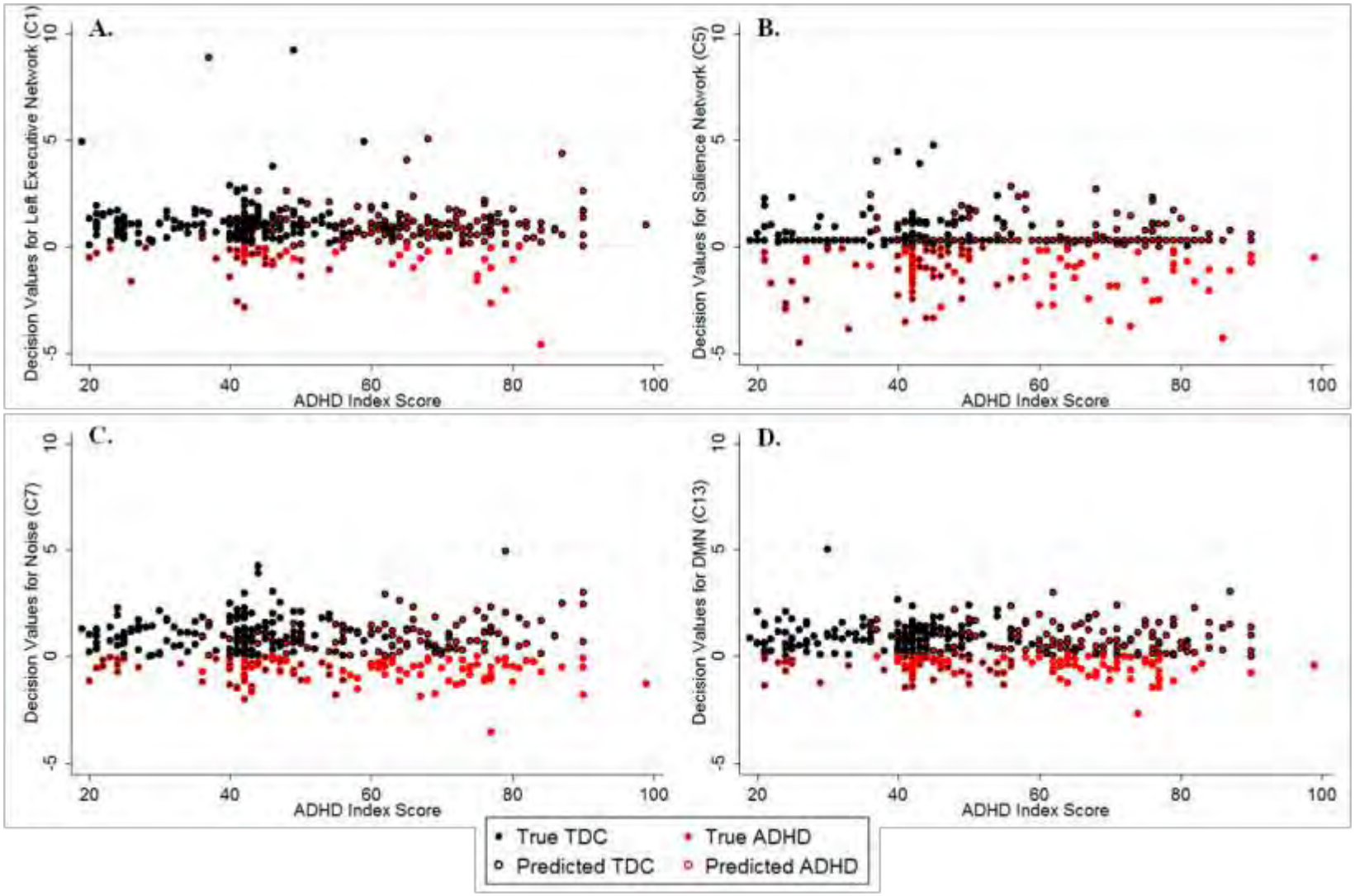
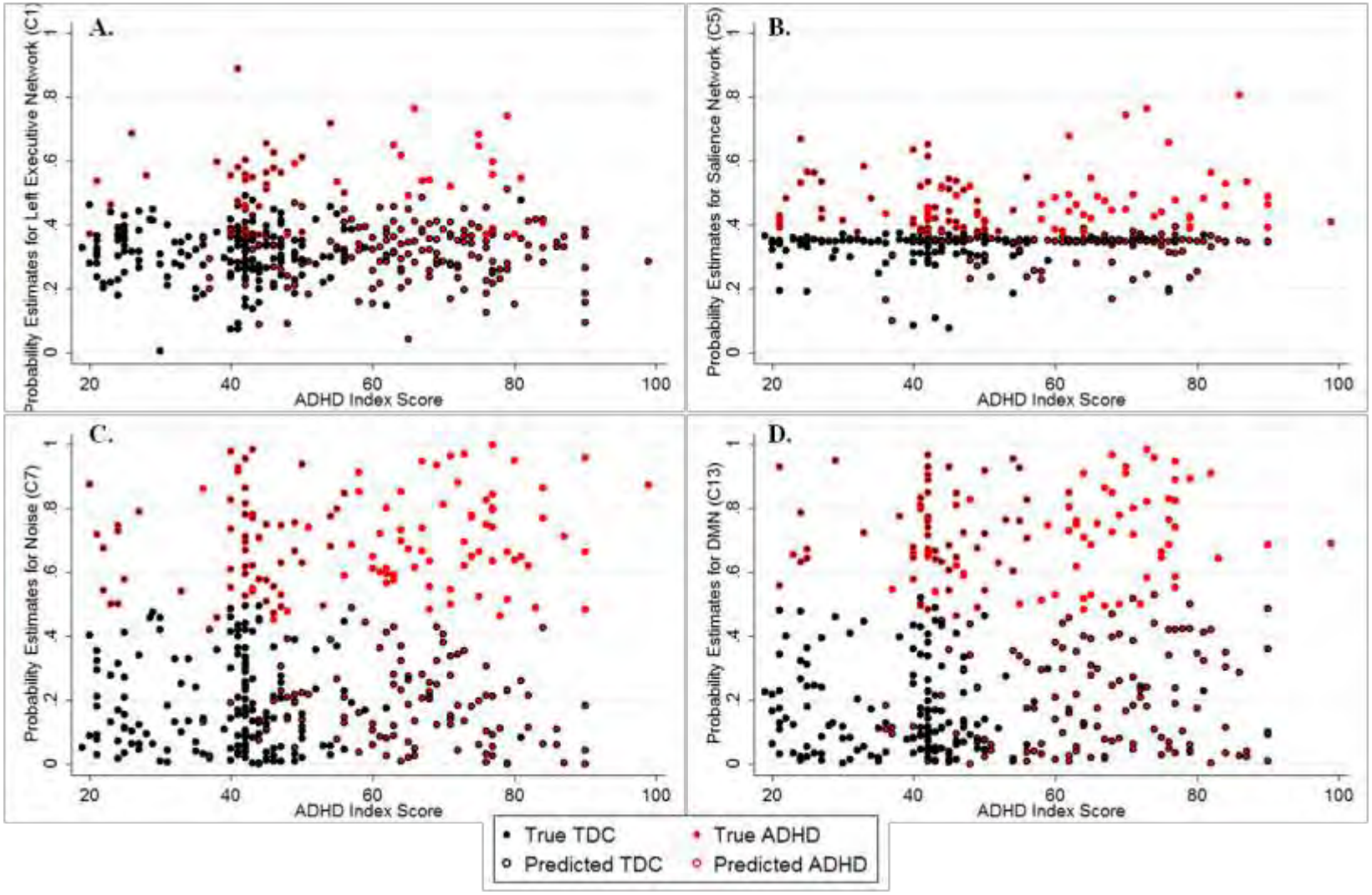

A.
Decision Values for Left Executive Network (C1)
ADHD Index Score
B.
Decision Values for Salience Network (C5)
ADHD Index Score
C.
Decision Values for Noise (C7)
ADHD Index Score
D.
Decision Values for DMN (C13)
ADHD Index Score
True TDC
Predicted TDC
True ADHD
Predicted ADHD

Figure 3.3. Plots of probability estimates for **A.** the left executive network (component 1), **B.** the salience network (component 5), **C.** a noise component (component 7), and **D.** the default mode network (DMN, component 13) versus ADHD index score measured by either the ADHD Rating Scale IV or Conner's Parent Rating Scale (Revised, Long Version). Filled black or red circles represent typically developing controls (TDC) or participants diagnosed with ADHD, respectively; open black or red circles represent the status of the participants as predicted by the network/component in question.


A.
Probability Estimates for Left Executive Network (C1)
ADHD Index Score
B.
Probability Estimates for Salience Network (C5)
ADHD Index Score
C.
Probability Estimates for Noise (C7)
ADHD Index Score
D.
Probability Estimates for DMN (C13)
ADHD Index Score
True TDC
Predicted TDC
True ADHD
Predicted ADHD

were used to determine the optimal cut point for each score. The sensitivity, specificity, and accuracy for each optimal cut point was also noted and reported here in parentheses. For the All Score (maximum = 20) the best cut point was 6 (sensitivity 35%, specificity 75%, accuracy 55%); for the RSN Score (maximum = 12), the best cut point was 4 (sensitivity 25%, specificity 82%, accuracy 54%); for the Attention Score (maximum = 5), the best cut point was 1 (sensitivity 40%, specificity 67%, accuracy 54%); and for the Noise Score (maximum = 8), the best cut point was 1 (sensitivity 64%, specificity 47%, accuracy 55%). Each participant was then plotted using each of the 4 IP scores versus their ADHD Index score (Figure 3.1E-H) to show the relationship between resting brain connectivity and behavioral measures as diagnostic tools.

For all of the following plots, each participant is represented as a filled circle, encapsulated by a boundary. The color of the filled circle corresponds to the "true" diagnosis of the patient, as provided on the ADHD-200 website (black = TDC, red = ADHD). This diagnosis was found using behavioral measures, including the ADHD Index. The color of the boundary represents the predicted diagnosis of the patient, based on the cut point of the score in question (black = TDC, red = ADHD). For each of the scores (IP, Decision Values, and Probability Estimates), if a participant scored greater than or equal to the cut point score in question, they were predicted to have ADHD; if they scored less than the cut point score, they were predicted to be TDC. Participants that appear as completely colored red dots therefore represent true positives, whereas completely colored black dots represent true negatives. Participants who appear as filled

black circles surrounded by a red boundary are false positives, and filled red circles surrounded by a black boundary are false negatives.

(2) Decision Values: Decision values, which are here presented as a proxy measure for how confident the SVM was in its predictions, are shown plotted versus ADHD Index score for the left executive network (RSN/component 1, Figure 3.2A), salience network (RSN/component 5, 2B), physiological noise (RSN/component 7, Figure 3.2C), and the DMN (RSN/component 13, 2D). The left executive, salience, and DM networks were chosen as examples because they were predicted to have the best discrimination between ADHD and TDC; physiological noise is shown because it was ultimately the most sensitive in discriminating between the two classes. Here, positive decision values correspond with a predicted diagnosis of TDC, whereas negative values correspond with predicted ADHD. The greater the magnitude of the positive value, the farther away from the separating hyperplane that participant fell, and the more "certain" the SVM was in the TDC diagnosis. Conversely, the greater the magnitude of the negative value, the more "certain" the SVM was in the ADHD diagnosis. While the left executive network was among the least accurate in predicting ADHD diagnoses, it generated some of the largest decision values for TDC participants (very confident in incorrect diagnoses). By contrast, the salience network produced many decision values that were close to zero (very unsure of incorrect diagnoses). Both physiological noise and the DMN produced conservative decision values (not overly confident in correct and incorrect diagnoses); at the same time these were not next to zero (not very unsure of diagnoses).

(3) Probability Estimates: The probability of having ADHD, ranging from 0 (0%) to 1 (100%), was assigned to each participant based on their individual component integrities. These estimates are shown plotted versus ADHD Index score for the left executive network (component 1, Figure 3.3A), salience network (component 5, Figure 3.3B), physiological noise (component 7, Figure 3.3C), and the DMN (component 13, Figure 3.3D). By comparing the upper-right hand quadrant of each of these 4 examples, one can see the relative number of true positive participants found using each component; here it is clear that physiological noise (component 7) had the greatest number of true positives (and thus the highest sensitivity).

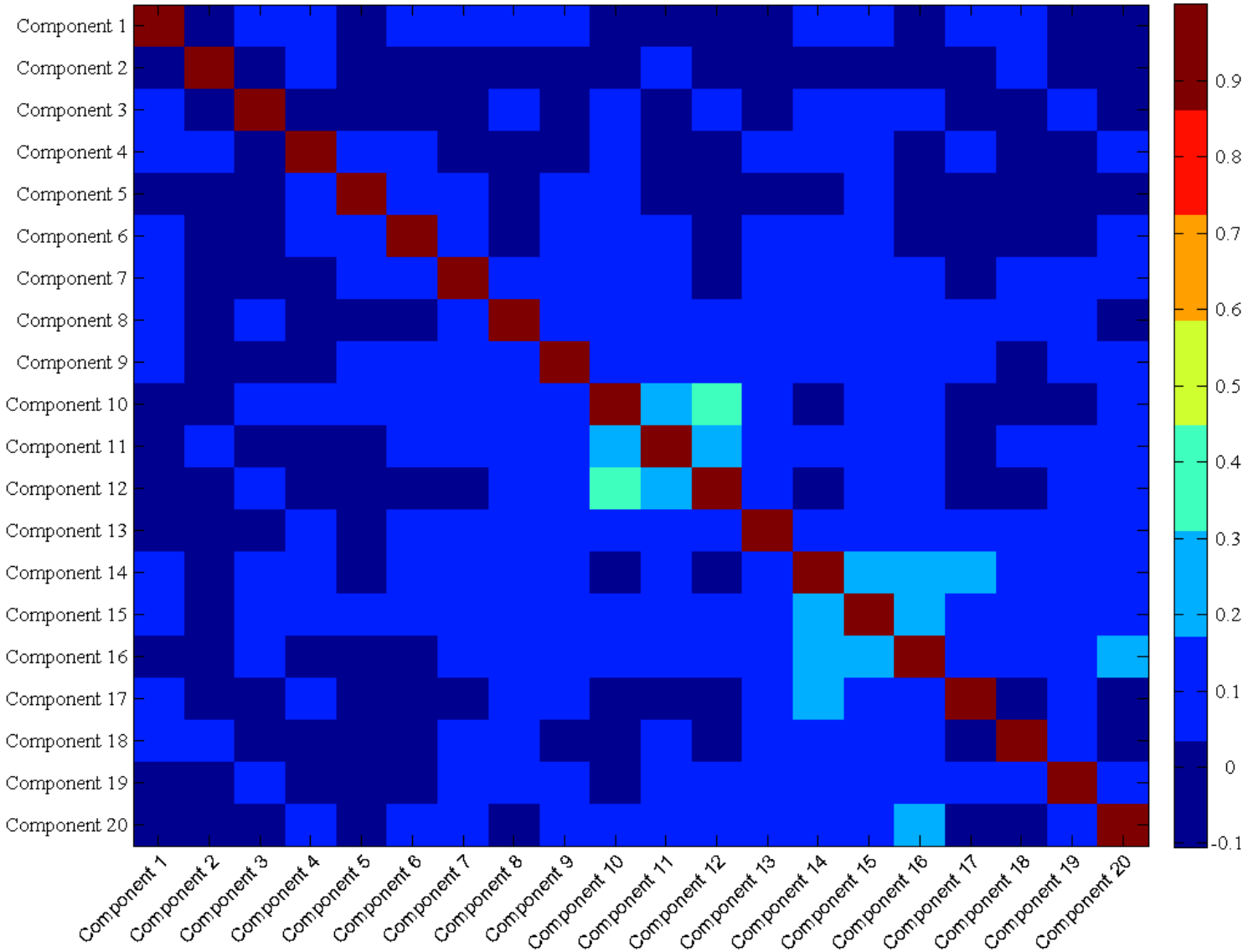
Discussion

The focus of this study was the integration of the output from the SVM approach described in the previous chapter. Here, we attempted to combine information from 20 putative resting state networks (12 of which matched networks previously identified in another study (Thomason et al., 2011) and 8 which remained unclassified) in 3 separate but related ways, by using: (1) individual prediction scores, based on each individual network's/component's diagnosis for each participant, (2) decision values, which represented each network's/component's confidence in its diagnosis, and (3) probability estimates, which assigned a likelihood of ADHD diagnosis between 0 and 100% for each network/component. We hypothesized that by including information from multiple networks/components, a diagnostic accuracy higher than that found using a single network/component (66.7% for the most accurate network) could be achieved. Contrary to our hypothesis, we discovered that combining information across

networks/components resulted in lower diagnostic accuracies (54-55%). To ensure that there were not better groupings of components that could be used for generating IP scores, we calculated pair-wise correlation coefficients for each component's diagnostic output (Figure 3.4). This analysis revealed that there were very few components that diagnosed participants similarly. The highest correlation was between component 11 (visual network) and component12 (unclassified) (R=0.37), while most components had correlations around zero. Together, these findings argue that the examination of individual networks/components may be more helpful than combining information across networks in the search for objective markers of ADHD.

It has been established that ADHD is a heterogeneous disorder; hence the proposed approach of assigning a binary diagnosis of ADHD or TDC did not result in high diagnostic accuracy. That is, the diagnosis of ADHD may actually include a range of sub-conditions, each of which may be the result of independent (but perhaps associated) underlying functional deficits. In this case, treating individual ADHD diagnoses as a single disorder would yield a high within-group variance in rs-FC measures and possible overlap with the TDC population, similar to what was found for neuropsychological testing (Homack and Cynthia A. Riccio, 2004; Van Mourik et al., 2005; Willcutt et al., 2005). In our analysis, this would make it extremely difficult, if not impossible, for the SVM to cleanly separate ADHD and TDC groups. This would explain our finding that diagnostic accuracy was not high for any given network/component. Furthermore, this would also explain why combining information across networks/components caused a large drop in accuracy: one ADHD sub-condition may involve certain RSNs, while

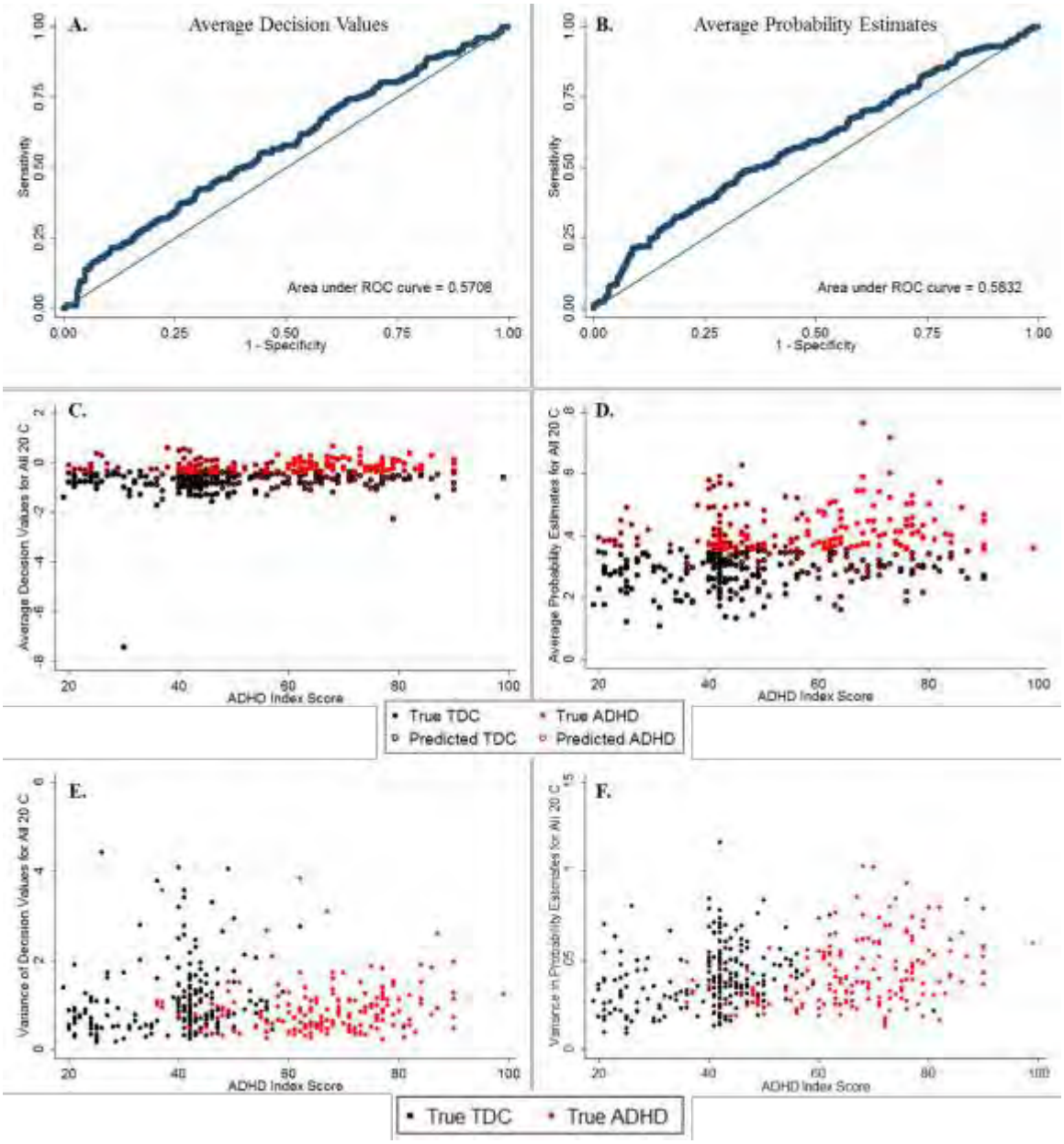
Figure 3.4. Correlation coefficient (cc) matrix for all 20 components. Colorbar corresponds to R-values, with warmer colors indicating a higher correlation and cooler colors indicating a lower correlation or negative correlation.


Component 1
Component 2
Component 3
Component 4
Component 5
Component 6
Component 7
Component 8
Component 9
Component 10
Component 11
Component 12
Component 13
Component 14
Component 15
Component 16
Component 17
Component 18
Component 19
Component 20
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1

another may involve different RSNs altogether. If RSN differences are combined for all "ADHD" participants, as was done in the present work, distinguishing RSN profiles for ADHD sub-conditions will possibly be lost and their overall accuracy in predicting diagnoses will drop. Evidence for this can be seen in Figure 3.5, which shows the ROC curves for combined decision values (Figure 3.5A) and combined probability estimates (Figure 3.5B). Here, the overall accuracy is about 57% for combined decision values and 58% for combined probability estimates, still lower than the 66.7% accuracy for the best individual predictive network. Combined decision values or probability estimates versus behavioral measures of ADHD (ADHD Index) are shown in Figures 3.5C and D, respectively. Perhaps most interestingly, the variance of decision values or probability estimates across all 20 components is shown in Figures 3.5E and F, respectively. Here, it can be seen that for many subjects, the confidence in diagnosis (decision values) and the likelihood of being assigned an ADHD diagnosis (probability estimates) varies hugely across the components used. This supports the idea some participants may show deficits in one set of RSNs, while other participants may have a completely different RSN profile, yet all are clinically diagnosed with the label ADHD.

The question, then, turns from "Can we use resting state connectivity measures to accurately diagnose ADHD?" to "Could these measures identify ADHD subgroups?" or perhaps more generally, "What is the best way to identify ADHD subgroups?" To answer this question, the Research Domain Criteria (RDoC) approach laid out by the NIMH may prove extremely suitable (Insel et al., 2010). The RDoC framework seeks to uncover

Figure 3.5. Receiver Operating Characteristic (ROC) curves for **A.** average decision values and **B.** average probability estimates. Plots of **C.** average decision values across all components for each participant, **D.** average probability estimates across all components for each participant, **E.** variance in decision values for each participant across all components, and **F.** variance in probability estimates for each participant across all components versus ADHD index score measured by either the ADHD Rating Scale IV or Conner's Parent Rating Scale (Revised, Long Version). Filled black or red circles represent typically developing controls (TDC) or participants diagnosed with ADHD, respectively. Open black or red circles represent the status of the participants as TDC or ADHD as determined by the cut point for each measure, respectively.


A.
Average Decision Values
Sensitivity
1 - Specificity
Area under ROC curve = 0.5708
B.
Average Probability Estimates
Sensitivity
1 - Specificity
Area under ROC curve = 0.5832
C.
Average Decision Values for All 20 C
ADHD Index Score
D.
Average Probability Estimates for All 20 C
ADHD Index Score
True TDC
True ADHD
Predicted TDC
Predicted ADHD
E.
Variance of Decision Values for All 20 C
ADHD Index Score
F.
Variance in Probability Estimates for All 20 C
ADHD Index Score
True TDC
True ADHD

links between quantifiable indices (such as genes, brain networks, and behavioral measures, to name a few) without regard for current disorder labels (such as ADHD). In this way, the focus of investigation goes beyond attempting to pinpoint the biological basis of a given disorder to understanding intrinsic brain-behavior relationships that may cut across diagnostic categories, or even be present at sub-threshold levels in the typically developing population. Following this scheme, investigators have recently attempted to relate neuropsychological deficits to underlying ADHD subgroups, either by using the pre-labeled inattentive and combination subtypes (Nikolas and Joel T. Nigg, 2013) or by grouping participants with ADHD based on their neuropsychological performance (Fair et al., 2012). Both of these approaches were successful in making this relationship, although the conclusions drawn from each group were very different. The first study found that ADHD patients of the combined subtype performed worse than those of the inattentive subtype on all neuropsychological measures examined, supporting the theory that ADHD may exist on a spectrum (Nikolas and Joel T. Nigg, 2013). Other investigators have taken this theory to the extreme, positing that these two subtypes of ADHD may represent completely separate disorders (Milich et al., 2001). By contrast, the second study found that participants' neuropsychological performance profiles could be sorted into 6 subgroups; this was true for both ADHD and typically developing children (Fair et al., 2012). This second study further found that symptom severity did not vary among the ADHD subgroups, arguing against a spectrum model of ADHD and instead supporting the notion that multiple cognitive subtypes exist not only within the ADHD patient population, but also for typically developing youths (Fair et al., 2012). Several

other studies have also found evidence of multiple cognitive profiles within the ADHD group, although the number of subgroups varies by study (Roberts et al., 2013; Thaler et al., 2013). The natural extension of this model would be to attempt to relate rs-fMRI measures, such as RSN integrity, within and across subgroups defined by their cognitive profiles.

A better understanding of what comprises ADHD (as well as how deficits associated with this disorder can be accurately identified) is of great relevance to public health. In addition to the large annual costs of ADHD in the US (estimated to be about $31 billion dollars, (Hanna, 2009)), there is the question of how a misdiagnosis affects the health of a child. First, consider the case of a patient who may not truly have the disorder, but is diagnosed with it (a false positive). Risks for this child include inappropriate medication with psychostimulants such as methylphenidate or dextroamphetamine. While these pharmacological agents have been found to be effective in reducing the core symptoms of ADHD, treatment with them carries the potential of experiencing such side effects as insomnia, anorexia, and increased anxiety (Barkley et al., 1990; Efron et al., 1997; Scheffler et al., 2009; The MTA Cooperative Group, 1999). In addition, stimulants are considered potential drugs of abuse, making their prescription to children (especially to a child who may not receive therapeutic benefit from them) a concern (Hanna, 2009). Furthermore, some ADHD symptoms are common with other disorders, introducing the possibility that a child misdiagnosed with ADHD may actually have another psychiatric or learning disorder (Berger, 2011). In this case, the misrecognition of ADHD may actually prevent appropriate treatment for the true disorder

experienced by the child. Conversely, one may also consider the case of a patient who truly does have ADHD, but the diagnosis is missed and he or she is considered typically developing (a false negative). This child may benefit from medication, but will not receive it. In addition, without the diagnosis of ADHD this child may not qualify for educational services that could increase his or her ability to perform in school (Gregg, 2000).

In conclusion, it is clear that a more objective marker of ADHD would be of great clinical value. It could help ensure the recognition of individuals with this disorder while simultaneously validating it as a "true" illness. It could help prevent the misdiagnosis of children who do not truly have ADHD. Eventually, an objective diagnostic aid may even be useful for distinguishing between ADHD and other psychiatric disorders of childhood. However, several challenges need to be overcome before an intrinsic marker of ADHD can be considered appropriate for use in the clinical setting. The foremost of these challenges is a better understanding the deficits that underlie "ADHD" as a diagnostic label, and how these deficits can be quantified and reliably related to the disorder. To this end, future studies may focus on linking different cognitive subgroups within the "ADHD" label to RSN integrities, following the RDoC paradigm.

SUMMARY AND CONCLUSION

In the current work, we examined how RSN integrity differed between children with ADHD and their typically developing peers. We found significant over- and under-connections not only within the DMN, which has been previously shown to be under-connected in ADHD, but also in three networks that underlie attention: the salience network, the left executive network, and the right executive network. Furthermore, differences in the integrity of these RSNs were found in two sub-analyses, the first using only patients with the combination subtype of ADHD and the second using only patients with the inattentive subtype of ADHD. From this analysis we concluded that attention networks are generally over-connected in ADHD, and that the DMN is generally under-connected in combination subtype ADHD, but over-connected in inattentive subtype ADHD.

Next, we investigated whether differences in each network could be used to predict ADHD status using a support vector machine. We found that the DMN, contrary to our hypothesis, was not the best predictor of ADHD. Instead, we found that two relatively unfamiliar networks (the IFG-middle temporal network, 66.8% accuracy and the parietal association network, 86.6% specificity and 48.5% PPV) and a physiological noise component (sensitivity 39.7% and NPV 69.6%) had the best classification performance. This led us to conclude that while there is information contained within RSN integrities that is useful for diagnostic predictions, the approach explored in this work did not produce a stand-alone diagnostic tool.

Finally, we attempted to combine information across networks to improve classification accuracy. We used three different methods – Individual Prediction Scores,

Decision Values, and Probability Estimates – to integrate information from multiple RSNs. Contrary to our hypothesis that this would improve classification, the accuracy actually decreased to 54% - 55% when information was combined. From this result we concluded that individual patients may have ADHD-related deficits that are expressed in a combination of RSNs unique to that patient, or a smaller subset of ADHD patients.

There are several limitations to the present work. First and foremost, there is the issue of how each participant's "true" diagnosis (taken from the ADHD-200 website) was made: diagnostic instruments varied from site to site, and all of these relied at least in part on parent report of symptoms or behaviors. This lack of a precise gold standard may have served as a source of error in determining the accuracy of RSNs in predicting the presence of ADHD. Second, the ADHD-200 database is cross-sectional, and so lacks information about the course of ADHD symptoms and response to medication over time. Lastly, most ADHD-200 sites reported that ADHD subjects were withdrawn from stimulant medication 1-3 days before their scans took place; however, several sites did not report whether medication was withdrawn. This introduces the possibility that some ADHD subjects included in our analysis were scanned while experiencing the effects of medication, which may have affected their RSN integrities, and so the accuracy of the SVM during classification.

While our approach here was largely unsuccessful in accurately classifying ADHD patients, this avenue of research still holds great potential for the development of diagnostic tools. Overall, these findings imply that while resting state connectivity is different between patients with ADHD and TDCs, how it differs may vary from patient to

patient. As it appears that there is no one common deficit or set of deficits in RSNs shared by the majority of patients with ADHD, the use of RSNs to predict ADHD status poses a challenging, though promising, problem. The solution may lie in the identification and characterization of patient subgroups within the ADHD diagnosis. To this end, future studies may focus on using resting state connectivity measures to sort both ADHD patients and TDCs into clusters that share similar connectivity profiles, with the goal of investigating how symptoms and functional impairments vary across these clusters. Ultimately, it may be possible to use resting state connectivity measures as a diagnostic aide not for the overarching diagnosis of "ADHD," but instead for membership to subgroups that may exist within the disorder.

REFERENCES

AACAP Official Action (2007): Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 46.

Allen EA, Erhardt EB, Damaraju E, Gruner W, Segall JM, Silva RF, Havlicek M, Rachakonda S, Fries J, Kalyanam R (2011): A baseline for the multivariate comparison of resting-state networks. Frontiers in systems neuroscience 5.

Ambroise C, McLachlan GJ (2002): Selection bias in gene extraction on the basis of microarray gene-expression data. Proceedings of the National Academy of Sciences 99: 6562-6566.

American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders (5th Ed.). Arlington, VA: American Psychiatric Publishing.

Bandettini PA, Wong EC, Hinks RS, Tikofsky RS, Hyde JS (1992): Time course EPI of human brain function during task activation. Magn Reson Med 25: 390-397.

Barbaresi WJ, Colligan RC, Weaver AL, Voigt RG, Killian JM, Katusic SK (2013): Mortality, ADHD, and psychosocial adversity in adults with childhood ADHD: a prospective study. Pediatrics 131: 637-644.

Barkley RA (1997): Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol Bull 121: 65.

Barkley RA, Fischer M, Edelbrock C, Smallish L (1991): The adolescent outcome of hyperactive children diagnosed by research criteria—III. Mother–child interactions, family conflicts and maternal psychopathology. Journal of Child Psychology and Psychiatry 32: 233-255.

Barkley RA, McMurray MB, Edelbrock CS, Robbins K (1990): Side effects of methylphenidate in children with attention deficit hyperactivity disorder: a systemic, placebo-controlled evaluation. Pediatrics 86: 184-192.

Bechara A, Tranel D, Damasio H (2000): Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain 123: 2189-2202.

Beckmann CF, DeLuca M, Devlin JT, Smith SM (2005): Investigations into resting-state connectivity using independent component analysis. Philos Trans R Soc Lond B Biol Sci 360: 1001-1013.

Bell AJ, Sejnowski TJ (1995): An information-maximization approach to blind separation and blind deconvolution. Neural Comput 7: 1129-1159.

Berger I (2011): Diagnosis of attention deficit hyperactivity disorder: much ado about something. IMAJ-Israel Medical Association Journal 13: 571.

Biederman J, Faraone S, Milberger S, Guite J, Mick E, Chen L, Mennin D, Marrs A, Ouellette C, Moore P (1996): A prospective 4-year follow-up study of attention-deficit hyperactivity and related disorders. Arch Gen Psychiatry 53: 437.

Biederman J, Kwon A, Aleardi M, Chouinard V, Marino T, Cole H, Mick E, Faraone SV (2005): Absence of gender effects on attention deficit hyperactivity disorder: findings in nonreferred subjects. Am J Psychiatry 162: 1083-1089.

Biederman J, Mick E, Faraone SV, Braaten E, Doyle A, Spencer T, Wilens TE, Frazier E, Johnson MA (2002): Influence of gender on attention deficit hyperactivity disorder in children referred to a psychiatric clinic. Am J Psychiatry 159: 36-42.

Biederman J, Monuteaux MC, Mick E, Spencer T, Wilens TE, Silva JM, Snyder LE, Faraone SV (2006): Young adult outcome of attention deficit hyperactivity disorder: a controlled 10-year follow-up study. Psychol Med 36: 167-179.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995): Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34: 537-541.

Booth JR, Burman DD, Meyer JR, Lei Z, Trommer BL, Davenport ND, Li W, Parrish TB, Gitelman DR, Mesulam MM (2005): Larger deficits in brain networks for response inhibition than for visual selective attention in attention deficit hyperactivity disorder (ADHD). J Child Psychol Psychiatry 46: 94-111.

Börger N, van der Meere J (2000): Motor control and state regulation in children with ADHD: a cardiac response study. Biol Psychol 51: 247-267.

Börger N, van Der Meere J, Ronner A, Alberts E, Geuze R, Bogte H (1999): Heart rate variability and sustained attention in ADHD children. J Abnorm Child Psychol 27: 25-33.

Boser BE, Guyon IM, Vapnik VN (1992): A training algorithm for optimal margin classifiers. : 144-152.

Brown MR, Sidhu GS, Greiner R, Asgarian N, Bastani M, Silverstone PH, Greenshaw AJ, Dursun SM (2012): ADHD-200 Global Competition: diagnosing ADHD using personal characteristic data can outperform resting state fMRI measurements. Frontiers in systems neuroscience 6.

Buchhorn R, Conzelmann A, Willaschek C, Störk D, Taurines R, Renner TJ (2012): Heart rate variability and methylphenidate in children with ADHD. ADHD Attention Deficit and Hyperactivity Disorders 4: 85-91.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008): The brain's default network. Ann N Y Acad Sci 1124: 1-38.

Calhoun V, Adali T, Pearlson G, Pekar J (2001): A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp 14: 140-151.

Calhoun VD, Liu J, Adali T (2009): A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. Neuroimage 45: S163-72.

Castellanos FX, Proal E (2012): Large-scale brain systems in ADHD: Beyond the prefrontal–striatal model. Trends Cogn Sci (Regul Ed) 16: 17-26.

Castellanos FX, Margulies DS, Kelly C, Uddin LQ, Ghaffari M, Kirsch A, Shaw D, Shehzad Z, Di Martino A, Biswal B, Sonuga-Barke EJ, Rotrosen J, Adler LA, Milham MP (2008): Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. Biol Psychiatry 63: 332-337.

Chang C, Lin C (2011): LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2: 27.

Cheng W, Ji X, Zhang J, Feng J (2012): Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. Frontiers in Systems Neuroscience 6.

Colby JB, Rudie JD, Brown JA, Douglas PK, Cohen MS, Shehzad Z (2012): Insights into multimodal imaging classification of ADHD. Frontiers in systems neuroscience 6.

Cole DM, Smith SM, Beckmann CF (2010): Advances and pitfalls in the analysis and interpretation of resting-state FMRI data. Front Syst Neurosci 4: 8.

Collett BR, Ohan JL, Myers KM (2003): Ten-year review of rating scales. V: scales assessing attention-deficit/hyperactivity disorder. Journal of the American Academy of Child & Adolescent Psychiatry 42: 1015-1037.

Conners CK, Sitarenios G, Parker JD, Epstein JN (1998): The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. J Abnorm Child Psychol 26: 257-268.

Cortes C, Vapnik V (1995): Support-vector networks. Mach Learning 20: 273-297.

Cortese S (2012): The neurobiology and genetics of Attention-Deficit/Hyperactivity Disorder (ADHD): What every clinician should know. Eur J Paediatr Neurol.

Dagli MS, Ingeholm JE, Haxby JV (1999): Localization of cardiac-induced signal change in fMRI. Neuroimage 9: 407-415.

Dai D, Wang J, Hua J, He H (2012): Classification of ADHD children through multimodal magnetic resonance imaging. Frontiers in systems neuroscience 6.

Damoiseaux J, Rombouts S, Barkhof F, Scheltens P, Stam C, Smith SM, Beckmann C (2006): Consistent resting-state networks across healthy subjects. Proceedings of the National Academy of Sciences 103: 13848-13853.

De Luca M, Beckmann C, De Stefano N, Matthews P, Smith SM (2006): fMRI resting state networks define distinct modes of long-distance interactions in the human brain. Neuroimage 29: 1359-1367.

Deligiannidis KM, Sikoglu EM, Shaffer SA, Frederick B, Svenson AE, Kopoyan A, Kosma CA, Rothschild AJ, Moore CM (2013): GABAergic neuroactive steroids and resting-state functional connectivity in postpartum depression: A preliminary study. J Psychiatr Res.

Dias TGC, Kieling C, Graeff-Martins AS, Moriyama TS, Rohde LA, Polanczyk GV (2013): Developments and challenges in the diagnosis and treatment of ADHD. Revista Brasileira de Psiquiatria 35: S40-S50.

Dickstein SG, Bannon K, Castellanos FX, Milham MP (2006): The neural correlates of attention deficit hyperactivity disorder: an ALE meta-analysis. J Child Psychol Psychiatry 47: 1051-1062.

DuPaul GJ, Power TJ, Anastopoulos AD, Reid R, McGoey KE, Ikeda MJ (1997): Teacher ratings of attention deficit hyperactivity disorder symptoms: Factor structure and normative data. Psychol Assess 9: 436.

Durston S, Davidson MC, Mulder MJ, Spicer JA, Galvan A, Tottenham N, Scheres A, Xavier Castellanos F, van Engeland H, Casey BJ (2007): Neural and behavioral correlates of expectancy violations in attention-deficit hyperactivity disorder. J Child Psychol Psychiatry 48: 881-889.

Durston S, Tottenham NT, Thomas KM, Davidson MC, Eigsti IM, Yang Y, Ulug AM, Casey BJ (2003): Differential patterns of striatal activation in young children with and without ADHD. Biol Psychiatry 53: 871-878.

Durston S, van Belle J, de Zeeuw P (2011): Differentiating frontostriatal and fronto-cerebellar circuits in attention-deficit/hyperactivity disorder. Biol Psychiatry 69: 1178-1184.

Edwards G, Barkley RA, Laneri M, Fletcher K, Metevia L (2001): Parent–adolescent conflict in teenagers with ADHD and ODD. J Abnorm Child Psychol 29: 557-572.

Efron D, Jarman F, Barker M (1997): Side effects of methylphenidate and dexamphetamine in children with attention deficit hyperactivity disorder: a double-blind, crossover trial. Pediatrics 100: 662-666.

Eloyan A, Muschelli J, Nebel MB, Liu H, Han F, Zhao T, Barber A, Joel S, Pekar JJ, Mostofsky S (2012): Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging.

Fabiano GA, Pelham Jr WE, Majumdar A, Evans SW, Manos MJ, Caserta D, Girio-Herrera EL, Pisecco S, Hannah JN, Carter RL (2013): Elementary and middle school teacher perceptions of attention-deficit/hyperactivity disorder prevalence. : 1-13.

Fair DA, Bathula D, Nikolas MA, Nigg JT (2012): Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. Proceedings of the National Academy of Sciences 109: 6769-6774.

Fair DA, Posner J, Nagel BJ, Bathula D, Dias TGC, Mills KL, Blythe MS, Giwa A, Schmitt CF, Nigg JT (2010): Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. Biol Psychiatry 68: 1084-1091.

Faraone SV, Biederman J, Mick E (2006): The age-dependent decline of attention deficit hyperactivity disorder: a meta-analysis of follow-up studies. Psychol Med 36: 159-165.

Fassbender C, Zhang H, Buzy WM, Cortes CR, Mizuiri D, Beckett L, Schweitzer JB (2009): A lack of default network suppression is linked to increased distractibility in ADHD. Brain Res 1273: 114-128.

Filippini N, MacIntosh BJ, Hough MG, Goodwin GM, Frisoni GB, Smith SM, Matthews PM, Beckmann CF, Mackay CE (2009): Distinct patterns of brain activity in young carriers of the APOE-ε4 allele. Proceedings of the National Academy of Sciences 106: 7209-7214.

Fox MD, Greicius M (2010): Clinical applications of resting state functional connectivity. Front Syst Neurosci 4: 19.

Freilich ER, Gaillard WD (2010): Utility of functional MRI in pediatric neurology. Curr Neurol Neurosci Rep 10: 40-46.

Gaub M, Carlson CL (1997): Gender differences in ADHD: a meta-analysis and critical review. Journal of the American Academy of Child & Adolescent Psychiatry 36: 1036-1045.

Gershon J, Gershon J (2002): A meta-analytic review of gender differences in ADHD. Journal of attention disorders 5: 143-154.

Gregg S. 2000. ADHD and school law. Appalachia Educational Lab.

Greicius M (2008): Resting-state functional connectivity in neuropsychiatric disorders. Curr Opin Neurol 21: 424-430.

Hanna N (2009): Attention Deficit Disorder (ADD) Attention Deficit Hyperactive Disorder (ADHD). American Journal of Clinical Medicine 6: 22-28.

Hart H, Radua J, Nakao T, Mataix-Cols D, Rubia K (2013): Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. JAMA psychiatry 70: 185-198.

Hebb D (1949): The organization of behavior: A neuropsychological theory.

Hinshaw SP, Owens EB, Sami N, Fargeon S (2006): Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. J Consult Clin Psychol 74: 489.

Homack S, Riccio CA (2004): A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. Archives of Clinical Neuropsychology 19: 725-743.

Huettel SA, Song AW, McCarthy G. 2009. Functional magnetic resonance imaging. Sinauer Associates Sunderland.

Huh Y, Choi I, Song M, Kim S, Hong SD, Joung Y (2011): A comparison of comorbidity and psychological outcomes in children and adolescents with attention-deficit/hyperactivity disorder. Psychiatry Investig 8: 95-101.

Hyvärinen A, Oja E (2000): Independent component analysis: algorithms and applications. Neural Networks 13: 411-430.

Insel TR, Cuthbert BN, Garvey MA, Heinssen RK, Pine DS, Quinn KJ, Sanislow CA, Wang PS (2010): Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 167: 748-751.

Jorna P (1992): Spectral analysis of heart rate and psychological state: A review of its validity as a workload index. Biol Psychol 34: 237-257.

Kohavi R (1995): A study of cross-validation and bootstrap for accuracy estimation and model selection. : 1137-1145.

Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, Turner R (1992): Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc Natl Acad Sci USA 89: 5675-5679.

Lee MH, Smyser CD, Shimony JS (2013): Resting-state fMRI: a review of methods and clinical applications. AJNR Am J Neuroradiol 34: 1866-1872.

Lezak MD. 2004. Neuropsychological Assessment 4 Ed. Oxford university press.

Liston C, Cohen MM, Teslovich T, Levenson D, Casey B (2011): Atypical prefrontal connectivity in attention-deficit/hyperactivity disorder: pathway to disease or pathological end point? Biol Psychiatry 69: 1168-1177.

Loe IM, Feldman HM (2007): Academic and educational outcomes of children with ADHD. J Pediatr Psychol 32: 643-654.

Malonek D, Grinvald A (1996): Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. Science 272: 551.

Mannuzza S, Klein RG (2000): Long-term prognosis in attention-deficit/hyperactivity disorder. Child Adolesc Psychiatr Clin N Am 9: 711-726.

McKeown MJ, Sejnowski TJ (1998): Independent component analysis of fMRI data: examining the assumptions. Hum Brain Mapp 6: 368-372.

Milham MP, Fair D, Mennes M, Mostofsky SH (2012): The ADHD-200 consortium: a model to advance the translational potential of neuroimaging in clinical neuroscience. Frontiers in Systems Neuroscience 6: 62.

Milich R, Balentine AC, Lynam DR (2001): ADHD combined type and ADHD predominantly inattentive type are distinct and unrelated disorders. Clinical psychology: science and practice 8: 463-488.

Mitchell TM (1997): Machine learning. 1997. Burr Ridge, IL: McGraw Hill 45.

Mostofsky SH, Rimrodt SL, Schafer JG, Boyce A, Goldberg MC, Pekar JJ, Denckla MB (2006): Atypical motor and sensory cortex activation in attention-deficit/hyperactivity disorder: a functional magnetic resonance imaging study of simple sequential finger tapping. Biol Psychiatry 59: 48-56.

Nikolas MA, Nigg JT (2013): Neuropsychological performance and attention-deficit hyperactivity disorder subtypes and symptom dimensions. Neuropsychology 27: 107.

Noble WS (2006): What is a support vector machine? Nat Biotechnol 24: 1565-1567.

Norman KA, Polyn SM, Detre GJ, Haxby JV (2006): Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci (Regul Ed) 10: 424-430.

Norvilitis JM, Fang P (2005): Perceptions of ADHD in China and the United States: A preliminary study. Journal of Attention Disorders 9: 413-424.

Ogawa S, Lee TM, Nayak AS, Glynn P (1990): Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. Magn Reson Med 14: 68-78.

Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K (1992): Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci USA 89: 5951-5955.

Oldehinkel M, Francx W, Beckmann CF, Buitelaar JK, Mennes M (2013): Resting state FMRI research in child psychiatric disorders. Eur Child Adolesc Psychiatry 22: 757-770.

Parens E, Johnston J (2009): Facts, values, and attention-deficit hyperactivity disorder (ADHD): an update on the controversies. Child Adolesc Psychiatry Ment Health 3: 1.

Parker A, Corkum P (2013): ADHD Diagnosis: As Simple As Administering a Questionnaire or a Complex Diagnostic Process? J Atten Disord.

Pelham J, William E, Fabiano GA, Massetti GM (2005): Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. Journal of Clinical Child and Adolescent Psychology 34: 449-476.

Pliszka SR, Glahn DC, Semrud-Clikeman M, Franklin C, Perez R,3rd, Xiong J, Liotti M (2006): Neuroimaging of inhibitory control areas in children with attention deficit hyperactivity disorder who were treatment naive or in long-term treatment. Am J Psychiatry 163: 1052-1060.

Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA (2007): The worldwide prevalence of ADHD: a systematic review and metaregression analysis. Am J Psychiatry 164: 942-948.

Posner J, Park C, Wang Z (2014): Connecting the Dots: A Review of Resting Connectivity MRI Studies in Attention-Deficit/Hyperactivity Disorder. Neuropsychol Rev: 1-13.

Pugzles Lorch E, Eastham D, Milich R, Lemberger CC, Polley Sanchez R, Welsh R, van den Broek P (2004): Difficulties in Comprehending Causal Relations Among Children With ADHD: The Role of Cognitive Engagement. J Abnorm Psychol 113: 56.

Qiu MG, Ye Z, Li QY, Liu GJ, Xie B, Wang J (2011): Changes of brain structure and function in ADHD children. Brain Topogr 24: 243-252.

Quinn PO (2008): Attention-deficit/hyperactivity disorder and its comorbidities in women and girls: an evolving picture. Curr Psychiatry Rep 10: 419-423.

Raichle ME, Snyder AZ (2007): A default mode of brain function: a brief history of an evolving idea. Neuroimage 37: 1083-1090.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001): A default mode of brain function. Proc Natl Acad Sci USA 98: 676-682.

Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S (2004): The role of the medial frontal cortex in cognitive control. Science 306: 443-447.

Roberts BA, Martel MM, Nigg JT (2013): Are There Executive Dysfunction Subtypes Within ADHD? Journal of attention disorders: 1087054713510349.

Rubia K, Overmeyer S, Taylor E, Brammer M, Williams SC, Simmons A, Bullmore ET (1999): Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: a study with functional MRI. Am J Psychiatry 156: 891-896.

Rubia K, Smith AB, Brammer MJ, Toone B, Taylor E (2005): Abnormal brain activation during inhibition and error detection in medication-naive adolescents with ADHD. Am J Psychiatry 162: 1067-1075.

Samanez-Larkin GR, D'Esposito M (2008): Group comparisons: imaging the aging brain. Soc Cogn Affect Neurosci 3: 290-297.

Sava S, Yurgelun-Todd DA (2008): Functional magnetic resonance in psychiatry. Top Magn Reson Imaging 19: 71-79.

Scheffler RM, Brown TT, Fulton BD, Hinshaw SP, Levine P, Stone S (2009): Positive association between attention-deficit/hyperactivity disorder medication use and academic achievement during elementary school. Pediatrics 123: 1273-1279.

Sciutto MJ, Eisenberg M (2007): Evaluating the evidence for and against the overdiagnosis of ADHD. Journal of Attention Disorders 11: 106-113.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD (2007): Dissociable intrinsic connectivity networks for salience processing and executive control. The Journal of neuroscience 27: 2349-2356.

Sibley MH, Evans SW, Serpell ZN (2010): Social cognition and interpersonal impairment in young adolescents with ADHD. Journal of Psychopathology and Behavioral Assessment 32: 193-202.

Simon V, Czobor P, Balint S, Meszaros A, Bitter I (2009): Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. Br J Psychiatry 194: 204-211.

Smith AB, Taylor E, Brammer M, Toone B, Rubia K (2006): Task-specific hypoactivation in prefrontal and temporoparietal brain regions during motor inhibition and task switching in medication-naive children and adolescents with attention deficit hyperactivity disorder. Am J Psychiatry 163: 1044-1051.

Song XW, Dong ZY, Long XY, Li SF, Zuo XN, Zhu CZ, He Y, Yan CG, Zang YF (2011): REST: a toolkit for resting-state functional magnetic resonance imaging data processing. PLoS One 6: e25031.

Storer JL, Evans SW, Langberg JM (2014): Organization Interventions for Children and Adolescents with Attention-Deficit/Hyperactivity Disorder (ADHD). In: Handbook of School Mental Health. Springer. p 385-398.

Suskauer SJ, Simmonds DJ, Fotedar S, Blankner JG, Pekar JJ, Denckla MB, Mostofsky SH (2008): Functional magnetic resonance imaging evidence for abnormalities in response selection in attention deficit hyperactivity disorder: differences in activation associated with response inhibition but not habitual motor response. J Cogn Neurosci 20: 478-493.

Tamm L, Menon V, Reiss A (2006): Parietal attentional system aberrations during target detection in adolescents with attention deficit hyperactivity disorder: event-related fMRI evidence. Am J Psychiatry 163: 1033-1043.

Tamm L, Menon V, Ringel J, Reiss AL (2004): Event-related FMRI evidence of frontotemporal involvement in aberrant response inhibition and task switching in

attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 43: 1430-1440.

Thaler NS, Bello DT, Etcoff LM (2013): WISC-IV profiles are associated with differences in symptomatology and outcome in children with ADHD. Journal of Attention Disorders 17: 291-301.

The MTA Cooperative Group (1999): A 14-Month Randomized Clinical Trial of Treatment Strategies for Attention-Deficit/Hyperactivity Disorder. Arch Gen Psychiatry 56: 1073-1086.

Thomason ME, Dennis EL, Joshi AA, Joshi SH, Dinov ID, Chang C, Henry ML, Johnson RF, Thompson PM, Toga AW, Glover GH, Van Horn JD, Gotlib IH (2011): Resting-state fMRI can reliably map neural networks in children. Neuroimage 55: 165-175.

Thulborn KR, Waterton JC, Matthews PM, Radda GK (1982): Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field. Biochimica et Biophysica Acta (BBA)-General Subjects 714: 265-270.

Tonhajzerova I, Ondrejka I, Adamik P, Hruby R, Javorka M, Trunkvalterova Z, Mokra D, Javorka K (2009): Changes in the cardiac autonomic regulation in children with attention deficit hyperactivity disorder (ADHD). Indian J Med Res 130: 44.

Uddin LQ, Kelly AM, Biswal BB, Margulies DS, Shehzad Z, Shaw D, Ghaffari M, Rotrosen J, Adler LA, Castellanos FX, Milham MP (2008): Network homogeneity reveals decreased integrity of default-mode network in ADHD. J Neurosci Methods 169: 249-254.

Vaidya CJ, Bunge SA, Dudukovic NM, Zalecki CA, Elliott GR, Gabrieli JD (2005): Altered neural substrates of cognitive control in childhood ADHD: evidence from functional magnetic resonance imaging. Am J Psychiatry 162: 1605-1613.

Valera EM, Spencer R, Zeffiro TA, Makris N, Spencer TJ, Faraone SV, Biederman J, Seidman LJ (2010): Neural substrates of impaired sensorimotor timing in adult attention-deficit/hyperactivity disorder. Biol Psychiatry 68: 359-367.

van den Heuvel MP, Hulshoff Pol HE (2010): Exploring the brain network: a review on resting-state fMRI functional connectivity. Eur Neuropsychopharmacol 20: 519-534.

Van Mourik R, Oosterlaan J, Sergeant JA (2005): The Stroop revisited: A meta-analysis of interference control in AD/HD. Journal of Child Psychology and Psychiatry 46: 150-165.

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008): Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. J Neurophysiol 100: 3328-3342.

Visser SN, Danielson ML, Bitsko RH, Holbrook JR, Kogan MD, Ghandour RM, Perou R, Blumberg SJ (2014): Trends in the Parent-Report of Health Care Provider-Diagnosed and Medicated Attention-Deficit/Hyperactivity Disorder: United States, 2003–2011. Journal of the American Academy of Child & Adolescent Psychiatry 53: 34-46. e2.

Wang L, Zhu C, He Y, Zang Y, Cao Q, Zhang H, Zhong Q, Wang Y (2009): Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. Hum Brain Mapp 30: 638-649.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF (2005): Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. Biol Psychiatry 57: 1336-1346.

Youden W (1950): Index for rating diagnostic tests. Cancer 3: 32-35.

Zhu CZ, Zang YF, Cao QJ, Yan CG, He Y, Jiang TZ, Sui MQ, Wang YF (2008): Fisher discriminative analysis of resting-state brain function for attention-deficit/hyperactivity disorder. Neuroimage 40: 110-120.

Zuo X, Kelly C, Adelstein JS, Klein DF, Castellanos FX, Milham MP (2010): Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage 49: 2163-2177.