No. of words abstract: 248; No. of words main text: 5007
No. of figures: 3; No. of tables: 2
Supplemental information: 1

Multivariate Pattern Analysis (MVPA) Reveals Distinct Neural Profiles of Frontoparietal Networks in Boys with Attention-Deficit/Hyperactivity Disorder and Boys with Persistent Depressive Disorder

Veronika Vilgis[1]*, Debbie Yee[2,3]*, Tim J. Silk,[1,4,5] and Alasdair Vance[1,5,6]

[1]Department of Paediatrics, University of Melbourne, Australia

[2]Washington University in St. Louis, United States

[3]Brown University, United States

[4]Murdoch Children's Research Institute, Australia

[5]School of Psychology, Deakin University, Australia

[6]Royal Children's Hospital, Parkville, Australia

*these authors contributed equally to this work

Correspondence to: Debbie Yee, Cognitive, Linguistic & Psychological Sciences, Brown University, Box 182, Metcalf Research Building, 190 Thayer Street, Providence, RI 02912
Phone: +15166033958
Email:debbie_yee@brown.edu

Short/running title (50 characters): MVPA of Working Memory in ADHD and Depression

Six Keywords: fMRI; MVPA: working memory; ADHD; depression; children

Abstract

Working memory deficits are common in attention-deficit/hyperactivity disorder (ADHD) and depression, two common neurodevelopmental disorders with overlapping cognitive profiles but distinct clinical presentation. Multivariate techniques have previously been utilized to understand working memory processes in functional brain networks in healthy adults, but have not yet been applied to investigate how working memory processes within the same networks differ within typical and atypical developing populations. We used multivariate pattern analysis (MVPA) to identify whether brain networks discriminated between spatial vs. verbal working memory processes in ADHD and Persistent Depressive Disorder (PDD). 36 male clinical participants and 19 typically developing (TD) boys participated in a fMRI scan while completing a verbal and a spatial working memory task. Within *a priori* functional brain networks (frontoparietal, default mode, salience) the TD group demonstrated differential response patterns to verbal and spatial working memory. Both clinical groups show less differentiation than TD, with neural profiles suggesting ADHD is associated with weaker differentiation in both frontoparietal and salience networks and PDD is associated with weaker differentiation in left frontoparietal and default mode networks. Whereas the TD group's neural profile indicates network response patterns that are sensitive to task demands, the neural profiles of the ADHD and PDD group suggest less specificity in neural representations of spatial and verbal working memory. We highlight within-group classification as innovative tool for understanding the neural mechanisms of how cognitive processes may deviate in clinical disorders, an important intermediary step towards improving translational psychiatry to inform clinical diagnoses and treatment.

Introduction

Half of all mental illnesses begin prior to the age of 15 making them the leading cause of disability in young people worldwide (Kessler et al., 2007). Among the most common diagnosed neuropsychiatric conditions are attention-deficit/hyperactivity disorder (ADHD) and depressive disorders. ADHD and depression present distinctly in everyday life, yet research suggests some commonalities, as both disorders have been associated with impairments in attention (Brooks, Iverson, Sherman, & Roberge, 2010; Cataldo, Nobile, Lorusso, Battaglia, & Molteni, 2005; Günther, Konrad, De Brito, Herpertz-Dahlmann, & Vloet, 2011; Wilding, 2005), executive function (Franklin et al., 2010; Martel, Nikolas, & Nigg, 2007; Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005; Matthews, Coghill, & Rhodes, 2008) and emotion regulation (Beauregard, Paquette, & Levesque, 2006; Kovacs, Joormann, & Gotlib, 2008; Shaw, Stringaris, Nigg, & Leibenluft, 2014). Additionally, neuroimaging studies have shown similar structural and functional anomalies in frontal, striatal, limbic, parietal and cerebellar brain regions in both disorders (Arnsten & Rubia, 2012; Cortese et al., 2012; Miller, Hamilton, Sacchet, & Gotlib, 2015). Although evidence suggests that ADHD and depression potentially share some common cognitive and neural deficits, the extent to which such deficits are similar or distinct across both disorders remains unclear. Insight into how specific impairments (e.g., working memory) differ between ADHD and depression would improve characterization of key cognitive and neural processes disrupted and refine delineation in their underlying neurobiology.

One promising delineation between the working memory deficits often associated with ADHD and depression may be the specific type of working memory impaired. Children and adolescents with ADHD consistently demonstrate working memory deficits and tend to have greater impairments in processing spatial stimuli compared to verbal stimuli in these tasks

(Martinussen et al., 2005; Sowerby, Seal, & Tripp, 2011; Willcutt, Doyle, Nigg, Faraone, & Pennington, 2005). Conversely, although adults with depressive disorders also demonstrate working memory deficits (Gohier et al., 2009; Rose, Simonotto, & Ebmeier, 2006), they exhibit more persistent impairments in the verbal domain even after remission and independent of age (Gruber, Zilles, Kennel, Gruber, & Falkai, 2011; Shilyansky et al., 2016). However, in contrast to the wealth of research implicating executive function deficits in children and adolescents with ADHD (Martel et al., 2007; Willcutt et al., 2005), much less is understood about the nature of executive function deficits in children and adolescents with depressive disorders (Vilgis, Silk, & Vance, 2015). Nevertheless, the few extant studies have observed a deficit in working memory more often (Franklin et al., 2010; Günther, Holtkamp, Jolles, Herpertz-Dahlmann, & Konrad, 2004; Klimkeit, Mattingley, Sheppard, Lee, & Bradshaw, 2005; Matthews et al., 2008; Vance & Winther, 2020) than not (Korhonen et al., 2002; Maalouf et al., 2011), though the extent to which verbal working memory deficits observed in adults also generalize to children and adolescents is unknown.

Both spatial and verbal working memory tasks consistently engage distinct neural representations in the frontoparietal network (FPN) (Daniel, Katz, & Robinson, 2016; Rottschy et al., 2012). In typically developing children and adults, the FPN exhibits hemispheric asymmetry, with greater right FPN activation for spatial working memory and left FPN activation for verbal working memory (Thomason et al., 2009). Children and adults with ADHD typically demonstrate less activation in the FPN during working memory tasks (Cortese et al., 2012), with hypoactivation over the right hemisphere compared to controls when processing visuospatial stimuli (e.g., Silk et al., 2005; Vance et al., 2007) and bilateral prefrontal hypoactivation during a verbal task (Cubillo et al., 2014). Conversely, several studies in adults with depression have observed atypical

activation of the left prefrontal cortex during working memory (Barch, Sheline, Csernansky, & Snyder, 2003; Vasic, Wolf, & Walter, 2007; Walter, Wolf, Spitzer, & Vasic, 2007), a pattern also observed in a pediatric sample of boys with Persistent Depressive Disorder (PDD) during a spatial task (Vilgis, Chen, Silk, Cunnington, & Vance, 2014). Because disorder-specific hemispheric asymmetries exist independent of the type of stimulus material being processed, it is difficult to determine the extent to which atypical activation in ADHD and depression arises from dysregulated processing of a specific type of working memory (e.g., spatial vs. verbal) versus a more general neurological deficit associated with either disorder. Therefore, in order to develop greater understanding into how the neural patterns that underlie spatial and verbal working memory are altered in this network, it is necessary to perform a systematic comparison of how left and right FPN are disrupted in both working memory tasks in both ADHD and depression.

An important additional consideration is that networks that interact with the FPN may contribute to working memory deficits or play a compensatory role in the two disorders. In particular, the FPN interacts with the salience and default-mode network to support key working memory processes (Liang, Zou, He, & Yang, 2015). The salience network (SN) plays a central role via continuously monitoring for endogenous and exogenous salient stimuli in order to facilitate the dynamic switching between internally directed, self-focused cognition carried out by the default-mode network (DMN) and externally directed cognition supported by the FPN (Menon & Uddin, 2010; Seeley, 2019). In children with ADHD, the absence of explicit task demands has been associated with weaker and more variable interconnectivity between FPN, SN, and DMN, including shorter and less persistent brain states (Cai, Chen, Szegletes, Supekar, & Menon, 2018). Similarly, in patients with depression, symptoms of rumination, emotional over-reactivity, and cognitive regulation of affective information have been associated with altered functional

connectivity in the DMN, SN and FPN respectively (Hamilton, Chen, & Gotlib, 2013; Manoliu et al., 2014). This compelling evidence suggests that the locus of working memory impairments in ADHD and depression likely arises from dysregulation of these core networks and their interactions. Yet, how neural patterns within each network support different types of working memory, and more critically, how these patterns putatively deviate in ADHD and depression, remains an enigma.

Thus, an outstanding question relates to understanding the precise mechanisms driving network-level differences in spatial and verbal working memory in ADHD and depression. Here, we present an innovative approach that combines network-based and multivariate analyses to investigate how neural patterns of spatial vs. verbal working memory putatively differ within clinical groups. Network-based approaches are uniquely powerful, as they provide a systematic approach for investigating neural systems underpinning cognitive function (Cole, Bassett, Power, Braver, & Petersen, 2014). Multivariate pattern analysis (MVPA) is a powerful machine learning method utilized to classify brain voxels to predict whether they carry information regarding distinct categories, tasks, or other cognitive states (Haxby, Connolly, & Guntupalli, 2014; Mumford, Turner, Ashby, & Poldrack, 2012; Yang, Fang, & Weng, 2012). Importantly, MVPA has been used to successfully dissociate between neural representations of cognitive tasks in prefrontal and parietal cortices (Etzel, Cole, Zacks, Kay, & Braver, 2016; Wisniewski, Reverberi, Momennejad, Kahnt, & Haynes, 2015; Woolgar, Afshar, Williams, & Rich, 2015). This precedent for utilizing MVPA to decode neural task representations strongly motivates the adaptation of this cutting-edge technique for examining networks that support working memory processes (e.g., FPN, SN, DMN).

Whereas MVPA has been previously adopted to investigate cognitive processes in typical development (Chow, Westphal, & Rissman, 2018), such approaches have not yet been applied to

understand how they diverge within clinical populations. In particular, MVPA may serve as a powerful tool that can systematically evaluate how distinctness of neural patterns underpinning different cognitive processes within the same brain networks may differ across typical vs atypical populations. Moreover, this data-driven method is aligned with computational psychiatry, an emerging field that promotes computationally rigorous approaches to improve the understanding, prediction, and treatment of mental disorders (Dwyer, Falkai, & Koutsouleris, 2018; Wiecki, Poland, & Frank, 2015). Thus, combining MVPA with a network-based approach for clinical neuroimaging data provides a promising novel direction for improving mechanistic understanding of the distinctness of neural representations underlying spatial and verbal working memory tasks in ADHD and depression.

In the current study, we utilize MVPA to decode neural patterns underpinning spatial vs. verbal working memory in *a priori* brain networks (Hebart & Baker, 2018) within three different groups: typical development (TD), ADHD and Persistent Depressive Disorder (PDD). We focus on PDD as a more chronic developmental disorder more similar to ADHD in its stability over time (versus major depressive disorder characterized by relatively shorter episodes of negative mood). An innovative feature of our within-group classification analysis entails neural decoding of spatial vs. verbal working memory processes in the same network. Given the known issues with how fMRI decoding results differ based upon the number of voxels in regions of interest (ROI) (Haynes, 2015), our approach of utilizing the same ROI for each decoding facilitates clear comparison of how neural representations of working memory processes are uniquely characterized within each clinical group. We predict distinct neural patterns in left and right FPN for spatial vs. verbal working memory processes in typical development. Moreover, we expect a double dissociation with an atypical pattern for the right but not left FPN in the ADHD group, and an atypical pattern

for the left but not right FPN in the PDD group. Additionally, we examine neural patterns in the SN and DMN, two prominent networks known to interact with the FPN and previously identified as functioning atypically in ADHD and depression (Menon, 2011). Our objective was to identify shared and unique network response patterns for each developmental disorder, which may inform disorder-specific neural profiles underpinning well-known behavioral deficits in spatial and verbal working memory.

Methods and Materials[1]

Participants

39 male children and adolescents were recruited through an outpatient pediatric psychiatry clinic, met inclusion criteria for the study, and agreed to participate in the MRI scan. Of those, 17 fulfilled DSM-IV (American Psychiatric Association, 2000) diagnostic criteria for PDD (labelled as Dysthymic Disorder in DSM-IV) and 22 for a diagnosis of ADHD - Combined Type (i.e. at least six symptoms of inattentive and hyperactive/impulsive symptoms were met). All young people and their caregivers were interviewed using the Anxiety Disorders Interview Schedule for Children (Silverman & Albano, 1996). Caregivers also completed the Conners' Parent Rating Scale - long version (Conners, Sitarenios, Parker, & Epstein, 1998) and the Child Behavior Checklist (CBCL) (Achenbach, 1991) and the young person completed the Children's Depression Inventory (Kovacs, 1992). An experienced child and adolescent psychiatrist (AV) confirmed each diagnosis. Participants were included if they were male, right-handed, free of metal implants (MRI compatibility) and with a full-scale IQ above 70. Exclusion criteria were the presence of an intellectual disability, learning disorder or known neurological or endocrine condition. Participants were also excluded if they had a previous diagnosis of an Autistic Spectrum Disorder, Bipolar

[1] All task scripts and analysis scripts are uploaded on our OSF repository: https://osf.io/a5349/

Disorder or Psychotic Disorder. All clinical participants in the PDD and ADHD group had that disorder as primary diagnosis, but comorbidities of Oppositional Defiant Disorder, Conduct Disorder and anxiety disorders were permitted. Participants with ADHD (n=6) who were treated pharmacologically at time of testing withheld their stimulant medication at least 48h before participating in the MRI scan. Only one participant with PDD had been started on antidepressant medication prior to scanning. After excluding two ADHD participants due to excessive head motion and one PDD participant due to task performance, the final sample comprised 16 boys with PDD and 20 boys with ADHD. 19 typically developing participants (TD) and their caregivers were recruited through local schools. They completed the same questionnaires and semistructured clinical interviews as the clinical participants to ascertain normal psychological functioning. None of the TD met diagnostic criteria for any psychiatric diagnosis. An abbreviated IQ test was conducted with two subtests of the Verbal Comprehension Index (Similarities and Vocabulary) and two subtests of the Perceptual Reasoning Index (Block Design and Matrix Reasoning) (Wechsler, 2003) and scores were calculated as previously described (Crawford, Anderson, Rankin, & MacDonald, 2010). All procedures were approved by the local Human Research Ethics Committee. All parents/guardians and young people gave informed written consent to participate in the study.

fMRI Data Acquisition

Images were acquired on a 3T Siemens TIM Trio scanner (Siemens, Erlangen Germany) at the Royal Children's Hospital, Melbourne. Participants lay supine with their head supported in a 32-channel head coil. High-resolution T1-weighted structural MRI images were acquired for each participant (TR=1900 ms, TE=2.24 s, FA= 90°, in-plane pixels=0.9 x 0.9 mm) prior to the

functional scan. For functional imaging, T2*-weighted gradient-echo echo-planar images (EPI) were acquired (TR=2700 ms, TE=40 ms, FA=90°, 39 axial slices (co-planar with AC-PC) with 3.0 mm isotropic resolution). A total of 158 image volumes were acquired per 7 min 10 s sequence. This was run twice, one for each working memory task. Participants' responses were recorded with a scanner compatible two button-box (Fibre-optic response pads, Current Designs, Philadelphia, PA).

fMRI Task

A Sternberg type working memory task divided into verbal and spatial working memory components was adapted from Thomason et al. (2009). In the verbal task, participants viewed an array of capital letters, two in the low load condition and six in the high load condition. In the spatial task, either one or five black dots were presented randomly. After a 3000 ms delay, a single letter in lower case (verbal task) or a circle (spatial task) was presented for 1500 ms while participants indicated by button-press whether the probe matched the identity (verbal task) or location (spatial task) of the initial cues. Both tasks included control conditions that matched experimental conditions as much as possible (motor response, decision-making, visual stimuli, luminance, total trial length), except for the 3000 ms delay which was shortened to 100 ms. Probes matched the target 50% of the time. For each task, a total of 64 trials (32 experimental, 32 control) were presented in 16 pseudorandomly alternating blocks of four trials. Here, we combined high and low load conditions to optimize the beta estimates used in the MVPA analysis. A more detailed task description is provided in the Supplementary Material and a schematic representation is shown in Figures 1a and 1b.

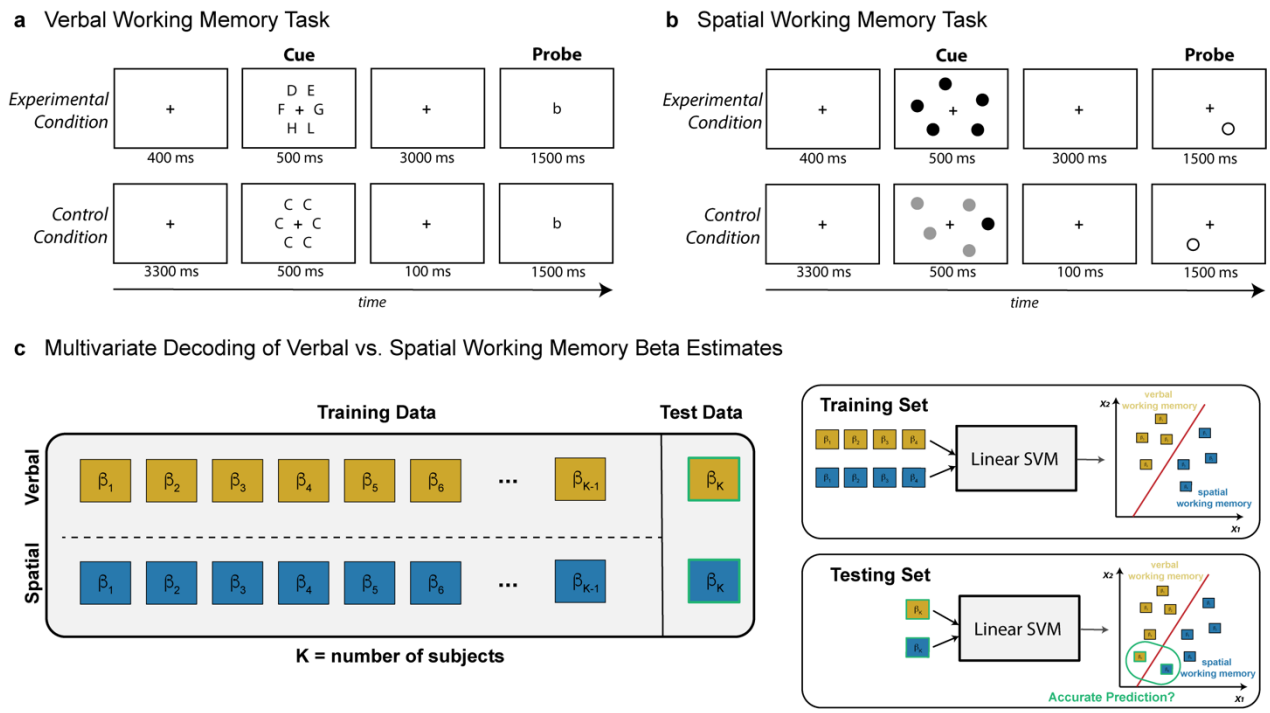

Figure 1 a) Schematic of the Verbal Working Memory Task. Participants were cued with an array of uppercase capital letters. After 3000 ms delay, a single lowercase letter was presented for 1500 ms while participants indicated by button-press whether the probe matched the identity of the initial cues. The control condition matched the experimental condition in as many elements as possible (motor response, decision-making, visual stimuli, luminance, total trial length), except for the 3000 ms delay which was shortened to 100 ms. **b)** Schematic of the Spatial Working Memory Task. This task was similar to the verbal task, except the cue was a set of black dots, and participants were tasked with indicating whether the circle probe matched the location of the initial task cues. **c)** Multivariate Decoding of Spatial vs. Verbal Working Memory Beta Estimates. A linear support vector machine (SVM) classifier with K-fold cross validation was implemented for each network within each group, with K equal to the number of subjects in each group. The classifier was trained on beta estimates for spatial and verbal working memory from K-1 subjects and tested on the remaining subject in that group. The process was repeated for the control conditions. Permutation tests (nsim=1000) evaluated whether observed classification accuracies for each network were statistically significant above chance.

fMRI Preprocessing

Subject data was preprocessed using AFNI (https://afni.nimh.nih.gov/) (Cox, 1996). The EPI volumes were realigned, slice-timing corrected, registered to an MNI152 atlas, and normalized. EPI volumes were smoothed using a 4 mm FWHM Gaussian kernel to improve signal to noise ratio, which has been demonstrated in previous work (Gardumi et al., 2016; Jackson & Woolgar, 2018; Woolgar et al., 2015). A general linear model was applied to each subject to extract block-level beta estimates for each of the four block types: spatial experimental, spatial control,

verbal experimental and verbal control conditions. The design matrix was estimated using AFNI's 3dDeconvolve function, and a generalized least squares time series fit with REML estimation of the temporal auto-correlation structure (3dREMLfit) was applied to estimate individual betas for the four regressors per subject. Each block lasted 21.6 seconds (8 TRs), and each run contained 16 blocks. These extracted beta estimates were fed into the linear Support Vector Machine (SVM) to classify spatial vs. verbal tasks for each of the 5 networks. Six motion parameters were included as GLM nuisance regressors. TRs (current and previous) were censored if the derivative values had a Euclidean norm above 3.5 mm.

To ensure the working memory tasks reliably activated frontoparietal regions in each clinical group, general linear tests were applied to the estimated beta weights to calculate task activation for spatial and verbal working memory independently for each participant. Within each group, AFNI's 3dttest++ function was applied to generate statistical maps for the average of the spatial and working memory conditions. Significant clusters for each task condition and group (thresholded at $p<.001$, cluster corrected with threshold $\alpha = .01$) are shown in Supplemental Figure S1.

ROI Masks

The ROI masks were selected from https://www.fmrib.ox.ac.uk/datasets/brainmap+rsns/ and have been described previously (Smith et al., 2009). We focused on the left and right FPN, key networks supporting working memory related activity and in accordance with our predictions that classification accuracy would differ between the groups given the previous literature. In addition, we included the DMN and the SN, as these networks dynamically interact with the FPN to support working memory. We also included the sensorimotor (SM) network as control network,

where we do not expect significant differences between tasks or groups. All brainmaps were resampled to match the dimensions of the data (3x3x3 mm isotropic), thresholded at 3.5, and binarized to create masks for each of the networks. The ROIs are shown in Figure 3a.

Multivariate Pattern Analysis (MVPA)

Multivariate pattern analysis (MVPA) was performed on each of the 5 networks (left and right FPN, SN, DMN, SM) using the Decoding Toolbox (Hebart, Gorgen, & Haynes, 2014), and custom MATLAB code. All classifications implemented a linear SVM classifier with a fixed cost parameter (c=1). K-fold cross-validation was performed for each group using a leave-one-run-out protocol, with K equal to the number of subjects in each group. Here, the classifier was trained on data from K-1 subjects (per group) and tested on the left out remaining subject in that group. The beta-estimates were used to train and test our classifier (Pereira, Mitchell, & Botvinick, 2009) on spatial vs. verbal working memory (See Figure 1c). Permutation tests (nsim=1000) evaluated whether the observed classification accuracies for each network were statistically significant above chance (Hebart et al., 2014). The procedure was repeated for the spatial and verbal control conditions.

Results

Clinical and Behavioral Results

We first characterized how clinical symptoms differed at the group level, which would inform subsequent within-group differences. Both clinical groups exhibited greater levels of externalizing and internalizing symptoms compared to the TD group and lower scores on IQ measures. As expected, both clinical groups differed from each other in scores of internalizing symptoms (PDD > ADHD) and related global ADHD and restless symptoms (ADHD > PDD). All clinical measures are listed in Table 1.

Table 1: Means (M) and standard deviations (SD) of demographic and clinical variables for each group.

	ADHD (n=20) *M (SD)*	**PDD (n=16)** *M (SD)*	**TD (n=19)** *M (SD)*	***F* (2,52)***	***p***	**Follow-up t-test (*p* Bonf. corrected)**
Age	12.77 (19.6)	12.94 (2.05)	13.04 (2.4)	0.0799	n.s.	
CDI Sum (self-report)	10.00 (6.59)	14.88 (10.37)	4.16 (2.89)	10.1 (*F* (2,51))	< 0.001	ADHD > TD (.04) DD > TD (< .001)
CBCL Internalizing (p)	62.35 (9.70)	72.5 (7.39)	45.68 (10.51)	36.726	< .001	DD > ADHD (.007) ADHD > TD (< .001) DD > TD (< .001)
CBCL Externalizing (p)	67.50 (11.38)	63.31 (7.32)	45.11 (7.69)	32.45	< .001	ADHD > TD (< .001) DD > TD (< .001)
CPRS Global Index (p)	75.37 (11.48)	65.6 (10.46)	47.26 (5.41)	43.465 (*F* (2,50))	< .001	ADHD > DD (.012) ADHD > TD (< .001) DD > TD (< .001)
CPRS Restless (p)	72.63 (14.83)	63.53 (8.75)	48.11 (5.87)	25.763 (*F* (2,50))	< .001	ADHD > DD (.05) ADHD > TD(< .001) DD > TD (< .001)
Verbal Comprehension Index	95.79 (10.20)	95.6 (10.11)	111.06 (8.34)	15.147 (*F* (2,49))	< .001	TD > ADHD (< .001) TD > DD(< .001)
Perceptual Reasoning Index	89.74 (8.68)	94.4 (13.02)	108.78 (12.17)	14.021 (*F* (2,49))	< .001	TD > ADHD (< .001) TD > DD (.002)

CDI = Child Depression Inventory, CBCL = Child Behavior Checklist; CPRS = Conners Parent Rating Scale; (p) = parent-rated
*adjusted degrees of freedom reported for measures that contained missing data

With regard to the behavioral performance on the two working memory tasks, within each group, reaction times (RT) and accuracy did not differ between the verbal and spatial task

components (Accuracy: PDD $t(15)=0.37$, $p=0.717$; ADHD $t(19)=0.61$, $p=0.547$; TD $t(17)=-0.31$, $p=0.758$; RT: PDD $t(15)=-1.21$, $p=0.246$; ADHD $t(19)=-0.97$, $p=0.342$; TD $t(18)=-1.27$, $p=0.222$). In the control conditions, behavioral performance was similar between both tasks, though there was a trend towards faster reaction times in the verbal than spatial control condition in the PDD group (Accuracy: PDD $t(15)=0.92$, $p=0.370$; ADHD $t(19)=-0.50$, $p=0.624$; TD $t(17)=0.89$, $p= 0.385$; RT: PDD $t(15)=-2.20$, $p=0.044$*, ADHD $t(19)= -1.43$; $p=0.169$; TD $t(18)=-1.62$, $p=0.123$). Task performance measures are illustrated in Figure 3 and provided in Table S1 in the Supplementary Material. Additionally, motion did not differ between verbal and spatial working memory conditions nor control conditions (See Table S2 in Supplementary Material).

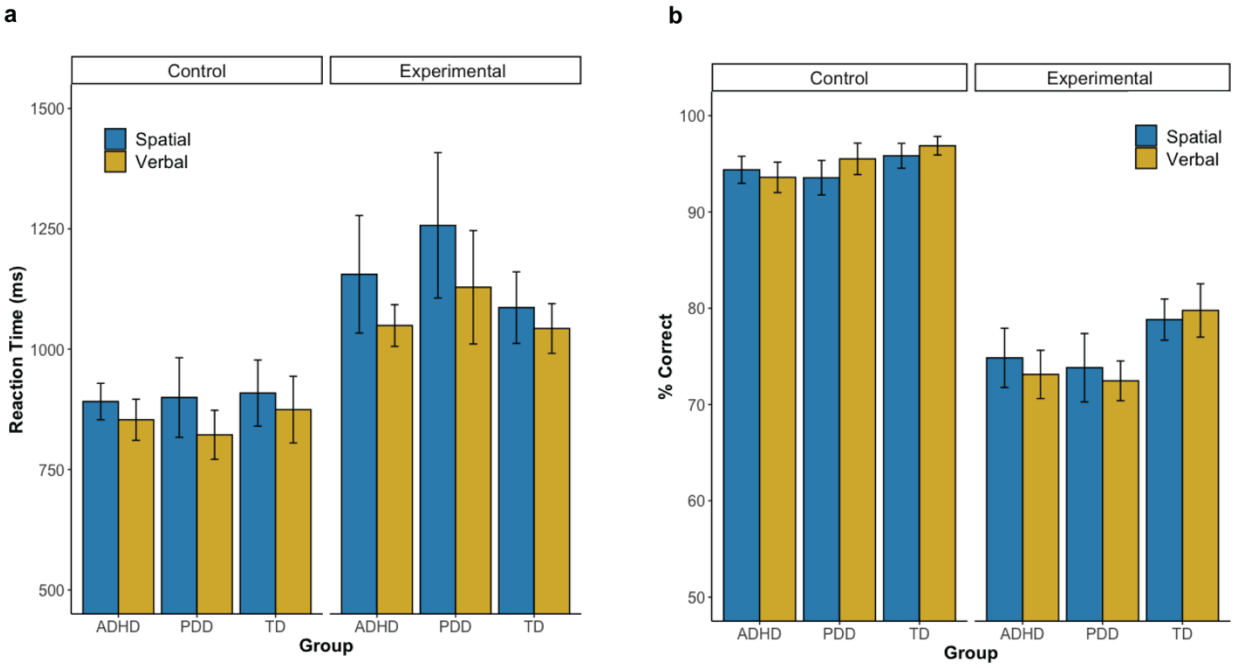

Figure 2: a) Reaction Time by Experimental and Control Condition. Participants were overall slower when performing the experimental conditions relative to the control conditions. Although RT performance was similar between verbal and spatial memory tasks, there was a slight trend towards faster RT in the verbal condition in the PDD group. For both plots, error bars indicate standard error. **b) Accuracy by Experimental and Control Condition.** Participants were overall less accurate when performing the experimental conditions relative to the control conditions. For each group, the accuracy did not differ between verbal and spatial memory task components.

Multivariate Decoding Results (MVPA)

We applied a linear SVM to *a priori* brain networks and identified disorder-specific classification accuracies that indicate the distinctness of spatial vs. verbal working memory task representations within each network. These networks are shown in Figure 3a. Table 2 presents the classification accuracy minus chance (chance = 50%) from the decoding analyses for each ROI. Our primary focus was on the left and right FPN within each group. For both networks, we expected to observe significant classification accuracies in the TD group indicative of differential neural representations of verbal and spatial working memory. Indeed, the TD group showed significantly above chance accuracies within both the left (p=0.001) and right (p=0.001) FPN. For the ADHD group, a trend towards significant differentiation was observed in the left (p = 0.059) FPN but classification accuracy in the right (p=0.129) FPN was not significant. For the PDD group, classification accuracy was not significant in the left (p=0.307) FPN, but was significantly high in the right (p=0.020) FPN.

Additionally, we included the SN and DMN as networks of interest due to their known interactions with the FPN and prior work indicating aberrant within and across network functions in both ADHD and PDD. For the TD group, both the SN (p=0.036) and DMN (p=0.007) exhibited significant classification accuracies. For the PDD group, only the SN (p=0.020) but not the DMN (p=0.089) exhibited significant classification accuracy. Conversely, for the ADHD group, classification accuracy was significant in the DMN (p=0.010) but not in the SN (p=0.317).

We also examined the classification accuracies of the spatial vs. verbal control conditions (trials that matched the experimental working memory trials in all aspects with the exception of the delay period, thus removing the memory component of the task). None of the classification accuracies for these control conditions were significant in the TD group, suggesting that spatial

and verbal information does not *per se* elicit differential processing in the right and left FPN, DMN and SN. In the PDD group, we observed the same pattern as in the TD group. For the ADHD group, the DMN (p=0.010) showed significant differential processing of verbal and spatial material in the control condition.

Lastly, we examined the SM as a control network, and predicted that classification accuracies would be similar in all three groups. As expected, in all three groups classification accuracy was significantly high for the working memory trials. For the control trials, we did not observe significant differentiation for TD and ADHD groups, but observed significant classification accuracy for the control condition in the PDD group (p=0.030). Figure 3b displays the classification accuracy minus chance across ROIs, delineated for each group.

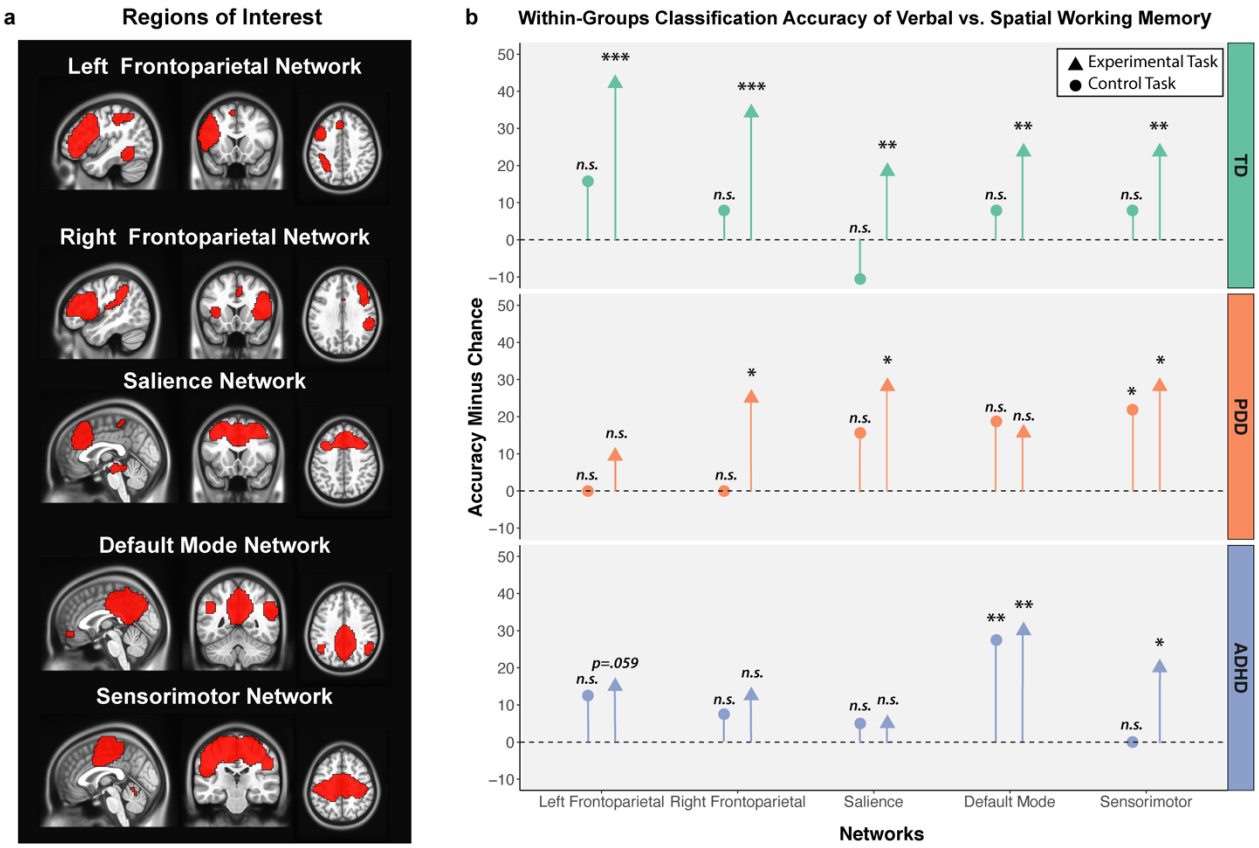

Figure 3 a) Regions of Interest. Left and Right Frontoparietal Network (LFPN and RFPN) forming as the primary ROIs for working memory function, followed by the Salience Network (SN) and Default Mode Network (DMN)

known to interact with the FPN. The sensorimotor network (SM) served as control ROI. b) Within-Groups Classification Accuracy of Spatial vs. Verbal Working Memory for the five selected networks. Here, the classification accuracy indicates the distinctness of spatial vs. verbal working memory task representations within each ROI. Importantly, utilization of the same ROI for each decoding facilitates clear comparison of how neural representations of working memory tasks may differ across clinical groups. High classification accuracies are observed across all networks but with group specific patterns. The TD group shows distinguishable patterns of verbal and spatial working memory processes in all networks. Classification accuracies are significantly high for the right FPN, salience and sensorimotor networks but not the left FPN or DMN in the PDD group. For the ADHD group we observed significant classification accuracies only in the DMN with a trend for the left FPN. Statistical significance for one-tailed permutation tests: *p<.05; **p<.01; ***p<.001.

Table 2: Within Groups Classification Accuracy Minus Chance by Network, Condition and Group.

	TD		**ADHD**		**PDD**	
	Exp	**Control**	**Exp**	**Control**	**Exp**	**Control**
Left Frontoparietal	42.11*** (p=0.001)	15.79 (p=0.066)	15.00 (p=0.059)	12.50 (p=0.139)	9.38 (p=0.307)	0.00 (p=0.574)
Right Frontoparietal	34.21*** (p=0.001)	7.89 (p=0.259)	12.50 (p=0.129)	7.50 (p=0.238)	25.00* (p=0.020)	0.00 (p=0.554)
Salience	18.42* (p=0.036)	-10.53 (p=0.882)	5.00 (p=0.317)	5.00 (p=0.337)	28.13* (p=0.020)	15.63 (p=0.168)
Default Mode	23.68** (p=0.007)	7.89 (p=0.258)	30.00* (p=0.010)	27.50* (p=0.010)	15.63 (p=0.089)	18.75 (p=0.069)
Sensorimotor	23.68** (p=0.003)	7.89 (p=0.245)	20.00* (p=0.020)	0.00 (p=0.535)	28.13* (p=0.020)	21.88* (p=0.030)

**p<.05; **p<.01; ***p<.001*

Whole Brain Univariate Results for Spatial and Verbal Working Memory Tasks

One possible explanation for the lower classification accuracy in the ADHD and PDD group may arise from reduced activation in spatial and verbal working memory tasks in this groups. That is, lower classification accuracy in these groups may be due to insufficient activation during each task condition. To rule out the possibility that lower decoding arose from lower task-relevant activation, we averaged whole-brain activations for spatial and verbal working memory experimental conditions for each group (TD, ADHD, PDD). Our results revealed that all groups show expected activation in left and right FPN and SN, as well as expected deactivation the DMN. Importantly, similar activation of these networks across all groups provide evidence that these key working memory networks are involved in both spatial and verbal tasks. Averaged beta estimations

for each group are shown in Figure S1, and MNI coordinates of peak activations for all conditions are shown in Table S3 in the Supplemental Material.

Discussion

In this study, we identified distinct neural profiles for spatial versus verbal working memory processes in frontoparietal, salience, and default mode networks in boys with ADHD, boys with PDD, and typically developing male youth. We used a novel approach that combined MVPA and network-based analyses to compare multivariate neural representations of verbal and spatial working memory within each group. Our approach revealed clear distinguishable neural patterns in the TD group in *a priori* networks not observed in either clinical group. Consistent with our predictions and previous literature, the PDD group differentially processed verbal and spatial material in right but not left FPN while in the ADHD group an opposite pattern was observed with greater differentiation in the left than right FPN. The neural profiles of the clinical groups also suggest differential contributions of the SN and DMN in the processing of spatial vs. verbal working memory with a clear demarcation in the SN (but not DMN) in PDD, and the DMN (but not SN) in ADHD.

The TD group's neural profile shows clear distinctions between verbal and spatial working memory processes across all networks. In contrast, no differential processing of verbal and spatial material was observed in control conditions. The highest classification accuracies were observed in the left and right FPN, demonstrating known lateralized processes such as verbal rehearsal (left > right) (Nagel, Herting, Maxwell, Bruno, & Fair, 2013) or spatial information processing (right > left) (Ray, Mackay, Harmer, & Crow, 2007). The significant classification accuracies in the DMN may be attributed to greater involvement of posterior DMN regions in mental imagery (i.e., precuneus), supporting retention of non-verbal material (Cavanna & Trimble, 2006). Lastly,

the significant distinctiveness found in the SN likely suggests that one type of information, verbal or spatial, may be more or less salient in the context of working memory.

The PDD group revealed a neural profile of significant differentiation between verbal and spatial working memory in the right FPN and SN, and non-significant differentiation in the left FPN and DMN. Interestingly, both left and right FPN show no differential processing of verbal and spatial material in the control condition, whereas only the right FPN revealed differential processing during the working memory condition. This pattern suggests that pediatric PDD may be associated with a failure to trigger specialized processing patterns in the left FPN during working memory, consistent with prior work implicating aberrant functioning of left FPN during working memory tasks in adults with depression (Vasic et al., 2007; Walter et al., 2007). Additionally, given the DMN's integral role in supporting self-monitoring processes (Davey, Pujol, & Harrison, 2016) linked to rumination in depression (Eshel & Roiser, 2010; Hamilton et al., 2011), one plausible explanation is that weak differentiation in the DMN might reflect deviant processing within that network. In other words, consistent depression-related rumination may interfere with the ability to concentrate and optimally perform on cognitively demanding tasks, which may be moderated by DMN activity (Wagner, Schachtzabel, Peikert, & Bär, 2015). Future studies should investigate the degree to which depressive symptoms (e.g., rumination) interface with working memory processes in the DMN, which can inform neurobiological mechanisms underpinning working memory deficits in pediatric PDD.

With the exception of the DMN, the ADHD group's neural profile indicates weak differentiation of verbal and spatial working memory processes in the two FPNs and the SN – though there was a trend towards differential processing in the left FPN. Nevertheless, this overall pattern of weaker differentiation is consistent with prior work suggesting ADHD is associated with

more diffusely connected networks and an overactive DMN (Castellanos & Proal, 2012; Mowinckel et al., 2017; Pruim et al., 2019) that contributes to more unreliable neural responses during cognitive tasks (Hoekzema et al., 2014). Moreover, the significant differentiation in the DMN is consistent with the "default mode hypothesis" of ADHD, which theorizes executive function deficits arise from DMN interference during tasks (Sonuga-Barke & Castellanos, 2007). As such, the DMN may be playing a compensatory role to support working memory processes typically carried out in the FPN. Alternatively, heterogeneity within the diagnostic group may account for the indistinct response to spatial and verbal processes (Cortese, Aoki, Itahashi, Castellanos, & Eickhoff, 2020; Nigg, 2012). Though we cannot explain the overall weaker differentiation in our ADHD group, we speculate future studies that incorporate the variability of symptom presentation could help elucidate the locus of working memory deficits in ADHD.

We included the sensorimotor network as a control analysis, expecting comparable classification accuracy across groups. Indeed, we observed similar classification accuracies for each group in the working memory condition, indicating that the SM is suitable as a control network. However, an unexpected result was the distinct response patterns across both working memory tasks observed in the working memory condition. Although the tasks were closely matched, and we did not observe significant behavioral differences in RT and accuracy, we cannot exclude the possibility that other systematic differences may have contributed to the distinct neural patterns observed in the SM and other networks.

We acknowledge important limitations, including the lack of spatial specificity (i.e., which voxels within each network are most distinct) or whether deviant responses to either the verbal or spatial task (or both) contributed to significant differentiation of neural patterns. Additionally, our sample was relatively small and focused on only male participants within a limited age range,

which may restrict the generalizability of the results. It is also not clear to what extent out results may apply to adults with ADHD or PDD. However, up to 94% of individuals with early onset PDD (i.e., before the age of 21) report a subsequent a major depressive episode at some point in their life (Klein, Schwartz, Rose, & Leader, 2000) suggesting neural alterations early on may constitute a vulnerability for later depression and likely stay constant or worsen with age. For those with ADHD, between 50% to 60% show persistence of symptoms into adulthood (Roy et al., 2016), particularly symptoms of inattention that are more closely linked to executive function deficits such as working memory. As such, while studies in adults are necessary to confirm the transferability of our findings, we believe the findings and methodology equally applies to adults with the disorder. Finally, although our small sample is well characterized clinically, it is not well-suited for diagnostic (i.e., between-subjects) classification. Future studies could use classification approaches to predict whether dissociable network patterns predict clinical diagnoses (Nielsen, Barch, Petersen, Schlaggar, & Greene, 2020; Saeed, 2018; Sundermann, Herr, Schwindt, & Pfleiderer, 2014) or subtypes within diagnosis (Fair, Bathula, Nikolas, & Nigg, 2012). Such investigations could focus on transdiagnostic measures, which entails predicting symptom dimensions rather than categorical diagnoses, as proposed by the Research Domain Criteria Framework (Buckholtz & Meyer-Lindenberg, 2012; Insel et al., 2010; Parkes, Satterthwaite, & Bassett, 2020).

Conclusion

In conclusion, we highlight within-group classification as an innovative approach to identify the distinctness of neural response patterns in *a priori* brain networks across two neurodevelopmental disorders. We found group-specific neural profiles comprising network-

specific spatial vs. verbal working memory processes. The neural profile of the TD group is characterized by distinct neural response patterns when retaining spatial vs. verbal information in working memory, suggesting greater specialization for and sensitivity to task demands in several key networks supporting essential cognitive, emotional and social functions. The neural profiles of both clinical groups markedly deviate from the control group, hinting at neural representations putatively less responsive to varying task demands. Identifying group-specific neural profiles of essential cognitive processes in health and disease is a much-needed intermediary step to advance translational psychiatry and to help inform dysregulated brain mechanisms (Huys, Maia, & Frank, 2016; Woo, Chang, Lindquist, & Wager, 2017). Critically, we demonstrate that data-driven multivariate approaches can advance further investigation into how brain networks underpinning cognitive processes may differ between clinical disorders (Yoon et al., 2008), ultimately improving the translation of research to inform clinical diagnosis and treatment.

Acknowledgments

This research was conducted within the Developmental Imaging research group, Murdoch Children's Research Institute and the Academic Child Psychiatry Unit, Department of Paediatrics, The University of Melbourne. Royal Children's Hospital, Melbourne, Victoria. It was supported by the Murdoch Children's Research Institute, the Royal Children's Hospital, The Royal Children's Hospital Foundation, Department of Paediatrics, The University of Melbourne, and the Victorian Government's Operational Infrastructure Support Program. TS was supported by a NHMRC Career Development Award. DY was supported by the National Institutes of Health NRSA Fellowship F31-DA042574. We would like to thank all children and families who participated in this research. We would also like to thank Dr. Lauren Richmond for helpful comments on the manuscript. DY and VV would also like to thank the Kavli Foundation for enabling participation in their Summer School in Cognitive Neuroscience (2016, Santa Barbara, USA), where this collaboration originated.

Disclosures

Dr. Yee, Dr. Vilgis, Dr. Silk and Dr. Vance all report no biomedical financial interests or potential conflicts of interest.

Open Practices Statement

The data and materials for all experiments are available on our OSF repository: https://osf.io/a5349/. The experiment was not preregistered.

References

Achenbach, T. M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR and TRF profiles.* Burlington: Department of Psychiatry, University of Vermont.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders : DSM-IV-TR*. Washington, DC: American Psychiatric Association.

Arnsten, A. F., & Rubia, K. (2012). Neurobiological Circuits Regulating Attention, Cognitive Control, Motivation, and Emotion: Disruptions in Neurodevelopmental Psychiatric Disorders. *Journal of the American Academcy of Child and Adolsecent Psychiatry, 51*(4), 356-367.

Barch, D. M., Sheline, Y. I., Csernansky, J. G., & Snyder, A. Z. (2003). Working Memory and Prefrontal Cortex Dysfunction: Specificity to Schizophrenia Compared with Major Depression. *Biological Psychiatry, 53*, 376-384. doi:10.1016/S0006-3223(02)01674-8

10.1016/S0006-3223(03)01674-8

Beauregard, M., Paquette, V., & Levesque, J. (2006). Dysfunction in the neural circuitry of emotinal self-regulation in major depressive disorder. *Learning and Memory, 17*(8), 843-846.

Brooks, B. L., Iverson, G. L., Sherman, E. M., & Roberge, M. C. (2010). Identifying cognitive problems in children and adolescents with depression using computerized neuropsychological testing. *Applied Neuropsychology, 17*(1), 37-43. doi:10.1080/09084280903526083

Buckholtz, Joshua W., & Meyer-Lindenberg, A. (2012). Psychopathology and the Human Connectome: Toward a Transdiagnostic Model of Risk For Mental Illness. *Neuron, 74*(6), 990-1004. doi:10.1016/j.neuron.2012.06.002

Cai, W., Chen, T., Szegletes, L., Supekar, K., & Menon, V. (2018). Aberrant Time-Varying Cross-Network Interactions in Children With Attention-Deficit/Hyperactivity Disorder and the Relation to Attention Deficits. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(3), 263-273. doi:https://doi.org/10.1016/j.bpsc.2017.10.005

Castellanos, F. X., & Proal, E. (2012). Large-scale brain systems in ADHD: beyond the prefrontal-striatal model. *Trends in Cognitive Sciences, 16*(1), 17-26. doi:10.1016/j.tics.2011.11.007

Cataldo, M. G., Nobile, M., Lorusso, M. L., Battaglia, M., & Molteni, M. (2005). Impulsivity in depressed children and adolescents: a comparison between behavioral and neuropsychological data. *Psychiatry Res, 136*(2-3), 123-133. doi:10.1016/j.psychres.2004.12.012

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain, 129*(Pt 3), 564-583. doi:10.1093/brain/awl004

Chow, T. E., Westphal, A. J., & Rissman, J. (2018). Multi-voxel pattern classification differentiates personally experienced event memories from secondhand event knowledge. *Neuroimage, 176*, 110-123. doi:10.1016/j.neuroimage.2018.04.024

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron, 83*(1), 238-251. doi:10.1016/j.neuron.2014.05.014

Conners, C. K., Sitarenios, G., Parker, J. D., & Epstein, J. N. (1998). The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *J Abnorm Child Psychol, 26*(4), 257-268.

Cortese, S., Aoki, Y. Y., Itahashi, T., Castellanos, F. X., & Eickhoff, S. B. (2020). Systematic Review and Meta-analysis: Resting State Functional Magnetic Resonance Imaging Studies of Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. doi:https://doi.org/10.1016/j.jaac.2020.08.014

Cortese, S., Kelly, C., Chabernaud, C., Proal, E., Di Martino, A., Milham, M. P., & Castellanos, F. X. (2012). Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies. *American Journal of Psychiatry, 169*(10), 1038-1055. doi:10.1176/appi.ajp.2012.11101521

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res, 29*(3), 162-173. doi:10.1006/cbmr.1996.0014

Crawford, J. R., Anderson, V., Rankin, P. M., & MacDonald, J. (2010). An index-based short-form of the WISC-IV with accompanying analysis of the reliability and abnormality of differences. *British Journal of Clinical Psychology, 49*(Pt 2), 235-258. doi:10.1348/014466509X455470

Cubillo, A., Smith, A. B., Barrett, N., Giampietro, V., Brammer, M., Simmons, A., & Rubia, K. (2014). Drug-specific laterality effects on frontal lobe activation of atomoxetine and methylphenidate in attention deficit hyperactivity disorder boys during working memory. *Psychol Med, 44*(3), 633-646. doi:10.1017/S0033291713000676

Daniel, T. A., Katz, J. S., & Robinson, J. L. (2016). Delayed match-to-sample in working memory: A BrainMap meta-analysis. *Biol Psychol, 120*, 10-20. doi:10.1016/j.biopsycho.2016.07.015

Davey, C. G., Pujol, J., & Harrison, B. J. (2016). Mapping the self in the brain's default mode network. *Neuroimage, 132*, 390-397. doi:10.1016/j.neuroimage.2016.02.022

Dwyer, D. B., Falkai, P., & Koutsouleris, N. (2018). Machine Learning Approaches for Clinical Psychology and Psychiatry. *Annual Review of Clinical Psychology, 14*(1), 91-118. doi:10.1146/annurev-clinpsy-032816-045037

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry, 68*(2), 118-124. doi:10.1016/j.biopsych.2010.01.027

Etzel, J. A., Cole, M. W., Zacks, J. M., Kay, K. N., & Braver, T. S. (2016). Reward Motivation Enhances Task Coding in Frontoparietal Cortex. *Cereb Cortex, 26*(4), 1647-1659. doi:10.1093/cercor/bhu327

Fair, D., Bathula, D., Nikolas, M., & Nigg, J. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America, 109*(17), 6769-6774.

Franklin, T., Lee, A., Hall, N., Hetrick, S., Ong, J., Haslam, N., . . . Vance, A. (2010). The association of visuospatial working memory with dysthymic disorder in pre-pubertal children. *Psychol Med, 40*(2), 253-261. doi:10.1017/S0033291709990365

Gardumi, A., Ivanov, D., Hausfeld, L., Valente, G., Formisano, E., & Uludağ, K. (2016). The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis. *Neuroimage, 132*, 32-42. doi:https://doi.org/10.1016/j.neuroimage.2016.02.033

Gohier, B., Ferracci, L., Surguladze, S. A., Lawrence, E., El Hage, W., Kefi, M. Z., . . . Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar depression. *J Affect Disord, 116*(1-2), 100-105. doi:10.1016/j.jad.2008.10.028

Gruber, O., Zilles, D., Kennel, J., Gruber, E., & Falkai, P. (2011). A systematic experimental neuropsychological investigation of the functional integrity of working memory circuits in major depression. *Eur Arch Psychiatry Clin Neurosci, 261*(3), 179-184. doi:10.1007/s00406-010-0165-3

Günther, T., Holtkamp, K., Jolles, J., Herpertz-Dahlmann, B., & Konrad, K. (2004). Verbal memory and aspects of attentional control in children and adolescents with anxiety disorders or depressive disorders. *J Affect Disord, 82*(2), 265-269. doi:10.1016/j.jad.2003.11.004

Günther, T., Konrad, K., De Brito, S. A., Herpertz-Dahlmann, B., & Vloet, T. D. (2011). Attentional functions in children and adolescents with ADHD, depressive disorders, and the comorbid condition. *Journal of Child Psychology and Psychiatry, 52*(3), 324-331. doi:10.1111/j.1469-7610.2010.02320.x

Hamilton, J. P., Chen, M. C., & Gotlib, I. H. (2013). Neural systems approaches to understanding major depressive disorder: An intrinsic functional organization perspective. *Neurobiology of Disease, 52*, 4-11. doi:https://doi.org/10.1016/j.nbd.2012.01.015

Hamilton, J. P., Furman, D. J., Chang, C., Thomason, M. E., Dennis, E., & Gotlib, I. H. (2011). Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination. *Biological Psychiatry, 70*(4), 327-333. doi:10.1016/j.biopsych.2011.02.003

Haxby, J. V., Connolly, A. C., & Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annu Rev Neurosci, 37*, 435-456. doi:10.1146/annurev-neuro-062012-170325

Haynes, J.-D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron, 87*(2), 257-270. doi:https://doi.org/10.1016/j.neuron.2015.05.025

Hebart, M. N., & Baker, C. I. (2018). Deconstructing multivariate decoding for the study of brain function. *Neuroimage, 180*(Pt A), 4-18. doi:10.1016/j.neuroimage.2017.08.005

Hebart, M. N., Gorgen, K., & Haynes, J. D. (2014). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front Neuroinform, 8*, 88. doi:10.3389/fninf.2014.00088

Hoekzema, E., Carmona, S., Ramos-Quiroga, J. A., Richarte Fernandez, V., Bosch, R., Soliva, J. C., . . . Casas, M. (2014). An independent components and functional connectivity analysis of resting state fMRI data points to neural network dysregulation in adult ADHD. *Hum Brain Mapp, 35*(4), 1261-1272.

Huys, Q. J., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci, 19*(3), 404-413. doi:10.1038/nn.4238

Insel, T. R., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry, 167*(7), 748-751.

Jackson, J. B., & Woolgar, A. (2018). Adaptive coding in the human brain: Distinct object features are encoded by overlapping voxels in frontoparietal cortex. *Cortex, 108*, 25-34. doi:10.1016/j.cortex.2018.07.006

Kessler, R. C., Angermeyer, M., Anthony, J. C., De Graaf, R., Demyttenaere, K., Gasquet, I., . . . Ustün, T. B. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative.

World psychiatry : official journal of the World Psychiatric Association (WPA), 6(3), 168-176.

Klein, D. N., Schwartz, J. E., Rose, S., & Leader, J. B. (2000). Five-year course and outcome of dysthymic disorder: A prospective, naturalistic follow-up study. *The American Journal of Psychiatry, 157*(6), 931-939.

Klimkeit, E. I., Mattingley, J. B., Sheppard, D. M., Lee, P., & Bradshaw, J. L. (2005). Motor preparation, motor execution, attention, and executive functions in attention deficit/hyperactivity disorder (ADHD). *Child Neuropsychology, 11*(2), 153-173. doi:10.1080/092970490911298

Korhonen, V., Laukkanen, E., Antikainen, R., Peiponen, S., Lehtonen, J., & Viinamaki, H. (2002). Effect of major depression on cognitive performance among treatment-seeking adolescents. *Nordic Journal of Psychiatry, 56*(3), 187-193.

Kovacs, M. (1992). *Children's Depression Inventory*. New York: Multi-Health Systems.

Kovacs, M., Joormann, J., & Gotlib, I. H. (2008). Emotion (Dys)regulation and Links to Depressive Disorders. *Child Development Perspectives, 2*(3), 149-155. doi:10.1111/j.1750-8606.2008.00057.x

Liang, X., Zou, Q., He, Y., & Yang, Y. (2015). Topologically Reorganized Connectivity Architecture of Default-Mode, Executive-Control, and Salience Networks across Working Memory Task Loads. *Cerebral Cortex, 26*(4), 1501-1511. doi:10.1093/cercor/bhu316

Maalouf, F. T., Brent, D., Clark, L., Tavitian, L., McHugh, R. M., Sahakian, B. J., & Phillips, M. L. (2011). Neurocognitive impairment in adolescent major depressive disorder: state vs. trait illness markers. *J Affect Disord, 133*(3), 625-632. doi:10.1016/j.jad.2011.04.041

Manoliu, A., Meng, C., Brandl, F., Doll, A., Tahmasian, M., Scherr, M., . . . Sorg, C. (2014). Insular dysfunction within the salience network is associated with severity of symptoms and aberrant inter-network connectivity in major depressive disorder. *Frontiers in Human Neuroscience, 7*(930). doi:10.3389/fnhum.2013.00930

Martel, M. M., Nikolas, M., & Nigg, J. T. (2007). Executive function in adolescents with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 46*(11), 1437-1444. doi:10.1097/chi.0b013e31814cf953

Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*(4), 377-384. doi:10.1097/01.chi.0000153228.72591.73

Matthews, K., Coghill, D., & Rhodes, S. (2008). Neuropsychological functioning in depressed adolescent girls. *J Affect Disord, 111*(1), 113-118. doi:10.1016/j.jad.2008.02.003

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences, 15*(10), 483-506. doi:10.1016/j.tics.2011.08.003

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function, 214*(5-6), 655-667. doi:10.1007/s00429-010-0262-0

Miller, C. H., Hamilton, J. P., Sacchet, M. D., & Gotlib, I. H. (2015). Meta-analysis of Functional Neuroimaging of Major Depressive Disorder in Youth. *JAMA Psychiatry, 72*(10), 1045-1053. doi:10.1001/jamapsychiatry.2015.1376

Mowinckel, A. M., Alnaes, D., Pedersen, M. L., Ziegler, S., Fredriksen, M., Kaufmann, T., . . . Biele, G. (2017). Increased default-mode variability is related to reduced task-

performance and is evident in adults with ADHD. *Neuroimage Clin, 16*, 369-382. doi:10.1016/j.nicl.2017.03.008

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage, 59*(3), 2636-2643.

Nagel, B. J., Herting, M. M., Maxwell, E. C., Bruno, R., & Fair, D. (2013). Hemispheric lateralization of verbal and spatial working memory during adolescence. *Brain Cogn, 82*(1), 58-68. doi:https://doi.org/10.1016/j.bandc.2013.02.007

Nielsen, A. N., Barch, D. M., Petersen, S. E., Schlaggar, B. L., & Greene, D. J. (2020). Machine Learning With Neuroimaging: Evaluating Its Applications in Psychiatry. *Biol Psychiatry Cogn Neurosci Neuroimaging, 5*(8), 791-798. doi:10.1016/j.bpsc.2019.11.007

Nigg, J. T. (2012). Future directions in ADHD etiology research. *J Clin Child Adolesc Psychol, 41*(4), 524-533. doi:10.1080/15374416.2012.686870

Parkes, L., Satterthwaite, T. D., & Bassett, D. S. (2020). Towards precise resting-state fMRI biomarkers in psychiatry: synthesizing developments in transdiagnostic research, dimensional models of psychopathology, and normative neurodevelopment. *arXiv preprint arXiv:2006.04728*.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage, 45*(1 Suppl), S199-S209. doi:10.1016/j.neuroimage.2008.11.007

Pruim, R. H. R., Beckmann, C. F., Oldehinkel, M., Oosterlaan, J., Heslenfeld, D., Hartman, C. A., . . . Mennes, M. (2019). An Integrated Analysis of Neural Network Correlates of

Categorical and Dimensional Models of Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging, 4*(5), 472-483. doi:10.1016/j.bpsc.2018.11.014

Ray, M. K., Mackay, C. E., Harmer, C. J., & Crow, T. J. (2007). Bilateral Generic Working Memory Circuit Requires Left-Lateralized Addition for Verbal Processing. *Cerebral Cortex, 18*(6), 1421-1428. doi:10.1093/cercor/bhm175

Rose, E. J., Simonotto, E., & Ebmeier, K. P. (2006). Limbic over-activity in depression during preserved performance on the n-back task. *Neuroimage, 29*(1), 203-215. doi:10.1016/j.neuroimage.2005.07.002

Rottschy, C., Langner, R., Dogan, I., Reetz, K., Laird, A. R., Schulz, J. B., . . . Eickhoff, S. B. (2012). Modelling neural correlates of working memory: A coordinate-based meta-analysis. *Neuroimage, 60*(1), 830-846. doi:10.1016/j.neuroimage.2011.11.050

Roy, A., Hechtman, L., Arnold, L. E., Sibley, M. H., Molina, B. S. G., Swanson, J. M., . . . Group, M. T. A. C. (2016). Childhood Factors Affecting Persistence and Desistence of Attention-Deficit/Hyperactivity Disorder Symptoms in Adulthood: Results From the MTA. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(11), 937-944.e934. doi:10.1016/j.jaac.2016.05.027

Saeed, F. (2018). Towards quantifying psychiatric diagnosis using machine learning algorithms and big fMRI data. *Big Data Analytics, 3*(1). doi:10.1186/s41044-018-0033-0

Seeley, W. W. (2019). The Salience Network: A Neural System for Perceiving and Responding to Homeostatic Demands. *The Journal of Neuroscience, 39*(50), 9878. doi:10.1523/JNEUROSCI.1138-17.2019

Shaw, P., Stringaris, A., Nigg, J., & Leibenluft, E. (2014). Emotion dysregulation in attention deficit hyperactivity disorder. *American Journal of Psychiatry, 171*(3), 276-293. doi:10.1176/appi.ajp.2013.13070966

Shilyansky, C., Williams, L. M., Gyurak, A., Harris, A., Usherwood, T., & Etkin, A. (2016). Effect of antidepressant treatment on cognitive impairments associated with depression: a randomised longitudinal study. *The Lancet Psychiatry, 3*(5), 425-435. doi:10.1016/s2215-0366(16)00012-2

Silk, T. J., Vance, A., Rinehart, N., Egan, G., O'Boyle, M., Bradshaw, J. L., & Cunnington, R. (2005). Fronto-parietal activation in attention-deficit hyperactivity disorder, combined type: functional magnetic resonance imaging study. *British Journal of Psychiatry, 187*(3), 282-283. doi:10.1192/bjp.187.3.282

Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV: Clinician's manual*: Psychological Corporation.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., . . . Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A, 106*(31), 13040-13045. doi:10.1073/pnas.0905267106

Sonuga-Barke, E. J., & Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. *Neurosci Biobehav Rev, 31*(7), 977-986. doi:10.1016/j.neubiorev.2007.02.005

Sowerby, P., Seal, S., & Tripp, G. (2011). Working memory deficits in ADHD: the contribution of age, learning/language difficulties, and task parameters. *J Atten Disord, 15*(6), 461-472. doi:10.1177/1087054710370674

Sundermann, B., Herr, D., Schwindt, W., & Pfleiderer, B. (2014). Multivariate classification of blood oxygen level-dependent FMRI data with diagnostic intention: a clinical perspective. *AJNR Am J Neuroradiol, 35*(5), 848-855. doi:10.3174/ajnr.A3713

Thomason, M. E., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G. H., & Gabrieli, J. D. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience, 21*(2), 316-332. doi:10.1162/jocn.2008.21028

Vance, A., Silk, T. J., Casey, M., Rinehart, N. J., Bradshaw, J. L., Bellgrove, M. A., & Cunnington, R. (2007). Right parietal dysfunction in children with attention deficit hyperactivity disorder, combined type: a functional MRI study. *Molecular Psychiatry, 12*(9), 826-832, 793. doi:10.1038/sj.mp.4001999

Vance, A., & Winther, J. (2020). Irritability, Depressed Mood, Inattention and Spatial Working Memory in Children and Adolescents with Major Depressive Disorder With/Without Persistent Depressive Disorder. *Child Psychiatry Hum Dev*. doi:10.1007/s10578-020-01061-x

Vasic, N., Wolf, R. C., & Walter, H. (2007). Executive functions in patients with depression. The role of prefrontal activation. *Nervenarzt, 78*(6), 628, 630-622, 634-626 passim. doi:10.1007/s00115-006-2240-6

Vilgis, V., Chen, J., Silk, T. J., Cunnington, R., & Vance, A. (2014). Frontoparietal function in young people with dysthymic disorder (DSM-5: Persistent depressive disorder) during spatial working memory. *J Affect Disord, 160*, 34-42. doi:10.1016/j.jad.2014.01.024

Vilgis, V., Silk, T. J., & Vance, A. (2015). Executive function and attention in children and adolescents with depressive disorders: a systematic review. *European Child and Adolescent Psychiatry, 24*(4), 365-384. doi:10.1007/s00787-015-0675-7

Wagner, G., Schachtzabel, C., Peikert, G., & Bär, K.-J. (2015). The neural basis of the abnormal self-referential processing and its impact on cognitive control in depressed patients. *Hum Brain Mapp, 36*(7), 2781-2794. doi:10.1002/hbm.22807

Walter, H., Wolf, R. C., Spitzer, M., & Vasic, N. (2007). Increased left prefrontal activation in patients with unipolar depression: an event-related, parametric, performance-controlled fMRI study. *J Affect Disord, 101*(1), 175-185.

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). San Antonio, TX: The Psychological Corporation.

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

Wilding, J. (2005). Is attention impaired in ADHD? *British Journal of Developmental Psychology, 23*, 487-505.

Willcutt, E. G., Doyle, A. E., Nigg, J. T., Faraone, S. V., & Pennington, B. F. (2005). Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry, 57*(11), 1336-1346. doi:10.1016/j.biopsych.2005.02.006

Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., & Haynes, J. D. (2015). The Role of the Parietal Cortex in the Representation of Task-Reward Associations. *J Neurosci, 35*(36), 12355-12365. doi:10.1523/JNEUROSCI.4882-14.2015

Woo, C.-W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: brain models in translational neuroimaging. *Nature Neuroscience, 20*(3), 365-377. doi:10.1038/nn.4478

Woolgar, A., Afshar, S., Williams, M. A., & Rich, A. N. (2015). Flexible Coding of Task Rules in Frontoparietal Cortex: An Adaptive System for Flexible Cognitive Control. *J Cogn Neurosci, 27*(10), 1895-1911. doi:10.1162/jocn_a_00827

Yang, Z., Fang, F., & Weng, X. (2012). Recent developments in multivariate pattern analysis for functional MRI. *Neuroscience bulletin, 28*(4), 399-408.

Yoon, J. H., Tamir, D., Minzenberg, M. J., Ragland, J. D., Ursu, S., & Carter, C. S. (2008). Multivariate pattern analysis of functional magnetic resonance imaging data reveals deficits in distributed representations in schizophrenia. *Biol Psychiatry, 64*(12), 1035-1041. doi:10.1016/j.biopsych.2008.07.025