Behavioral/Cognitive

Electroencephalography Correlates of Spatial Working Memory Deficits in Attention-Deficit/Hyperactivity Disorder: Vigilance, Encoding, and Maintenance

Agatha Lenartowicz,[1,2] **Arnaud Delorme,**[3,4,5] **Patricia D. Walshaw,**[1,2] **Alex L. Cho,**[1,2] **Robert M. Bilder,**[1,2] **James J. McGough,**[1,2] **James T. McCracken,**[1,2] **Scott Makeig,**[3] **and Sandra K. Loo**[1,2]

[1]Semel Institute for Neuroscience and Human Behavior, and [2]David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California 90095, [3]Swartz Center for Computational Neuroscience, University of California San Diego, La Jolla, California 92093-0559, and [4]Centre de Recherche Cerveau et Cognition (CERCO), Paul Sabatier University and [5]CERCO, CNRS, 31062 Toulouse, France

In the current study we sought to dissociate the component processes of working memory (WM) (vigilance, encoding and maintenance) that may be differentially impaired in attention-deficit/ hyperactivity disorder (ADHD). We collected electroencephalographic (EEG) data from 52 children with ADHD and 47 typically developing (TD) children, ages 7–14 years, while they performed a spatial Sternberg working memory task. We used independent component analysis and time-frequency analysis to identify midoccipital alpha (8–12 Hz) to evaluate encoding processes and frontal midline theta (4–7 Hz) to evaluate maintenance processes. We tested for effects of task difficulty and cue processing to evaluate vigilance. Children with ADHD showed attenuated alpha band event-related desynchronization (ERD) during encoding. This effect was more pronounced when task difficulty was low (consistent with impaired vigilance) and was predictive of memory task performance and symptom severity. Correlated with alpha ERD during encoding were alpha power increases during the maintenance period (relative to baseline), suggesting a compensatory effort. Consistent with this interpretation, midfrontal theta power increases during maintenance were stronger in ADHD and in high-load memory conditions. Furthermore, children with ADHD exhibited a maturational lag in development of posterior alpha power whereas age-related changes in frontal theta power deviated from the TD pattern. Last, subjects with ADHD showed age-independent attenuation of evoked responses to warning cues, suggesting low vigilance. Combined, these three EEG measures predicted diagnosis with 70% accuracy. We conclude that the interplay of impaired vigilance and encoding in ADHD may compromise maintenance and lead to impaired WM performance in this group.

Key words: ADHD; alpha; EEG; ICA; theta

Introduction

Deficits of executive functioning are generally agreed to play a key role in neuropsychological models of attention-deficit/ hyperactivity disorder (ADHD) (Barkley, 1997; Tannock, 1998; Castellanos and Tannock, 2002). Meta-analyses identify working memory (WM) performance as one of the most impaired of executive processes in individuals with ADHD and have been noted to respond inconsistently to drug therapy. WM problems are

widely interpreted to be due to frontostriatocerebellar dysfunctions, reflecting catecholaminergic dysregulation of prefrontal cortex (PFC; McCracken, 1991; Arnsten, 2006).

Despite the consensus in the field for the association of WM deficits with ADHD, there is much still unanswered about the neural mechanisms underlying WM dysfunction. For instance, WM impairments are not ubiquitous (Biederman and Faraone, 2005; Nigg, 2005; Nigg, et al., 2005; Loo et al., 2008). Although individuals with ADHD perform more poorly than typically developing (TD) controls when assessed for storage and manipulation of information in short-term memory (STM), they appear to effectively use top-down control (another executive process that relies on PFC function) to filter distractors (Mason et al., 2003, 2004; Huang-Pollock et al., 2005). These findings are contrary to the prediction of the predominant model of frontostriatal dysfunction in ADHD, which would predict impairment across PFC-related functions. This implies that the occurrence of WM impairments must encompass multinetwork interactions that require clarification.

One possibility is that WM deficits in ADHD may be secondary to a problem with vigilance (van der Meere and Sergeant, 1988; Biederman and Spencer, 1999; Robertson and Garavan,

Table 1. Sample demographics, symptoms, and performance

	N	Age (y)	Males	FSIQ	$SWAN_{inatt}$	$SWAN_{hyper}$	ACC	RT (s)	RT_{SD}
ADHD	37	10.2	24	104	15.2**	22.7**	0.76*	1.4†	0.44*
TD	43	10.6	26	108	36.3	37.6	0.81*	1.3†	0.39

**p < 0.01, *p < 0.05, and †p < 0.1 indicate group difference significance. FSIQ, Full-scale IQ; SWAN, Strengths and Weaknesses of ADHD symptoms and Normal behavior scale; inatt, inattention subscale; hyper, hyperactivity/impulsivity subscale; ACC, accuracy; RT, reaction time, RT_{SD}, SD of reaction time. Note that high SWAN scores suggest low levels of the problem behavior.

2004; Huang-Pollock et al., 2006; Friedman-Hill et al., 2010), thought to compromise the activation of control functions associated with PFC, such as WM. Another possibility is that PFC-sensory cortex interactions are atypical in ADHD, which would interfere with visual encoding of information that is to be maintained in STM. However, if interactions between the PFC and parietal cortex were compromised, we may predict a problem in maintaining information in STM, regardless of encoding success. All three scenarios could deteriorate performance on a typical WM task, although by different mechanisms, contributing to variability with respect to tests of PFC function.

In the current study, we sought to identify which contributing process of WM (vigilance, encoding, or maintenance) is differentially impaired in ADHD. To answer this question we measured electroencephalographic (EEG) neural responses in children with and without ADHD, while they performed the Sternberg spatial working memory (SWM) task (Sternberg, 1966). We analyzed EEG signals using independent component analysis (ICA) and time-frequency analysis, leveraging the dissociation between parieto-occipital alpha (8–12 Hz) and frontal-midline theta (4–7 Hz), whose event-related sShifts in mean amplitude have been associated with encoding and maintenance, respectively (Onton et al., 2005; Klimesch et al., 2007). We used EEG responses to a warning cue and a difficulty manipulation to gauge vigilance. Our results provide the first event-related EEG assay of spectral brain dynamics underlying SWM performance deficits in ADHD. We show group differences in vigilance, encoding, and maintenance that differ in developmental trajectories, and suggest that variability in encoding efficacy contributes to impaired WM performance in ADHD.

Materials and Methods

Participants and diagnoses. A total of 102 children (52 with ADHD, 62 males) were recruited from the community through radio and newspaper advertisements, community organizations (CHADD; www.chadd.org), local schools, and primary care physicians. After receiving verbal and written explanations of study requirements, and before any study procedures, all parents/participants provided written informed permission/assent as approved by the Institutional Review Board. We note that the group comparisons reported here are from a subsample, reduced by our ICA procedures (described below) and the requirement that accuracy exceeded 60%, of 80 participants (50 males, 62.5% of sample) aged 7–14 years; 37 with ADHD (24 males, 64.9% of sample) and 43 age-matched (26 males, 60.5% of sample) TD peers (Table 1).

All participants underwent extensive phenotypic assessment including diagnostic interviews and EEG recording. We evaluated children for ADHD and other psychiatric disorders through a semistructured diagnostic interview with the primary caretaker (usually the mother) and a direct interview with the child (if 8 years of age or older) using the Schedule for Affective Disorders and Schizophrenia for School-Age Children (KSADS-PL; Kaufman et al., 1997). In addition, we obtained teacher reports that supplemented data obtained during clinical interviews. Psychiatric disorders were considered "present" if the participant currently met full DSM-IV diagnostic criteria for any ADHD subtype. Clinical psychologists or highly trained interviewers with extensive experience in psychiatric diagnoses and training in using the KSADS-PL con-
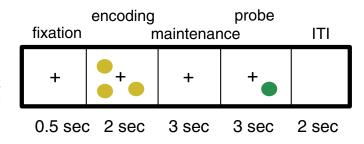

Figure 1. Participants performed a Sternberg visual working memory task during EEG recordings. In each trial, appearance of an alerting fixation-cross cued trial onset. Participants then viewed an encoding stimulus containing either 1 or 3 dots (low-loads), or 5 or 7 dots (high-loads), attending the spatial positions of the dots for 2 s. The dots then disappeared from screen, beginning a 3 s maintenance period. Upon presentation of the probe stimulus, participants were asked to indicate, by button press, whether or not the location of the probe disc matched the location of any of the encoding stimulus discs. During the ensuing intertrial interval (ITI) the screen was blank.

ducted all interviews. Senior clinicians (J.J.M., J.T.M.) confirmed psychiatric diagnoses after individual review of symptoms, developmental course, and impairment level. Estimated intelligence (IQ) was assessed using the Wechsler Abbreviated Scale of Intelligence (WASI). Additionally, severity of ADHD symptoms was assessed using the Strengths and Weaknesses of ADHD symptoms and Normal (SWAN) Behavior Scale (http://www.adhd.net/; Lakes et al., 2012).

Subjects were excluded from participation if they were positive for any of the following: neurological disorder, head injury resulting in concussion, diagnoses of autism, chronic tic disorder, major depression, bipolar disorder, panic disorder or psychosis, or estimated Full-scale IQ <80. In addition, TD control subjects were excluded if they had any major Axis I diagnosis, with the exception of oppositional defiant disorder or simple phobia, or were on any type of psychoactive medication. No participants, including those diagnosed with ADHD, were on medication, as these data were collected as part of a baseline assessment for a clinical trial.

Task and procedures. We used a computerized version of the SWM task (Sternberg, 1966; Glahn et al., 2002) to assess encoding and maintenance of WM (Fig. 1). Participants were presented with a fixation cross for 500 ms, followed by an encoding display containing 1, 3, 5, or 7 yellow dots whose locations they were instructed to remember. After 2 s, the screen turned blank and remained blank for a 3 s maintenance interval. Next, a single dot (probe) was presented and children were asked to decide whether this probe stimulus was in a location previously shown (match) or not (nonmatch). The probe remained on the screen for either 3 s or until a response was recorded (whichever was earlier). This task yields as dependent variables: accuracy, reaction time (RT), and SD of reaction time (RT_{SD}) as an index of response variability (Table 1).

We were interested in three phases of the task trials: alerting (during fixation), encoding, and maintenance. During fixation, participants were alerted by a fixation cross to the impending stimulus presentation. Analysis of EEG data during the fixation period allowed us to assess how vigilant participants were in attending to the task. During the encoding phase, participants had to encode the locations of the presented stimuli. During the maintenance phase, they had to hold these locations in STM. Additionally, we varied the task load (i.e., the number of presented dots), and thereby the degree of WM challenge. Impairments in encoding or in maintenance specific to WM should be more prominent at high-load. Impairments associated with low vigilance might be most prominent at low-load because at low-load the encoding process is less engaging and the participant must exert more vigilance to remain on task (Friedman-Hill et al., 2010). The load manipulation therefore added an additional test of dissociation between impairments of WM circuits versus impairment in their activation through a deficit in task vigilance.

Participants performed a total of 48 SWM trials in each experimental block, and two blocks total. In each block there were equal numbers of trials for each load and match/no-match response type; the order of these types were randomized within each block. Stimuli were presented on a Dell PC and responses were collected on a QWERTY keyboard. The left

and right arrow keys were assigned to match and nonmatch responses, respectively. The stimulus presentation and response collection was controlled by E-Prime software (v1.1b5; Psychology Software Tools).

A training block preceded the testing session. Here children were shown eight trials in which the encoding and probe stimuli appeared side-by-side, then another eight trials in which these were shown consecutively, omitting the maintenance interval. In a final eight trials, children performed the full task including the maintenance interval. A requirement of >50% accuracy during practice was required to continue to the two testing blocks. Each block lasted ~7 min, for a total testing time of ~17 min, including practice.

Behavioral analysis. We analyzed subject task accuracy, RT and RT_{SD} (an index of response variability) using a repeated-measures ANOVA implemented in SPSS (IBM). Each analysis included three factors. The between-subject factor of GROUP (ADHD vs TD controls) tested for overall differences in performance by the two diagnostic groups. The within-subject factor of LOAD (1 and 3 dots at low-load vs 5 and 7 dots at high-load) distinguished between group differences at high-load (requiring more WM engagement for successful performance) and low-load (requiring more vigilance to remain on task). If group differences were specific to either vigilance or WM, a significant Group × Load interaction would be expected. We also included the between-subjects factor of age to test for developmental changes. Two age subgroups were defined: 7–10 years and 11–14 years.

EEG recording. While participants performed the SWM task, EEG recordings were collected using an Electrocap (Electro-cap International), containing 40 silver chloride electrodes positioned in accordance with the 10/20 System (AEEGS, 1991). Electrode impedances were brought <4 KΩ before task recording. Electrical signals were recorded using MANSCAN hardware and recording software (SAM Technology). EEG was recorded at 256 Hz with linked-ears reference. Electrode locations were recorded before the EEG session by measuring the pairwise distances between electrodes and landmarks (preauricular points and nasion), using Fowler calipers. These distances were transformed within the MANSCAN software to 3-D spherical coordinates.

EEG analysis. We adopted a three-step analysis approach. First, we applied ICA (Makeig et al., 1996) to single-subject EEG data. The independent components (ICs) produced by this analysis have temporally near-independent activity time courses and near-dipolar scalp projections, and have been proposed to represent spatially coherent local field activity within a single cortical area (Makeig et al., 1996, 2004; Onton et al., 2006; Delorme et al., 2012). Each IC represents an EEG signal source that projects with fixed polarity and proportions across the scalp electrodes. In this way, ICA "unmixes" the time courses of brain and non-brain source activities that contribute to the scalp recorded signals, potentially unmasking effects that would not be detectable in the mixed signals at individual electrodes. This approach differs from standard channel space analyses in that the latter speak to the combined effects of multiple sources at individual electrodes. The objective of the current analysis was to evaluate group effects in the unmixed sources of EEG signals.

Treating the ICs as putative "cortical area electrodes," we computed mean event-related potentials (ERPs) and event-related spectral perturbations (ERSPs) for IC activity epochs time locked to task events (stimuli cueing onset of fixation, encoding, maintenance). In the second step, for all subjects, we fit single equivalent dipole source models to the scalp map of each IC, and ICs that were compatible with an origin in cortex were gathered into functionally distinct source clusters across subjects using a *k*-means clustering algorithm. From these, we identified clusters corresponding to midoccipital alpha and midfrontal theta (also used for analysis of alerting responses). In the third step, we conducted group and task load comparisons of the computed epoch-mean IC measures on the identified cluster ICs.

Step 1: ICA analysis. Offline EEG processing and analyses were performed using custom MATLAB (Mathworks) scripts using functions from the EEGLAB environment (Delorme and Makeig, 2004; v 11.0.3b). The EEG data were high-pass filtered (>1 Hz), and then inspected for noisy electrodes which were excluded from further analysis. The data were re-referenced to average reference. Data from the two task blocks were concatenated to form a continuous time series. This time series was then inspected, in 0.5 s bins, for outlier epochs encompassing gross movements and muscle artifact, and these time bins were removed.

The trimmed data were then decomposed into maximally independent component processes using temporal ICA decomposition using extended infomax (Lee et al. 1999). ICA linearly unmixes the original EEG channel data into a channel-weighted sum of maximally temporally independent component activity time courses that each has a fixed spatial projection pattern. In our analyses, we used the default parameters of the *binica* program in EEGLAB, with stopping weight change set to 1e-7 and the maximum learning steps set to 1000. For each individual, the algorithm produced as many components as there were data channels. A single IC process comprised an activity time course throughout the dataset plus a spatial scalp map giving its relative projection strengths and polarities to each of the scalp electrode channels.

The resulting ICs were processed as follows. First, we inspected the spatial, spectral, and temporal properties of each IC to identify those components corresponding to nonbrain sources; eye blinks, lateral eye movement, cardiac artifacts, single-channel artifacts, and high-frequency line noise. These were excluded from further analyses. Additionally, we excluded from further analysis component processes that could not be attributed to a cortical source. To do this, we calculated a single-equivalent current dipole model for each IC scalp topography using the DIPFIT toolbox in EEGLAB (http://sccn.ucsd.edu/wiki/A08:_DIPFIT). For the source modeling, each participant's electrode positions were registered to a template MNI (Montreal Neurological Institute) head model and an electrical forward solution was calculated using the boundary element method. A source model consisting of a single-equivalent current dipole was then estimated for each remaining IC source, and the projection pattern of its equivalent dipole model to the scalp electrodes was linearly fit to its scalp map. Any IC whose equivalent dipole model accounted for <85% of the variance of its scalp map (typically low-amplitude ICs with multifocal or "messy" scalp maps) were excluded from further analyses, as were ICs whose equivalent dipole was located outside the brain volume (including ICs accounting for neck or muscle activities, single electrode noise, etc). This procedure left an average of 17.3 (of a maximum of 40) ICs per subject (SD = 3), and 1766 ICs across all subjects, whose single-equivalent dipole models on average accounted for 93.3% (SD = 1.4%) of the IC scalp map. Between groups, 17.6 (SD = 3.2) ICs were retained for TD subjects, 17.1 (SD = 3.5) for ADHD subjects (nonsignificant by *t* test, $t_{(100)} < 1$). Additionally, we calculated how much of the variance of the whole-channel EEG time series was accounted for by the projection of our retained ICs into channel space. Across subjects, the retained ICs accounted for an average of 67.8% (SD = 16.3%) of the EEG data. Note that because the IC projections summed by the scalp electrodes may have sign differences, the total percentage variance accounted for by the sum of the retained ICs (pvaf, here 67.8%) may be less than the summed percent variance accounted for (pvaf) values for the individual ICs (within subjects) or IC clusters (across subjects).

For these remaining ICs, the primary functional measure we evaluated was the mean event-related spectral perturbation (ERSP) time/frequency image time locked to either the encoding stimuli or the succeeding maintenance intervals (Makeig, 1993). IC activity epochs were extracted time-locked to the onset of the encoding stimulus, beginning 1.6 s before and ending 8 s after stimulus presentation. This 9.5 s time interval encompassed the entire duration of the trial (Fig. 1): the 500 ms fixation period (Fixation) that alerted the trial, the encoding period (Encoding) during the 2 s on-screen stimulus display, and the 3 s memory maintenance interval (Maintenance) that followed, ending with presentation of the response query. Using these relatively long epochs also served to circumvent edge effects in the time/frequency decomposition. Epochs containing any remaining artifacts or followed by incorrect responses were removed. Remaining epochs used in the analyses were as follows: 65 (SD = 13) per subject for the TD group and 60 (SD = 12) for the ADHD group.

To calculate the ERSPs, we applied Morlet wavelet decomposition, as implemented in EEGLAB *newtimef()*, to the epoch component time series. We calculated power for 100 log-spaced frequencies ranging from 3

to 128 Hz, and along 200 linearly spaced time bins (advanced in 47 ms steps) across the epoch. To account for the trade-off between frequency and temporal resolution, the wavelets were modified, such that three cycles (comprising 1000 ms) were used at the lowest frequency (3 Hz), increasing linearly to 25.6 cycles (lasting 200 ms) at the highest frequency (128 Hz). For each IC of each subject, these trial spectrograms were averaged and converted to decibel units. Finally, we subtracted the log-power of the baseline period (also in dB) preceding onset of the fixation cross (-600 to -100 ms) at each frequency from the log spectrogram values at each latency, and used the resulting ERSP in both clustering and subsequent group analyses

For analysis of mean responses to fixation stimulus onsets we also calculated stimulus-locked ERPs. We did this by averaging the time-domain epochs and subtracting the baseline preceding the fixation cue (-100 to 0 ms). Again, these epochs were averaged separately for each subject IC. ERPs were selected for fixation analysis because we were interested in possible differences in the height of the P2 response peak, a positivity appearing on frontocentral scalp $\sim$200 ms after cue stimulus presentation; P2 amplitude has been shown to correlate with the processing of task cue content (Potts, 2004; Lenartowicz et al. 2010; D'Ardenne et al., 2012). We reasoned that if children with ADHD were less vigilant during the task, they would be less likely to process the alerting cues and so should show a weaker P2 response to the fixation stimulus.

Step 2: ICA clusters. Because ICA decomposition produces multiple ICs for each subject, the resulting ICs need to be compared across subjects to identify those subsets of ICs that capture functionally equivalent or near-equivalent source activities. We did this using a k-means approach that clustered the full set of ICs from all subjects based on a combination of spatial (dipole location and scalp topography) as well as functional (ERP, ERSP) IC characteristics (as described by Onton and Makeig, 2006), based on distances between ICs taking into account dissimilarities in all these measures. Note that this method of clustering ICs is not guaranteed to return the same clusters as other possible definitions, as discussed by Bigdely-Shamlo et al. (2013). However, for these data adjusting the cluster distance metric had relatively small effects on the clustering, which was (intentionally) dominated by similarity in IC equivalent dipole locations. Here, the number of clusters was established heuristically as the number of ICs (12) of the subject who had the fewest ICs that passed all exclusion criteria. During this clustering process, we also removed ICs whose centroids were >3 SDs from the centroid of any cluster in cluster metric space. Each cluster therefore included one or more positionally and functionally more similar ICs from most participants [range, 12 participant/cluster (cluster containing eye blinks) to 97 participants/cluster of 102].

Finally we inspected the 12 clusters to identify those sets of ICs that showed a midoccipital maximal projection (for analysis of 8–12 Hz alpha band activity during encoding) or midfrontal projection (for analysis of 4–7 Hz theta band activity during maintenance). For analysis of alerting responses (ERPs) during fixation, we used the same midfrontal cluster as for theta, as the spatial topography is comparable across both of these metrics and the medial frontal cluster contributed the most variance (30.4%) to the grand mean ERP during fixation. Although we had prior expectations for clusters of interest, all clusters were evaluated for group differences and are illustrated and described fully in Figure 2. We present results here only for those clusters whose activities exhibited significant effects.

Step 3: Group analysis. We conducted the group comparisons on spectral and ERP measures extracted from cluster ICs. For clusters with midoccipital scalp topography, we extracted alpha-band ERSP values during encoding. From each IC in a cluster, ERSP values from the encoding interval (0–2 s) in the (8–12 Hz) alpha frequency band were averaged. We identified one cluster with a midfrontal scalp topography from which we extracted ERSP values from the maintenance (2–5 s) in the (4–7 Hz) theta frequency band. From these midfrontal clusters, we also extracted peak ERP amplitudes of the P2 in response to fixation. If a participant in a given cluster contributed more than one IC, the ERSPs from these ICs were averaged such that each individual contributed a single ERSP metric to the group analysis. The resulting measures, occip-

ital encoding-alpha ERSPs, frontal maintenance-theta ERSPs, and frontal fixation-related ERPs, were analyzed separately using repeated-measures ANOVA, including the same three factors as in the behavior analysis (Group, Load, Age).

EEG predictors of behavior and symptoms. To test whether the ERSP neural indices were linked to subsequent trial motor responses we constructed a multiple regression model for each of the performance measures (accuracy, RT, RT_{SD}). The component measures were entered as continuous independent variables. Because performance is known to improve with age, we included age as a continuous covariate of no interest. The model significance was tested using an F test; individual regression coefficients were evaluated using t tests.

EEG predictors of ADHD diagnosis. In a final analysis, we tested whether the component measures had predictive value in distinguishing between children with and without ADHD. Toward this aim, we conducted a binary logistic regression analysis, including the component measures as independent variables and diagnosis status as the dependent variable. In this analysis, we collapsed across task loads to maximize power. The analysis was conducted using a sixfold cross-validation scheme. On each of six iterations, we built a logistic regression model using data from 83% (5/6) of the subjects and then validated the predictive accuracy on the remaining 17% (1/6) of the datasets. Each testing set was unique and contained equal proportions of ADHD and TD subjects of each age group. Each testing validation analysis gave one contingency table. These six classification tables were then summed to create a cross-validation confusion table for the whole dataset.

We assessed the predictive value of the component measures by calculating, from the cross-validation confusion table, the mean diagnostic prediction accuracy and comparing it against its 95% confidence interval (CI) bounds, as well as against the maximum by-chance accuracy rate. To evaluate the diagnostic accuracy of the model, we also computed its sensitivity [p(ADHD′|ADHD; where ADHD′ = predicted ADHD status] and specificity [p(TD′|TD; where TD′ = predicted TD status]. To assess the predictive value of the models, we computed both positive likelihood [p(ADHD′|ADHD)/p(ADHD′|TD)] and negative likelihood [p(TD′|ADHD)/p(TD′|TD)]. The predictive odds ratio, describing how much greater the odds are that a child has ADHD given a positive diagnosis than given a negative diagnosis, was derived as the ratio of these two values.
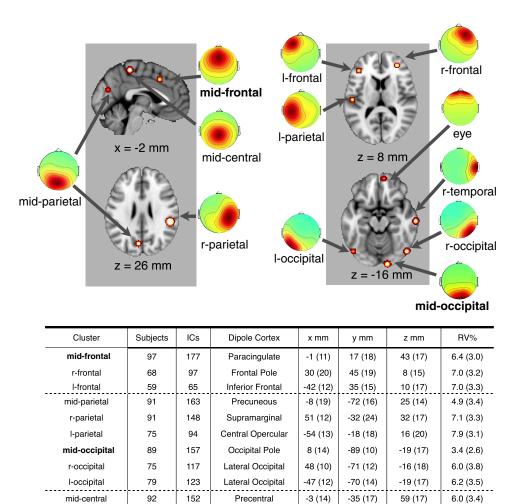
Results

Sample demographics and performance

The sample demographics are presented in Table 1. The ADHD and TD groups did not differ in mean age ($t_{(78)} < 1$) or FSIQ ($t_{(78)} = 1.5$, p < 0.14) or the proportion of males to females in each group [$\chi^2(1, N = 80) = 0.16, p < 0.7$]. In addition, for the ADHD group we used the one-sample binomial test to evaluate the proportion of individuals of combined and inattentive subtypes. There were 22 (59%) children with combined type ADHD and 15 (41%) children with inattentive type ADHD (none with hyperactive-impulsive only type), and these proportions did not differ significantly ($z = 1.12, p < 0.32$) from chance.

The groups did however differ in performance (Table 1). The ADHD group was overall less accurate ($t_{(78)} = 2.15, p < 0.03$) and more variable in their RTs ($t_{(78)} = 2.2, p < 0.03$). RTs in the ADHD group were also slower, but this was significant at the trend level only ($t_{(1.67)}, p < 0.1$). These findings are consistent with previous reports that SWM is impaired in children with ADHD.

ICA clusters

The complete set of 12 IC clusters identified in these analyses is presented in Figure 2. Six of the IC clusters had single-equivalent dipole source models centered in occipital or parietal cortex. Of these clusters only one (midoccipital) showed significant group effects within the alpha frequency band (8–12 Hz) during the

Cluster	Subjects	ICs	Dipole Cortex	x mm	y mm	z mm	RV%
mid-frontal	97	177	Paracingulate	-1 (11)	17 (18)	43 (17)	6.4 (3.0)
r-frontal	68	97	Frontal Pole	30 (20)	45 (19)	8 (15)	7.0 (3.2)
l-frontal	59	65	Inferior Frontal	-42 (12)	35 (15)	10 (17)	7.0 (3.3)
mid-parietal	91	163	Precuneous	-8 (19)	-72 (16)	25 (14)	4.9 (3.4)
r-parietal	91	148	Supramarginal	51 (12)	-32 (24)	32 (17)	7.1 (3.3)
l-parietal	75	94	Central Opercular	-54 (13)	-18 (18)	16 (20)	7.9 (3.1)
mid-occipital	89	157	Occipital Pole	8 (14)	-89 (10)	-19 (17)	3.4 (2.6)
r-occipital	75	117	Lateral Occipital	48 (10)	-71 (12)	-16 (18)	6.0 (3.8)
l-occipital	79	123	Lateral Occipital	-47 (12)	-70 (14)	-19 (17)	6.2 (3.5)
mid-central	92	152	Precentral	-3 (14)	-35 (17)	59 (17)	6.0 (3.4)
r-temporal	70	98	Middle Temporal	65 (12)	-17 (14)	-12 (18)	7.4 (3.9)
eye	25	25	Frontal Pole	6 (14)	61 (8)	-19 (9)	10.2 (3.9)

mid = middle, r = right, l = left; RV = residual variance; SD = standard deviation

Figure 2. Retained ICs from all subjects, grouped into 12 clusters using k-means clustering, represent the identified IC sources that, in sum, account for the majority ($M = 67.8\%$) of the channel data. The equivalent dipole centroids for these 12 clusters are shown in the top and described in the table. The dipole centroids are presented as hot spots with radius indicating the 95% CI. Clusters reported in the results where those that showed significant group effects (bold type). The cluster labeled as "eye" captured 25 ICs accounting for eye blink artifacts that had escaped our thresholding procedure.

Encoding epoch (0–2 s), as shown below. The mean of the equivalent dipole source models for the ICs in this cluster was localized to near (just below) the right occipital pole (MNI: $x = 8$ mm, $y = -89$ mm, $z = -19$ mm), and whose scalp projection accounted for 96.7% of the variance in the mean component scalp map. The cluster ICs accounted for just 45.5% (SE = 4.3) of TD EEG time course variance and for 40.7% (SE = 3.6) of ADHD EEG time course variance at the maximally projected scalp channel (Iz). This cluster contained ICs from 89 of 102 subjects.

The cluster solution also identified three clusters with single-equivalent dipole source models localized to frontal cortex. Only the midfrontal cluster (Fig. 2) was associated with significant group effects in our analyses during the maintenance interval (2–5 s), as described below. The mean of the equivalent dipole source models for the ICs in this cluster was localized to medial cortex near paracingulate gyrus (MNI: $x = -1$ mm, $y = 17$ mm, $z = 43$ mm; between Brodmann areas 6/24), within the cognitive division of the anterior cingulate as identified in the meta-analysis of Bush et al. (2000). This model accounted for 93.7% of the variance in the mean scalp projection of this midfrontal IC cluster, which was maximal at frontal and central electrodes. The

cluster ICs accounted for just 52.2% (SE = 2.5) of TD EEG time course variance and for 45.3% (SE = 2.9) ADHD EEG time course variance at the maximally projected scalp channel (FCz). This cluster contained ICs from 97 of 102 participants.

Of the remaining clusters, none showed group effects and are not discussed further here. One cluster (eye) that appeared to capture 25 nonbrain IC processes associated with eye blinks and that escaped our residual-variance based procedure for excluding nonneuronal sources, suggesting that the clustering algorithm effectively separated these remaining nonbrain sources from other component processes.

Group differences in alerting response during fixation
Evoked responses to the fixation cross, the alerting cue, are presented in Figure 3. The repeated-measures ANOVA conducted on the amplitude of the P2 peak in the midfrontal cluster ERP, revealed a main effect of Group ($F_{(1,74)} = 10.2$, $p < 0.002$). Children with ADHD had a smaller midfrontal cluster-component P2 (0.21 μV) than TD children (0.56 μV). Neither the Load effect nor the Group-by-Load interaction were significant ($F_{(1,74)} < 1$, $p > 0.30$) as expected given that LOAD was randomized across
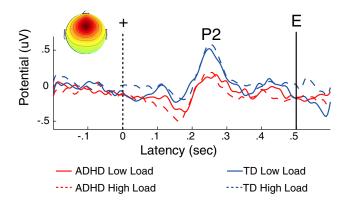
Figure 3. Children with ADHD had a smaller midfrontal-cluster alerting-response to the fixation ($+$) stimulus that cued the imminent (500 ms later) onset of the encoding stimulus (E). A larger-amplitude P2 (positive peak near 240 ms following fixation) occurred in TD children, suggesting that this group was more actively anticipating the trials and that they thus were more generally vigilant during the task than children with ADHD. This effect was not attributable to impaired perceptual responding in this group.

trials and could not be predicted during fixation. These results indicate that children with ADHD had a weaker midfrontal cortical evoked response to the alerting fixation stimulus, consistent with weakened vigilance during the task.

We also tested whether the alerting response to fixation stimuli response varied by age. The amplitude of the midfrontal cluster P2 peak did not vary as a function of Age ($F_{(74)} < 1$), and this variable did not interact with Load or Group ($F_{(74)} < 1$). Hence, the Group difference in P2 amplitude was stable across task load and age group.

An alternate interpretation of the P2 result might be that the ADHD group showed weaker sensory responses to the fixation stimuli, which then lead to a smaller frontal P2. We tested this explicitly by comparing the peak amplitude of the occipital P1, a positive potential occurring ~100 ms following stimulus onset and indexing sensory visual processing, in response to both fixation and the encoding stimuli. A clear P1 was present in channels Oz, O_1, and O_2 and in the midoccipital cluster during fixation and encoding; however, Group effects were not significant at either interval ($F_{(74)} < 1$).

Group differences during encoding

The repeated-measures ANOVA on mean changes in alpha (8–12 Hz) log-power during the encoding interval (0–2 s), revealed effects of both Group and Load (Fig. 4a, top). The main effect of Group was significant ($F_{(1,67)} = 12.0$, $p < 0.001$), as was the main effect of Load ($F_{(1,67)} = 26.5$, $p < 0.001$. An event-related desynchronization (ERD) or "blocking" of mean alpha band power during Encoding was more pronounced in TD children (-0.81 dB) relative to children with ADHD (-0.14 dB), and at high-load (-0.75 dB) relative to low-load (-0.19 dB). The interaction between these variables was also significant ($F_{(1,67)} = 5.3$, $p < 0.02$); the Group difference was over twice as large at low-load (Δ 0.92 dB; TD: -0.66 dB, ADHD: 0.26 dB) than at high-load (Δ 0.42 dB; TD: -0.96 dB, ADHD: -0.54 dB). *Post hoc* t tests revealed that the Group difference was significant at low-load ($t_{(69)} = 3.91$, $p < 0.001$) but not at high-load ($t_{(69)} = 1.69$, $p < 0.1$).

These results indicate that children with ADHD had an attenuated attention-related modulation of their ongoing EEG activity to the encoding stimulus, especially when the associated task difficulty was low. This is conceptually consistent with the analysis
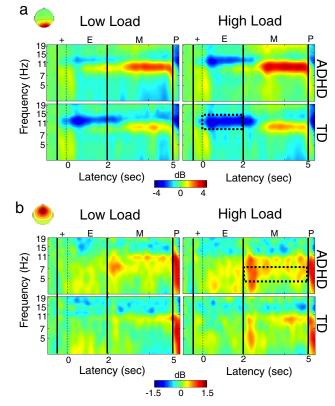

Figure 4. *a*, The ERSPs for the mid-occipital cluster were characterized by an alpha band (8–12 Hz) ERD during encoding (E) and probe (P) periods, and by a somewhat lower-frequency ERS during memory maintenance (M). Our a priori analyses examined the encoding period (broken-line box), during which alpha ERD was enhanced in TD children and during high-load. The alpha ERD group difference was stronger at low-load than at high-load, consistent with poor WM engagement. Additional analyses revealed a similar pattern of effects for alpha ERS during maintenance. *b*, The ERSP for the midfrontal cluster showed a (4–7 Hz) theta band ERS during the maintenance interval (broken-line box) that was stronger in children with ADHD and during high-load trials.

of the alerting response during fixation and suggests that lowered or more variable vigilance may compromise the ability to encode information in this group. We note however that the peak amplitude of the fixation-related P2 was not significantly correlated, controlling for effects of age, with the mean encoding alpha ERD in either the combined sample or within either group ($RS_{max} = -0.13$, $p_{min} = 0.43$, two-tailed). This could mean that the group differences in alerting during fixation and encoding are two independent impairments, or that their relationship is nonlinear (e.g., mediated by another network).

However, it is possible that the group difference was attributable to differences in resting state alpha, which can be suppressed in individuals with ADHD (Loo et al., 2009) and which therefore could artificially result in weaker alpha blocking during stimulus encoding. To test this hypothesis we extracted absolute alpha power during the baseline period and repeated our analysis. The Group effect was not significant ($F_{(1,67)} = 0.001$, $p > 0.97$) and neither was its interaction with Load ($F_{(1,67)} = 0.21$, $p < 0.21$) suggesting that the alpha effect was attributable to differences in how fully the children focused attention on the encoding stimulus when presented, rather than reflecting differences in alpha power in the prestimulus baseline.

We also observed significant difference in development with respect to alpha ERD (Fig. 5a). The effect of AGE on alpha ERD during stimulus encoding was significant ($F_{(1,67)} = 19.8$, $p <$
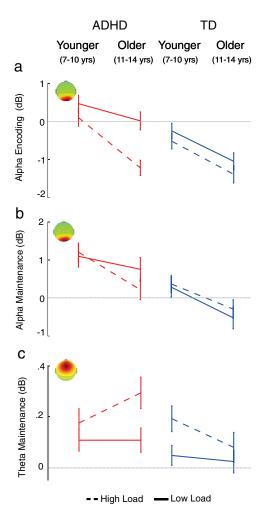
Figure 5. Maturational delay was observed in the ADHD group for the (*a*) mid-occipital cluster alpha ERD during encoding and the (*b*) midoccipital ERS during maintenance. Although in both groups and task intervals, alpha ERD increased with age, the increase in the ADHD group lagged behind that of the TD group. In addition, (*a*) alpha ERD during encoding revealed an interaction, the ADHD group showed a weaker maturation effect in the low-load condition, consistent with poor vigilance. *c*, Unlike for alpha power, the midfrontal cluster theta power increase during maintenance showed a maturational disparity. In the TD group, theta power showed a load effect in younger children but not in older children. In the ADHD group, the opposite effect was observed. In these children a load effect emerged with age.

0.001). Alpha ERD was stronger in older children (-0.91 dB) than in younger children (-0.04 dB). The interaction between Age and Load was also significant ($F_{(1,67)} = 4.7, p < 0.03$) and the three-way interaction between Age, Load, and Group showed a strong trend ($F_{(1,67)} = 3.38, p < 0.07$).

These interactions were attributable to differences in encoding stimulus-related alpha ERD trends with age across groups. In the TD group only the effects of Load ($F_{(1,36)} = 9.1, p < 0.01$) and Age ($F_{(1,36)} = 8.8, p < 0.01$) were significant, meaning that alpha ERD responses were stronger in older children and at high-load. The interaction was not significant ($F_{(1,36)} < 1$. In the ADHD group, in addition to these two main effects ($F_{(1,31)} > 11.6, p < 0.002$) their interaction was also significant ($F_{(1,31)} = 4.7, p < 0.04$). In this group, the maturation effect on alpha ERD was present only for high-load stimuli. This suggests that, with age, children with ADHD do develop mechanisms necessary for stimulus encoding but, consistent with the reduced-vigilance hypothesis, it appears that this mechanism is not (consistently) recruited when task difficulty is low.

Group differences during maintenance

The result of the repeated-measures ANOVA on log mean (4–7 Hz) theta band power during the maintenance interval (2–5 s), revealed significant effects of both Group ($F_{(1,74)} = 4.7, p < 0.03$) and Load ($F_{(1,74)} = 9.8, p < 0.002$; Fig. 4b, bottom). These main effects occurred because during the maintenance interval, theta power showed an event-related increase or synchronization (ERS), relative to baseline, that was more pronounced in children with ADHD (0.17 dB) relative to TD children (0.09 dB) regardless of load, and at high-load (0.19 dB) relative to low-load (0.07 dB) regardless of group. To the extent that these theta band spectral changes are associated with WM maintenance-related processes, these results indicate that during the maintenance interval children with ADHD activated these processes more strongly than TD children. The interaction between Load and Group was not significant ($F_{(1,74)} = 0.13, p < 0.72$). In particular, we did not observe a larger Group difference at high-load than at low-load, which would be expected if the deficit in the ADHD group were specific to WM maintenance. We additionally tested whether the theta ERS during maintenance, which we interpreted as extra activation of maintenance processes in ADHD, was predicted by the attenuated alpha ERD during encoding. Indeed, in the ADHD group, alpha encoding ERD and theta maintenance ERS were significantly correlated ($rs_{(32)} = 0.35, p < 0.05$, two-tailed), suggesting that the efficacy of maintenance is related to efficacy of encoding processes. This correlation was not significant in the TD group ($rs_{(37)} = 0.01, p < 0.98$).

Effects of AGE on the theta ERS during maintenance are shown in Figure 5c. Unlike alpha ERD, the theta increase did not show a significant AGE effect ($F_{(1,74)} = 0.02, p < 0.9$). The only near-significant effect of Age (a strong trend) was an interaction with Group ($F_{(1,74)} = 3.2, p < 0.08$); theta synchronization decreased with age in the TD group but increased with age in the ADHD group, resulting in a significant Group difference in older children ($t_{(32)} = 2.6, p < 0.02$, two-tailed) but not in younger children ($t_{(42)} < 1$). In other words, the stronger midfrontal theta response in the ADHD group emerged with development, a trajectory that differed from that of TD controls. This pattern also differed from the observed development effects for midoccipital alpha ERD, which showed a maturational delay.

Although our planned analyses were to evaluate alpha effects during encoding and theta during maintenance, an inspection of the maintenance interval ERSPs in the midoccipital cluster (Fig. 4a) showed clear increases in mean alpha band spectral power (i.e., ERS) during the maintenance period that could be functionally significant to the group analyses. An occipital alpha ERS can be consistent with a sensory attention inhibition mechanism (Klimesch, 1997, 2007), such as one might engage during active memory maintenance of a preceding sensory scene so as to inhibit the interference of external sensory signals on other sensory or cognitive processes; here, maintaining the to-be-encoded object locations. We therefore decided to also explore these alpha ERS for differences across Group, Load, and Age.

As for the midoccipital alpha during encoding, this maintenance interval analysis revealed a significant effect of Group ($F_{(1,67)} = 7.91, p < 0.006$), but unlike the former it did not show a main effect of Load ($F_{(1,67)} < 1$). Alpha ERS during maintenance was more pronounced in children with ADHD (0.82 dB) than the TD group (-0.01 dB), but did not differ between low-load (0.39 dB) and high-load (0.36 dB). However, the interaction between these variables was significant ($F_{(1,67)} = 4.1, p < 0.05$); the Group difference at low-load (Δ 1.03 dB; TD: -0.13 dB, ADHD: 0.90 dB) was 150% of that at high-load (Δ 0.66 dB; TD:

0.03 dB, ADHD: 0.69 dB). This analysis revealed an effect that was qualitatively quite similar to that during encoding, namely a load-dependent difference. The similarity of this finding across encoding and maintenance may imply that attenuated alpha ERD and exaggerated alpha ERS during the two respective intervals are functionally related and may represent a processing profile more often adopted by the ADHD group. They may also indicate that impaired encoding leads to compensatory efforts during maintenance. Indeed, the correlation between the strength of the encoding alpha ERD and maintenance alpha ERS was significant in both the TD group ($r_{S(36)} = 0.47$, $p < 0.004$, two-tailed), and the ADHD group ($r_{S(31)} = 0.74$, $p < 0.001$, two-tailed).

Additionally, effects of age on the alpha ERS during maintenance (Fig. 5b) followed trends very similar to those for alpha ERD during encoding, which is consistent with the strong correlation that we observed between these variables. The main effect of Age was also significant ($F_{(1,67)} = 17.1$, $p < 0.006$). Alpha synchronization was weaker (or absent) in older children (0.02 dB) compared with younger children (0.72 dB), which is consistent with the interpretation that the ERS reflects effort to engage WM maintenance mechanisms and disengage sensory attention; a need that may decrease with age as attentional mechanisms mature. Of the interactions, only the three-way interaction between Age, Load, and Group was significant ($F_{(1,67)} = 4.76$, $p < .03$). As for the encoding interaction, this three-way effect was attributable to differential changes with age across groups. TD children showed only an effect of Age ($F_{(1,36)} = 10.98$, $p < 0.002$). In this group, older children had a weaker (or absent) alpha ERS during maintenance (-0.42 dB) compared with younger children (0.32). By contrast, in the ADHD group only the interaction between Load and Age was significant ($F_{(1,31)} = 4.16$, $p < 0.05$) because only in the high-load condition was the maturational decrease in alpha recruitment observed (younger: 1.18 dB; older: 0.18 dB). At low-load, children with ADHD showed only a modest decrease in alpha ERS with age (younger: 1.07 dB; older: 0.73 dB). Post hoc tests confirmed that only in the high-load condition was this age effect significant ($t_{(31)} = 2.45$, $p < 0.02$, two-tailed); it was not significant at low-load ($t_{(31)} < 1$). These data suggest that children with ADHD exhibit maturational delay in occipital mechanisms underlying both encoding and maintenance of spatial locations such that only older children with ADHD (and only at high-load) demonstrate the typical patterns of occipital alpha ERD/ERS.

Frequency shifts in alpha during encoding and maintenance

Interestingly, as Figure 4a shows, the maximal increase in alpha band power during WM maintenance appeared at a lower frequency than the maximal alpha band decrease during encoding. This suggested the possibility that the observed changes in amplitude could be attributed to a shift in peak alpha frequency during the task trial relative to baseline. For instance, if the task peak frequency shifted down (to a lower frequency) relative to baseline, we would observe a power decrease at higher frequencies of alpha and a power increase at lower frequencies of alpha. In our data, this pattern of ERD/ERS distribution across frequency bands was present. The maximum of the alpha ERD during encoding was at 11.1 Hz (SE = 0.16 Hz) whereas the maximum of the alpha power ERS during maintenance was at 9.5 Hz (SE = 0.13 Hz), and this difference was significant across subjects ($t_{(70)} = 7.5$, $p < 0.001$) consistent with a peak frequency shift down (from higher to lower frequency) between baseline and task.

To test this hypothesis directly, we extracted the peak of absolute power (rather than ERD/ERS) in the alpha range (8–12 Hz)

Table 2. Multiple regression models of behavioral measures on EEG source measures

	t_{beta}				
	ACC	RT (s)	RT_{SD}	$SWAN_{inatt}$	$SWAN_{hyper}$
Regression variables					
$Alpha_e$	−1.92*	2.05*	2.09*	−2.75**	−1.79[†]
$Alpha_{m \perp e}$	<1	<1	<1	−1.34	−1.19
Theta	<1	−1.42	<1	<1	<1
P2	1.03	−1.53	<1	1.88[†]	<1
Age	2.75**	−3.35**	−2.64*	<1	<1
Model statistics					
R_{sq} Adj	0.25	0.31	0.22	0.14	0.05
F	5.38**	6.95**	4.86**	3.21*	1.65

**$p < 0.01$, *$p < 0.05$, [†]$p < 0.1$ indicate group difference significance. SWAN, Strengths and Weaknesses of ADHD symptoms and Normal behavior scale; inatt, inattention subscale; hyper, hyperactivity/impulsivity subscale; ACC, accuracy; RT, reaction time, RT_{SD}, SD of reaction time.

for each of baseline (-1.1 to -0.6 s relative to stimulus onset), encoding (0–2 s), and maintenance (2–3 s). A repeated-measure ANOVA with Group (ADHD vs TD) and Time interval (baseline, encoding, maintenance) showed neither a significant effect of Group ($F_{(1,69)} < 1$, or an interaction between GROUP and TIME ($F_{(1,69)} = 1.78$, $p < 0.19$) implying that the peak frequency of alpha did not differ by group. However the main effect of TIME was significant ($F_{(1,69)} = 5.3$, $p < 0.02$). *Post hoc t* tests (two-tailed), with Bonferroni corrected alpha set at .017, showed that peak frequency during baseline (9.87 Hz) was significantly higher than during encoding (9.51 Hz; $t_{(70)} = 2.81$, $p < 0.006$) and maintenance (9.65 Hz; $t_{(70)} = 2.38$, $p < 0.02$); the peak frequencies during encoding and maintenance did not differ ($t_{(70)} = 1.42$, $p < 0.16$). Hence, the observed decrease in alpha power during encoding and increase during maintenance were accompanied by a decrease in alpha peak frequency relative to baseline. This peak frequency shift may partially account for the observed alpha ERD and ERS during task.

EEG predictors of behavior and symptoms

To evaluate whether our component measures (midfrontal P2 peak amplitude, midoccipital alpha ERD, midfrontal theta ERS, and midoccipital alpha ERS during maintenance) could predict subject WM performance, we constructed a multiple linear regression model with these measures as independent variables, as well as age as a covariate of no interest (Table 2). Because the two alpha measures were strongly correlated, we orthogonalized the midoccipital maintenance alpha ERS with respect to the encoding alpha, using the Gram–Schmidt process (Golub and van Loan, 1989), which amounts here to calculating of the residuals from the regression of maintenance alpha ERS onto encoding alpha. Thus, we attributed all shared variance to the temporally preceding encoding measure. The midoccipital maintenance independent variable tested only whether this maintenance alpha predicted performance over and above the contribution from encoding alpha. A separate model, with independent variables as above, was constructed for each of the dependent variables: accuracy, RT, RT_{SD}, as well as the inattention and hyperactivity submeasures of the SWAN scale of behavioral ADHD symptoms. Note that SWAN scores correlate negatively with symptom severity (higher scores indicate lower severity of symptoms).

The strongest relationship between behavioral and component EEG measures across subjects was observed for midoccipital alpha ERD during encoding. A stronger mean alpha ERD (more desynchronization) predicted quicker subject responses to the subsequent query stimulus ($b_{RT} = 0.24$, $t_{(62)} = 2.1$, $p < 0.05$) less

variable response times ($b_{RTSD} = 0.26$, $t_{(62)} = 2.1$, $p < 0.04$) and higher scores (lower symptom severity) on the inattention subscale of the SWAN ($b_{SWANinatt} = -0.36$, $t_{(62)} = -2.7$, $p < 0.008$). Moreover, at the trend level, stronger alpha ERD also predicted higher accuracy ($b_{accuracy} = -0.24$, $t_{(62)} = -1.9$, $p < 0.06$) and higher scores (lower symptom severity) on the hyperactive subscale of the SWAN ($b_{SWANhyper} = -0.25$, $t_{(62)} = 1.8$, $p < 0.08$). Hence, stronger alpha ERD during stimulus encoding predicted both better performance on the task and lower symptom severity.

The midoccipital alpha ERS during maintenance had no additional significant relationship with performance ($t_{(62)} < -1.34$, $p > 0.19$) over and above its correlation with the encoding alpha ERD.

The midfrontal cluster measures, theta ERS and P2 peak amplitude, showed a weaker relationship with performance. P2 peak amplitude was weakly predictive of scores on the inattentive subscale of the SWAN ($b = 0.2$; $t_{(62)} = 1.7$, $p < 0.10$); however, no other relationships (accuracy, RT, RT_{SD}, SWN_{inatt}) were significant ($t_{(62)} < 1.53$, $p > 0.13$) for either measure.

Overall-model statistics indicated that these models accounted for a significant amount of variance in accuracy ($R^2_{adj} > 0.25$; $F_{(5,62)} = 5.4$, $p < 0.001$), RT ($R^2_{adj} > 0.31$; $F_{(5,62)} = 7.0$, $p < 0.001$), and RT_{SD} ($R^2_{adj} > 0.22$; $F_{(5,62)} = 4.9$, $p < 0.001$). Model fit was likely influenced by the strong relationship between each of these measures and age. Age was associated with higher accuracy ($b_{accuracy} = 0.34$; $t_{(62)} = 2.8$, $p < 0.009$) faster responses ($b_{RT} = -0.4$; $t_{(62)} = -3.4$, $p < 0.001$) and less variable responses ($b_{RTsd} = 0.33$; $t_{(62)} = 2.6$, $p < 0.01$). The SWAN analysis revealed that the model predicted a significant portion of SWAN score variability for the inattentiveness subscale ($R^2_{adj} > 0.14$; $F_{(5,62)} = 3.2$, $p < 0.01$), but not the hyperactive scale ($R^2_{adj} > 0.05$; $F_{(5,62)} = 1.7$, $p < 0.16$). In neither of the SWAN models was age a significant predictor ($b < -0.09$; $t_{(62)} < 1$) likely because of the restricted age range.

EEG predictors of ADHD diagnosis
We included three measures, midfrontal fixation P2 peak amplitude, midoccipital encoding period alpha ERD, and midfrontal maintenance period theta ERS, in a logistic regression model attempting to predict group membership. Here, only subjects who had ICs contributing to both midoccipital and midfrontal clusters were considered ($N = 64$). We omitted midoccipital alpha maintenance ERS because of its strong correlation with encoding alpha ERD. The three clustered-component measures jointly predicted diagnosis in our cross-validation sample (Fig. 6). The sixfold logistic regression classifier predicted diagnosis with 70% accuracy, exceeding the upper 95% CI ($CI_{0.95} = 0.38 - 0.62$) for by-chance accuracy (0.5). Moreover, diagnostic accuracy was nearly identical across diagnostic groups (ADHD: 71%, TD: 70%). Thus, sensitivity of ADHD detection was significant (0.69, $CI_{0.95} = 0.50 - 0.83$), as was test specificity (0.72, $CI_{0.95} = 0.53 - 0.86$). The positive likelihood of the test was 2.44, meaning that the probability of a positive test being correct was 2.44 times stronger than of it raising a false alarm. The negative likelihood was 0.44, meaning that the probability of a negative test correctly identifying a TD child was 2.3 times stronger than of it missing an ADHD diagnosis. The odds ratio (2.44/0.44) was 5.56. This implies that the component measure-based model has diagnostic potential, with the odds being over five times greater that a child has ADHD given a positive test result than given a negative test result. A natural follow-up to this result would be to test, in stepwise analysis, how the three variables alone predict group membership. Within our current sample we lacked the power to conduct such an analysis as we would require ~150 sample
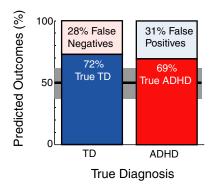

Figure 6. Prediction of diagnostic group membership by logistic regression on measures of midfrontal P2 amplitude during fixation, midoccipital alpha ERD during encoding, and midfrontal theta ERS during maintenance. The model predicted diagnosis correctly for 70% of individuals, exceeding the 95% CI (gray shaded region) for (50% chance-level) random guessing (black horizontal line).

points to conduct a stepwise analysis with three variables (Hosmer and Laemeshow, 2000); however, we are actively pursuing this question in continuing research.

Discussion
In this study, we set out to characterize group differences across alerting responses to fixation (as an index of vigilance), encoding and maintenance during the Sternberg WM task so as to better understand which cortical source processes contributed to group differences in WM performance; a key predictor of diagnosis. Our results indicate an age invariant deficit in vigilance in the ADHD group, supported by reduced alerting responses at fixation and enhanced group differences during encoding and maintenance (for alpha ERD/ERS) during low-load. The midoccipital encoding deficit, measured by alpha ERD, was predictive of the ensuing maintenance effect (alpha ERS), but also task performance and symptom severity. Coupled with the elevated occipital alpha ERS and midfrontal theta ERS during maintenance in the ADHD group, this finding suggests that problems during encoding in the ADHD group lead to compensatory efforts during maintenance. In particular, we did not find evidence for a maintenance-only deficit. Together, these data indicate that WM performance deficits in ADHD are at least in part the result of poor vigilance and atypical encoding; possibly from a deficit and/or variability in degree of focus of attention first to the alerting fixation cue and then to the ensuing encoding stimulus.

ADHD deficits of encoding and vigilance
The dependence of encoding alpha ERD magnitude on task load in the current study replicates the pattern of behavioral deficits in ADHD reported by Friedman-Hill et al. (2010) who suggested that poor responses at low-loads reflect "sluggish recruitment" of WM processes (Robertson and Garavan, 2004). The occipital source-level alpha ERD during encoding that we report may be an effective neural correlate of this phenomenon. This conclusion is consistent with the common interpretation of alpha band activities as supporting active inhibition of processing of task-irrelevant visual or other stimuli (Worden et al., 2000; Pfurtscheller, 2003; Klimesch et al., 2007; Foxe and Snyder, 2011); conversely, alpha desynchronization or alpha blocking (Berger, 1930; Jasper and Penfield, 1949; Chatrian et al., 1959) is thought to reflect a release of such inhibition, allowing enhanced processing of attended stimuli. Alpha is therefore an index of a sort of "gating" or "filtering" of stimuli to be encoded (Klimesch, 1997,

2007) a process that is a likely constituent of the "WM recruitment" described by Friedman-Hill et al. (2010).

It is important to note, however, that in this task increases in load involved increases in stimulus richness (more dots on the screen), so poorer WM performance at high-load could therefore have also been associated with a deficit in sensory processing of more complex stimuli. Thus, the encoding process might have been impaired in ADHD by top-down or bottom-up deficits. Prior studies indicating preserved visuospatial orienting in ADHD (Huang-Pollock and Nigg, 2003) hint at a top-down problem being the likely culprit, and are consistent with the absence of a group difference in the (more bottom-up) P1 responses to encoding stimuli in the current study. The presence of group differences in the (more top-down) P2 component during fixation, without a deficiency in low-level visual stimulus perception, is additional evidence. That this occurred during an initial alerting stimulus, complements theories that posit ADHD to be largely a problem with "vigilance" (Huang-Pollock et al., 2005, 2006; Sergeant, 2005), whereby the inability to sustain attention leads to frequent attention lapses that also impact WM (Manly et al., 2001; Johnson et al., 2008). That this difference was present in both younger and older children suggests this is an early trait of ADHD that may not normalize with age. Together these results point to deficient recruitment of WM in ADHD, accompanied by or possibly produced by task disengagement.

ADHD deficits of maintenance

If stimuli were neither effectively attended nor encoded, one would expect compromises both to WM maintenance and accuracy of responses. Indeed, we found that encoding alpha ERD was predictive of changes during maintenance as well as in subsequent recognition performance. Furthermore, occipital alpha ERS and also midfrontal theta ERS during maintenance were elevated in the ADHD group rather than suppressed, suggesting that these children performed more intensive cognitive processing during the maintenance interval, perhaps to compensate for inefficient encoding. This is consistent with the observation that medial frontal theta power during memory maintenance increases with task difficulty (Gevins et al., 1997; Jensen and Tesche, 2002; Onton et al., 2005). That the midfrontal theta ERS group difference increased with age (Chabot and Serfontein, 1996; Dickstein et al., 2006; maturational deviation), whereas encoding alpha ERD lagged with age (Kinsbourne, 1973; Mann et al., 1992; Clarke et al., 1998; Rubia, 2007; maturational delay), may further indicate that maintenance impairment emerges from dysfunctional encoding processes. The presence of both developmental trends is consistent with cortical thickness patterns reported by Shaw et al. (2007, 2010, 2013), suggesting that frontal cortex maturity is deviant in ADHD (persisting into adulthood), whereas other regions (including parietal and occipital functions that might subserve encoding alpha) show a maturational delay.

Notably the midfrontal theta effects in our data clearly spanned frequencies above the traditional definition of theta (4–7 Hz) and into the lower alpha band (Fig. 4b). Our frontal theta ERS was strongest, on average, in bursts at the onset of the maintenance interval. This was followed by a significant theta-like maintenance alpha band increase in occipital cortex during maintenance. This unique alpha-ERS biased pattern may be indicative of maintenance mechanisms during our spatial Sternberg task that differ from those typically engaged with increases of theta. For instance in the study by Onton et al. (2005), increases in theta were observed while participants rehearsed, and simultaneously updated a memorized letter subset while ignoring dis-

tractors. Such updating is reminiscent of a WM process called "manipulation" of information stored in STM. In our Sternberg task the maintenance interval required rehearsal but not manipulation. The weaker midfrontal theta results in our study may be reflective of relatively weak demands for manipulation. Furthermore, the contents in memory in the current task were spatial rather than verbal, unlike Onton et al. (2005). Spatial maintenance may have engaged different maintenance mechanisms than verbal maintenance, in particular occipital-parietal interactions. These differences may explain the strong effect of occipital alpha ERS, and a weaker midfrontal theta effect in our results. At the same time, our findings indicate that manipulating the contents of WM during maintenance is not required to observe ADHD deficits in performance, consistent with the interpretation that other system impairments (such as, vigilance and encoding) may contribute to the WM impairments.

To our knowledge, our findings are a first report of source level event-related potential and power spectral differences between ADHD and TD groups during SWM. Most prior studies of alpha and theta band power in ADHD examined absolute EEG frequency power and only at individual or averaged scalp channels (Barry et al., 2003; Mitchell et al., 2008). We are aware of one other study (Gomarus et al., 2009) in which event-related changes in scalp alpha and theta power were measured while children with ADHD memorized a set of letters for later retrieval. There, no group differences in alpha were found. That study, however, used a verbal rather than SWM task and only children with the combined subtype of ADHD, both conditions less reliably associated with WM deficits (van der Meere and Sergeant, 1988; Martinussen et al., 2005). Using a different cued-attention paradigm, Mazaheri et al. (2010) reported attenuated posterior scalp alpha ERD following presentations of attended versus to-be-ignored stimuli in ADHD individuals, in agreement with our findings.

Significance to diagnosis of ADHD

What emerges from these data are evidence suggesting that attention deficits in our ADHD sample reflect interactions between brain systems supporting WM, including those for vigilance, encoding, and maintenance. Logistic regression on measures of these processes alone proved able to detect ~70% of clinically established ADHD or non-ADHD cases, notable in relation to behavioral rating scales that claim sensitivities of 38–79% (vs here 69%) and specificities of 13–61% (Snyder et al., 2008; vs here 72%). This finding is encouraging and begs testing of its generalizability in a larger sample, which would also provide sufficient power to test the predictive value of individual measures. Promising follow-ups may include studying cortical source interactions and the relationship of source-level EEG features to ADHD symptom dimensions. Perhaps of greatest clinical importance is the possibility of such refined investigations to be applied to studies of interventions designed to ameliorate WM deficits, which show weaker and more variable responses to drug therapy than ADHD symptoms. The impact of these WM deficits is strongly supported by findings that WM performance in childhood is related to ADHD symptom severity, academic attainment, does not normalize over time, and predicts social functioning and the emergence of other psychopathology over long-term follow-up.

Conclusion

In this study we present a first analysis of cortical source-level event-related EEG dynamics underlying often-reported deficits

of SWM in ADHD. Our results imply that deficits in (here) visual vigilance and encoding may contribute to WM deficits observed in ADHD, suggesting that these components of WM may be a worthwhile target for future intervention and study.

Notes

Supplemental material (evaluation and comparison of independent component and channel analyses) for this article is available at http://alenarto.bol.ucla.edu/page2/page2.html. This material has not been peer reviewed.

References

AEEGS (1991) American Electroencephalographic Society guidelines for standard electrode position nomenclature. J Clin Neurophysiol 8:200–202. Medline

Arnsten AF (2006) Fundamentals of attention-deficit/hyperactivity disorder: circuits and pathways. J Clin Psychiatry 67:7–12. CrossRef Medline

Barkley RA (1997) Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol Bull 121:65–94. CrossRef Medline

Barry RJ, Clarke AR, Johnstone SJ (2003) A review of electrophysiology in attention-deficit/hyperactivity disorder: I. Qualitative and quantitative electroencephalography. Clin Neurophysiol 114:171–183. CrossRef Medline

Berger H (1930) Electroencephalogram of humans. J Fur Psychologie Und Neurologie 40:160–179.

Biederman J, Spencer T (1999) Attention-deficit/hyperactivity disorder (ADHD) as a noradrenergic disorder. Biol Psychiatry 46:1234–1242. CrossRef Medline

Biederman J, Faraone SV (2005) Attention-deficit hyperactivity disorder. Lancet 366:237–248. CrossRef Medline

Bigdely-Shamlo N, Mullen T, Kreutz-Delgado K, Makeig S (2013) Measure projection analysis: a probabilistic approach to EEG source comparison and multi-subject inference. Neuroimage 72, 287–303 CrossRef Medline

Bush G, Luu P, Posner MI (2000) Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 4:215–222. CrossRef Medline

Castellanos FX, Tannock R (2002) Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. Nat Rev Neurosci 3:617–628. CrossRef Medline

Chabot RJ, Serfontein G (1996) Quantitative electroencephalographic profiles of children with attention deficit disorder. Biol Psychiatry 40:951–963. CrossRef Medline

Chatrian GE, Petersen MC, Lazarte JA (1959) The blocking of the rolandic wicket rhythm and some central changes related to movement. Electroencephalogr Clin Neurophysiol 11:497–510. CrossRef Medline

Clarke AR, Barry RJ, McCarthy R, Selikowitz M (1998) EEG analysis in attention-deficit/hyperactivity disorder: a comparative study of two subtypes. Psychiatry Res 81:19–29. CrossRef Medline

D'Ardenne K, Eshel N, Luka J, Lenartowicz A, Nystrom LE, Cohen JD (2012) Role of prefrontal cortex and the midbrain dopamine system in working memory updating. Proc Natl Acad Sci U S A 109:19900–19909. CrossRef Medline

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21. CrossRef Medline

Delorme A, Palmer J, Onton J, Oostenveld R, Makeig S (2012) Independent EEG sources are dipolar. Plos One 7:e30135. CrossRef Medline

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. CrossRef Medline

Foxe JJ, Snyder AC (2011) The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. Front Psychol 2:154. CrossRef Medline

Friedman-Hill SR, Wagman MR, Gex SE, Pine DS, Leibenluft E, Ungerleider LG (2010) What does distractibility in ADHD reveal about mechanisms for top-down attentional control? Cognition 115:93–103. CrossRef Medline

Gevins A, Smith ME, McEvoy L, Yu D (1997) High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex 7:374–385. CrossRef Medline

Glahn DC, Kim J, Cohen MS, Poutanen VP, Therman S, Bava S, Van Erp TG, Manninen M, Huttunen M, Lönnqvist J, Standertskjöld-Nordenstam CG, Cannon TD (2002) Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. Neuroimage 17:201–213. CrossRef Medline

Golub GH, van Loan CF (1989) Matrix computations, Ed 2. Baltimore: John Hopkins UP.

Gomarus HK, Wijers AA, Minderaa RB, Althaus M (2009) Do children with ADHD and/or PDD-NOS differ in reactivity of alpha/theta ERD/ERS to manipulations of cognitive load and stimulus relevance? Clin Neurophysiol 120:73–79. CrossRef Medline

Hosmer DW, Laemeshow S (2000) Applied logistic regression, Ed 2. New York: Wiley.

Huang-Pollock CL, Nigg JT (2003) Searching for the attention deficit in attention deficit hyperactivity disorder: the case of visuospatial orienting. Clin Psychol Rev 23:801–830. CrossRef Medline

Huang-Pollock CL, Nigg JT, Carr TH (2005) Deficient attention is hard to find: applying the perceptual load model of selective attention to attention deficit hyperactivity disorder subtypes. J Child Psychol Psychiatry 46:1211–1218. CrossRef Medline

Huang-Pollock CL, Nigg JT, Halperin JM (2006) Single dissociation findings of ADHD deficits in vigilance but not anterior or posterior attention systems. Neuropsychology 20:420–429. CrossRef Medline

Jasper HH, Penfield W (1949) Electrocortigrams in man: effect of voluntary movement upon the electrical activity of the precentral gyrus. Eur Arch Psychiatry Clin Neurosci 183:163–174.

Jensen O, Tesche CD (2002) Frontal theta activity in humans increases with memory load in a working memory task. Eur J Neurosci 15:1395–1399. CrossRef Medline

Johnson KA, Robertson IH, Barry E, Mulligan A, Dáibhis A, Daly M, Watchorn A, Gill M, Bellgrove MA (2008) Impaired conflict resolution and alerting in children with ADHD: evidence from the attention network task (ANT). J Child Psychol Psychiatry 49:1339–1347. CrossRef Medline

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, Williamson D, Ryan N (1997) Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. J Am Acad Child Adolesc Psychiatry 36:980–988. CrossRef Medline

Kinsbourne M (1973) Minimal brain dysfunction as a neurodevelopmental lag. Ann N Y Acad Sci 205:268–273. CrossRef Medline

Klimesch W (1997) EEG-alpha rhythms and memory processes. Int J Psychophysiol 26:319–340. CrossRef Medline

Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev 53:63–88. CrossRef Medline

Lakes KD, Swanson J, Riggs M (2012) The reliability and validity of the English and Spanish Strengths and Weaknesses of ADHD and Normal Behavior Rating Scales in a preschool sample: continuum measures of hyperactivity and inattention. J Atten Disord 16:510–516. CrossRef Medline

Lee TW, Girolami M, Sejnowski TJ (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Comput 11:417–441. CrossRef Medline

Lenartowicz A, Escobedo-Quiroz R, Cohen JD (2010) Updating of context in working memory: an event-related potential study. Cogn Affect Behav Neurosci 10:298–315. CrossRef Medline

Loo SK, Rich EC, Ishii J, McGough J, McCracken J, Nelson S, Smalley SL (2008) Cognitive functioning in affected sibling pairs with ADHD: familial clustering and dopamine genes. J Child Psychol Psychiatry 49:950–957. CrossRef Medline

Loo SK, Hale TS, Macion J, Hanada G, McGough JJ, McCracken JT, Smalley SL (2009) Cortical activity patterns in ADHD during arousal, activation and sustained attention. Neuropsychologia 47:2114–2119. CrossRef Medline

Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr Clin Neurophysiol 86:283–293. CrossRef Medline

Makeig S, AnlloVento L, Jung P, Bell AJ, Sejnowski TJ, Hillyard SA (1996) Independent component analysis of event-related potentials during a selective attention task. Soc Neurosci Abstr 22:1698.

Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. Trends in Cognitive Sciences 8:204–210. CrossRef Medline

Manly T, Anderson V, Nimmo-Smith I, Turner A, Watson P, Robertson IH (2001) The differential assessment of children's attention: the test of everyday attention for children (TEA-Ch), normative sample and ADHD performance. J Child Psychol Psychiatry 42:1065–1081. CrossRef Medline

Mann CA, Lubar JF, Zimmerman AW, Miller CA, Muenchen RA (1992) Quantitative-analysis of eeg in boys with attention-deficit-hyperactivity disorder-controlled-study with clinical implications. Pediatr Neurol 8:30–36. CrossRef Medline

Martinussen R, Hayden J, Hogg-Johnson S, Tannock R (2005) A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. J Am Acad Child Adolesc Psychiatry 44:377–384. CrossRef Medline

Mason DJ, Humphreys GW, Kent LS (2003) Exploring selective attention in ADHD: visual search through space and time. J Child Psychol Psychiatry 44:1158–1176. CrossRef Medline

Mason DJ, Humphreys GW, Kent L (2004) Visual search, singleton capture, and the control of attentional set in ADHD. Cogn Neuropsychol 21:661–687. CrossRef Medline

Mazaheri A, Coffey-Corina S, Mangun GR, Bekker EM, Berry AS, Corbett BA (2010) Functional disconnection of frontal cortex and visual cortex in attention-deficit/hyperactivity disorder. Biol Psychiatry 67:617–623. CrossRef Medline

McCracken JT (1991) A two-part model of stimulant action on attention-deficit hyperactivity disorder in children. J Neuropsychiatry Clin Neurosci 3:201–209. Medline

Mitchell DJ, McNaughton N, Flanagan D, Kirk IJ (2008) Frontal-midline theta from the perspective of hippocampal "theta". Prog Neurobiol 86:156–185. CrossRef Medline

Nigg JT (2005) Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. Biol Psychiatry 57:1424–1435 CrossRef Medline

Nigg JT, Willcutt EG, Doyle AE, Sonuga-Barke EJ (2005) Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? Biol Psychiatry 57:1224–1230. CrossRef Medline

Onton J, Delorme A, Makeig S (2005) Frontal midline EEG dynamics during working memory. Neuroimage 27:341–356. CrossRef Medline

Onton J, Makeig S (2006) Information-based modeling of event-related brain dynamics. Prog Brain Res 159:99–120. CrossRef Medline

Onton J, Westerfield M, Townsend J, Makeig S (2006) Imaging human EEG dynamics using independent component analysis. Neurosci Biobehav Rev 30:808–822. CrossRef Medline

Pfurtscheller G (2003) Induced oscillations in the alpha band: functional meaning. Epilepsia 44:2–8. CrossRef Medline

Potts GF (2004) An ERP index of task relevance evaluation of visual stimuli. Brain Cogn 56:5–13. CrossRef Medline

Robertson IH, Garavan H (2004) Vigilant attention. In: The cognitive neurosciences, Ed 3 (Gazzaniga MS, ed), pp. 563–578. Cambridge: MIT.

Rubia K (2007) Neuro-anatomic evidence for the maturational delay hypothesis of ADHD. Proc Natl Acad Sci U S A 104:19663–19664. CrossRef Medline

Sergeant JA (2005) Modeling attention-deficit/hyperactivity disorder: a critical appraisal of the cognitive-energetic model. Biol Psychiatry 57:1248–1255. CrossRef Medline

Shaw P, Eckstrand K, Sharp W, Blumenthal J, Lerch JP, Greenstein D, Clasen L, Evans A, Giedd J, Rapoport JL (2007) Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. Proc Natl Acad Sci U S A 104:19649–19654. CrossRef Medline

Shaw P, Gogtay N, Rapoport J (2010) Childhood psychiatric disorders as anomalies in neurodevelopmental trajectories. Hum Brain Mapp 31:917–925. CrossRef Medline

Shaw P, Malek M, Watson B, Greenstein D, de Rossi P, Sharp W (2013) Trajectories of cerebral cortical development in childhood and adolescence and adult attention-deficit/hyperactivity disorder. Biol Psychiatry 74:599–606. CrossRef Medline

Snyder SM, Quintana H, Sexson SB, Knott P, Haque AF, Reynolds DA (2008) Blinded, multi-center validation of EEG and rating scales in identifying ADHD within a clinical sample. Psychiatry Res 159:346–358. CrossRef Medline

Sternberg S (1966) High-speed scanning in human memory. Science 153:652–654. CrossRef Medline

Tannock R (1998) Attention deficit hyperactivity disorder: advances in cognitive, neurobiological, and genetic research. J Child Psychol Psychiatry 39:65–99. CrossRef Medline

van der Meere J, Sergeant J (1988) Focused attention in pervasively hyperactive children. J Abnorm Child Psychol 16:627–639. CrossRef Medline

Worden MS, Foxe JJ, Wang N, Simpson GV (2000) Anticipatory biasing of visuospatial attention indexed by retinotopically specific alpha-band electroencephalography increases over occipital cortex. J Neurosci 20:RC63. Medline