Supplementary information

Participant demographics split by diagnosis.

In this manuscript, we chose to use a dimensional approach to studying ADHD and its interaction with anxiety in order to retain maximum information from our measurements. This approach is statistically warranted; checks confirmed that the assumptions of linear regression were not violated. Because of this dimensional approach, in the main paper we only report demographic information based on the sample as a whole. However, participants for NeuroIMAGE were recruited based on diagnosis (ADHD or control) and, despite the loss of statistical power compared to dimensional approaches, categorical approaches to studying ADHD are common. We therefore report demographic information on the participants split by diagnostic status in Table S1. Further, the values of the clusters found in the main analysis were re-analysed within participants with ADHD and control subjects in order to check whether findings are specific to diagnostic categories.

Table S1. Demographic information on the participants, split by diagnostic status.

Variable	*Total*	*ADHD*	*Subthreshold*	*Controls*	*Statistic*	*DF*	*P-value*
Sample size	371	112	49	200			
Age (SD)	17.1 (3.4)	17.0 (3.3)	17.8 (3.9)	16.9 (3.3)	F=2.77	369	.10
Male sex	52.3%	64.3%	55.1%	45%	X^2=10.1	2	.006
SES (SD)	12.2 (2.5)	11.4 (1.9)	12.0 (2.6)	12.7 (2.6)	F=9.8	369	.002
Amsterdam Location	63.6%	49.1%	65.3%	71.5%	X^2=15.2	2	<.001
CPRS score (SD)	11.3 (11.5)	23.6 (10.3)	10.0 (7.9)	4.0 (4.4)	F=98.3	369	<.001
SDQ-E score (SD)	4.16 (2.03)	5.46 (1.94)	4.53 (2.00)	3.23 (1.55)	F=62.7	369	<.001
Medication use (SD)	447.6 (904.6)	1163.3 (1138.6)	477.6 (951.4)	3.6 (50.8)	F=61.8	359	<.001

Note: SD= Standard deviation; SES= Socioeconomic status, as measured by the parents' averaged years of education. CPRS= Conners Parent Rating Scale. SDQ-E= Strengths and Difficulties Questionnaire-Emotion subscale; DF= Degrees of Freedom. Medication use is average number of days of prescribed stimulant medication.

Full results from task accuracy analysis.

Table S2. All fixed effects from the regression of task accuracy.

Predictor	*Coefficient*	*Error*	*Degrees of freedom*	*T-value*	*P-value*
Intercept	0.10	0.23	1.1	0.44	.73
Anxiety	-0.05	0.04	662.4	-1.21	.26
ADHD	-0.10	0.05	613.6	-2.01	.04
Age	0.20	0.04	517.1	5.25	$2.2 * 10^{-7}$
Nijmegen location	-0.06	0.09	218.7	-0.69	.49
Male sex	-0.20	0.07	662.4	-2.67	.008
Parents' years of education	0.11	0.04	216.3	2.55	.01
Baseline accuracy	0.25	0.03	727.1	7.51	$1.8 * 10^{-13}$
ADHD * anxiety	0.01	0.03	722.7	0.43	.77

ADHD diagnostic algorithm

To determine psychiatric diagnoses, all participants (children and parents alike) were assessed with a combination of ADHD rating scales and a semi-structured diagnostic interview. In order to determine ADHD diagnoses, a diagnostic algorithm was applied based on the behavioural questionnaires (typically filled in by parents as well as a second observer) and the diagnostic interview, using DSM V criteria (American Psychiatric Association, 2000). Inconsistent cases were reviewed by a team of trained experts, in order to derive a consensus diagnosis.

Measures

Participants were assessed with a parent rating scale (CPRS-R:L; 1998a), and either a teacher rating scale (CTRS-R:L; 1998b), applied for participants < 18 years, or a self-report (CAARS-S:S; 1999), applied for participants ≥ 18 years. A semi-structured diagnostic interview (KSADS-PL; Kaufman et al., 1997) was administered to both the participants (if ≥ 12 years old) and their parents separately. Initially, all participants were only administered the screening interview. Participants with elevated scores on any of the screen items were administered the full ADHD section.

Parents were assessed similarly with an observer ADHD rating scale (CAARS-O:SV; 1999), typically filled in by their partner. The KSADS-PL was administered to all parents, who were, if possible, interviewed together with their partner.

Of the Conners' ADHD questionnaires the following scales were used:

- DSM Inattentive behaviour
- DSM Hyperactive/Impulsive behaviour
- DSM Total

For all participants using medication, ratings were done of the participant's functioning off medication.

The diagnostic algorithm

The diagnostic algorithm applied to all participants was based on a combination of symptom counts on the ADHD rating scales and the KSADS-PL, both providing operational definitions of each of the 18 behavioural symptoms of ADHD defined by the DSM V. Combined counts for each symptom were determined based on the KSADS-PL scores combined with scores on either the teacher rating scale (for participants <18 years), the self-report (for participants ≥18), or the observer rating (for parents).

Based on the algorithm, participants were given either an 'affected' (ADHD diagnosis) status or 'unaffected' status. The following criteria were used to classify ADHD ('affected' status):

- Combined symptom count of ≥ 6 symptoms of inattentive or hyperactive/impulsive behaviour
- T-score ≥ 63 on at least one of the ADHD subscales on at least one of the available Conners' ADHD rating scales
- Age of onset before 12
- Symptoms cause clinical impairment
- Symptoms are not better accounted for by another disorder

For participants ≥18 years and parents, criteria were slightly adapted, such that a combined symptom count of 5 symptoms and age of onset before 12 years were sufficient for an 'affected' status.

Participants were labelled 'unaffected' if they received a T<63 on each of the scales of the Conners' rating scales, and if they had ≤ 3 symptoms (or ≤ 2 symptoms for participants of ≥18 years and parents), derived from the combined symptom counts.

For analysis purposes, participants who did not meet criteria for either affected or unaffected status, were labelled 'subthreshold ADHD'.

References

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

Conners, C. K., Erhardt, D., & Sparrow, E. P. (1999). *Conner's Adult ADHD Rating Scales: CAARS*: Multi-Health Systems, North Tonawanda, NY.

Conners, C. K., Sitarenios, G., Parker, J. D. A., & Epstein, J. N. (1998a). The revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *Journal of abnormal child psychology, 26*(4), 257-268.

Conners, C. K., Sitarenios, G., Parker, J. D. A., & Epstein, J. N. (1998b). Revision and restandardization of the Conners Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *Journal of abnormal child psychology, 26*(4), 279-291.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (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*(7), 980-988.

Strengths and Difficulties Questionnaire – Emotion Subscale Scores.

Table S3 gives an overview of the mean scores per item on the SDQ-E. As can be seen from this table, Item "Often unhappy, down-hearted, tearful" has the lowest mean score. This item relates more to depression than the other items. Low scores on this item is in accordance with Bekker *et al.* who found that individuals with ADHD who had elevated scores on the SDQ-E (4+) have almost nine times more chance of developing a comorbid anxiety disorder than those with normal scores (0-3). Odds-ratios for depression based on this scale did not differ significantly from one.

Table S3. An overview of the items of the Strengths and Difficulties Questionnaire – Emotion subscale, and average scores of the sample on a scale of 0-2.

Item	Average score (SD)
Often gets headaches, stomach-aches, sickness	0.87 (0.72)
Worries a lot	1.32 (0.59)
Often unhappy, down-hearted, tearful	0.40 (0.59)
Nervous in new situations, loses confidence	1.04 (0.74)
Has many fears, is easily scared	0.53 (0.66)

We based our score on a combination of parent and child informants, in order to be maximally sensitive to often underreported anxiety symptoms. In Figure S1, besides the histogram of this composite score, we present histograms of the scores separately per informant to enable comparisons with other studies. The means and standard deviations are given in the caption.

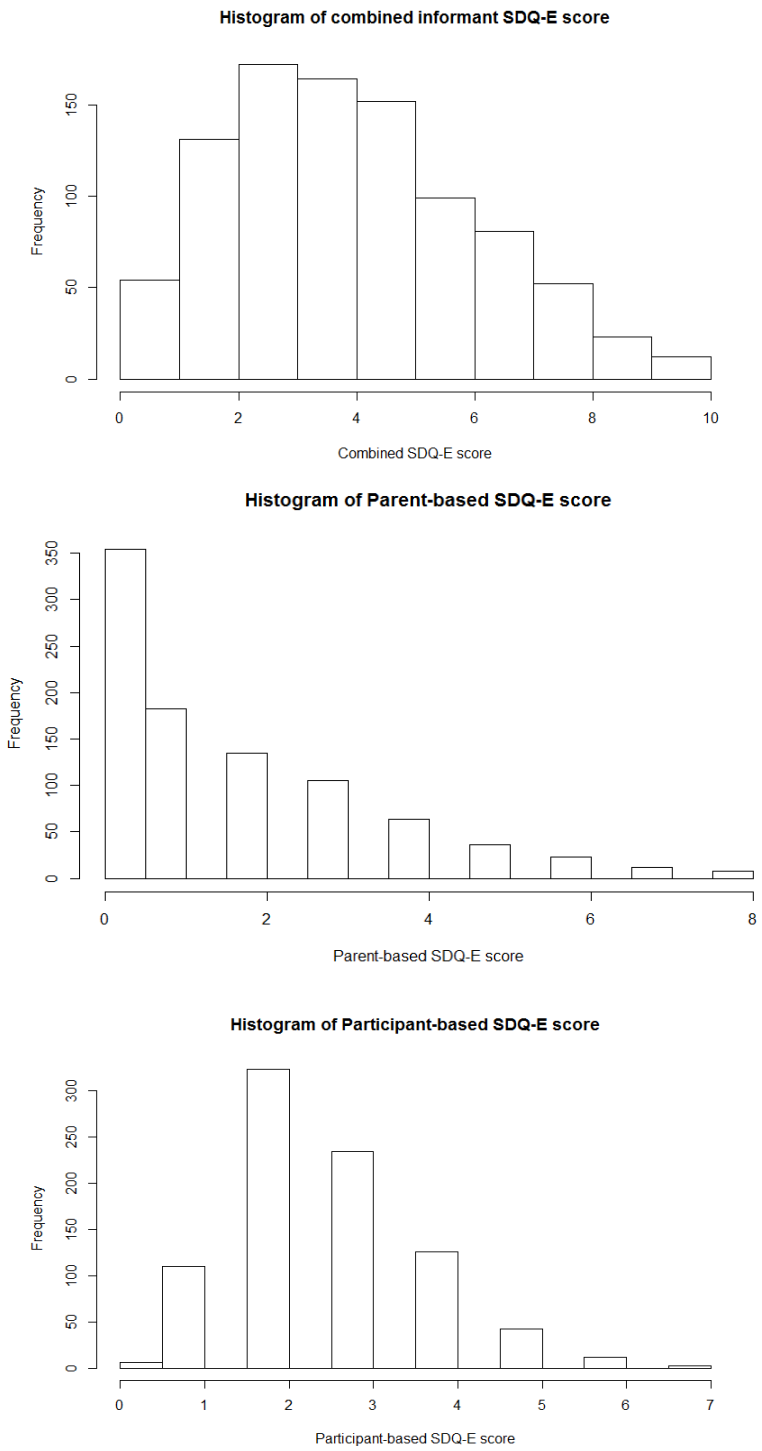

Figure S1. Top histogram shows the distribution of scores on the Strengths and Difficulties Questionnaire-emotion subscale (SDQ-E), when based on both parent and self-report. Mean score was 4.17, standard deviation (SD) 2.06. The middle histogram shows the distribution of scores, when only based on parent-report (mean 1.47, SD 1.77). The bottom histogram shows the distribution of scores, when only based on self-report (Mean 2.51, SD 1.10).

Sensitivity Analyses

Effect of location

In order to test whether our findings were different between testing locations (Amsterdam vs. Nijmegen), we added an interaction term with the 'location' covariate (Amsterdam coded as '0', Nijmegen as '1') to the model and analyzed its effect on the suprathreshold clusters from the main analysis. As can be seen in Table S4, there were no significant interaction effects with location in any of the clusters. As in the main analysis, the effect of ADHD severity on neural activity remained significant for the cluster in the frontal pole (a) and the effect of ADHD co-occurring anxiety remained significant for the other clusters (b-d).

Table S4. Results from the clusters found to be significant in the main analysis, with an additional interaction term between location and our predictors of interest (ADHD, anxiety, and their interaction).

a) *Cluster in the frontal pole found for mean working memory fMRI contrast, with a significant effect of ADHD severity*

	Coefficient	SE	T-value	P-value
ADHD	-0.34	0.06	-5.33	<.00001
Anxiety	0.08	0.06	1.28	0.2
Location	0.06	0.12	0.49	0.63
ADHD * anxiety	0.00	0.04	-0.01	0.99
ADHD * location	0.15	0.1	1.41	0.16
Anxiety * location	0.09	0.09	0.92	0.36
ADHD * anxiety * location	0.02	0.07	0.28	0.78

b) *Cluster in the cerebellum found for mean working memory fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	-0.15	0.06	-2.42	0.02
Anxiety	0.09	0.06	1.54	0.12
Location	0.05	0.12	0.43	0.67
ADHD * anxiety	-0.22	0.04	-5.11	<.00001
ADHD * location	0.13	0.10	1.28	0.20
Anxiety * location	-0.04	0.09	-0.45	0.66
ADHD * anxiety * location	0.00	0.07	0.03	0.97

c) *Cluster in the left basal ganglia found for the memory load fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	0.05	0.06	0.79	0.43
Anxiety	0.01	0.06	0.09	0.92
Location	-0.03	0.11	-0.26	0.79
ADHD * anxiety	-0.35	0.05	-7.63	<.00001
ADHD * location	0.06	0.10	0.62	0.54
Anxiety * location	0.08	0.10	0.86	0.39
ADHD * anxiety * location	0.12	0.07	1.70	0.09

d) *Cluster in the right basal ganglia found for the memory load fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	-0.06	0.06	-0.96	0.34
Anxiety	-0.01	0.06	-0.17	0.86
Location	0.02	0.12	0.16	0.87
ADHD * anxiety	-0.38	0.04	-8.41	<.00001
ADHD * location	0.23	0.10	2.24	0.03
Anxiety * location	0.04	0.09	0.44	0.66
ADHD * anxiety * location	0.10	0.07	1.42	0.16

Effect of medication

In order to test whether our findings were influenced by stimulant medication use, we added a 'treatment duration' covariate (number of days on which stimulant medication was prescribed to the participant) to the model and analyzed its effect on the suprathreshold clusters from the main analysis. As practically only participants with ADHD have a treatment duration above zero, we ran this analysis only in the subset of individuals with an ADHD diagnosis. As can be seen in Table S5, while treatment duration did have a significant effect on brain activity in some of the clusters, the regression coefficient and p-value of the predictor of interest (ADHD severity for cluster in the frontal pole (a), and ADHD severity * anxiety severity for the other clusters (b-d)) remained significant.

Table S5. Results from the clusters found to be significant in the main analysis, with an additional 'treatment duration' covariate.

a) *Cluster in the frontal pole found for the mean working memory fMRI contrast, with a significant effect of ADHD severity*

	Coefficient	SE	T-value	P-value
ADHD	-0.31	0.09	-3.48	.0006
Anxiety	0.19	0.09	2.23	0.03
Treatment duration	-0.23	0.09	-2.66	0.008
ADHD * anxiety	-0.23	0.08	-3.03	.003

b) *Cluster in the cerebellum found for the mean working memory fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	-0.14	0.07	-1.85	.07
Anxiety	-0.17	0.07	-2.36	.02
Treatment duration	-0.50	0.07	-7.04	<.00001
ADHD * anxiety	-0.37	0.06	-5.98	<.00001

c) *Cluster in the left basal ganglia found for the memory load fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	0.04	0.08	0.51	.61
Anxiety	-0.43	0.08	-5.46	<.00001
Treatment duration	-0.13	0.08	-1.65	0.10
ADHD * anxiety	-0.27	0.07	-3.93	.0001

d) *Cluster in the right basal ganglia found for the memory load fMRI contrast, with a significant effect of the interaction between ADHD severity and anxiety level*

	Coefficient	SE	T-value	P-value
ADHD	0.05	0.08	0.61	.55
Anxiety	-0.46	0.07	-6.17	<.00001
Treatment duration	-0.21	0.07	-2.85	.005
ADHD * anxiety	-0.23	0.07	-3.58	.0004

MRI data acquisition and preprocessing

All subjects were scanned with either a Siemens MAGNETOM Sonata 1.5 Tesla (at VU UMC in Amsterdam) or a Siemens MAGNETOM Avanto 1.5 Tesla (at Donders Centre for Cognitive Neuroimaging in Nijmegen) MRI scanner (Siemens, Erlangen, Germany), using identical protocols. Four functional runs were acquired (GE EPI, TR=2340 ms, TE=40 ms, FOV=224x224 mm, voxel size=3.5x3.5x3.0 mm, 38 slices, 107 volumes per run). For spatial localization and normalization, we included each participant's high resolution T1 scan (MPRAGE, TR=2730 ms, TE=2.95 ms, TI=1000 ms, voxel size=1x1x1 mm, FOV=256x256 mm, 176 slices).

We made use of FSL FEAT (FMRI Expert Analysis Tool; FMRIB Analysis group, Oxford, UK) for preprocessing, which consisted of removal of the first three volumes of each run, within-run motion correction to the middle volume, slice-timing correction, spatial smoothing with a 6 mm Gaussian kernel, high-pass temporal filtering (0.01 Hz), and transformation to the participant's T1 anatomical image using linear boundary-based registration. Runs with more than 3 mm absolute displacement were excluded from analysis, and individuals with less than three runs were excluded.

Starting out with 427 participants with complete behavioral and MRI data, 14 participants were excluded from analysis due to incidental findings after visual inspection (e.g. enlarged ventricles or unexpected hypo-intensities), 12 due to scan quality (e.g. artefacts, missing volumes, insufficient coverage of the entire brain), 12 due to below chance performance on the task, 18 due to excessive motion in multiple runs, and one subject was excluded due to extreme outliers on BOLD activation parameters. This led to our final sample size of N=371.

Whole-brain task activation

Figure S2 shows whole-brain task activation of all participants combined for the mean working memory contrast, thresholded at z=6 to break up the otherwise very large clusters. Table S6 provides details on these clusters. Figure S3 and Table S7 provide this same information for the memory load contrast.

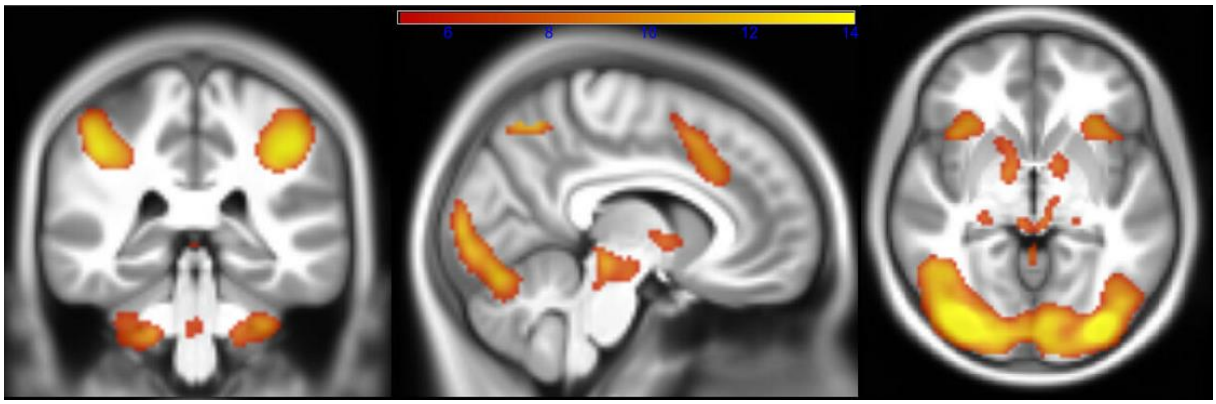

Figure S2. Task activation, thresholded at z=6 across all participants, for the mean visuospatial working memory contrast. Statistical image is overlain on the MNI152 template at X=6, Y=-36, Z=0.

Table S6. Info on the location and size of the clusters significant at P<.01, family-wise error corrected, for the mean working memory contrast. Note: X, Y, Z coordinates are in MNI-space in mm, and represent the peak of the cluster. The anatomical labels are according to the Harvard-Oxford atlas. MNI=Montreal Neurological Institute.

Regions	*X*	*Y*	*Z*	*Size*
Lateral Occipital Cortex, Cerebellum	34	-86	8	27981
Superior Frontal Gyrus, Middle Frontal Gyrus	28	-4	56	1675
Thalamus, Pallidum	8	-26	-10	1138
Insular Cortex, Frontal Orbital Cortex	32	22	4	538
Insular Cortex, Frontal Orbital Cortex	-32	22	2	515
Thalamus, Caudate, Putamen, Pallidum	12	0	0	321

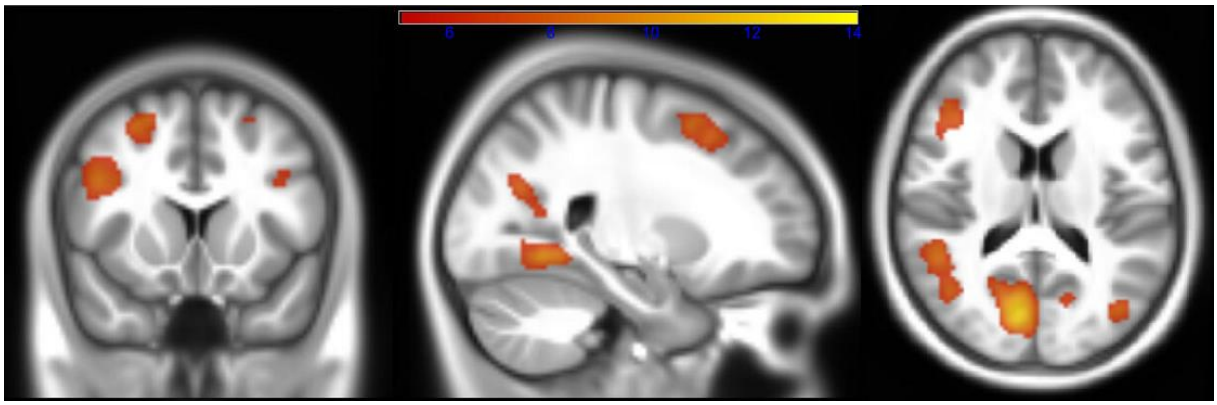

Figure S3. Task activation, thresholded at Z=6 across all participants, for the memory load contrast. Statistical image is overlain on the MNI152 template at X=22, Y=-14, Z=18.

Table S7. Info on the location and size of the clusters significant at p<.01, family-wise error corrected, for the memory load contrast. Note: X, Y, Z coordinates are in MNI-space in mm, and represent the peak of the cluster. The anatomical labels are according to the Harvard-Oxford atlas. MNI=Montreal Neurological Institute.

Regions	*X*	*Y*	*Z*	*Size*
Precuneus Cortex	12	-68	-2	3267
Middle Frontal Gyrus, Inferior Frontal Gyrus	46	16	28	925
Precuneus Cortex	6	-50	50	525
Superior Frontal Gyrus	26	12	54	496
Middle Temporal Gyrus	50	-44	16	398
Lateral Occipital Cortex	-40	-74	14	171