SUPPLEMENTS

SECTION 1: SAMPLE COMPOSITION DETAILS

This study is part of our larger RCT project that examined individual differences in executive functioning in college students with ADHD and their peers as well as the effectiveness of computerized working memory treatment. Two major samples can be distinguished: the earlier Cogmed sample (e.g., Gropper et al., 2014) and the later Engage sample (which added the low intensity treatment arm and focused more on ADHD). The present study, which used the visuospatial working memory capacity (VWMC) task, uses the Engage sample. The Engage sample, can be broken down into two independent subsamples: the pilot-Engage sample (see Mawjee et al., 2017) and the main Engage sample (see Mawjee et al., 2015, for more detail). For the present analysis of neural and behavioral data from the change detection task before and after CMWT, we had planned *a priori* to combine participants from two Engage samples to maximize statistical power in finding neural effects and create a buffer for any potential data loss due to EEG artifacts.

Participants from both Engage samples were postsecondary education students with ADHD registered with disability services that underwent the exact same CMWT treatment program, and used identical outcome measures. The VWMC-task was left unchanged across these samples. The Engage pilot sample was intended to optimize procedures, make potential changes to, and estimate required sample sizes for, the larger Engage sample. Differences between samples were minor and consisted of improved standardization of the coach calls for the waitlist group, the recommended use of a planner to reduce study attrition, and the addition of

post-training interviews (see also, Mawjee et al. 2017, for details). The pattern of the results and the overall conclusion was similar after excluding the pilot sample from the analysis.

In the context of treatment with neural outcomes, these samples were also combined in the previously published Liu et al., 2016, and Liu et al. 2017 studies. Liu et al., 2016, examined neural correlates of working memory through examining alpha power and in a task unable to determine effects of attentional control. Liu et al., 2017, examined neural effects of transfer, i.e., whether inhibitory control is affected after intense working memory training. The present study, which was considered our most comprehensive study, utilizes the VWMC-task and examines effects of WM-training on the CDA waves (both tapping into effects of attention and capacity).

Gropper, R. J., Gotlieb, H., Kronitz, R., & Tannock, R. (2014). Working memory training in college students with ADHD or LD. Journal of Attention Disorders, 18(4), 331-345.

Liu, Z., Glizer, D., Tannock, R., & Woltering, S. (2016). EEG alpha power during maintenance of information in working memory in adults with ADHD and its plasticity due to working memory training: a randomized controlled trial. *Clinical Neurophysiology*, *127* (2), 1307-1320.

Liu, Z. X., Lishak, V., Tannock, R., & Woltering, S. (2017). Effects of working memory training on neural correlates of Go/Nogo response control in adults with ADHD: A randomized controlled trial. *Neuropsychologia*, *95*, 54-72.

Mawjee, K., Woltering, S., & Tannock, R. (2015). Working memory training in post-secondary students with ADHD: a randomized controlled study. *PloS one*, *10*(9), e0137173.

Mawjee, K., Woltering, S., Lai, N., Gotlieb, H., Kronitz, R., & Tannock, R. (2017). Working memory training in ADHD: controlling for engagement, motivation, and expectancy of improvement (pilot study). *Journal of attention disorders*, *21*(11), 956-968.

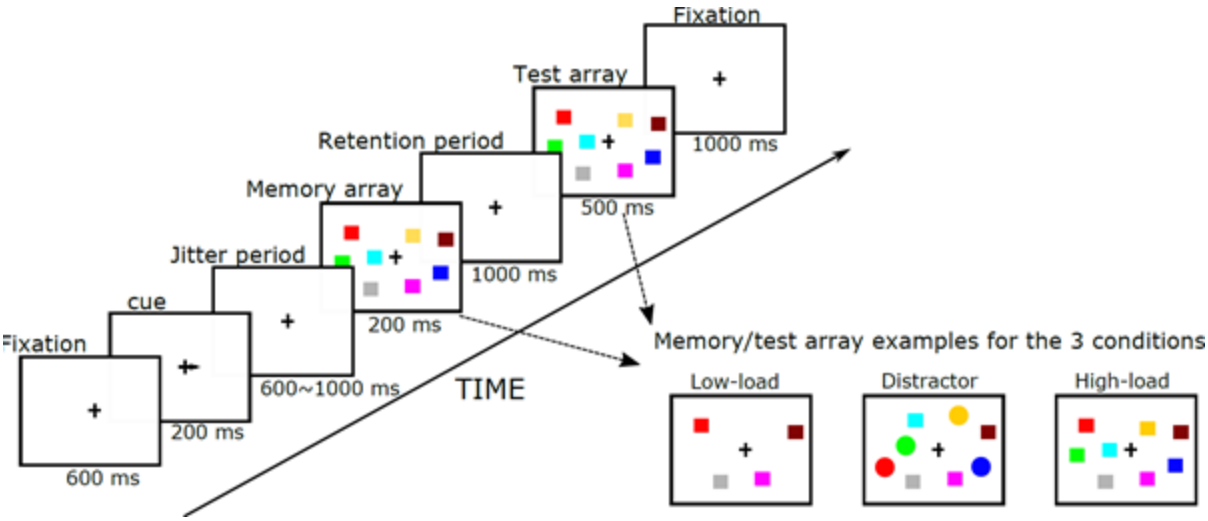
SECTION 2: CHANGE DETECTION TASK DETAILS

Participants were seated in front of a 17-inch VGA monitor at a distance of approximately 80 cm. Our paradigm had several phases, starting with a fixation, a cue, a jitter period, a memory array, retention stage, and a test array, which are all shown in Figure S1. The fixation cross appeared for 600 ms in the center of the screen after which a cue appeared for 200 milliseconds in the form of an arrow pointing either to the left or right (50% chance) to inform participants as to which side of their visual field (i.e., left or right hemifield) they should pay attention. After a 400 – 900 ms jitter period, a memory array appeared of 2 or 4 colored shapes in both hemifields for 200 ms in which participants were required to only memorize the colored squares in the cued hemifield. After the memory array disappeared, there was a 1000 ms retention period during which participants needed to maintain the memory array in their WM. Then, the test array appeared and participants were asked to indicate whether or not the test array was identical to the memory array by clicking one of two keys on a keyboard using the index and middle fingers of their dominant hand. On half of the trials, the test array was identical to the memory array, but on the other half, the color of one shape on the attended side was changed. As soon as the response was made, the test array disappeared and the current trial ended. If participants did not make a response within 1500 ms, the trial ended automatically. The test array

was shown for 500ms after which a fixation appears for 1000 ms (for task protocol, see Figure S1).

The task had 3 conditions: in the low memory load condition (LL), the memory array contained 2 colored squares in each hemifield; in the high memory load condition (HL), the memory array contained 4 colored squares. In the distraction condition (DL), the memory array contained 2 colored squares and 2 colored circles (see also Figure S1). Participants were required to only memorize the squares (targets) in the attended hemifield in each condition and ignore the circles (distractors) in the distraction condition. For each condition, these memory items, i.e., squares (size:0.57°× 0.57°) and circles (diameter: 0.57°), were randomly presented within 5.7°× 8.53° rectangular regions that were centered 2.86°4 cm to the right and left side of the fixation cross. The color of each item was randomly chosen from 10 easily discriminable colors and a given color only appeared once in the attended hemifield of memory and test arrays. The task had 312 trials, with 104 trials for each condition. Trials of different conditions and of different hemifield-attended sides were randomly mixed and divided within blocks. There were 13 blocks with 24 trials in each block. At the end of each block, participants' accuracy, as well as progress through the task, was shown.

Figure S1: Simplified flowchart of the visuospatial change detection task with distractors.



Special care was taken with task instructions considering the nature of our study population. Instructions were presented on several subsequent screens and read aloud by a research assistant who would verify that the participant understood the instructions. For example, participants had to repeat verbally that they were to ignore the circles in the distractor condition. Participants were also told that it was important to keep their eyes focused on the fixation cross in the center of the screen for the duration of the trial and to direct their attention, and not their eyes, to the array on the side they were instructed to attend to. Also, we asked them to sit in a comfortable, relaxed position and minimize eye blinks, verbalizations, and movements as much as they could. During the practice block, research assistants were actively checking for eye-movements and would correct behavior until participants understood and complied when acting out instructions.

SECTION 3: EEG PROCESSING DETAILS

EGI's Netstation software package was used to filter (.05 -30Hz) and segment the data for correct trials (400 ms before stimulus onset memory array and 1000 ms post stimulus). Blinks, vertical and horizontal eye-movements were measured by bipolar electrodes placed above and below the left eye and at the outer canthi of both eyes. Data were transferred to Matlab 9.1 (The Mathworks, Inc.) for further processing. EEGLAB, an open source analysis tool, was used to preprocess the recorded ECG data. Bad channels were detected manually and replaced by interpolation function provided in EEGLAB. ICA functions from EEGLAB were executed to decompose EEG data into several independent components (ICs). The noisy components were firstly located by the SemiAutomatic Selection of Independent Components for Artifact correction (SASICA) plugin for EEGLAB and then manually reviewed and subtracted from the preprocessed data.

Table S1 shows the trial counts separated by group, session, and condition. A mixed model 2 (Time: pre, post) by 3 (Group: Standard, short and waitlist) by 3 (Condition: Low, High, Distractor load) repeated measures ANOVA found no differences in trial count between main and interaction effects of Group and Time. There was a main effect of condition ($p < .001$) whereby the low load had higher trial counts than the distractor load which, in turn, had higher trial counts than the high load. Trial count was not a significant factor when it was added as a covariate in the main analyses.

Table S1. Average number of trial counts and standard deviation in neural task before and after training for each training group per condition.

Trial counts	Pre-training	Post-training
Standard-length group		
Low load	74.2 (8.9)	76.0 (12.3)
High load	64.3 (9.7)	63.9 (11.9)
Distractor load	71.4 (9.3)	71.3 (12.7)
Shortened-length group		
Low load	72.6 (7.9)	75.0 (8.7)
High load	63.3 (9.2)	66.4 (9.2)
Distractor load	68.4 (8.4)	72.5 (8.6)
Waitlist group		

Low load	72.3 (8.1)	74.2 (8.9)
High load	65.0 (8.5)	63.1 (8.6)
Distractor load	69.3 (8.5)	69.3 (12.5)

After average-referencing, lateral-posterior sites were chosen to calculate the CDA. This selection of electrodes was based on previous studies (e.g., Woodman & Luck, 1999) and inspection of the grand average waveform of all subjects. As indexed using the standard EGI system nomenclature, electrode sites 52, 51, 59, 66, 61, 60, and 65 on the left hemisphere and electrode sites 92, 97, 91, 84, 78, 85, and 90 on the right hemisphere were selected (see figure S2). First, for each condition and at the 14 electrode sites, ERPs were calculated separately for left and right hemifield-attended trials. Next, ERPs on the right hemisphere electrode sites were subtracted from the left sites when participants were cued to memorize the right hemifield of memory arrays; and the ERPs on the left hemisphere sites were subtracted from the right sites when participants were cued to the left hemifield. At this point, for each condition and at each electrode site, there was a separate CDA for attended left and attended right hemifield. Then the two sides of the CDA were averaged to produce a single CDA waveform. This procedure was repeated for each of the 3 conditions at each electrode site. Finally, waveforms from different electrode sites were averaged to produce a final CDA waveform for each condition.

Figure S2: Electrode site selection for the CDA