**The Key Role of Daytime Sleepiness in Cognitive Functioning of Adults with ADHD**

Bartosz Helfer1\*, Natali Bozhilova2, Ruth E. Cooper2,3, Joanna Ismene Douzenis2, Stefanos Maltezos2,4, Philip Asherson2

1 National Heart and Lung Institute, Faculty of Medicine, Imperial College London, UK

2 Institute of Psychiatry, Psychology and Neuroscience, King’s College London, UK

3 Newham Centre for Mental Health, Unit for Social and Community Psychiatry, Blizard Institute, Queen Mary University of London, UK

4 Adult ADHD Service, South London and Maudsley NHS Foundation Trust, UK

\* bartosz.helfer@gmail.com

**Online Supplemental Material**

**Supplemental Methods**

### The Sustained Attention to Response Task (SART)

In the SART task (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997), a type of a go/ no-go task, participants looked at a computer screen where a string of 225 single digits, from one to nine, appeared in random order. Each of the single digits was shown 25 times. Participants were requested to push a response button when they saw any digit, apart from when they saw the target digit three, in which case they should have ceased the response (proportion of go to no-go stimuli was 8:1). Subjects used their preferred hand. This procedure was repeated over three blocks of five minutes each. Digits were presented above a central fixation cross for 150ms, with an inter-stimulus interval of 1000ms. Task duration was 15min divided in 5min blocks.

### Cued Continuous Performance Task with flankers (CPT-OX)

In the CPT-OX task(McLoughlin et al., 2010) participants looked at a computer screen where a string of 400 black letters was centrally presented, including cue letter ‘O’, target letter ‘X’, and distractors ‘H’, ‘B’, ‘C’, ‘D’, ‘E’, ‘F’, ‘G’, ‘J’, and ‘L’. Letters were flanked on either side by the letters ‘X’ or ‘O’, and cue and target letters (O and X) were flanked by the incompatible letter (X and O). Participants were instructed to ignore the flanker letters and respond as quickly as possible to cue and target letters (O and X). 80 cues (O) were followed by the target letter (X) in 40 trials (go condition), and neutral distractors in the remaining trials (no-go condition). In 40 trials a letter X was not preceded by a cue ‘O’ and was supposed to be ignored. Letters were presented every 1650ms for 150ms in pseudo-random order. The task duration was 11min and it was administered just after the SART.

Here, we used not only the SART task, but also the CPT-OX data, even though no sleepiness measurement was taken during the latter task. The rationale was that CPT-OX was administered directly after SART (only with a short break), so the on‑task sleepiness would have been similar or develop even further in the next minutes.

### The Observer‑Rated Drowsiness Scale (ORD)

A systematic assessment protocol (Wiegand, McClafferty, McDonald, & Hanowski, 2009), including the ORD scale(Wierwille & Ellsworth, 1994), was used to evaluate participants sleepiness during the three blocks of the SART. Sleepiness was assessed on a continuous scale from 0-100 and divided into six categories with detailed behavioural descriptions (ranging from “not drowsy” to “extremely drowsy”). An ORD Behaviour and Mannerism Checklist (Wiegand et al., 2009) was additionally provided to support quantitative evaluation. This included a list of items across five categories: eyes/ eyebrows (with items such as “blank/ fixed stare”, “squinting”, “excessive/ hard blinking” or “slow closure”); mouth (“yawning” or “tongue motion”); body (“slumping/ slouching/ leaning”, “sighing” or “slack muscle tone”); face (“facial contortions”); and neck/ head (“leaning”, “holding”, “position change” or “nodding/ drooping”).

The last minute of each of the three SART blocks for each participant was extracted to reduce data quantity. Recordings from eleven participants were incomplete, damaged or missing, so a total of 100 participants videos were evaluated (that is, 300 one‑minute recordings). A systematic data extraction and assessment protocol was developed to ensure reliability of the observer-rated sleepiness data. Two observers underwent a training session before commencing the independent evaluation of the videos. All extracted recordings were randomised and anonymised to avoid bias. All observers were blind to the diagnostic status. Once the two raters provided ratings for all the videos the data was collated. If the ratings fell into the same category (that is “not sleepy”: 0 – 12.49 points; “slightly sleepy”: 12.5 – 37.49 points; “moderately sleepy”: 37.5 – 62.49 points; “very sleepy”: 62.5 – 87.49 points; or “extremely sleepy”: 87.5 – 100 points) an average of the two scores was taken. If not, a third observer made another blind independent assessment and an average of all three observers’ scores was taken. This procedure was repeated separately for t1, t2 and t3 scores, which were then averaged to form and overall sleepiness score for each study participant. Inter-rater reliability was established by correlating the total score for each participant between observer one and two (Wiegand et al., 2009).

### Quantitative EEG Slowing

EEG data from the resting-state eyes-open baseline condition was used for this analysis. The data was recorded at the beginning of the EEG session using a 62-channel direct-current-coupled recording system called ActiCap in 10-20 montage (Brain Products, Munich, Germany). Electrode impedance was below 10 kΩ. Data were pre-processed in Brain Vision Analyzer 2.0 (Brain Products, Munich, Germany).

Initially, the recording was resampled to 256Hz, screened manually and selected channels were interpolated. Overload regions were removed, and the Independent Component Analysis (ICA) was run. Data was epoched into 2 second segments and components were selected manually for exclusion. The data was then un-epoched and the ocular channels were removed before the data was manually inspected again to remove any remaining noise. Fast Fourier Transform (FFT) Analysis was used and the grand-average of the FTT was computed for each group. EEG frequency power values were standardised by mean centring. A z-score of 3 or above was considered an outlier. EEG power density within alpha (7.5–12.5Hz), beta (12.5–30Hz), theta (4‑7.5Hz) and delta (0.5–3.5Hz), frequency bands was averaged across the frontal (Fz, F1, F2, F3, F4, F5, F6, F7, F8), central (Cz, C1, C2, C3, C4, C5, C6) and parietal (Pz, P3, P4, P7, P8) EEG scalp regions for each study participants. An additional global measure for each frequency band was calculated by averaging across these regions. EEG slowing was calculated as a ratio of the slow frequency bands (delta + theta) to the fast frequency bands (alpha + beta) for each scalp region as well as across regions. This spectral calculation of EEG slowing ensures that the ratio stays sensitive to both simultaneous changes across bands as well as in any individual frequency (Petit, Lorrain, Gauthier, & Montplaisir, 1993).

EEG data were recorded at the beginning of each cognitive testing session. Therefore, they represent baseline sleepiness related to individual and/ or psychopathological differences in sleepiness, rather than the effect of any of the tasks. We consider this beneficial for the analysis as otherwise the data would be additionally confounded by the sleepiness developing during the cognitive testing session or EEG noise resulting from cognitive processing. The existing qEEG measures used in psychiatry are controversial, as both specific frequency bands and scalp regions has been linked to multiple processes; and because of the high heterogeneity of these processes, as well as psychiatric disorders itself. This is one reason why we decided to use EEG slowing in this analysis, as it is rather robustly linked to both normal sleepiness and the psychopathology of ADHD.

### Supplemental Analysis and Results

All data are reported as means ± SD unless indicated otherwise. The two-tailed alpha was set at 0.05 and Bonferroni‑corrected for multiple comparisons when necessary. For the estimation of effect sizes standard 95% confidence intervals were used when appropriate. Up to eighteen participants per model (16%) had to be removed due to a combination of missing data in the included measures.

### Observer‑rated sleepiness

To investigate the development of the observer‑rated sleepiness over the three 5-minute time‑points during SART (t1, t2, t3) in the ADHD and neurotypical group we used 3x2 mixed ANOVA with time as a within-subject factor and group as the between‑subject factor. Inter‑rater reliability between observer one and two was high, *r*(101) = 0.74, p < 0.001. The on-task sleepiness data were log‑transformed to conform to normality. No outliers were identified, as assessed by examination of studentized residuals for values greater than ±3. Sleepiness scores at t1, t2 and t3 were normally distributed, as assessed by Normal Q-Q Plots. There was homogeneity of variances, as assessed by Levene's test of homogeneity (p > 0.05) and homogeneity of covariances, as assessed by Box's test of equality of covariance matrices (p > 0.001). Mauchly's test of sphericity indicated that the assumption of sphericity was violated for the two-way interaction, χ2(2) = 25.534, p = 0.001. As the estimated epsilon was greater than 0.75 (ε = 0.8) the Huynh-Feldt correction was used (Maxwell & Delaney, 2003).

As most study participants were not visibly sleepy during the cognitive testing procedure and because the reliability of the observer‑rated sleepiness tool is higher at the higher levels of sleepiness, we created an additional group for participants falling into the “at least slightly sleepy” category (sleepiness scores 12.5 and higher, see Supplement) to further investigate how sleepiness relates to cognitive performance and results from clinical scales and to be able to compare the results of the more sleepy subjects with those who: 1) are not sleepy and are diagnosed with ADHD (group 2); and 2) are not sleepy and have no ADHD (group 3). The “at least slightly sleepy” group consisted of six neurotypical people and 25 people with ADHD (this higher prevalence is in line with the previous analysis showing higher levels of sleepiness in the ADHD group). For the selected clinical scales and cognitive measures, we run general linear models to determine the effect of sleepiness on error rates, reaction time variability, ADHD symptom severity and sleep quality. For the analysis of cognitive variables, we additionally controlled for IQ.

After adjustment for IQ, there was a statistically significant difference between the three groups in the following cognitive measures:

* Omission errors in the SART, F(2, 94) = 10.924, p < 0.001, partial η2 = 0.189; with highest error rates in in the sleepy group (OE = 5.43, SE = 0.65; 95% CI 4.17 to 6.76), followed by the ADHD group (OE = 2.45, SE = 0.55; 95% CI 1.34 to 3.65), followed by neurotypical group (OE = 1.09, SE = 0.74; 95% CI -0.38 to 2.56);
* Omission errors in the CPT-OX, F(2, 93) = 4.369, p = 0.015, partial η2 = 0.086; with highest error rates in the sleepy group (OE = 2.12, SE = 0.40; 95% CI 1.32 to 2.91), followed by the ADHD group (OE = 1.77, SE = 0.34; 95% CI 1.11 to 2.44), followed by neurotypical group (OE = 0.43, SE = 0.45; 95% CI -0.47 to 1.32);
* Reaction time variability in SART, F(2, 94) = 3.594, p = 0.031, partial η2 =0 .071; with highest variability in the sleepy group (RTV = 88.0, SE = 7.7; 95% CI 72.7 to 103.3), followed by the ADHD group (RTV = 71.4 SE = 6.6; 95% CI 58.4 to 84.4), followed by neurotypical group (RTV = 56.3 SE = 4.2; 95% CI 47.9 to 64.7).

Commission errors in both tasks and RTV in the CPT-OX were not statistically different between groups (p > 0.05).

A multiple regression was run to predict omission errors in the SART from sleepiness, EEG‑slowing and ADHD symptom severity. There was independence of residuals, as assessed by a Durbin-Watson statistic of 2.01. There was no evidence of multicollinearity, as assessed by tolerance values greater than 0.1. The assumption of normality was met, as assessed by a Q-Q Plot.

### EEG Slowing

A mixed 3x2 ANOVA was used to investigate the effect of EEG slowing, with scalp location (frontal, parietal, central) as within‑subject factor and group (ADHD vs neurotypical) as between‑subject factor.

To aggregate differences in EEG slowing across the different scalp regions and reduce the data into a single measure, a global EEG slowing ratio was calculated (see Supplementary Methods) and used for further analyses. An independent-samples t-test was run to confirm the case-control difference in global EEG slowing. Inspection of the boxplot revealed an outlier in the neurotypical group, however we decided not to remove or modify the data as the value was well within the range of the ADHD group and most likely reflected true individual differences. Data were normally distributed as assessed by Normal Q-Q Plot. The assumption of homogeneity of variances was violated, as assessed by Levene's test for equality of variances (p < .001).

Binomial logistic regression was used to investigate the effect of EEG slowing on ADHD diagnostic status and linear regression to investigate the effect of EEG slowing on ADHD symptom severity.

EEG slowing and ADHD diagnosis: Box-Tidwell procedure (Box & Tidwell, 1962) for assessing linearity of the continuous variable with respect to the logit of the dependent variable yielded a negative result. To correct this problem, the EEG slowing variable was log‑transformed. No outliers were identified using studentized residuals. The logistic regression model was statistically significant, χ2(4) = 4.112, p = 0.042, explained 5% (Nagelkerke R2) of the variance and correctly classified 76.5% of cases. EEG slowing was a statistically significant predictor of ADHD, p = 0.047.

EEG slowing and ADHD symptom severity: Linearity was established by visual inspection of the scatterplot. Residuals were normally distributed as assessed by visual inspection of a normal probability plot. EEG slowing accounted for 4% of the variation in ADHD symptom severity with adjusted R2 = 3%, a small size effect according to Cohen (Cohen, 1988). EEG slowing statistically significantly predicted ADHD symptom severity, F(1, 100) = 3.978, p = 0.049.

**Supplemental References**

Box, G. E., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics, 4*(4), 531-550.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd. In: Hillsdale, NJ: erlbaum.

Maxwell, S. E., & Delaney, H. D. (2003). *Designing experiments and analyzing data: A model comparison perspective*: Routledge.

Petit, D., Lorrain, D., Gauthier, S., & Montplaisir, J. (1993). Regional spectral analysis of the REM sleep EEG in mild to moderate Alzheimer's disease. *Neurobiol Aging, 14*(2), 141-145.

Wiegand, D. M., McClafferty, J., McDonald, S. E., & Hanowski, R. J. (2009). *Development and evaluation of a naturalistic observer rating of drowsiness protocol*. Retrieved from

Wierwille, W. W., & Ellsworth, L. A. (1994). Evaluation of driver drowsiness by trained raters. *Accident Analysis & Prevention, 26*(5), 571-581.