https://doi.org/10.1016/j.neuropsychologia.2021.107954

Prefrontal activation and pupil dilation during *n*-back task performance: A combined fNIRS and pupillometry study

Michael K. Yeung[1], Tsz L. Lee[2], Yvonne, M. Y. Han[1], and Agnes S. Chan[2,3*]

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong SAR, China

[2]Neuropsychology Laboratory, Department of Psychology, The Chinese University of Hong Kong, Hong Kong SAR, China

[3]Research Center for Neuropsychological Well-being, The Chinese University of Hong Kong, Hong Kong SAR, China

[*]Address correspondence to:

Agnes S. Chan, Ph.D.

aschan@cuhk.edu.hk

Department of Psychology, The Chinese University of Hong Kong

Shatin, N.T., Hong Kong SAR

Abstract

The *n*-back task is one of the most commonly used working memory (WM) paradigms in cognitive neuroscience. Converging evidence suggests activation in the lateral prefrontal cortex (PFC) and pupil dilation [a proxy for locus coeruleus (LC) activation] during this task. However, it remains unclear whether the lateral PFC and the LC are functionally associated during *n*-back task performance. This study's aim was to examine the relationship between changes in lateral PFC activity and the pupil diameter (a proxy for LC activity) and to evaluate the effect of WM load on such relationship during the *n*-back task. Thirty-nine healthy young adults (10 males, 29 females) underwent a number *n*-back paradigm with 0- and 3-back conditions. Their prefrontal hemodynamics and changes in pupil size during task performance were simultaneously measured using a 16-channel functional near-infrared spectroscopy (fNIRS) device and a wearable eye tracker. Young adults exhibited significant activation in the bilateral lateral PFC and significant increases in pupil size when the WM load was high (i.e., 3-back) but not low (i.e., 0-back) compared with the resting period. Interestingly, significant positive correlations were found between changes in lateral PFC activity and pupil size during the 0-back task only. These correlations tended to be stronger during the 0-back than the 3-back condition. Thus, the functional relationship between the lateral PFC and the LC may vary at different load levels during the *n*-back task. Our findings have important implications for neuropsychiatric research and support concurrent fNIRS and pupillometric measurements for a better understanding of the mechanisms underlying WM processing.

Keywords: working memory, *n*-back, prefrontal cortex, near-infrared spectroscopy, pupillometry

1. Introduction

Human short-term memory has a limited capacity (Baddeley, 2012; Miller, 1956). Therefore, the ability to replace old, irrelevant information with information that is relevant to the current context is a core aspect of executive function essential for the pursuit of goals (Miyake et al., 2000). In cognitive neuroscience, the *n*-back task has been one of the most widely used working memory (WM) paradigms (Kirchner, 1958). Studies using this task have shown that accuracy and reaction time (RT) worsen with increasing WM load across ages and materials (Boisgueheneuc et al., 2006; Bopp & Verhaeghen, 2020). Additionally, meta-analyses of the positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) literature have identified activity in the bilateral dorsolateral and ventrolateral prefrontal cortex (PFC), premotor cortex, supplementary motor area, and parietal cortex that increases with increasing WM load during the *n*-back task (Mencarelli et al., 2019; Owen et al., 2005; Rottschy et al., 2012; Yaple et al., 2019). Studies using functional near-infrared spectroscopy (fNIRS) have also generated converging evidence of the involvement of the bilateral lateral PFC during this task (Ito et al., 2011; Koike et al., 2013; Kopf et al., 2013; Yeung et al., 2016, 2019). It is generally assumed that activation in various PFC subregions during the *n*-back task reflects the active operation of the cognitive processes associated with the executive control of WM, such as the storage and manipulation of internal representations, monitoring incoming information, the implementation of stimulus response mapping, and the use of organizational strategies (Mencarelli et al., 2019; Owen et al., 2005).

Accumulated evidence indicates that the pupil size of the eye increases in response to an increasing cognitive load during a variety of tasks (Beatty, 1982; Kahneman & Beatty, 1966; van der Wel & van Steenbergen, 2018). In the context of the *n*-back task, the pupil size increases with an increasing WM load (Belayachi et al., 2015; Karatekin et al., 2007;

Niezgoda et al., 2015; Scharinger et al., 2015). It is well known that the pupil size is influenced by the sympathetic and parasympathetic branches of the autonomic nervous system, which is modulated by the locus-coeruleus (LC) norepinephrine system (Joshi et al., 2016; McDougal & Gamlin, 2011; Murphy et al., 2014; Samuels & Szabadi, 2008). The LC has been implicated in the regulation of arousal and vigilance (Sara & Bouret, 2012). Therefore, the task-evoked pupillary response has been and can be employed as a proxy measure for arousal and mental effort (i.e., LC activity; Bradley et al., 2008; Mathot, 2018). In the present study, arousal refers to the degree of vigilance and alertness during periods of wakefulness, manifesting as motor activation, responsiveness to sensory inputs, emotional reactivity, and enhanced cognitive processing (Carter, de Lecea, & Adamantidis, 2013). Arousal is expected to occur during the *n*-back task because individuals have to detect incoming sensory stimuli, operate on the information, and generate an efferent response in a timely manner.

The LC extensively innervates the cerebral cortex of all hemispheric lobes, including the dorsolateral and the dorsomedial PFC (Arnsten & Goldman-Rakic, 1984; Aston-Jones & Cohen, 2005). Neurons in this small pontine nucleus are the primary source of cortical norepinephrine (Samuels & Szabadi, 2008). Thus, the LC is implicated in the regulation of arousal and mental effort across many tasks, playing a general, modulatory role in neocortical functions. In contrast, the lateral PFC is implicated in executive control and goal-directed behavior (Petrides, 2005). This region represents a primary site of WM and plays a crucial role in *n*-back task performance (Barbey, Koenigs, & Grafman, 2013; Levy & Goldman-Rakic, 2000; Tsuchida & Fellows, 2009). Although lateral PFC activation and pupil dilation have been separately observed during *n*-back task performance, whether they are functionally related to each other remains elusive. In addition, the effect of WM demand on the relationship between activity in the lateral PFC and the LC during the *n*-back task, if any,

remains unclear. A high WM load may enhance the functional integration of the two systems due to the increased general, modulatory effects exerted by the LC system on the cortex. Alternatively, it may reduce the functional coupling between the lateral PFC and the LC, or remove the co-dependence of the two systems, due to the specialized role of the lateral PFC in WM and to a LC or norepinephrine-induced increase in local neural communications within network during task performance (Eldar, Cohen, & Niv, 2013).

The aim of this study was to examine the relationship between changes in lateral PFC activity and the pupil diameter (a proxy for LC activity) during *n*-back task performance. In addition, the effect of WM load on such relationship was investigated by evaluating the correlation between the moment-to-moment changes in lateral PFC activity and the pupil size at two WM load levels (i.e., 0- and 3-back conditions). A significant positive correlation between the observed [oxy-Hb] changes and the [oxy-Hb] changes predicted from the pupil signal would suggest that lateral PFC activity and pupil dilation were subtended by common mechanisms. In contrast, a nonsignificant zero correlation between the observed and the predicted [oxy-Hb] changes would indicate that the two proxies might be underpinned by distinct mechanisms, and more evidence is needed to elucidate the relationship. Clarifying the functional relationship between the lateral PFC and the LC during the *n*-back task would give us insights into the meaning of lateral PFC activity (e.g., lateral PFC activity correlating with arousal level) during this task.

There is growing recognition that fNIRS is relatively tolerant to motion and is free from environmental interference (Ferrari & Quaresima, 2012). Thus, fNIRS is a suitable tool to use in conjunction with eye trackers, and it is an excellent tool for monitoring brain activity in a natural setting (Hosseini et al., 2017; İşbilir et al., 2019). For these reasons, we used fNIRS to probe prefrontal hemodynamic changes. Due to the use of a blocked design

and fast trial presentation to increase power, the distinction between a shift in tonic and phase LC activity (Aston-Jones & Cohen, 2005) in response to increased WM load would not and could not be made in the present study.

2. Method

2.1. Participants

Thirty-nine right-handed college students (10 males, 29 females) aged between 18 and 24 years (Mean = 19.8 years; *SD* = 1.5 years) were recruited from the subject pool of the Chinese University of Hong Kong. This sample size was larger than those in most of the previous fNIRS and pupillometry studies using the *n*-back task (e.g., Ito et al., 2011; Karatekin et al., 2007; Niezgoda et al., 2015; Yeung et al., 2016). No participant reported a history of psychiatric or neurological disorder, and none was taking any medication. All participants self-reported normal or corrected-to-normal vision or else had no difficulty identifying *n*-back test stimuli. All participants also gave written informed consent prior to the study. The study protocol received approval from the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee.

2.2. Procedure and Materials

All participants underwent a digit version of the *n*-back paradigm with 0- and 3-back conditions under simultaneous fNIRS and pupillometric measurements. The paradigm was adapted from previous fNIRS studies using the *n*-back task, all of which involved 1–3 task blocks and 28–60 trials (12–15 target trials) for each condition (Ehlis et al., 2008; Koike et al., 2013; Yeung et al., 2016). Figure 1 illustrates the flow of the experimental paradigm. The paradigm started and ended with a rest block and involved alternations between 0- and 3-back blocks interleaved with rest blocks. Each task block was presented twice, and the order of the task blocks was counterbalanced across the participants. All participants carried out four

blocks. Although half of the participants underwent the 0-3-0-3 sequence, others underwent the 3-0-3-0 sequence. Blocks of a given load were not presented consecutively to avoid habituation effects. Each task block lasted for 47 s, and each rest block lasted for 30 s. The entire task lasted for 338 s.

Each task block started with an instruction cue presented for 5 s, followed by seven target and 21 nontarget trials (i.e., 28 trials per block). In each trial, a number chosen from 0 to 9 was first presented at the center of a computer screen for 0.5 s and then replaced with a blank screen for 1 s. Participants were asked to press one of the two buttons on a computer mouse as accurately and quickly as possible upon seeing each number. During the 0-back task, participants pressed the left button if the number they saw was "0" but the right button for any other number. During the 3-back task, participants pressed the left button if the number they were seeing was the same as the number they saw three trials before but the right button for any other number. Participants had 1.5 s to respond during each trial. No feedback was given. All individuals practiced the 0- and 3-back tasks before the *n*-back paradigm actually began. The test stimuli were presented using E-Prime 1.2 software (Psychology Software Tools, Pittsburgh, PA).

2.3. Functional Near-Infrared Spectroscopy Measurement

A 16-channel continuous-wave fNIRS device was used to measure hemodynamic changes in the bilateral PFC (OEG-SpO2; Spectratech Inc., Tokyo, Japan). This machine emitted two wavelengths of near-infrared lights: 770 and 840 nm. The probe consisted of six light-emitting and six light-receiving optodes arranged in a 2 × 6 matrix. The emitter-detector spacing was 3 cm, hence allowing a measurement of hemodynamic changes that took place 1.5–2 cm below the scalp (Cui et al., 2011; Schroeter et al., 2006). The probe was placed in the forehead region and was centered at Fpz in accordance with the international 10–20

system (Jasper, 1958). Accordingly, the optodes at the bottom left and right corners were placed around F7 and F8, respectively.

Samples were collected at a rate of 12.21 Hz at 16 locations (i.e., measurement channels) midway between each pair of emitters and detectors. The arrangement of the 16 measurement channels is shown in Figure 2. For the purpose of increasing the signal-to-noise ratio (Plichta et al., 2006), channels 13–16, 7–10, and 1–4 were grouped to represent the left, medial, and right prefrontal regions, respectively. The same number of channels was used for each region to ensure a comparable signal-to-noise ratio across regions, hence reducing biases. Based on the anatomical locations of the International 10–20 cortical projection points (Okamoto et al., 2004; Koessler et al., 2009), it was estimated that the outermost channels were located around the dorsolateral and the ventrolateral PFC, and the medial channels were located around the medial frontal pole.

2.4. Pupillometry Recording

Wearable eye-tracking glasses were used to measure changes in pupil diameter (Pupil Labs, Berlin, Germany). This eye tracker had two video cameras and two infrared cameras facing toward both eyes for binocular tracking (see Figure 2). The two eye cameras monitored the pupil center and provided the absolute pupil diameter as measurement data by using the dark pupil method in reference to a three-dimensional geometric model of the eyeball. The system also generated a confidence index ranging from 0–1, with 1 indicating the highest level of confidence in pupil detection during recording. The sampling rate was 60 Hz. Moreover, the eye tracker had an integrated world camera capturing the participant's field of view. The image recorded with this camera was used to synchronize external events (e.g., block onset) with the pupillometry data.

Although both the fNIRS sensor and the eye tracker relied on (near-)infrared light, several features of the setup ensured no interference between these two modalities. First, the

fNIRS sensor was fixed by a light-insulating headband to ensure close contact between the sensor band and skin, as well as to minimize the effects of outside (near-)infrared sources. Second, a separation existed between the infrared cameras and the fNIRS sensor such that the infrared cameras were oriented toward the pupil center at an angle that was away from the forehead. More importantly, we checked using a visual inspection that turning the eye tracker on and off had no effect on fNIRS signals, and vice versa.

2.5. Data Processing and Analysis

Behavioral performance during the *n*-back task was first analyzed. Then, fNIRS and pupillometry data were analyzed before the relationship between the two modalities was investigated. Normality was checked for all primary fNIRS and pupillometric measures using Shapiro–Wilk tests. Only a few variables significantly deviated from normality (ps < 0.014). For all variables, parametric (*t*-tests) and nonparametric tests (Wilcoxon signed-rank tests) yielded similar results, and the same conclusions could be reached (see Supplementary Table 1). Therefore, parametric test results are reported throughout. Preprocessing was performed using customized scripts in Matlab® R2018a (MathWorks, Natick, NA). Statistical analysis was conducted using SPSS 22.0 software (IBM Corporation, Armonk, NY, USA). All *t*-tests and correlation tests were two-tailed. The alpha level was set at 0.05. When multiple comparisons were involved, the Bonferroni correction was applied.

2.5.1. Behavioral Performance

Task performance was examined by analyzing the sensitivity (*A*') (Snodgrass & Corwin, 1988) and mean reaction time (RT) in the 0- and 3-back conditions. We analyzed *A*' rather than accuracy because *A*' considered both the hit and the false alarm rate, and it corrected for response bias. This index ranged from 0–1, with a higher value indicating better discriminability. In addition, the mean RT was calculated by averaging the RT across the correct trials, excluding trials with RTs < 150 ms or 2.5 *SD*s above the individual mean

because RTs < 150 ms would be too fast to reflect a real response to the stimulus (Stuss et al., 2005). Paired *t*-tests comparing the 0- and 3-back conditions were conducted to examine the effect of the WM load on behavioral performance.

2.5.2. Prefrontal Activation

The preprocessing of fNIRS data took place as follows: First, raw time course data were converted into changes in optical density. Then, a 0.10-Hz lowpass filter at 60 dB/octave was applied to remove cardiac and respiratory artifacts, as well as high-frequency noise. Then, the filtered optical density data were converted into changes in [oxy-Hb] and [deoxy-Hb] via the application of the modified Beer–Lambert law (Delpy et al., 1988). Next, correlation-based signal improvement (CBSI), which is based on the negative correlation between the dynamics of [oxy-Hb] and [deoxy-Hb], was applied to correct for motion artifacts (Cui et al., 2010). The time courses of [oxy-Hb] and [deoxy-Hb] became mirror images of each other after CBSI. Hence, only changes in the CBSI-corrected [oxy-Hb] were analyzed. Next, linear fitting based on the 10 s of rest before and after each task block was performed to correct for slow drifts. Then, raw changes in the CBSI-corrected [oxy-Hb] were divided by the *SD* of signals over the 10-s rest period before the corresponding task block to generate the normalized (i.e., *Z*; effect size) scores to account for individual differences in the differential pathlength factor (Schroeter, Zysset, Kruggel, & von Cramon, 2003). Finally, normalized changes in the CBSI-corrected [oxy-Hb] were averaged across the four channels within each prefrontal region.

Prefrontal activation during the *n*-back task was examined by analyzing the mean changes in the CBSI-corrected [oxy-Hb], which were calculated by averaging all values across the 42-s task period (i.e., excluding the 5-s cue period) and the two blocks for each condition. One-sample *t*-tests were first performed to examine whether the mean change in the CBSI-corrected [oxy-Hb] in each prefrontal region was above zero in the 0- and 3-back

conditions. Then, a repeated-measures analysis of variance (ANOVA) with Condition (0-back, 3-back) and Region (left, medial, right) as factors was performed to assess the effects of the WM load and region on prefrontal hemodynamics.

2.5.3. Pupil Dilation

For each participant, only data from the eye with the highest cumulated confidence level was considered for further analysis (Schwab et al., 2019). Consequently, the analysis of pupil dilation was based on the left pupil in 18 participants and on the right pupil in 21 participants. All participants achieved a cumulative confidence value of 0.68 (Mean = 0.92, *SD* = 0.06), which met the requirement of useful data (Pupil Labs, 2020). The preprocessing of pupillometry data took place as follows: First, all data points carrying confidence values < 0.60 were set to 0 (Pupil Labs, 2020). Note that eye blink signals also carried a confidence value of 0. Then, a spline interpolation was implemented to correct for zero confidence points before a 0.10 Hz low-pass butterworth filter was applied to exclude high-frequency noise. Next, linear fitting, which was based on the 10 s of rest before and after each task block, was applied to the absolute pupil diameter to correct for possible slow signal drifts. Finally, individual differences in the fluctuation of pupil size were controlled for by dividing the raw pupil diameter signals by the *SD* of signals over the 10-s rest period before the corresponding task block to generate the normalized (i.e., *Z*) scores.

Pupil dilation during the *n*-back task was examined by analyzing the mean change in the pupil diameter, which was calculated by averaging all values across the 42-s task period and the two blocks for each condition. One-sample *t*-tests were first performed to examine whether the mean change in the pupil diameter was above zero in the 0- and 3-back conditions. A paired *t*-test comparing the 0- and 3-back conditions was then performed to assess the mean change in the pupil diameter specific to increasing WM load.

We focused on task-evoked pupillary responses that occurred at the block level. In the present study, the intertrial duration was 1500 ms, which was shorter than that in previous pupillometry studies using the *n*-back task (i.e., ≥ 2000 ms; Karatekin et al., 2007; Scharinger et al., 2015). This rapid task design was adopted to increase detection power. Because a prestimulus baseline that was not influenced by the previous trial could not be derived, and changes in luminance were too rapid, task-evoked changes in the pupil diameter were not examined at the trial level.

2.5.4. Relationship Between Prefrontal Activation and Pupil Dilation

Finally, the relationship between changes in PFC activity and the pupil size during the *n*-back task was examined. We calculated the Pearson's correlation between the time courses of the hemodynamic response predicted by the pupillary response and of the observed hemodynamic response for each 42-s task period after the pupil diameter data were resampled at 12.21 Hz. The hemodynamic response predicted by the pupillary response was constructed by convolving the vector of the normalized pupil diameter data with the canonical hemodynamic response function implemented in SPM8 (Friston et al., 1998). Because the pupillary response was delayed by approximately 1 s after an event, the hemodynamic response function moved forward 1 s.

After downsampling and convolution, there were 513 points for the [oxy-Hb] changes predicted by the pupillary response and for the observed [oxy-Hb] changes during each 42-s task period. Pearson's correlation between the predicted and the observed [oxy-Hb] changes was then run on that number of points, yielding a correlation coefficient for each block and participant. Mean 0- and 3-back correlation coefficients were then generated by averaging the two correlation coefficients for each condition, and statistical analyses (one-sample *t*-tests) were conducted on the mean correlation coefficients. Notably, precision of the correlation coefficients was determined by the number of trials, whereas statistical power was dependent

on the number of participants. One-sample *t*-tests were conducted to determine whether the mean Pearson's correlations for the 0- and 3-back conditions were above 0.

3. Results

3.1. Behavioral Performance

First, performance in terms of *A*' and the mean RT during the *n*-back task was examined. Table 1a presents the means and *SD*s of these two variables in the 0- and 3-back conditions. Paired *t*-tests comparing the two conditions showed that *A*' was significantly lower in the 3-back condition than in the 0-back condition, $t(38) = 8.56$, $p < 0.001$, $d = 1.37$. Additionally, the mean RT was significantly slower in the 3-back condition than in the 0-back condition, $t(38) = 7.85$, $p < 0.001$, $d = 1.26$. These results indicate decrements in task performance when the WM load increased.

3.2. Prefrontal Activation

Next, PFC activation during the *n*-back task was investigated. Table 1b presents the means and SDs of mean changes in the CBSI-corrected [oxy-Hb] during the *n*-back task. Figure 3 illustrates the change in the CBSI-corrected [oxy-Hb] in each condition. Notably, there was little change in [oxy-Hb] in the medial prefrontal region across conditions and in all regions during the 0-back block. In contrast, there were increases in [oxy-Hb] in the bilateral lateral prefrontal regions following block onset during the 3-back block. These increases reached their plateau approximately 20 s after block onset and did not drop until 5 s after the end of the block.

First, one-sample *t*-tests (against zero) were performed to determine PFC activation in the 0- and 3-back conditions compared with the resting period. We found no significant mean changes in [oxy-Hb] in any prefrontal region in the 0-back condition, *t*s < 1.14, *p*s > 0.26, *d*s < 0.16. In contrast, we found significant mean changes in [oxy-Hb] in the left, $t(38) = 3.88$, $p < 0.001$, $d = 0.62$, and right, $t(38) = 3.32$, $p = 0.002$, $d = 0.53$, prefrontal regions in the 3-back

condition. Mean changes in [oxy-Hb] in the medial prefrontal region was not significant, $t(38) = 1.07$, $p = 0.29$, $d = 0.17$.

To determine the effect of the WM load and the regional specificity of PFC activation, a repeated-measures ANOVA with Condition (0-back, 3-back) and Region (left, medial, right) as factors was conducted on mean changes in the CBSI-corrected [oxy-Hb]. The main effect of condition was significant, $F(1, 76) = 10.43$, $p = 0.003$, $\eta_p^2 = 0.22$, whereas the main effect of region was not, $F(2, 76) = 1.61$, $p = 0.21$, $\eta_p^2 = 0.04$. More importantly, a significant condition × region interaction, $F(2, 76) = 6.07$, $p = 0.004$, $\eta_p^2 = 0.14$, was found. Post-hoc *t*-tests showed significantly larger mean changes in [oxy-Hb] in both the left, $t(38) = 4.37$, $p < 0.001$, $d = 0.70$, and right, $t(38) = 3.20$, $p = 0.003$, $d = 0.51$, prefrontal regions in the 3-back than the 0-back condition. The effect sizes ranged from medium to large. Additionally, there was no significant mean change in [oxy-Hb] in the medial prefrontal region between the two conditions, $t(38) = 0.88$, $p = 0.39$, $d = 0.14$.

3.3. Pupil Dilation

Pupil dilation during the *n*-back task was then examined. Table 1c presents the means and *SD*s of mean changes in the pupil diameter during the *n*-back task. Figure 4 illustrates the time course of changes in the pupil diameter in the 0- and 3-back conditions. Notably, the pupil diameter increased immediately following block onset in both conditions. During the 0-back block, the pupil diameter reached the peak 7 s after block onset, or 2 s after task onset. Then, it gradually decreased and returned to the baseline level midway through the block. During the 3-back block, the pupil diameter reached the peak 13 s after block onset, or 8 s after task onset, and then gradually decreased throughout the block. The pupil diameter in this condition did not completely return to the baseline level until 2 s after the end of the task. Because the pupil diameter decreased in amplitude immediately after reaching the peak value, restricted pupil variations were unlikely to have occurred during the 3-back task. If a

ceiling effect was present, then the pupillary response should have sustained in its amplitude after reaching the peak.

One-sample *t*-tests (against zero) were first performed to determine pupil dilation in the 0- and 3-back conditions compared with the resting period. We found no significant mean change in the pupil diameter in the 0-back condition relative to the resting period, $t(38) = 1.14$, $p = 0.26$, $d = 0.18$. In contrast, we found a significant mean change in the pupil diameter in the 3-back condition relative to the resting period, $t(38) = 5.33$, $p < 0.001$, $d = 0.85$. More importantly, paired *t*-tests comparing the 0- and 3-back conditions revealed a significantly larger mean change in the pupil diameter in the 3-back condition than in the 0-back condition, $t(38) = 4.31$, $p < 0.001$, $d = 0.69$. The effect size was medium to large.

3.4. Relationship Between Prefrontal Activation and Pupil Dilation

Next, the relationship between PFC activation and pupil dilation was examined. Because WM-induced activation was evident only in the left and right lateral prefrontal regions, all subsequent analyses focused on these two regions. Pearson's correlations between the time courses of the hemodynamic response predicted by the pupillary response and the observed hemodynamic response were examined (Figure 5). None of the variables significantly deviated from normality as revealed by Shapiro–Wilk tests (ps > 0.45), and no outliers were identified according to Grubbs' tests (Grubbs, 1969). Thus, one-sample *t*-tests were performed to determine whether the mean Pearson's correlations differed from zero, indicating the existence of a relationship at the within-subject level. Overall, the correlation between the time courses of the predicted and observed hemodynamic responses in the 0-back condition was significantly above zero in both the left lateral PFC ($M = 0.11$, $SD = 0.27$), $t(38) = 2.57$, $p = 0.014$, $d = 0.41$, and the right lateral PFC ($M = 0.10$, $SD = 0.29$), $t(38) = 2.23$, $p = 0.032$, $d = 0.36$. The effect sizes that represent the amplitudes of the difference from zero were medium. The mean correlation between the time courses of the predicted and

observed hemodynamic responses in the 3-back condition, however, was not significant in either the left lateral (M = -0.023, SD = 0.35), $t(38)$ = -0.41, p = 0.68, d = -0.07, or the right lateral PFC (M = 0.009, SD = 0.37), $t(38)$ = 0.02, p = 0.99, d = 0.00. The effect sizes that represent the amplitudes of the difference from zero were very small, if not negligible.

Next, a repeated-measures ANOVA with Condition (0-back, 3-back) and Region (left, right) as factors was conducted to test whether the mean Pearson's correlations between the predicted and the observed hemodynamic response differed between the two conditions. The main effect of condition approached significance, $F(1, 38) = 3.38$, p = 0.07, η_p^2 = 0.08, indicating a tendency for a stronger relationship between the moment-to-moment changes in lateral PFC activity and pupil size during the 0-back than the 3-back task. The effect size was medium-to-large. Neither the main effect of region nor the interaction between region and condition was significant, ps > 0.56.

3.5. Time-Related Physiological Activity

WM processing is expected to be constant (or increase) over time due to constant task demand and proactive interference from previous trials (Badre & Wagner, 2005; Gray et al., 2003), whereas perceptual arousal should decrease over time as the same (or similar) non-affective content is presented (Antikainen & Niemi, 1983; Stelmack & Siddle, 1982). Thus, additional analyses were performed to clarify the mechanisms underlying changes in lateral PFC activity and the pupil size over time (see Supplementary Table 2 for complete results). First, a repeated-measures ANOVA with Condition (0-back, 3-back), Time (0–21 s, 21–42 s), Block (first, second), Region (left, right) as factors was conducted on the mean CBSI-corrected [oxy-Hb] changes. Only the main effect of condition and the condition × time and condition × time × block interactions were significant, ps < 0.05. Thus, two follow-up repeated-measures ANOVAs with Time, Block, and Region as factors were carried out for the 0- and 3-back conditions separately, with the alpha level adjusted to 0.025 (i.e., two

conditions; 0.05/2 = 0.025). For the 3-back condition, only the main effect of time was significant, $p = 0.006$, indicating a larger mean [oxy-Hb] increase during the second half than the first half of the 3-back blocks. None of the effects were significant for the 0-back condition, ps > 0.025.

Next, a repeated-measures ANOVA with Condition, Time, and Block as factors was conducted on mean change in pupil size. Only the main effects of condition and time and the condition × time interaction were significant, ps < 0.05. Thus, two follow-up repeated-measures ANOVAs with Time and Block as factors were conducted for the 0- and 3-back conditions separately, with the alpha level set at 0.025 (i.e., two conditions; 0.05/2 = 0.025). For both conditions, only the main effect of time was significant (0-back: $p < 0.006$; 3-back: $p < 0.001$). In contrast to [oxy-Hb] changes, mean change in the pupil diameter was smaller during the second half than the first half of the blocks.

4. Discussion

This study's aim was to determine the relationship between changes in lateral PFC activity and the pupil size (a proxy for LC activity), as well as the effect of WM load on such relationship. It is the first to examine changes in PFC activity and the pupil size simultaneously measured during the *n*-back task. In keeping with the existing fMRI (Mencarelli et al., 2019; Owen et al., 2005) and fNIRS (Koike et al., 2013; Yeung et al., 2016, 2019) literature, we found negligible PFC activation at a low WM load (i.e., 0-back) but significant moderate-to-strong activation in the bilateral lateral PFC at a high WM load (i.e., 3-back). A contrast between the two conditions revealed significant WM-related increases in the lateral PFC. This regional specificity is consistent with the crucial role of the dorsolateral PFC in WM (Barbey et al., 2013; Levy & Goldman-Rakic, 2000), and more importantly, rules out the possibility that the observed [oxy-Hb] increases were merely driven by systemic influences. Also consistent with the existing pupillometry literature (Belayachi et

al., 2015; Karatekin et al., 2007; Niezgoda et al., 2015; Scharinger et al., 2015), we found negligible changes in the pupil size at a low WM load, but we witnessed strong pupil dilation at a high WM load, with a significant WM-related increase in the pupil size. More interestingly, significant positive correlations were found between the observed [oxy-Hb] changes and the [oxy-Hb] changes predicted from the pupil signal only at the low WM load. There was a statistical trend for these relationships to be stronger during the processing of low than high WM demand. Altogether, our findings suggest a load-dependent relationship between changes in lateral PFC activity and the pupil size in the context of WM.

Our findings suggest that changes in lateral PFC activity and the pupil size have a shared physiological mechanism at a low WM load. It is generally agreed that the pupil diameter is a proxy measure for arousal and mental effort in the context of cognitive tasks (Bradley et al., 2008; Mathot, 2018). Thus, hemodynamic changes in the lateral PFC may represent changes in the levels of arousal and task engagement during the processing of low WM demand, perhaps mediated by at least two mechanisms— systemic changes, such as blood pressure and heart rate (Kahnerman, Tursky, Shapiro, & Crider, 1969; Wang et al., 2018), and intrinsic interactions between the LC norepinephrine system and the PFC. For the latter, tracing studies on monkeys have shown that the LC receives projections from various parts of the frontal lobe, notably the dorsolateral and the dorsomedial PFC (Arnsten & Goldman-Rakic, 1984). Activity in the LC also positively correlates with activity in the anterior cingulate cortex in monkeys (Joshi et al., 2016) and humans (Murphy et al., 2014). Thus, the structural and functional connections between the LC, which regulates the pupil size via interaction with the autonomic nervous system (Samuels & Szabadi, 2008), and PFC subregions may underlie the positive correlation between changes in lateral PFC activity and the pupil size during an easy task requiring little effort.

While the magnitude of the correlation between changes in lateral PFC activity and the pupil size may seem small (mean $r \approx 0.10$), it falls in the range of values that are typically seen when the observed hemodynamic changes are predicted by certain events or factors based on the canonical hemodynamic response function. For example, one recent fMRI study found that the mean beta (or *r*) of BOLD changes (i.e., magnitude of the relationship between the observed BOLD changes and the BOLD changes predicted by an event) in the left lateral PFC was around 0.10 during the 3-back task (Lamichhane et al., 2020). Because this value was significantly different from zero, it was concluded that WM load was related to activation in the left lateral PFC. Similarly, we can conclude that lateral PFC activity and pupil dilation are associated during the 0-back condition because the correlation ($r \approx 0.10$) was significantly different from zero.

On the other hand, we found a considerable increase in the pupil size, indicating increased levels of arousal and mental effort at a high WM load. Although this increase was accompanied by an increase in bilateral lateral PFC activity, no significant relationship existed between the two. The dissociation between changes in lateral PFC activity and the pupil size during the 3-back task can also be observed by comparing the time courses of these changes. The increase in lateral PFC activity was sustained throughout the second half of the block, such that the mean lateral PFC activity was significantly larger during the second half than the first half of the block. This phenomenon may be related to (constant) WM demand and proactive interference throughout the 3-back task (Badre & Wagner, 2005; Gray et al., 2003). In contrast, the increase in the pupil diameter gradually dropped throughout the last two-thirds of the block, such that the mean pupil diameter was significantly smaller during the second half than the first half of the block. This pattern, which was also observable at 0-back, may imply a gradual decrease in perceptual arousal following repeated exposure to

same or similar non-affective stimuli (i.e., habituation; Antikainen & Niemi, 1983; Stelmack & Siddle, 1982).

The pupil is known to dilate with cognitive load (van der Wel & van Steenbergen, 2018) and proactive interference (Johansson et al., 2018; Johansson & Johansson, 2020), and our finding of an overall increase in pupil size from 0- to 3-back is consistent with this literature. In addition, our finding of a temporal reduction in pupil size at 3-back is compatible with the co-existence of (constant) proactive interference and perceptual habituation during the 3-back task. That is, (constant) WM processing and perceptual habituation would predict sustained and decreased pupil dilation over time, respectively. Thus, if both phenomena were present, the net effect would be a reduction in pupil dilation as time on task increased. In summary, lateral PFC activity appears to represent cognitive processing only, whereas pupil dilation may represent both cognitive processing and perceptual arousal.

Therefore, based on the physiological basis of the pupillary response, lateral PFC activity when facing a high level of WM demand does not seem associated with LC activity. Performing the 3-back task requires a high level of executive control capacity and draws on different types of cognitive resources. These resources include the storage and manipulation of internal representations, the monitoring of incoming information, and the use of organizational strategies (Mencarelli et al., 2019; Owen et al., 2005). Thus, lateral PFC activation during the processing of high WM demand appears to be a manifestation of the active operation of various cognitive processes supporting the performance of this difficult task. As such, there may be a decrease in the functional relationship between the lateral PFC and the LC at high WM load due to the specialization of the lateral PFC in WM (Barbey et al., 2013; Levy & Goldman-Rakic, 2000; Tsuchida & Fellows, 2009). Alternatively, the relatively uncoupled lateral PFC and LC activity may be attributable to an increase in

clustered neuronal communications within the lateral PFC network as a result of increased LC activity or norepinephrine level during the 3-back task (Eldar et al., 2013).

A functional segregation between the lateral PFC and LC systems in the presence of high WM demand may or may not be beneficial for WM task performance. It may be merely a side product of the modular organizational structure in the brain (Meunier et al., 2010). Alternatively, separate processing within the two systems while executing a demand WM task may allow WM processes to operate constantly in face of momentary fluctuations in arousal level, and efficiently in times of under- or over-arousal. It should however be noted that the effect of WM demand on the relationship between lateral PFC activity and pupil size changes was found to be only marginally significant. More evidence is needed to confirm the moderating effect of WM load.

Activity in the lateral PFC (Callicott et al., 1999) and LC (Aston-Jones & Cohen, 2005) are known to increase with increasing WM load in an inverted U-shaped manner. This study included only two WM load levels, and therefore could not determine or compare the inverted U-shaped functions of lateral PFC and LC activity. In any case, a difference in the inverted U-shaped relationship could not explain the absence of relationships during the 3-back task, because the Pearson's correlations between the time courses of two signals were not affected by differences in the mean levels of signals at a given load level.

Like most previous studies, reliability was not estimated for the fNIRS and pupillometric measures, and our study was not designed to evaluate that. However, it seems that these measures have been observed reliably. The *n*-back task has been extensively used in fNIRS and pupillometry studies, and robust lateral PFC activation and pupillary responses have been separately observed during *n*-back task performance (e.g., van der Wel & Steenbergen, 2018; Yeung et al., 2016). Also, the trial number in the present study was comparable to that in previous fNIRS studies. More importantly, significant positive

correlations were found between the [oxy-Hb] changes predicted by the pupillary response and the observed [oxy-Hb] changes for the 0-back task. These findings would not have been possible if any of the signals was unreliable.

Similarly to the present study, one study combined neuroimaging (i.e., fMRI) and pupillometry to concurrently measure changes in brain activity and the pupil size during the *n*-back task. Belayachi et al. (2015) identified WM-related (i.e., 3-back > 0-back) activation in the dorsolateral and the dorsomedial PFC, the insula, and the intraparietal sulcus. They also identified a WM-related increase in the pupil diameter during task performance. However, the relationship between brain activation and pupil dilation was not examined. Our data thus extend the literature by clarifying the functional role of the lateral PFC during the executive control of the WM. This brain region represents information in a context-dependent fashion—activity in this region represents the arousal level under little WM or cognitive demand, whereas it primarily reflects the operation of cognitive processes when the WM or cognitive demand is high.

Our findings have important implications for neuropsychiatric research. Over the past two decades, numerous studies have identified reduced WM-induced PFC activation during the *n*-back task in a wide variety of neuropsychiatric conditions, such as attention-deficit/hyperactivity disorder (Ehlis et al., 2008), schizophrenia (Koike et al., 2013), depression (Matsuo et al., 2007), mild cognitive impairment (Yeung et al., 2016), and Alzheimer's disease (Lim et al., 2008). The present study found a very weak, if not negligible, association between prefrontal hemodynamics and pupil size changes during the processing of high WM demand. Thus, simultaneous fNIRS and pupillometric assessments may provide unique and complementary information that informs the psychological and physiological mechanisms underlying WM processing, which could be useful for improving intervention strategies. For example, patients who show PFC hypoactivation in the presence

of pupil dilation may have WM or PFC dysfunction while being engaged and motivated to perform the task. Intervention for these patients may focus on improving their WM or PFC functions. Patients who exhibit a lack of pupil dilation despite intact PFC activation may have blunted autonomic arousal that does not necessarily affect their frontal executive functions (Critchley et al., 2003). Intervention for these patients may target at normalizing their autonomic nervous system functions. In addition, patients who demonstrate PFC hypoactivation and reduced pupil dilation may experience task engagement difficulties (e.g., fail to understand the task) or have motivation problems, requiring motivational enhancement therapies.

Interest has been growing in combining neuroimaging and pupillometry to elucidate the physiological mechanisms of cognitive functions. Murphy et al. (2014) combined fMRI and pupillometry and found that the pupil diameter was associated with blood-oxygen-level-dependent (BOLD) signals in the dorsal pontine and in the visual primary cortex during change detection. Schneider et al. (2018) combined fMRI and pupillometry, and they identified a positive correlation between pupil dilation and activity in the dorsal anterior cingulate cortex and in the bilateral insula during reward anticipation. Hosseini et al. (2017) combined fNIRS and pupillometry, and they found that pupil dilation was associated with right superior parietal activation during specific aspects of visuomotor behavior. The present study extends the existing literature by demonstrating a positive association between changes in lateral PFC activity and the pupil size in the context of WM. Like Hosseini et al. (2017), the present study demonstrates the promising combined application of a portable fNIRS device and eye-tracking glasses to elucidate the physiological processes underlying cognitive tasks in a naturalistic environment.

To conclude, the lateral PFC and the LC are functionally related during the processing of low WM demand, which may be due to common systemic influences or intrinsic

interactions between the two systems. In contrast, activity in these two regions are relatively uncoupled during the processing of high WM demand, which may be due to the specialized role of the lateral PFC in WM and to the tightly clustered pattern of neural interactions within the lateral PFC network when LC activity is high. However, more evidence is needed to confirm the load-dependent relationship between the lateral PFC and the LC. Based on these findings, we propose that the simultaneous application of fNIRS (i.e., functional neuroimaging) and pupillometry as probes for the central and autonomic nervous systems would allow a fuller understanding of the physiological mechanisms underlying WM processing in both healthy and clinical populations.

5. Conflict of Interest Statement

The authors have no conflict of interest to declare.

6. References

Arnsten, A. F. T., & Goldman-Rakic, P. S. (1984). Selective prefrontal cortical projections to the region of the locus coeruleus and raphe nuclei in the rhesus monkey. *Brain research*, *306*(1-2), 9-18.

Antikainen, J., & Niemi, P. (1983). Neuroticism and the pupillary response to a brief exposure to noise. *Biological psychology*, *17*(2-3), 131-135.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403-450.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual review of psychology*, *63*, 1-29.

Badre, D., & Wagner, A. D. (2005). Frontal lobe mechanisms that resolve proactive interference. *Cerebral Cortex*, *15*(12), 2003-2012.

Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. C*ortex*, *49*(5), 1195-1205.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, *91*(2), 276-292.

Belayachi, S., Majerus, S., Gendolla, G., Salmon, E., Peters, F., & Van der Linden, M. (2015). Are the carrot and the stick the two sides of same coin? A neural examination of approach/avoidance motivation during cognitive performance. *Behavioural brain research*, *293*, 217-226.

Boisgueheneuc, F. D., Levy, R., Volle, E., Seassau, M., Duffau, H., Kinkingnehun, S., ... & Dubois, B. (2006). Functions of the left superior frontal gyrus in humans: a lesion study. *Brain*, *129*(12), 3315-3328.

Bopp, K. L., & Verhaeghen, P. (2020). Aging and n-back performance: A meta-analysis. *The Journals of Gerontology: Series B*, *75*(2), 229-240.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602-607.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., ... & Weinberger, D. R. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral cortex*, *9*(1), 20-26.

Carter, M. E., de Lecea, L., & Adamantidis, A. (2013). Functional wiring of hypocretin and LC-NE neurons: implications for arousal. *Frontiers in behavioral neuroscience*, *7*, 43.

Critchley, H. D., Mathias, C. J., Josephs, O., O'Doherty, J., Zanini, S., Dewar, B. K., ... & Dolan, R. J. (2003). Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. *Brain*, *126*(10), 2139-2152.

Cui, X., Bray, S., & Reiss, A. L. (2010). Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage*, *49*(4), 3039-3046.

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *Neuroimage*, *54*(4), 2808-2821.

Delpy, D. T., Cope, M., van der Zee, P., Arridge, S. R., Wray, S., & Wyatt, J. S. (1988). Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine & Biology*, *33*(12), 1433.

Ehlis, A. C., Bähne, C. G., Jacob, C. P., Herrmann, M. J., & Fallgatter, A. J. (2008). Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study. *Journal of psychiatric research*, *42*(13), 1060-1067.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature neuroscience*, *16*(8), 1146-1153.

Ferrari, M., & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage*, *63*(2), 921-935.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A. N. D. R. E. W., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: characterizing differential responses. *Neuroimage*, *7*(1), 30-40.

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature neuroscience*, *6*(3), 316-322.

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics*, *11*(1), 1-21.

Hosseini, S. H., Bruno, J. L., Baker, J. M., Gundran, A., Harbott, L. K., Gerdes, J. C., & Reiss, A. L. (2017). Neural, physiological, and behavioral correlates of visuomotor cognitive load. *Scientific reports*, *7*(1), 1-9.

İşbilir, E., Çakır, M. P., Acartürk, C., & Tekerek, A. Ş. (2019). Towards a Multimodal Model of Cognitive Workload Through Synchronous Optical Brain Imaging and Eye Tracking Measures. *Frontiers in human neuroscience*, *13*, 375.

Ito, H., Yamauchi, H., Kaneko, H., Yoshikawa, T., Nomura, K., & Honjo, S. (2011). Prefrontal overactivation, autonomic arousal, and task performance under evaluative pressure: A near-infrared spectroscopy (NIRS) study. *Psychophysiology*, *48*(11), 1563-1571.

Jasper, H. H. (1958). The 10/20 international electrode system. *EEG and Clinical Neurophysiology*, *10*(2), 370-375.

Johansson, R., & Johansson, M. (2020). Gaze position regulates memory accessibility during competitive memory retrieval. *Cognition*, *197*, 104169.

Johansson, R., Pärnamets, P., Bjernestedt, A., & Johansson, M. (2018). Pupil dilation tracks the dynamics of mnemonic interference resolution. *Scientific reports*, *8*(1), 1-8.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron*, *89*(1), 221-234.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*(3756), 1583-1585.

Kahneman, D., Tursky, B., Shapiro, D., & Crider, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *Journal of experimental psychology*, *79*(1p1), 164.

Karatekin, C., Marcus, D. J., & Couperus, J. W. (2007). Regulation of cognitive resources during sustained attention and working memory in 10-year-olds and adults. *Psychophysiology*, *44*(1), 128-144.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of experimental psychology*, *55*(4), 352-358.

Koessler, L., Maillard, L., Benhadid, A., Vignal, J. P., Felblinger, J., Vespignani, H., & Braun, M. (2009). Automated cortical projection of EEG sensors: anatomical correlation via the international 10–10 system. *Neuroimage*, *46*(1), 64-72.

Koike, S., Takizawa, R., Nishimura, Y., Kinou, M., Kawasaki, S., & Kasai, K. (2013). Reduced but broader prefrontal activity in patients with schizophrenia during n-back working memory tasks: a multi-channel near-infrared spectroscopy study. *Journal of psychiatric research*, *47*(9), 1240-1246.

Kopf, J., Dresler, T., Reicherts, P., Herrmann, M. J., & Reif, A. (2013). The effect of emotional content on brain activation and the late positive potential in a word n-back task. *PloS one*, *8*(9).

Lamichhane, B., Westbrook, A., Cole, M. W., & Braver, T. S. (2020). Exploring brain-behavior relationships in the N-back task. *NeuroImage*, *212*, 116683.

Levy R., Goldman-Rakic P.S. (2000) Segregation of working memory functions within the dorsolateral prefrontal cortex. In: Schneider W.X., Owen A.M., Duncan J. (eds) *Executive Control and the Frontal Lobe: Current Issues* (pp. 23-32). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-59794-7_

Lim, H. K., Juh, R., Pae, C. U., Lee, B. T., Yoo, S. S., Ryu, S. H., ... & Lee, C. U. (2008). Altered verbal working memory process in patients with Alzheimer's disease. *Neuropsychobiology*, *57*(4), 181-187.

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, *1*(1).

Matsuo, K., Glahn, D. C., Peluso, M. A. M., Hatch, J. P., Monkul, E. S., Najt, P., ... & Fox, P. T. (2007). Prefrontal hyperactivation during working memory task in untreated individuals with major depressive disorder. *Molecular psychiatry*, *12*(2), 158-166.

McDougal, D. H., & Gamlin, P. D. (2011). Autonomic control of the eye. *Comprehensive Physiology*, *5*(1), 439-473.

Mencarelli, L., Neri, F., Momi, D., Menardi, A., Rossi, S., Rossi, A., & Santarnecchi, E. (2019). Stimuli, presentation modality, and load-specific brain activity patterns during n-back task. *Human brain mapping*, *40*(13), 3810-3831.

Meunier, D., Lambiotte, R., & Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, *4*, 200.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81-97.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human brain mapping*, *35*(8), 4140-4154.

Niezgoda, M., Tarnowski, A., Kruszewski, M., & Kamiński, T. (2015). Towards testing auditory–vocal interfaces and detecting distraction while driving: A comparison of eye-movement measures in the assessment of cognitive workload. *Transportation research part F: traffic psychology and behaviour*, *32*, 23-34.

Okamoto, M., Dan, H., Sakamoto, K., Takeo, K., Shimizu, K., Kohno, S., ... & Dan, I. (2004). Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10–20 system oriented for transcranial functional brain mapping. *Neuroimage*, *21*(1), 99-111.

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human brain mapping*, *25*(1), 46-59.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 781-795.

Plichta, M. M., Herrmann, M. J., Baehne, C. G., Ehlis, A. C., Richter, M. M., Pauli, P., & Fallgatter, A. J. (2006). Event-related functional near-infrared spectroscopy (fNIRS): are the measurements reliable?. *Neuroimage*, *31*(1), 116-124.

Pupil Labs (2020). *Raw Data Exporter.* Retrieved from https://docs.pupil-labs.com/core/software/pupil-player/#raw-data-exporter on April 30, 2020.

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

Samuels, E. R., & Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part I: principles of functional organisation. *Current neuropharmacology*, *6*(3), 235-253.

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: the locus coeruleus mediates cognition through arousal. *Neuron*, *76*(1), 130-141.

Scharinger, C., Soutschek, A., Schubert, T., & Gerjets, P. (2015). When flanker meets the n-back: What EEG and pupil dilation data reveal about the interplay between the two central-executive working memory functions inhibition and updating. *Psychophysiology*, *52*(10), 1293-1304.

Schneider, M., Leuchs, L., Czisch, M., Sämann, P. G., & Spoormaker, V. I. (2018). Disentangling reward anticipation with simultaneous pupillometry/fMRI. *Neuroimage*, *178*, 11-22.

Schroeter, M. L., Kupka, T., Mildner, T., Uludağ, K., & von Cramon, D. Y. (2006). Investigating the post-stimulus undershoot of the BOLD signal—a simultaneous fMRI and fNIRS study. *Neuroimage*, *30*(2), 349-358.

Schroeter, M. L., Zysset, S., Kruggel, F., & Von Cramon, D. Y. (2003). Age dependency of the hemodynamic response as measured by functional near-infrared spectroscopy. *Neuroimage*, *19*(3), 555-564.

Schwab, B. C., Misselhorn, J., & Engel, A. K. (2019). Modulation of large-scale cortical coupling by transcranial alternating current stimulation. *Brain stimulation*, *12*(5), 1187-1196.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: applications to dementia and amnesia. *Journal of experimental psychology: General*, *117*(1), 34.

Stelmack, R. M., & Siddle, D. A. (1982). Pupillary dilation as an index of the orienting reflex. *Psychophysiology*, *19*(6), 706-708.

Tsuchida, A., & Fellows, L. K. (2009). Lesion evidence that two distinct regions within prefrontal cortex are critical for n-back performance in humans. *Journal of Cognitive Neuroscience*, *21*(12), 2263-2275.

van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review*, *25*(6), 2005-2015.

Wang, C. A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., & Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in neurology*, *9*, 1029.

Yaple, Z. A., Stevens, W. D., & Arsalidou, M. (2019). Meta-analyses of the n-back working memory task: fMRI evidence of age-related changes in prefrontal cortex involvement across the adult lifespan. *NeuroImage*, *196*, 16-31.

Yeung, M. K., Lee, T. L., & Chan, A. S. (2019). Right-lateralized frontal activation underlies successful updating of verbal working memory in adolescents with high-functioning autism spectrum disorder. *Biological psychology*, *148*, 107743.

Yeung, M. K., Sze, S. L., Woo, J., Kwok, T., Shum, D. H., Yu, R., & Chan, A. S. (2016). Reduced frontal activations at high working memory load in mild cognitive impairment: near-infrared spectroscopy. *Dementia and geriatric cognitive disorders*, *42*(5-6), 278-296.

Table 1.

Descriptive and inferential statistics of (a) task performance, (b) prefrontal activation, and (c) pupil dilation during the n-back task (n = 39).

	Mean	***SD***	***t*-value**	***p*-value**	**Cohen's *d***
(a) Task performance					
Sensitivity A' (unit)					
0-back	0.985	0.018	-	-	-
3-back	0.859	0.093	-	-	-
3-back > 0-back	-0.126	0.092	-8.56	< 0.001***	1.37
Mean RT (ms)					
0-back	342.5	38.4	-	-	-
3-back	478.6	127.5	-	-	-
3-back > 0-back	136.1	108.3	7.85	< 0.001***	1.26
(b) Mean change in the CBSI-corrected [oxy-Hb] (*Z*)					
Left					
0-back > Rest	-0.46	3.61	-0.80	0.43	-0.13
3-back > Rest	3.72	6.00	3.88	< 0.001***	0.62
3-back > 0-back	4.19	5.99	4.37	< 0.001***	0.70
Medial					
0-back > Rest	0.12	4.15	0.18	0.86	0.03
3-back > Rest	1.10	6.38	1.07	0.29	0.17
3-back > 0-back	0.98	6.96	0.88	0.39	0.14
Right					
0-back > Rest	-0.87	4.79	-1.14	0.26	-0.18
3-back > Rest	2.80	5.27	3.32	0.002**	0.53
3-back > 0-back	3.67	7.16	3.20	0.003**	0.51
(c) Mean change in pupil diameter (*Z*)					
0-back > Rest	0.49	2.66	1.14	0.26	0.18
3-back > Rest	2.89	3.39	5.33	< 0.001***	0.85
3-back > 0-back	2.41	3.49	4.31	< 0.001***	0.69

Note. CBSI = correlation-based signal improvement. Asterisks indicate the significance level of one-sample or paired *t*-tests (two-tailed). ***p* < 0.01, ****p* < 0.001

Figure Captions

Figure 1. Flow of the *n*-back paradigm.

Figure 2. Setup of the functional near-infrared spectroscopy and pupillometry systems.

Figure 3. Mean changes in the correlation-based signal improvement (CBSI)-corrected [oxy-Hb] during the *n*-back task. Dotted lines indicate the end of the 5-s instruction cue. A mean change in [oxy-Hb] refers to the average of all values across the 42-s task period (i.e., excluding the 5-s cue period). Asterisks indicate the significance level of one-sample or paired *t*-tests (two-tailed). ** $p < 0.01$, *** $p < 0.001$

Figure 4. Mean changes in the pupil diameter during the *n*-back task. Dotted lines indicate the end of the 5-s instruction cue. A mean change in the pupil diameter refers to the average of all values across the 42-s task period (i.e., excluding the 5-s cue period). Asterisks indicate the significance level of one-sample or paired *t*-tests (two-tailed). *** $p < 0.001$

Figure 5. Pearson's correlations between the time courses of the hemodynamic response predicted by the pupillary response and the time course of the observed hemodynamic response throughout the task period (i.e., excluding the 5-s cue period) during the *n*-back task. Asterisks indicate the significance level of one-sample *t*-test (two-tailed). * $p < 0.05$

Figure 1.

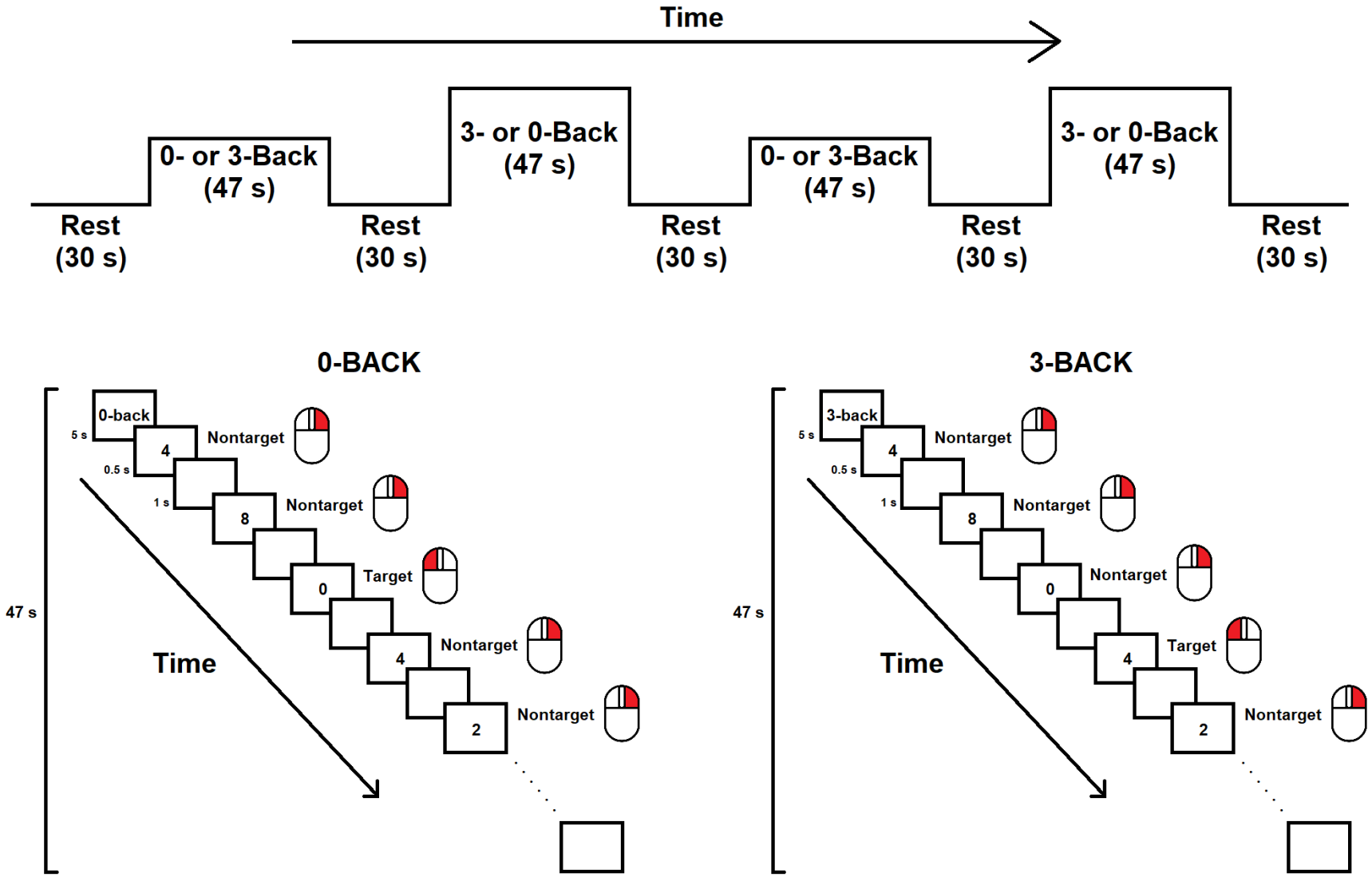

Time
0- or 3-Back (47 s)
3- or 0-Back (47 s)
0- or 3-Back (47 s)
3- or 0-Back (47 s)
Rest (30 s)
Rest (30 s)
Rest (30 s)
Rest (30 s)
Rest (30 s)
0-BACK
0-back
5 s
4
Nontarget
0.5 s
1 s
8
Nontarget
0
Target
4
Nontarget
2
Nontarget
47 s
Time
3-BACK
3-back
5 s
4
Nontarget
0.5 s
1 s
8
Nontarget
0
Nontarget
4
Target
2
Nontarget
47 s
Time

Figure 2.

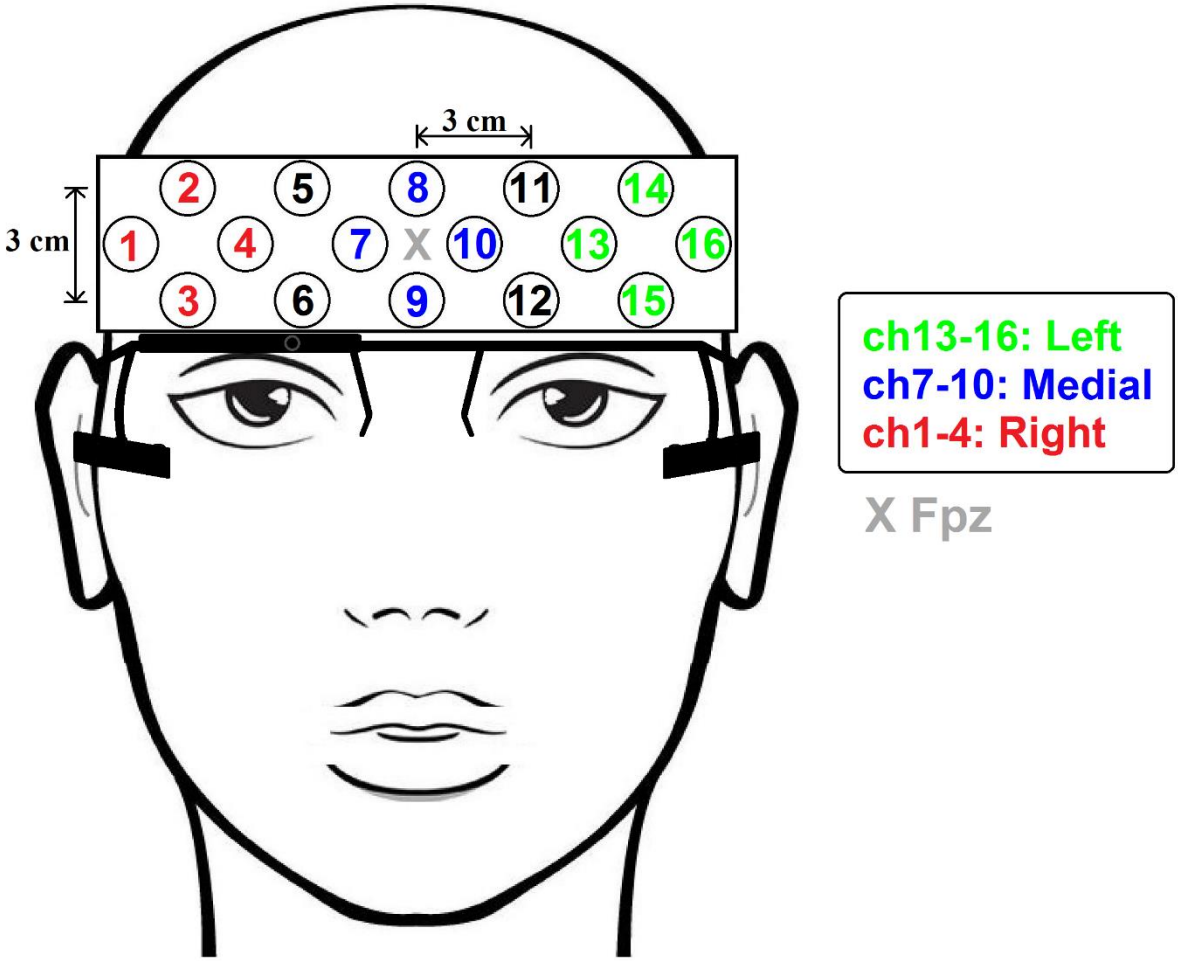
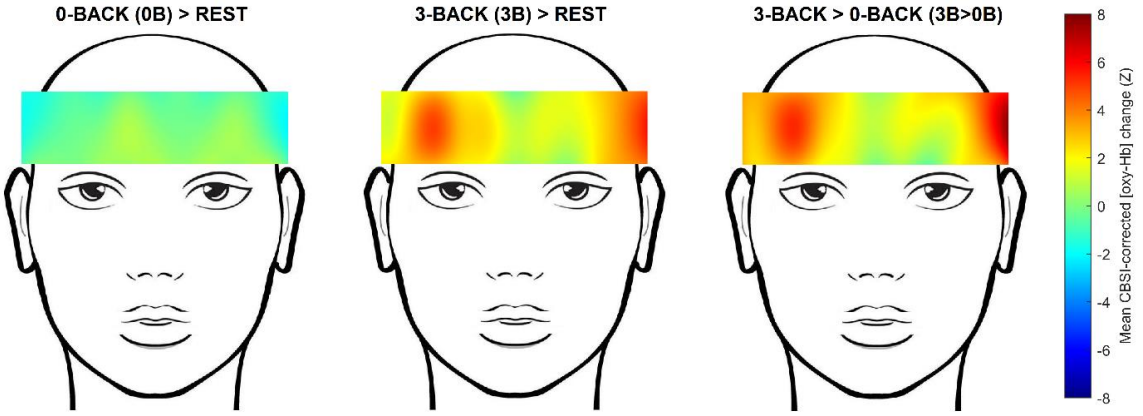
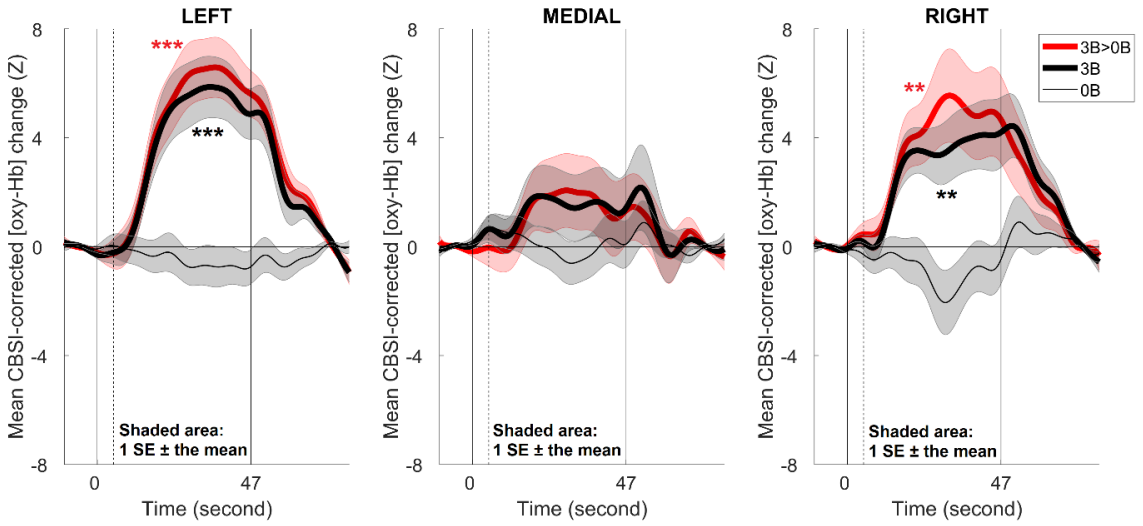

3 cm
3 cm
2
5
8
11
14
1
4
7
X
10
13
16
3
6
9
12
15
ch13-16: Left
ch7-10: Medial
ch1-4: Right
X Fpz

Figure 3.

(a)


0-BACK (0B) > REST
3-BACK (3B) > REST
3-BACK > 0-BACK (3B>0B)
Mean CBSI-corrected [oxy-Hb] change (Z)
8
6
4
2
0
-2
-4
-6
-8

(b)


LEFT
MEDIAL
RIGHT
3B>0B
3B
0B
***
***
**
**
Mean CBSI-corrected [oxy-Hb] change (Z)
Shaded area:
1 SE ± the mean
0
47
Time (second)

(c)

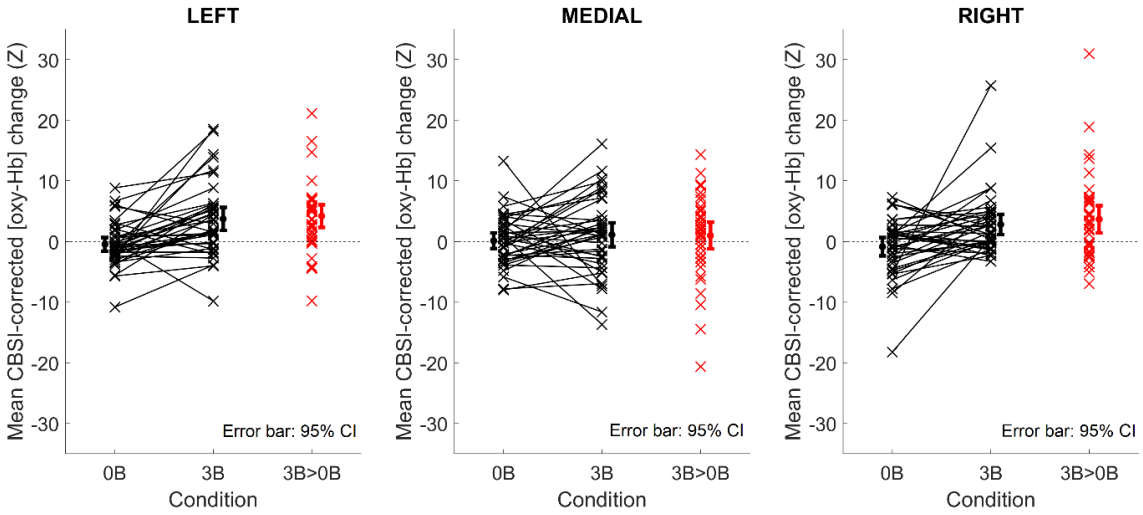

LEFT
MEDIAL
RIGHT
Mean CBSI-corrected [oxy-Hb] change (Z)
Error bar: 95% CI
0B
3B
3B>0B
Condition

Figure 4.

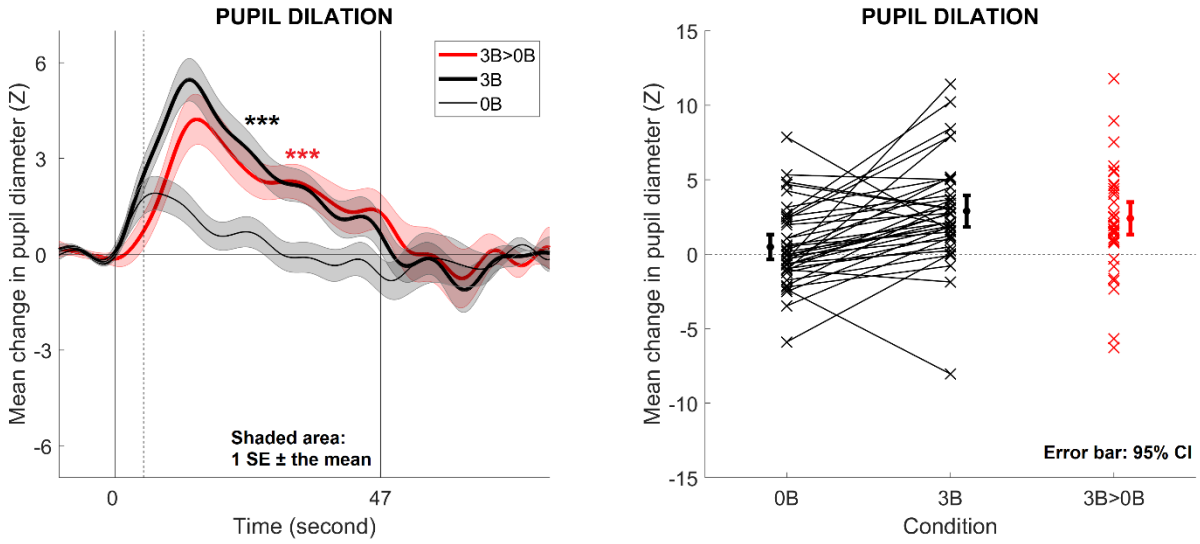
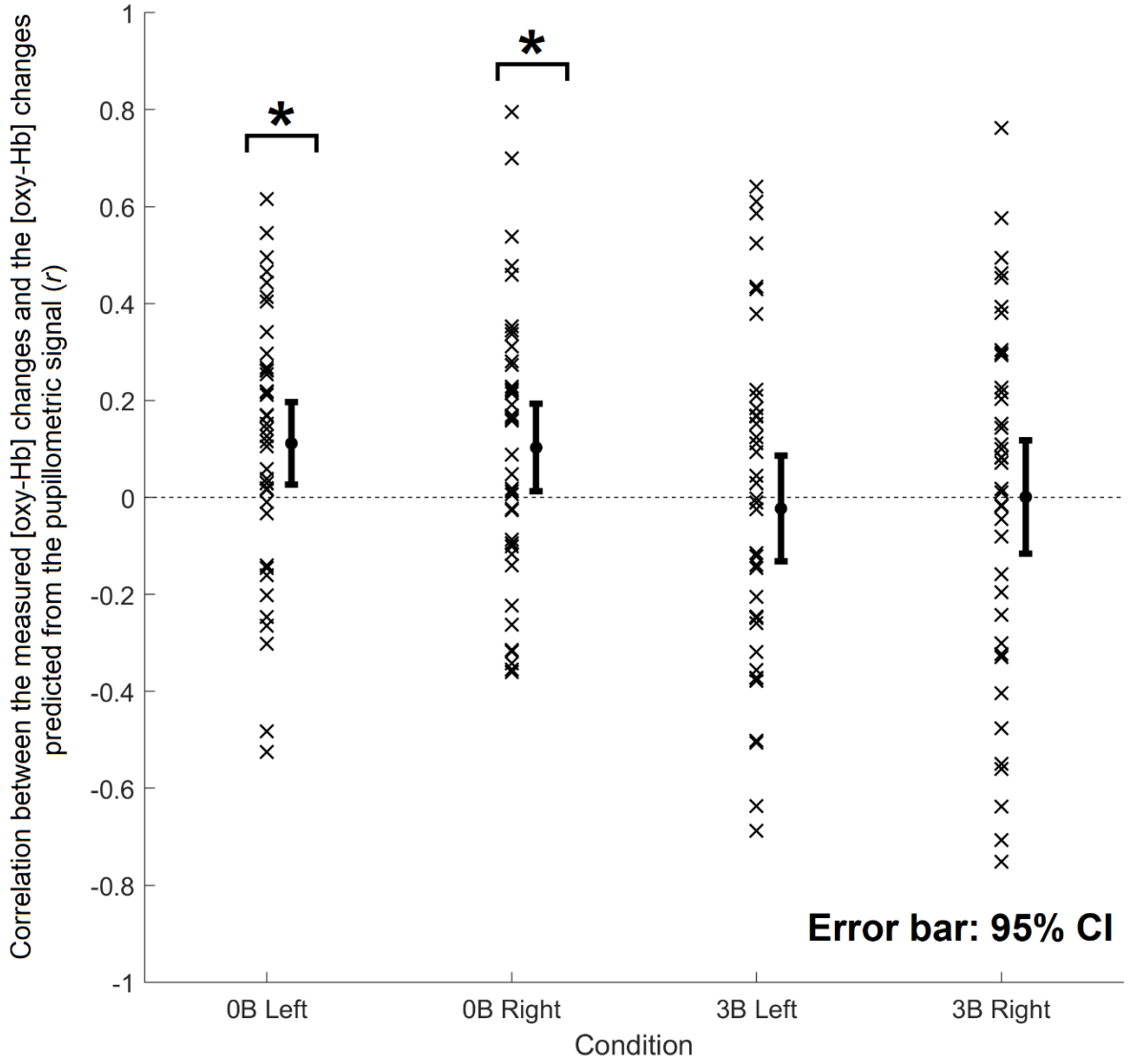

PUPIL DILATION
3B>0B
3B
0B
***
***
Mean change in pupil diameter (Z)
6
3
0
-3
-6
Shaded area:
1 SE ± the mean
0
47
Time (second)
PUPIL DILATION
15
10
5
0
-5
-10
-15
Mean change in pupil diameter (Z)
Error bar: 95% CI
0B
3B
3B>0B
Condition

Figure 5.


Correlation between the measured [oxy-Hb] changes and the [oxy-Hb] changes predicted from the pupillometric signal (r)
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
*
*
Error bar: 95% CI
0B Left
0B Right
3B Left
3B Right
Condition