






## Article

# When Eyes Wander Around: Mind-Wandering as Revealed by Eye Movement Analysis with Hidden Markov Models

Hsing-Hao Lee <sup>1,†</sup> , Zih-Ling Chen <sup>2,†</sup> , Su-Ling Yeh <sup>1,2,3,4,5,\*</sup> , Janet Huiwen Hsiao <sup>6,7</sup>   
and An-Yeu (Andy) Wu <sup>8</sup> 

<sup>1</sup> Department of Psychology, College of Science, National Taiwan University, Taipei City 10617, Taiwan; hsinghaolee@nyu.edu

<sup>2</sup> Graduate Institute of Brain and Mind Sciences, College of Medicine, National Taiwan University, Taipei City 10051, Taiwan; r06454014@ntu.edu.tw

<sup>3</sup> Neurobiology and Cognitive Science Center, National Taiwan University, Taipei City 10617, Taiwan

<sup>4</sup> Center for Artificial Intelligence and Advanced Robotics, National Taiwan University, Taipei City 10617, Taiwan

<sup>5</sup> Center for Advanced Study in the Behavioral Sciences, Stanford University, Stanford, CA 94305, USA

<sup>6</sup> Department of Psychology, The University of Hong Kong, Pok Fu Lam, Hong Kong; jhsiao@hku.hk

<sup>7</sup> State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Pok Fu Lam, Hong Kong

<sup>8</sup> Graduate Institute of Electronics Engineering, National Taiwan University, Taipei City 10617, Taiwan; andywu@ntu.edu.tw

\* Correspondence: suling@ntu.edu.tw; Tel.: +886-2-33663097

† Co-first author.



**Citation:** Lee, H.-H.; Chen, Z.-L.; Yeh, S.-L.; Hsiao, J.H.; Wu, A.-Y. When Eyes Wander Around: Mind-Wandering as Revealed by Eye Movement Analysis with Hidden Markov Models. *Sensors* **2021**, *21*, 7569. <https://doi.org/10.3390/s21227569>

Academic Editor: José Antonio Ruipérez-Valiente

Received: 16 October 2021

Accepted: 10 November 2021

Published: 14 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Mind-wandering has been shown to largely influence our learning efficiency, especially in the digital and distracting era nowadays. Detecting mind-wandering thus becomes imperative in educational scenarios. Here, we used a wearable eye-tracker to record eye movements during the sustained attention to response task. Eye movement analysis with hidden Markov models (EMHMM), which takes both spatial and temporal eye-movement information into account, was used to examine if participants' eye movement patterns can differentiate between the states of focused attention and mind-wandering. Two representative eye movement patterns were discovered through clustering using EMHMM: centralized and distributed patterns. Results showed that participants with the centralized pattern had better performance on detecting targets and rated themselves as more focused than those with the distributed pattern. This study indicates that distinct eye movement patterns are associated with different attentional states (focused attention vs. mind-wandering) and demonstrates a novel approach in using EMHMM to study attention. Moreover, this study provides a potential approach to capture the mind-wandering state in the classroom without interrupting the ongoing learning behavior.

**Keywords:** mind-wandering; sustained attention; eye movement analysis with hidden Markov models (EMHMM); fixation; learning

## 1. Introduction

Mind-wandering (MW), the shift of attention from the current task to task-unrelated thoughts, is a universal experience that occupies 47% of adults' daily thinking time [1]. We live in an era full of distractions where modern technology and social media have become a pervasive part of our lives. The increase of distractions has caused people to have more difficulty in concentrating on tasks. Although MW benefits creativity, imagination, and plans for the future [2,3], it also accompanies negative emotional feelings [4]. Moreover, MW is negatively correlated with task performance. For example, Stothart et al. [5] showed that cellphone notifications disrupted task performance in an attention demanding task, even when participants did not check their phones. Other studies also showed that MW impaired the extent of text comprehension [6,7], and even jeopardized safety during

driving [8]. In the educational scenario, MW can impact learning efficiency [9] and yield significant class performance cost [10]. Therefore, understanding when and what kind of people tend to mind-wander is a critical issue in modern society [11].

Detecting MW using a wearable device can help resolve this issue, and viewers' eye movements can be used as a good index of their attentional states. Indeed, as remote classes have become more mainstream recently due to the COVID-19 pandemic, one of the biggest challenges of online learning is struggling with staying focused on a screen for long periods of time. With online learning, there is also a higher chance for students to be distracted by social media, advertisements, or other websites. An eye-tracking detection system could thus be an alternative strategy for supervision without directly interfering with classes. By applying a wearable eye-tracking system for detecting attentional states, it is possible for instructors to notice when students have lapses of attention and adjust content accordingly. Additionally, with the development of imaging processing techniques, it is possible to capture people's eye movements using a low-cost camcorder (e.g., [12]), which makes it more plausible to use eye movement as an index of attention in remote learning scenarios.

It has been shown that eye movement and attention are strictly coupled, both temporally and spatially [13,14]. Thus, observers' eye movements are often used to investigate the deployment of attention [15,16]. Eye fixations allow people to focus on the target and maintain high acuity of the target on the fovea, and thus can serve as an index of attention [17]. The viewer's fixational behavior contains abundant spatial information while the transition of fixations includes the temporal information of eye movements. However, most studies investigating the relationships between eye movements and sustained attention emphasized spatial information rather than considering spatial and temporal information jointly. More specifically, most studies analyzed the fixation duration and time points when fixations lie in the pre-defined regions of interest (ROIs) as the indices of attention (e.g., [18–20]). Other studies have found an increased number of fixations and longer durations prior to reporting MW in the pre-defined ROIs during a reading task [21,22]. Nevertheless, these studies ignored an important aspect that might provide clues for the relationship between gaze and MW, namely the transition of fixation in the temporal domain.

The transition of fixation in the temporal domain can reflect the planning and strategy processes of the human mind [23] and where the eyes intend to land [24]. Indeed, both are highly correlated with attentional deployment. In addition to attention, individual differences in how likely people will deploy their eye movements to specific regions were identified to be correlated with individuals' cognitive performances [25]. Yet, this line of approach, by taking both spatial and temporal information of fixations into account, was missing in the research field of sustained attention. In addition, pre-defined ROIs might be subjective and arbitrary in determining the crucial ROIs for detecting sustained attention. Since every researcher has different pre-defined ROIs and different experimental stimuli, when it comes to counting the number of fixations lying in a targeted region (a pre-defined ROI), such as the lecturer in a lecture video, researchers might not be able to generalize the results to other scenarios that do not include a lecturer.

Eye movement analysis with hidden Markov models (EMHMM) can solve such a problem of arbitrarily pre-defined ROIs. This approach can determine ROIs based on transition information between fixations, in addition to fixation locations, and hence provide a data-driven approach for defining the ROIs. In addition, by calculating the probability of transition across ROIs, this approach can take into account both spatial and temporal information as well as individual differences in the viewing paths [26]. For example, Chan et al. [27] discovered that participants who used a similar movement pattern (focusing on the eye region of target faces) to view faces with angry and neutral expressions had higher social anxiety symptoms than those who transitioned their viewing strategies from focusing on the eyes to focusing on the nose. This study showed that relationships between viewing patterns and psychopathology could be revealed by the

EMHMM approach. Furthermore, in a face recognition task, Chan et al. [28] found that, in the aging population, those older adults who used more analytic viewing patterns scored higher on their cognitive test performance. Hence, EMHMM can also reveal individual differences in cognitive functions. However, the traditional eye movement data analysis approach cannot achieve these results because the usage of heat maps and fixation counts in pre-defined ROIs can only reveal the frequency that participants focus on the eyes of a face instead of identifying dynamic eye movement patterns. Therefore, by adding the temporal information using EMHMM when determining the relationship between MW and eye movements, we can quantify to what extent participants tend to deploy their eye movements with a specific pattern. This is likely to be related to MW given the close relationship between attention and eye movements. In the near future, this can thus become an index of MW applied to the educational scenario where focused attention is the key to efficient learning [29].

The aim of the current study is to use EMHMM by taking both spatial and temporal information of eye movements to identify specific eye movement patterns that can serve as indices of sustained attention. We hypothesized that MW measured with response to the target (i.e., an objective measure of MW) and subjective report (i.e., a subjective measure of MW) can be revealed by different eye movement patterns.

## 2. Materials and Methods

### 2.1. Participants

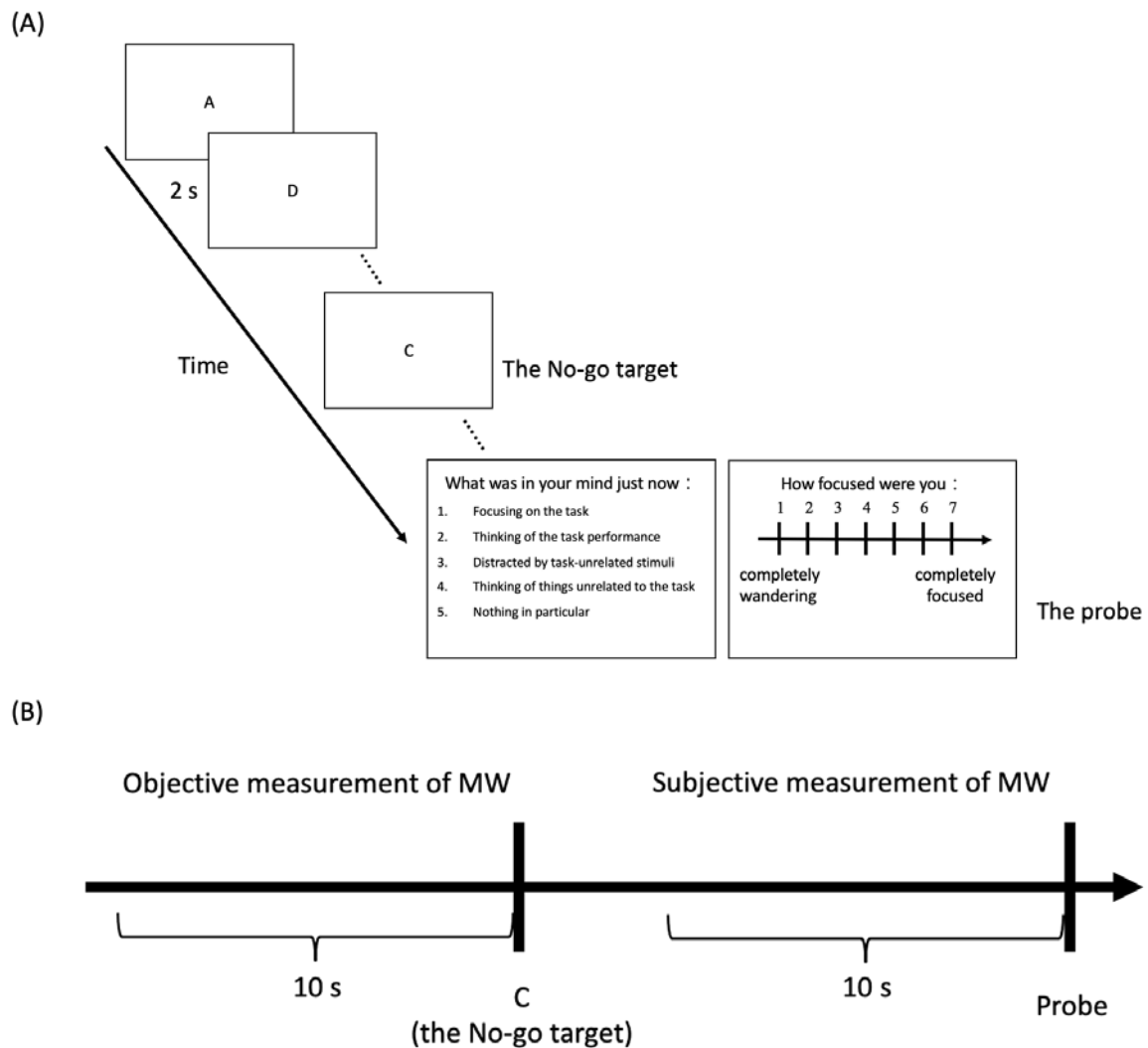
The targeted sample size was determined using the effect size (Cohen's  $d = 1.18$ ) of Chuk et al. [26] where two different viewing patterns in face recognition were found using EMHMM. According to the G-Power 3.1.9.6 software [30], 13 participants for each eye movement viewing group (based on their eye movement patterns; i.e., 26 participants) were required to reach decent statistical power (0.8). To be more conservative, we recruited 20% more participants than needed to verify our results. Therefore, 31 healthy adults were recruited to complete this study (mean age = 22.77 years, SD = 2.87 years, 18 females). All participants were right-handed and free from psychological and neurological disorders. All had normal or corrected-to-normal vision. Participants were naïve to the goal of the experiment. Participants signed the informed consent before the experiment and were rewarded 400 NTD for their participation.

### 2.2. Apparatus and Stimuli

Eye movement data were recorded by Tobii Pro Glasses 2 with Tobii SDK program sampled at 100 Hz. Saccades were defined as eye velocity signals exceeding 100 deg/s, and fixation as events where eye velocity was lower than 100 deg/s and maintained for at least 60 ms.

Stimuli were shown in black against a gray background and presented using the program E-prime (Psychology Software Tools, Pittsburgh, PA, USA). We employed the sustained attention to response task (SART; [31,32]) to measure MW. The SART is a Go/No-go task that has been widely used to induce MW and measure the state of attention. In the SART, 25 English letters (A–Y, except for Z) were presented pseudo-randomly at the center of the screen (extending approximately  $0.72^\circ$  horizontally and vertically), with a target letter (the letter C), presented between the 6th and 15th trial in a block. Each letter was presented for 2 s or until the participant responded. The inter-trial interval (ITI) varied with the reaction time of participants so that each trial (including the ITI) lasted for 2000 ms. For example, if the participant's reaction time was 300 ms, then the ITI would be 1700 ms to equate the duration of each trial. After 25 trials (with one No-go trial, the target), at the end of each block, participants were asked to answer a probe that asked them to subjectively report their state of attention (Figure 1A). They were instructed to answer the thought probe, "What was in your mind just now?" first with five options consisting of the following: 1. Focusing on the task; 2. Thinking of the task performance; 3. Distracted by task-unrelated stimuli; 4. Thinking of things unrelated to the task; 5. Nothing in particular.

Then, participants were asked to rate how focused they were from 1 (completely wandering) to 7 (completely focused) for the moment right before seeing the thought probe. There were 40 blocks in total. Participants were instructed to press 8 on the number-pad to respond to a Go trial and answer the probe question with the corresponding number buttons. After the probe, participants were told to take a short break and press the number 9 to initiate the next block at their own pace.



**Figure 1.** (A) Experimental procedure of the SART (the actual background color was grey). Participants were instructed to press the number 8 on the number-pad as quickly as possible whenever they saw an English letter but to withhold their response when they saw the target letter C. After 25 trials (at the end of a block), participants were instructed to answer the two probe questions. (B) The analyzed time windows of the SART. The objective measurement of MW was analyzed in the time window 10 s before the English letter C (i.e., the No-go target); the subjective measurement of MW was analyzed in the time window 10 s before the probe.

### 2.3. Procedure

The experiment was conducted in a sound-attenuated room. Participants were seated with their eyes approximately 80 cm away from the monitor and were instructed to do the one-point calibration for the Tobii Pro Glasses 2. Participants were given a detailed description of the thought probe and task content before beginning the experiment. They were told that there was no correct answer regarding the probe so that they could answer truthfully. The main experiment was preceded by three blocks of practice trials (25 trials per block). Participants were required to press a button (i.e., the Go trials) as soon as seeing

any English letter other than the letter C, but to withhold a response when the target (i.e., the No-go target, the letter C) was presented. The sensitivity ( $d'$ ) towards the target serves as the main performance index of MW [32,33]. In addition to such an objective measure of MW, we inserted a probe question as the thought sampling method to ask participants what they were thinking about at the moment right before the question appeared. Immediately after the probe question, participants were asked to rate how focused they were from a 7-point Likert scale at the moment right before the probe appeared. Their answer to the probe question and the rating scale about their attentional state (MW or focused, from 1 to 7) were used as the subjective measures of MW. On another note, this study is a portion of a bigger project that includes other physiological measurements, which will not be elaborated here, but see Chen et al. [34].

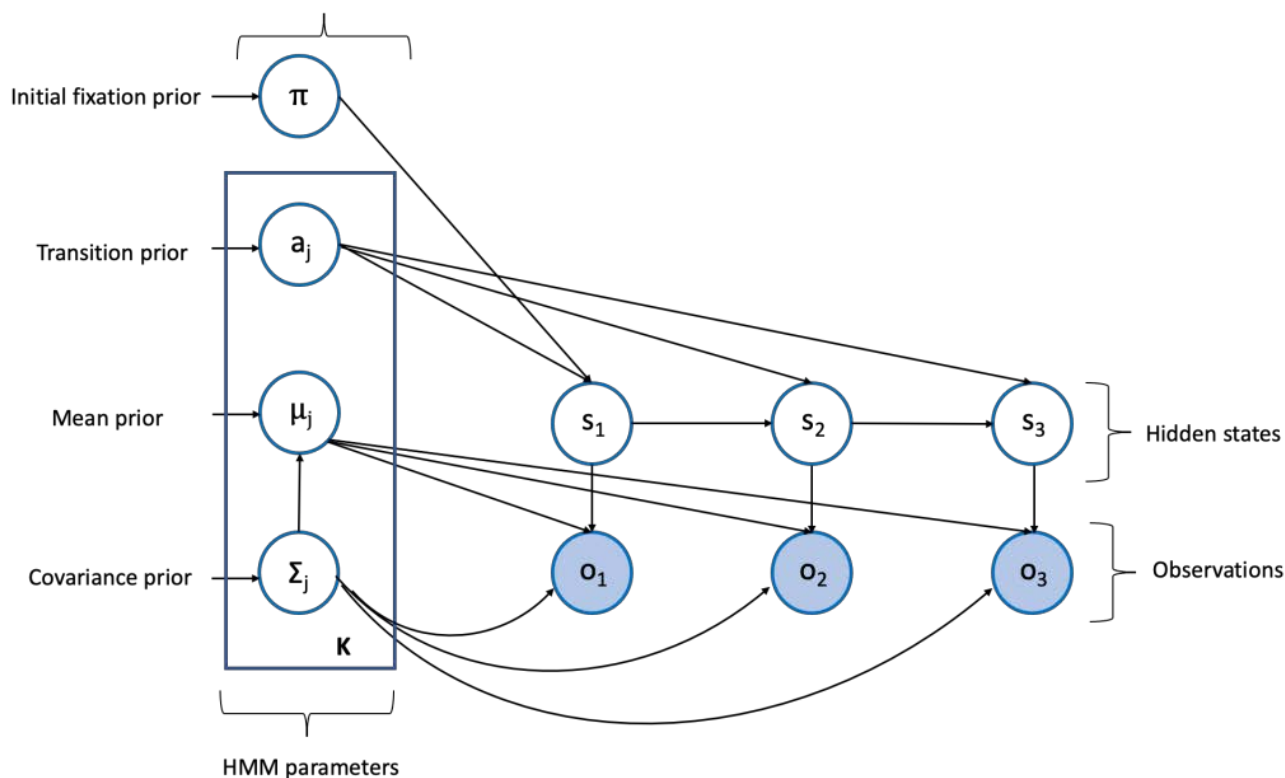
#### 2.4. Data Analysis

We conducted data analysis on the objective and subjective measures of MW separately (Figure 1B). In terms of the objective measure, the 10-s pre-target intervals preceding the No-go target trial were categorized as focused attention (FA) or MW based on the participant's sensitivity toward the No-go target ( $d'$ ; see below). In terms of the subjective measure, the 10-s pre-probe intervals were categorized as FA or MW based on the participant's subjective responses to the two probe questions. Subjective FA required fulfilling two criteria, namely responding with options 1, 2, and 5 for the first question (1: Focused on the task, 2: Thinking of the task performance, and 5: Nothing in particular) and having rating scores of 5–7 for the focus rating question on the 7-point scale. Subjective MW also required fulfilling two criteria, namely responding with options 3, 4, and 5 (3: Distracted by task-unrelated stimuli, 4: Thinking of things unrelated to the task, and 5: Nothing in particular) and having rating scores 1–3 for the focus rating question on the 7-point. The 5th option, "Nothing in particular", was defined as a neutral state and could be considered as either FA or MW, for the following reason. First, given that the SART is a relatively low-demand task, people with high working memory capacity can complete the task by devoting much fewer resources (i.e., nothing in particular in their mind) compared to people with low working memory capacity. Additionally, people tend not to be able to qualify their thought content all the time because it requires the ability to monitor one's own mental state [35] and might find the other four categories unfit. Thus, we provided the fifth option and categorized the trials selected as "Nothing in particular" as FA or MW based on the response of the subjective rating scale. Since the numeric quantification is more instinctive than content report, trials with focused rating scores greater than 4 were categorized as FA trials, and trials with focused rating scores lower than 4 were categorized as MW trials. Trials with a focused rating score of 4 on the Likert scale were defined as an ambiguous state because people can simultaneously be unfocused but also not mind-wandering (i.e., the gap between MW and FA).

For the objective measure of MW, we quantified participants' performance of sustained attention based on signal detection theory (SDT) in the time window of 10 s prior to target onset. The 10-s time window was determined according to Christoff et al. [36] who used functional magnetic resonance imaging (fMRI) to reveal the MW-related neural network. A similar time window was also used in other studies (e.g., [8,37]). If participants successfully withheld their response to the target, it was counted as a hit, and if not, the response would be counted as a miss. If participants failed to respond to a non-target letter, it was counted as a false alarm, or it was counted as a correct rejection. We then calculated  $d'$  based on the hit rate and false alarm rate. With respect to the subjective measure, the proportion of rating FA (out of the 40 probes) was adopted as the dependent variable.

Fixations with durations above three standard deviations of the individual's mean were excluded (2% for the pre-target session and 2% for the pre-probe session). For the pre-probe intervals, as the rating score 4 is an ambiguous state, either FA or MW, responses with a rating score of 4 were excluded from data analysis (15.81%). Overall, 17.81% of the data were excluded in the pre-probe analysis.

Eye movements were analyzed using EMHMM ([26]; toolbox: <http://visal.cs.cityu.edu.hk/research/emhmm/> accessed on 1 September 2019). Figure 2 underlines the logic of the EMHMM. The model took the x-y-coordinates for the fixations across time. The time windows of the modeling data were taken from the 10-s pre-target intervals and 10-s pre-probe intervals for objective and subjective measures of MW, respectively. The hidden states of the HMM represented the regions of interest (ROIs) for fixations. Each ROI is a Gaussian, and thus the HMM is a time series of mixtures of Gaussians [38]. We set the possible number of hidden states (ROIs) to be from three ( $K = 3$ ) to six ( $K = 6$ ) and chose the one with the highest data log-likelihood using the variational Bayesian method for the pre-target and the pre-probe intervals separately. This allowed us to select the model within this range that had the highest data log-likelihood in a bottom-up (data-driven) way. The parameters of each individual HMM were estimated using the variational Bayesian expectation-maximization (VBEM) algorithm [39], which places a prior distribution on each parameter and then approximates its posterior distribution using a factorized variational distribution.



**Figure 2.** Structure and parameters of an HMM. The  $O_n$  indicates the observed fixation data; The  $S_n$  represents the hidden states. The prior distributions of the HMM parameters were presented on the left where  $K$  is the number of the hidden states.

We then cluster the individuals' HMMs into groups and form representative HMMs for each group, which summarize each group's eye movements. The number of clusters was predetermined, which followed previous EMHMM studies where participants' eye movement patterns could be quantified along the dimension between two contrasting patterns [25,27,28,40–45]. To cluster the HMMs into two groups, so as to reveal common patterns among individuals, we used the variational hierarchical expectation-maximization (VHEM) algorithm [46], which clustered HMMs into groups in a bottom-up way based on their similarities and further produced the representative HMMs for each group to describe the ROIs and transitional information in the cluster [26]. More specifically, the algorithm first initialized the Gaussian emissions and transition matrix of each representative HMM using a randomly selected input HMM. Then, it iterated between the E-step and the M-step until convergence. At the E-step, it estimated the expectation of the log-

likelihood (similarity) of the representative HMMs with respect to the input HMMs. At the M-step, it grouped the input HMMs according to their similarity to the representative HMMs, and then updated the parameters of the representative HMMs using these cluster assignments [38]. Following previous EMHMM studies, we set the number of ROIs in the representative HMMs to the median number of ROIs in the individual models, performed the VHEM algorithm 100 times, and used the clustering results with the highest expected log-likelihood. We quantified the degree of similarity between individual HMMs and the two representative HMMs using data log-likelihoods. Here, we termed the two representative eye movement patterns distributed and centralized patterns hereafter in the present study based on their characteristics (cf. [26,42,43,45]). The mean-log-likelihood (MLL) of each participant's eye movement data given the representative HMMs of the distributed and centralized patterns of eye movements was calculated. We defined the D-C scale as the difference in MLL between using the distributed and the centralized patterns [45], which was calculated as the following:

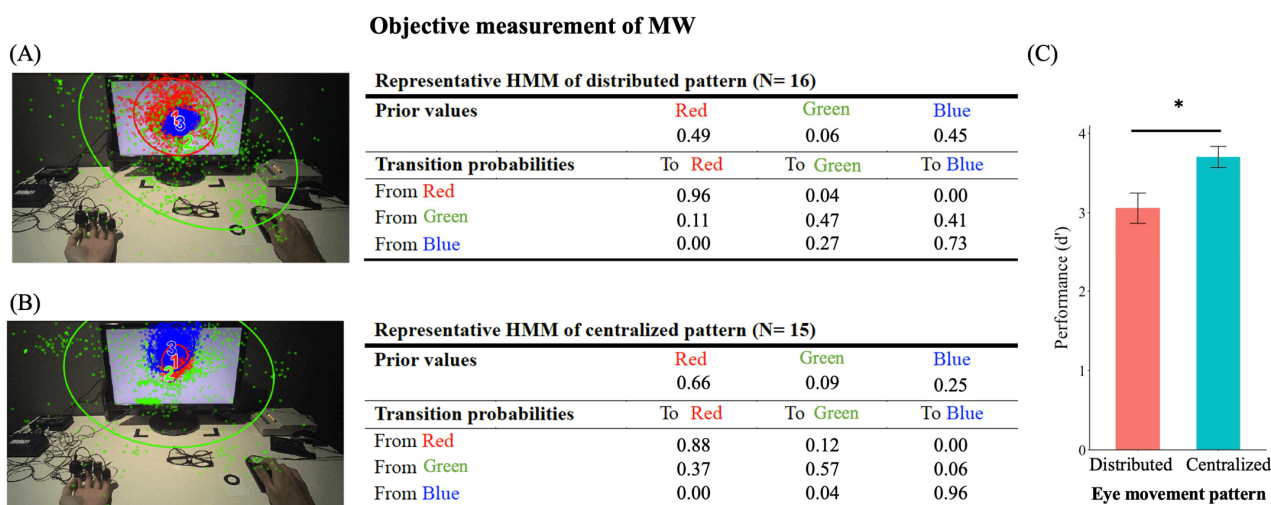
$$\frac{D \text{ MLL} - C \text{ MLL}}{|D \text{ MLL}| + |C \text{ MLL}|} \quad (1)$$

where D MLL indicates the MLL given the distributed pattern, and C MLL indicates the MLL given the centralized pattern. A more positive value represents a viewing pattern more similar to the distributed pattern, and a more negative value represents a viewing pattern more similar to the centralized pattern. We then used the D-C scale as a quantitative measure of participants' eye movement patterns during the task [28].

### 3. Results

#### 3.1. Eye Movement Data during the 10-s Pre-Target Period

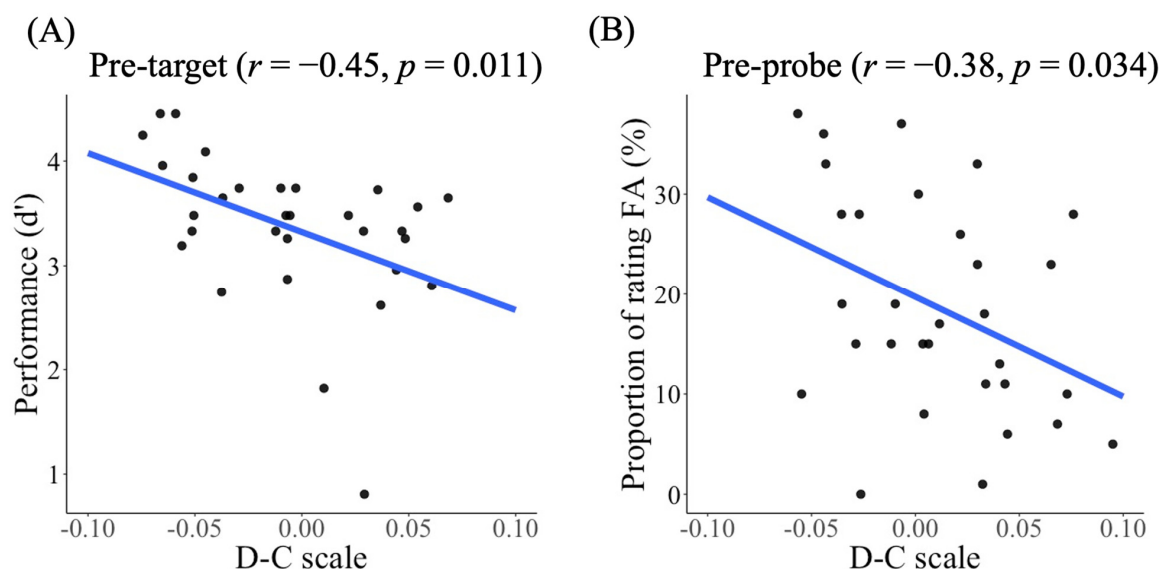
The two representative HMMs are shown in Figure 3A,B, which were the distributed pattern and the centralized pattern respectively based on the distributions of their ROIs. To evaluate if the centralized pattern is different from the distributed pattern, we calculated the mean log-likelihoods of the fixation sequences from the distributed pattern using the distributed and the centralized HMMs. The pairwise *t*-test showed that distributed participants' fixation sequences were more likely to be generated by the distributed HMM than the centralized HMM,  $t(15) = 4.33$ ,  $p < 0.001$ . The same procedure was used for the centralized group, with similar findings obtained. Namely, centralized participants' fixation sequences were more likely to be generated by the centralized HMM than the distributed HMM,  $t(14) = 7.63$ ,  $p < 0.001$ . According to the D-C scale, the distributed pattern (group) consisted of 16 participants and the centralized pattern (group) consisted of 15 participants. Based on the reported ROI locations, orders, and probabilities, people with the distributed pattern had a similar prior probability to start a fixation sequence from the red ROI and the blue ROI, as shown in Figure 3A. They most likely first scanned a wide range across the screen (the red ROI), then looked elsewhere away from the stimuli (the green ROI), and finally scanned back to the central region (the blue ROI). Participants with the centralized pattern most likely focused on the specific central region first (the red ROI), then scanned the left and right sides of the stimuli (the green ROI), and finally returned to the specific central region (the red or blue ROI).



**Figure 3.** The two representative eye movement patterns during the 10-s pre-target period for (A) the distributed eye movement pattern and (B) the centralized eye movement pattern. The tables in the middle panel show the transition matrix among the ROIs. Prior values refer to the probability of the first fixation of a trial landed at a specific ROI. (C)  $d'$  of participants with the two eye movement patterns as the objective measurement of MW. Error bars represent one S.E.M. \*  $p < 0.05$ .

### 3.2. Behavioral Performance during the 10-s Pre-Target Period

We compared the task performance using  $d'$  (see Methods). Figure 3C shows the results of  $d'$  across groups. Participants with the centralized pattern performed better (i.e., higher  $d'$ ) than those with the distributed pattern,  $t(29) = -2.74$ ,  $p = 0.01$ ,  $d = 0.99$ . Furthermore,  $d'$  was negatively correlated with the D-C scale,  $r = -0.45$ ,  $p = 0.011$ , suggesting that the more distributed the eye movement pattern, the poorer the performance (Figure 4A). Other behavioral performances and eye movement indices are summarized in Appendix A Figure A1.



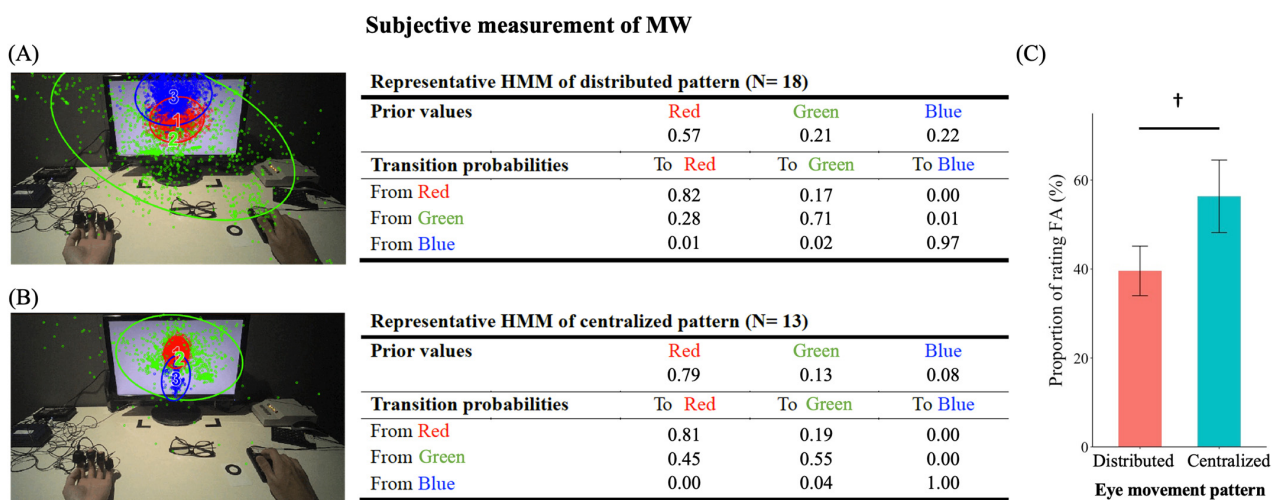
**Figure 4.** The correlations between (A) the performance ( $d'$ ) and the D-C scale during the pre-target phase (B) the proportion of rating focused attention (FA) and the D-C scale during the pre-probe phase.

### 3.3. Eye Movement Data during the 10-s Pre-Probe Period

Figure 5A,B show the HMMs of the two representative eye movement patterns. The distributed pattern (group) consisted of 18 participants and the centralized pattern (group)



consisted of 13 participants. The pairwise  $t$ -test showed that distributed participants' fixation sequences were more likely to be generated by the distributed HMM than the centralized HMM,  $t(17) = 5.79$ ,  $p < 0.001$ . The same procedure was used on the centralized group, with similar findings obtained. Centralized participants' fixation sequences were more likely to be generated by the centralized HMM than the distributed HMM,  $t(12) = 5.51$ ,  $p < 0.001$ . The results here suggest that the distributed and centralized HMMs represent two distinctive eye movement patterns. Based on the reported ROI locations, orders, and probabilities, participants with the distributed pattern demonstrated a wider range of viewing, whereas participants with a centralized pattern showed a high probability to look at the center and continued to view the central region. The ROIs for the centralized pattern were all inside the monitor whereas the ROIs for the distributed pattern were expanded across the entire visual field.



**Figure 5.** The two representative eye movement patterns during the 10-s pre-probe period for (A) the distributed eye movement pattern and (B) the centralized eye movement pattern. The tables in the middle panel show the transition matrix. Prior values refer to the probability of the first fixation of a trial landed at a specific ROI. (C) The proportion of rating focused attention (FA) for participants with the two eye movement patterns as a subjective measurement of MW. Error bars represent one S.E.M. †  $p < 0.1$ .

### 3.4. Behavioral Performance during the 10-s Pre-Probe Period

We compared the task performance of participants using the two eye movement patterns in the proportion of rating FA. Participants with the centralized pattern tended to rate themselves as more focused than those with the distributed pattern,  $t(29) = -1.76$ ,  $p = 0.089$ ,  $d = 0.629$  (Figure 5C). The proportion of rating FA was negatively correlated with the D-C scale,  $r = -0.38$ ,  $p = 0.034$ , suggesting that the more distributed the pattern, the lower the proportion of self-rated FA (Figure 4B). Other behavioral performance data and eye movement indices were summarized in Appendix A Figure A2.

### 3.5. Trial by Trial Analysis

To examine if our model can work on a trial-by-trial level instead of being limited to classifying people who are more prone to MW from those who are not, following Zhang et al. [44], the eye movement pattern for each trial across all participants was classified into the centralized or distributed pattern according to the log-likelihood generated by the representative model (Appendix A Tables A3 and A4). Here, we performed the likelihood ratio chi-squared statistical analysis (the  $G^2$  test) to see if trials belonging to the centralized pattern would be more likely to have correct no-responses to the target and also have higher proportions of rating FA. The  $G^2$  test is a maximum likelihood statistical test that provides an approximation of the theoretical chi-squared distribution better than the Pearson's chi-squared test [47]. For pre-target intervals, the odds ratio for trials in the

centralized pattern to withhold successfully and trials in the distributed pattern to fail to withhold were 1.88 times more likely than vice versa ( $G^2 = 24.01, p < 0.001$ ). For the pre-probe intervals, the odds ratio given that trials in the centralized pattern were scored as FA and trials in the distributed pattern were scored as MW was 2.39 times more likely than vice versa ( $G^2 = 33.56, p < 0.001$ ). The results support that the eye movement pattern measures quantified using EMHMM are not just limited to the participants' trait level, but can also be applied trial-by-trial. Namely, by analyzing distributed trials and centralized trials, we can apply the model to a trial-by-trial basis. Therefore, we are able to instantly detect people's attentional state based on the eye movement pattern in real-time rather than only classifying attentional states based on subsequent data analyses.

#### 4. Discussion

The current study found that MW can be revealed by eye movement patterns, as we categorized eye movement patterns into the distributed pattern and the centralized pattern via EMHMM. More importantly, we discovered that participants with the distributed pattern were more prone to MW than people using the centralized pattern. We drew this conclusion after analyzing an objective measure regarding performance towards the target and subjective ratings that included lower sensitivity ( $d'$ ) towards the target (withholding keypress) and a lower proportion of rating FA.

##### 4.1. The Relationships between Mind-Wandering and Eye Movement Patterns

In line with the discoveries in the current study, people who used a more centralized (less distributed) eye movement pattern as their strategy had better cognitive performances. For example, Chan et al. [28] found that older adults whose eye movement patterns showed better concentration on facial features in a face recognition task (i.e., the analytic pattern) had higher scores in the Montreal Cognitive Assessment (MoCA), which is a well-established neuropsychological test examining people's language, executive, visuospatial processing, and memory functions [48]. Chan et al. [49] also showed that people using the analytic pattern demonstrated more activation in brain regions related to top-down control compared to people using the holistic viewing pattern in a face recognition task, such as the frontal eye field (FEF), dorsolateral prefrontal cortex (DLPFC), and intraparietal sulcus (IPS). Therefore, it is possible that participants who adopted the centralized pattern engaged more top-down control of attention that helped filter out irrelevant information and improved the efficiency of information processing [49]. Additionally, in the education scenario, Zheng et al. [45] showed that participants who looked more at the center of the screen (i.e., the centralized pattern) had better comprehension of the lesson materials than those who looked around more (i.e., the distributed pattern). These studies along with our results verified that people using the centralized pattern have better cognitive performance in general compared to people using the distributed pattern.

In addition, both objective and subjective indices ( $d'$  and proportion of rating FA) were negatively correlated with the D-C scale, suggesting that the more distributed the eye movement pattern, the worse the performance in both measurements. The likelihood ratio test further verified that trials with the centralized pattern were more likely to be FA and trials with the distributed pattern were more likely to be MW. Therefore, in the future, we could possibly replace task performance index and subjective reports of attentional states with eye movement behaviors as a real-time indicator of MW. To be more specific, if the distributed pattern is detected from one's eye movements, there is a high possibility of disengagement from one's current task. As sustained attention plays a critical role in learning and memory [50], understanding when people tend to mind-wander may help them direct their focused attention back to the learning materials.

Some may argue that the SART commission error and probe-based response cannot represent the MW state. For instance, looking at the commission error as the objective measurement of MW, Head and Helton [51] suggested that the commission error indicates the failure of executive control but not MW. However, according to Robertson et al. [32],

the failure of executive control is the “consequence” of MW but not the main cause of the commission error. Indeed, Seli [52] has shown that SART errors do reflect MW even when controlling the RTs for go trials. As for the probe-based responses, despite some studies questioning its validity (e.g., [53]) and its characteristic to interrupt the task [54], others have proposed that thought probe is relatively robust to the variability of task parameters, and hence suitable to examine the thought content during a MW-related task [55]. Notwithstanding the limitations of the thought probe response, we investigated the MW based on the objective and subjective measure as Faber et al. [56] suggested, and found similar eye movement pattern for objective and subjective MW (i.e., the distributed pattern) as well as objective and subjective FA (i.e., the centralized pattern).

#### 4.2. Applications and Future Works

The eye movement detection system for MW can be applied to educational scenarios. Indeed, the commission error in the SART is a sensitive measure of the attention that is associated with the focused state of children during class [57]. Previous studies using behavioral tasks to measure MW, such as a detection system that pops out a window to check students’ attentional states after 10 min of idle time (when no mouse movement or keyboard activity has been detected), might interfere with the learning process. Such a method of monitoring student performance is disturbing and may cause dual-task interference (i.e., taking notes and moving the mouse), which leads to a more significant cognitive overload and might not be the best way to capture students’ attentional states [58]. Using an eye-tracking system to detect MW can thus avoid causing the extra dual-task demand. More importantly, the SART here is essentially analogous to the scenario when we are listening to lectures. Imagine that we are in a lecture, we are more likely to miss the content that the lecturer refers to when our mind wanders, which is parallel with the commission error found in the SART (i.e., skipping the No-go target). Meanwhile, we tend to start to retrospectively evaluate our own states and our understanding of the lecture, specifically when the lecturer calls on us during class. This phenomenon is similar to the probe question used in our task. To sum up, we consider the task we used here suitable to examine the states of attention and thus it can be applied to the educational scenario. We expect that eye movement patterns can assist teachers in observing inattentive behaviors directly in the classroom without interfering with students’ learning.

We have shown that people with the distributed pattern are more prone to MW compared to people with the centralized pattern, either with objective or subjective measures. As the centralized pattern is associated with being focused and having more top-down control of attention, future studies can develop a training program to examine if altering people’s viewing pattern from the distributed pattern to the centralized one can enhance cognitive ability and thus task performance. Not only can this help people who are easily distracted, but it can also aid older adults in performing cognitive-demanding tasks, as older adults tend to use the distributed pattern for cognitive processes [28]. This was not revealed by previous studies using traditional eye movement indices to analyze or classify MW state [59]. Specifically, Faber et al. [59] proposed that the eye movement pattern indices associated with MW might vary across tasks and that fixation is not a robust index for MW in centralized tasks, such as the SART used here and other audiobook tasks. In contrast, we demonstrated that when combining the transition matrix of eye movement in the temporal domain, fixation can still effectively predict the mental state, either being MW or FA. Furthermore, as MW during long-term driving occurs very often and can cause negative influences on safety [8,60], future developers can consider installing an eye movement pattern detection system in a car. With the aid of the detection system, drivers can reacquire the FA state (centralized pattern) whenever the distributed pattern is detected. Future studies can further examine if this pattern also applies to tasks with different spatial, visual, or discourse demands and other modalities (e.g., auditory stimuli) [61].

What is the benefit of using HMM rather than the deep neural net (DNN) or other alternative approaches to detect MW? Since HMM is a probabilistic time-series model, it

works well with a limited amount of data, which is in contrast to deep learning methods that require large amounts of data to train the model effectively. In addition, when there is a large pool of participants, learning individual models can be done in parallel, which makes it scalable. Furthermore, the VHEM algorithm for clustering HMMs is based on the parameters of the individual HMMs instead of the actual data, and the clustering can be done efficiently, too. For alternative learning methods, as compared with the recurrent DNN for predicting sequential information, one advantage of using HMMs is that it can make the learning model(s) more interpretable, which is an important trend in the current artificial intelligence research (e.g., [62,63]). For example, we have recently developed a computational model that combines a DNN with an HMM to learn eye movement strategies for object recognition [64]. The DNN learns optimal perceptual representations under the guidance of an attention mechanism summarized in an HMM, and the HMM learns optimal eye movement strategies through feedback from the DNN. The resulting HMM of the model is immediately interpretable and can be directly used for data analysis.

## 5. Conclusions

Our results suggest that eye movement patterns are associated with MW, where both objective and subjective measures of MW can be distinguished from focused attention state by the viewer's more distributed eye movement pattern. The current study is important both technically and practically. First, we provide a novel approach to utilizing EMHMM to study sustained attention. Second, we show that eye movements can be a potential way to detect people's state of attention, which can be used in either in-person or remote classes so that instructors can have better ideas about students' attention states and find ways to regain their attention once MW is detected.

**Author Contributions:** Conceptualization, S.-L.Y., H.-H.L. and J.H.H.; methodology, H.-H.L., S.-L.Y. and J.H.H.; software, Z.-L.C. and J.H.H.; validation, H.-H.L., Z.-L.C. and J.H.H.; formal analysis, H.-H.L. and Z.-L.C.; investigation, S.-L.Y., H.-H.L. and Z.-L.C.; resources, S.-L.Y. and A.-Y.W.; data curation, H.-H.L. and Z.-L.C.; writing—original draft preparation, H.-H.L. and Z.-L.C. writing—review and editing, S.-L.Y., J.H.H. and A.-Y.W.; visualization, Z.-L.C. and H.-H.L.; supervision, S.-L.Y.; project administration, Z.-L.C., H.-H.L. and S.-L.Y.; funding acquisition, S.-L.Y. and A.-Y.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by grants from the Ministry of Science and Technology, Taiwan (MOST 108-2622-8-002-012-TA, 109-2622-8-002-012-TA, 107-2410-H-002-129 -MY3, 108-2420-H-492 -001-MY3, 110-2218-E-002-034-MBK, and 110-2634-F-002-042).

**Institutional Review Board Statement:** This study was reviewed and approved by the Research Ethics Committee at National Taiwan University (NTU REC: 201812HM004) and implemented accordingly.

**Informed Consent Statement:** All participants had provided written informed consent before the experiment started.

**Data Availability Statement:** The data in this study are available from the link: [https://osf.io/zy3v7/?view\\_only=3075989e8eb140a8ab7d1ac7cb409887](https://osf.io/zy3v7/?view_only=3075989e8eb140a8ab7d1ac7cb409887).

**Acknowledgments:** The authors would like to thank Joshua Oon Soo Goh for his advice on an earlier draft. The authors also thank Yi-Ta Chen and Win-Ken Beh for their help with the technical issues when setting up the experiment.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

To examine the differences between the two distinct eye movement patterns in behavioral performance and eye movement indices, we used the lme4 package in R [65] to conduct linear mixed-effect models (LMM) for explorative data analysis. To the non-targets (Go trials), we analyzed the mean reaction time (RT). To the targets (the No-go targets) and the probes, we analyzed the coefficient of variance of RT (RTCV) from the 10-s time windows prior to the presentation of the target and the probe, respectively. For calculating

pupil baseline, the data was first down-sampled from 100 Hz to 50 Hz, then the data from the left and right eyes were averaged. Lastly, the pupil baseline was defined as the average pupil diameter 500-ms before the onset of each stimulus.

In the LMM, the target response or the probe response and the eye movement patterns served as fixed effects, and participants served as a random effect. The approximation of the  $p$ -values came from the lmerTest package.

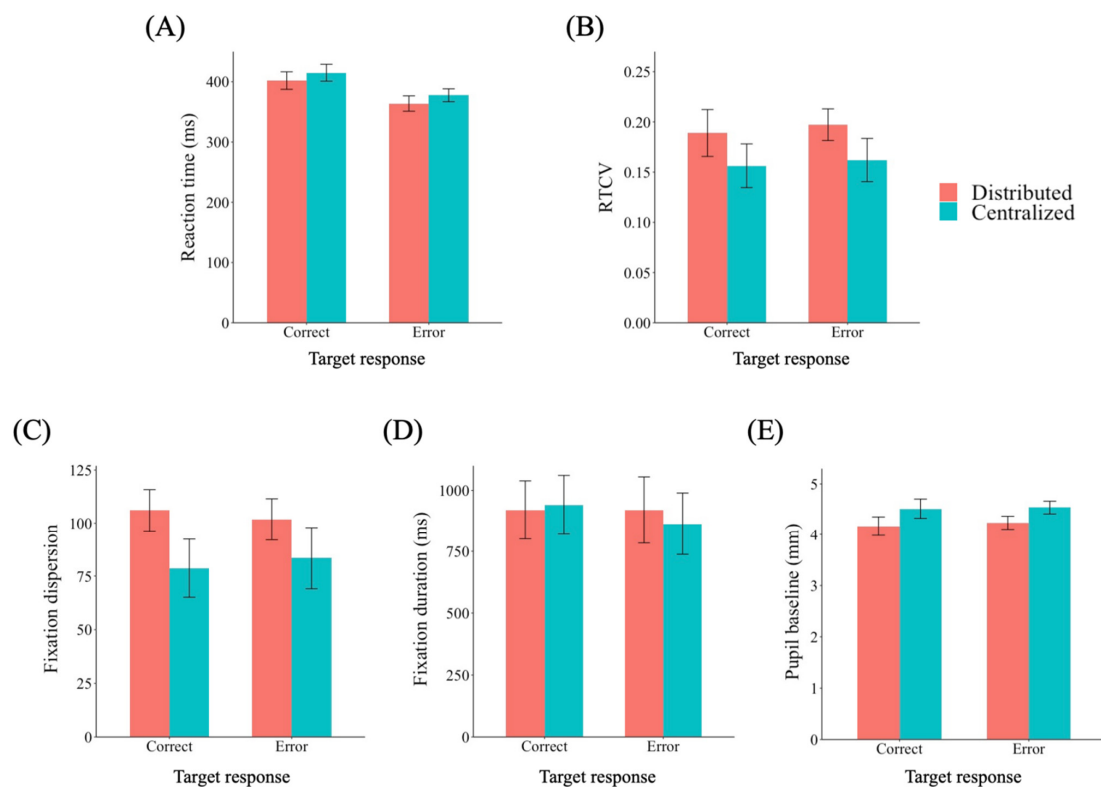
For the pre-target period (Table 1 and Figure A1), slower RTs to Go-stimuli ( $t = 6.7$ ,  $p < 0.001$ ) and a smaller pupil baseline ( $t = -2.29$ ,  $p = 0.029$ ) were found in the correct versus error target response (i.e., response to the No-go targets), which was in line with previous studies [34,66,67]. No other effects were found,  $ps > 0.05$ .

For the pre-probe period (Table A2 and Figure A2), a significant interaction between probe response and eye movement pattern was found in RTCV ( $t = -2.58$ ,  $p = 0.015$ ). Post-hoc analysis revealed that in the self-rated FA condition, participants with the centralized pattern showed smaller RTCVs than those with the distributed pattern,  $t = 3.15$ ,  $p = 0.004$ . In addition, participants with the centralized pattern also showed smaller fixation dispersion ( $t = -3.11$ ,  $p = 0.003$ ), longer fixation duration ( $t = 2.16$ ,  $p = 0.037$ ), and a larger pupil baseline ( $t = 2.38$ ,  $p = 0.024$ ) compared to participants with the distributed pattern. These results suggested that people with the centralized pattern tended to pay more attention overall. Additionally, the fixation dispersion was smaller when participants rated themselves as FA compared to MW ( $t = 2.73$ ,  $p = 0.012$ ).

**Table 1.** Behavioral response data and eye movement data for the 10-s pre-target period (the objective measurement of MW) analyzed with LMM.

	Estimate	SE	$t$	$p$
<b>RT</b>				
(Intercept)	363.83	12.94	28.12	<0.001 ***
Target response	38.12	6.63	5.70	<0.001 ***
Eye movement pattern	13.78	18.60	0.74	0.464
Target response $\times$ Eye movement pattern	-0.74	9.62	-0.08	0.939
<b>RTCV</b>				
(Intercept)	0.20	0.02	9.57	<0.001 ***
Target response	-0.01	0.02	-0.55	0.590
Eye movement pattern	-0.04	0.03	-1.19	0.243
Target response $\times$ Eye movement pattern	0.003	0.02	0.12	0.908
<b>Fixation dispersion</b>				
(Intercept)	101.72	11.87	8.57	<0.001 ***
Target response	4.29	8.43	0.51	0.615
Eye movement pattern	-18.26	17.07	-1.07	0.291
Target response $\times$ Eye movement pattern	-9.06	12.12	-0.75	0.461
<b>Fixation duration</b>				
(Intercept)	892.75	123.08	7.25	<0.001 ***
Target response	26.60	33.22	0.80	0.430
Eye movement pattern	-29.22	176.94	-0.17	0.870
Target response $\times$ Eye movement pattern	50.28	47.75	1.05	0.301
<b>Pupil baseline</b>				
(Intercept)	4.24	0.16	26.45	<0.001 ***
Target response	-0.07	0.03	-2.29	0.029 *
Eye movement pattern	0.30	0.23	1.31	0.201
Target response $\times$ Eye movement pattern	0.05	0.04	1.05	0.302

\*  $p < 0.05$ ; \*\*\*  $p < 0.001$ .

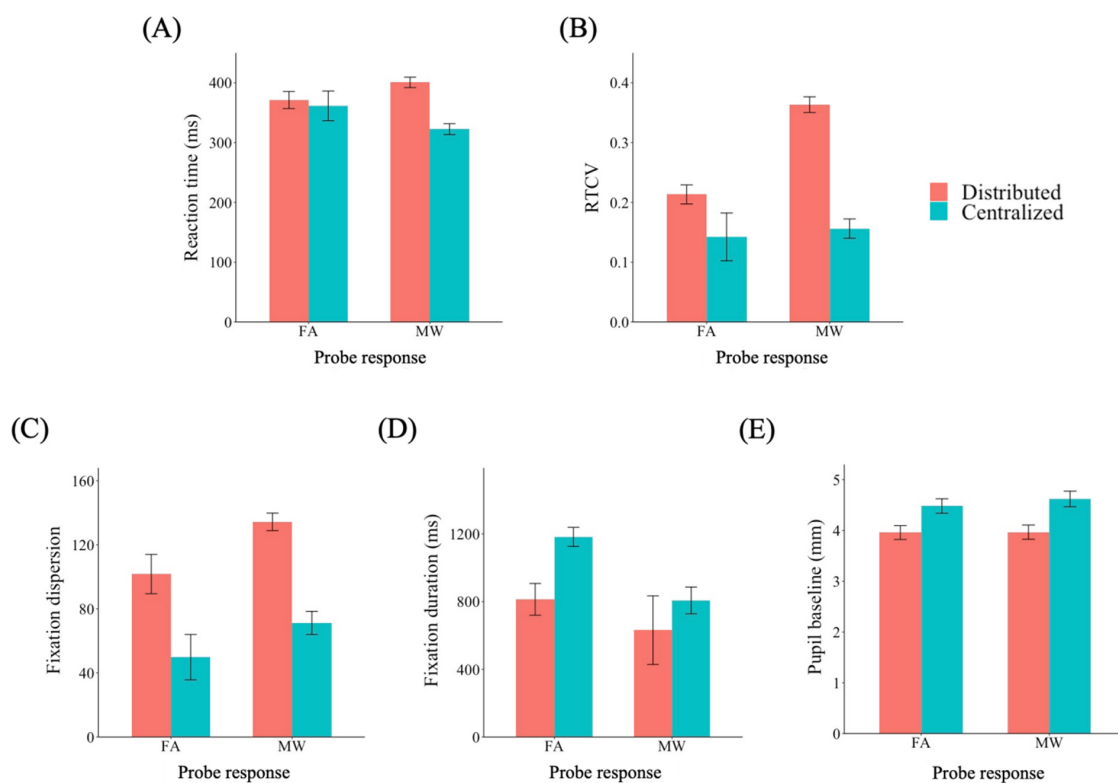


**Figure A1.** Behavioral performance and eye movement indices in the pre-target intervals. (A) Reaction time (RT). (B) Coefficient of variance of RT (RTCv). (C) Fixation dispersion. (D) Fixation duration. (E) Pupil baseline. The number in the bar graph denotes the mean value in that condition. Error bars represent one S.E.M.

**Table A2.** Behavioral response data and eye movement data for the 10-s pre-probe period (the subjective measurement of MW) analyzed with LMM.

	Estimate	SE	<i>t</i>	<i>p</i>
<b>RT</b>				
(Intercept)	363.27	17.81	20.40	<0.001 ***
Probe response	21.07	17.03	1.24	0.229
Eye movement pattern	−3.69	27.92	−0.13	0.896
Probe response × Eye movement pattern	−58.81	30.49	−1.93	0.065
<b>RTCv</b>				
(Intercept)	0.24	0.03	8.17	<0.001 ***
Probe response	0.17	0.03	5.18	<0.001 ***
Eye movement pattern	−0.09	0.05	−1.97	0.055
Probe response × Eye movement pattern	−0.15	0.06	−2.58	0.015 *
<b>Fixation dispersion</b>				
(Intercept)	101.73	10.70	9.51	<0.001 ***
Probe response	30.52	11.16	2.73	0.012 *
Eye movement pattern	−52.25	16.80	−3.11	0.003 **
Probe response × Eye movement pattern	−5.94	19.50	−0.31	0.763
<b>Fixation duration</b>				
(Intercept)	813.67	108.01	7.53	<0.001 ***
Probe response	−192.01	106.36	−1.81	0.087
Eye movement pattern	365.86	169.40	2.16	0.037 *
Probe response × Eye movement pattern	−211.32	186.65	−1.13	0.270
<b>Pupil baseline</b>				
(Intercept)	3.96	0.13	30.16	<0.001 ***
Probe response	0.09	0.06	1.36	0.189
Eye movement pattern	0.48	0.20	2.38	0.024 *
Probe response × Eye movement pattern	0.11	0.12	0.97	0.344

\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .



**Figure A2.** Behavioral performance and eye movement indices in the pre-probe intervals. (A) Reaction time (RT). (B) Coefficient of variance of RT (RTCVC). (C) Fixation dispersion. (D) Fixation duration. (E) Pupil baseline. The number in the bar graph denotes the mean value in that condition. Error bars represent one S.E.M. FA: Focused attention. MW: Mind-wandering.

**Table A3.** Trial numbers across conditions in the objective measure of MW (response to the No-go target).

	Centralized Pattern	Distributed Pattern
Correct	514 (58.68%)	362 (41.32%)
Error	144 (42.99%)	191 (57.01%)

*Note.* Numbers in the table indicate trials that belonged to Correct (successful stop) or Error (fail-to-stop) responses to No-go targets, which were classified as centralized or distributed patterns based on the model. The numbers in the parenthesis indicate the proportion of trials that belonged to either the centralized or the distributed pattern in the correct and error responses respectively.

**Table A4.** Trial numbers across conditions in the subjective measure of MW (response to the probe).

	Centralized Pattern	Distributed Pattern
Self-rated FA	329 (55.02%)	269 (44.98%)
Self-rated MW	90 (33.83%)	176 (66.17%)

*Note.* Numbers in the table indicate trials that belong to self-rated FA or MW, which were classified as centralized or distributed patterns based on the model. The numbers in the parenthesis indicate the proportion of trials that belonged to the centralized or the distributed pattern in the self-rated FA and self-rated MW respectively.

## References

- Smallwood, J.; Schooler, J.W. The restless mind. *Psychol. Bull.* **2006**, *132*, 946–958. [[CrossRef](#)]
- Mooneyham, B.W.; Schooler, J.W. The costs and benefits of mind-wandering: A review. *Can. J. Exp. Psychol.* **2013**, *67*, 11–18. [[CrossRef](#)] [[PubMed](#)]
- Ottaviani, C.; Couyoumdjian, A. Pros and cons of a wandering mind: A prospective study. *Front. Psychol.* **2013**, *4*, 524. [[CrossRef](#)]
- Killingsworth, M.A.; Gilbert, D.T. A wandering mind is an unhappy mind. *Science* **2010**, *330*, 932. [[CrossRef](#)]

5. Stothart, C.; Mitchum, A.; Yehner, C. The attentional cost of receiving a cell phone notification. *J. Exp. Psychol. Hum. Percept. Perform.* **2015**, *41*, 893–897. [[CrossRef](#)] [[PubMed](#)]
6. Feng, S.; D’Mello, S.; Graesser, A.C. Mind wandering while reading easy and difficult texts. *Psychon. Bull. Rev.* **2013**, *20*, 586–592. [[CrossRef](#)]
7. Schooler, J.W.; Smallwood, J.; Christoff, K.; Handy, T.C.; Reichle, E.D.; Sayette, M.A. Meta-awareness, perceptual decoupling and the wandering mind. *Trends. Cogn. Sci.* **2011**, *15*, 319–326. [[CrossRef](#)]
8. He, J.; Becic, E.; Lee, Y.-C.; McCarley, J.S. Mind wandering behind the wheel: Performance and oculomotor correlates. *Hum. Factors* **2011**, *53*, 13–21. [[CrossRef](#)] [[PubMed](#)]
9. Szpunar, K.K.; Moulton, S.T.; Schacter, D.L. Mind wandering and education: From the classroom to online learning. *Front. Psychol.* **2013**, *4*, 495. [[CrossRef](#)]
10. Wammes, J.D.; Seli, P.; Cheyne, J.A.; Boucher, P.O.; Smilek, D. Mind wandering during lectures II: Relation to academic performance. *Scholarsh. Teach. Learn. Psychol.* **2016**, *2*, 33–48. [[CrossRef](#)]
11. Ju, Y.-J.; Lien, Y.-W. Who is prone to wander and when? Examining an integrative effect of working memory capacity and mindfulness trait on mind wandering under different task loads. *Conscious. Cogn.* **2018**, *63*, 1–10. [[CrossRef](#)] [[PubMed](#)]
12. Saito, T.; Sudo, R.; Takano, Y. The gaze bias effect in toddlers: Preliminary evidence for the developmental study of visual decision-making. *Dev. Sci.* **2020**, *23*, e12969. [[CrossRef](#)] [[PubMed](#)]
13. Deubel, H.; Schneider, W.X. Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vis. Res.* **1996**, *36*, 1827–1837. [[CrossRef](#)]
14. Rizzolatti, G.; Riggio, L.; Dascola, I.; Umiltá, C. Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia* **1987**, *25*, 31–40. [[CrossRef](#)]
15. Ikkai, A.; Dandekar, S.; Curtis, C.E. Lateralization in alpha-band oscillations predicts the locus and spatial distribution of attention. *PLoS ONE* **2016**, *11*, e0154796. [[CrossRef](#)]
16. Lee, H.-H.; Yeh, S.-L. Blue-light effects on saccadic eye movements and attentional disengagement. *Atten. Percept. Psychophys.* **2021**, *83*, 1713–1728. [[CrossRef](#)] [[PubMed](#)]
17. Rolfs, M. Microsaccades: Small steps on a long way. *Vision Res.* **2009**, *49*, 2415–2441. [[CrossRef](#)]
18. Hutt, S.; Hardey, J.; Bixler, R.; Stewart, A.; Risko, E.F.; D’Mello, S. Gaze-based detection of mind wandering during lecture viewing. In Proceedings of the International Educational Data Mining Society, Wuhan, China, 25–28 June 2017.
19. Mills, C.; Bixler, R.; Wang, X.; D’Mello, S.K. Automatic gaze-based detection of mind wandering during narrative film comprehension. In Proceedings of the International Educational Data Mining Society, Raleigh, NC, USA, 29 June–2 July 2016.
20. Zhang, H.; Miller, K.F.; Sun, X.; Cortina, K.S. Wandering eyes: Eye movements during mind wandering in video lectures. *Appl. Cogn. Psychol.* **2020**, *34*, 449–464. [[CrossRef](#)]
21. Mills, C.; Graesser, A.; Risko, E.F.; D’Mello, S.K. Cognitive coupling during reading. *J. Exp. Psychol. Gen.* **2017**, *146*, 872–883. [[CrossRef](#)]
22. Uzzaman, S.; Joordens, S. The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Conscious. Cogn.* **2011**, *20*, 1882–1886. [[CrossRef](#)]
23. Van Opheusden, B.; Galbiati, G.; Kuperwajs, I.; Bnaya, Z.; Ma, W.J. Revealing the impact of expertise on human planning with a two-player board game. *PsyArXiv* **2021**. [[CrossRef](#)]
24. Radach, R.; Heller, D. Relations between spatial and temporal aspects of eye movement control. In *Reading as a Perceptual Process*; Kennedy, A., Heller, D., Pynte, J., Radach, R., Eds.; Elsevier: Amsterdam, The Netherlands, 2000; pp. 165–191.
25. Chuk, T.; Chan, A.B.; Hsiao, J.H. Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling. *Vision Res.* **2017**, *141*, 204–216. [[CrossRef](#)]
26. Chuk, T.; Chan, A.B.; Hsiao, J.H. Understanding eye movements in face recognition using hidden Markov models. *J. Vis.* **2014**, *14*, 8. [[CrossRef](#)] [[PubMed](#)]
27. Chan, F.H.; Barry, T.J.; Chan, A.B.; Hsiao, J.H. Understanding visual attention to face emotions in social anxiety using hidden Markov models. *Cogn. Emot.* **2020**, *34*, 1704–1710. [[CrossRef](#)]
28. Chan, C.Y.; Chan, A.B.; Lee, T.M.; Hsiao, J.H. Eye-movement patterns in face recognition are associated with cognitive decline in older adults. *Psychon. Bull. Rev.* **2018**, *25*, 2200–2207. [[CrossRef](#)]
29. Lodge, J.M.; Harrison, W.J. Focus: Attention science: The role of attention in learning in the digital age. *Yale J. Biol. Med.* **2019**, *92*, 21–28. [[PubMed](#)]
30. Faul, F.; Erdfelder, E.; Buchner, A.; Lang, A.-G. Statistical power analyses using G\* Power 3.1: Tests for correlation and regression analyses. *Behav. Res. Methods* **2009**, *41*, 149–160. [[CrossRef](#)] [[PubMed](#)]
31. Hu, N.; He, S.; Xu, B. Different efficiencies of attentional orienting in different wandering minds. *Conscious. Cogn.* **2012**, *21*, 139–148. [[CrossRef](#)] [[PubMed](#)]
32. Robertson, I.H.; Manly, T.; Andrade, J.; Baddeley, B.T.; Yiend, J. Oops!: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia* **1997**, *35*, 747–758. [[CrossRef](#)]
33. Smallwood, J.; Beach, E.; Schooler, J.W.; Handy, T.C. Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *J. Cogn. Neurosci.* **2008**, *20*, 458–469. [[CrossRef](#)]
34. Chen, Y.-T.; Lee, H.-H.; Shih, C.-Y.; Chen, Z.-L.; Beh, W.-K.; Yeh, S.-L.; Wu, A.Y. An effective entropy-assisted mind-wandering detection system with EEG signals based on MM-SART database. *arXiv*, 2020; arXiv:2005.12076.



35. Ibaceta, M.; Madrid, H.P. Personality and Mind-Wandering Self-Perception: The role of meta-awareness. *Front. Psychol.* **2021**, *12*, 1247. [[CrossRef](#)]
36. Christoff, K.; Gordon, A.M.; Smallwood, J.; Smith, R.; Schooler, J.W. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 8719–8724. [[CrossRef](#)]
37. Braboszcz, C.; Delorme, A. Lost in thoughts: Neural markers of low alertness during mind wandering. *Neuroimage* **2011**, *54*, 3040–3047. [[CrossRef](#)]
38. Chuk, T.; Crookes, K.; Hayward, W.G.; Chan, A.B.; Hsiao, J.H. Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition* **2017**, *169*, 102–117. [[CrossRef](#)] [[PubMed](#)]
39. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: Berlin, Germany, 2006.
40. An, J.; Hsiao, J.H. Modulation of mood on eye movement and face recognition performance. *Emotion* **2020**, *21*, 617–630. [[CrossRef](#)]
41. Hsiao, J.H.; An, J.; Zheng, Y.; Chan, A.B. Do portrait artists have enhanced face processing abilities? Evidence from hidden Markov modeling of eye movements. *Cognition* **2021**, *211*, 104616. [[CrossRef](#)] [[PubMed](#)]
42. Hsiao, J.H.; Chan, A.B.; An, J.; Yeh, S.-L.; Jingling, L. Understanding the collinear masking effect in visual search through eye tracking. *Psychon. Bull. Rev.* **2021**. [[CrossRef](#)]
43. Hsiao, J.H.; Lan, H.; Zheng, Y.; Chan, A.B. Eye Movement analysis with Hidden Markov Models (EMHMM) with co-clustering. *Behav. Res. Methods* **2021**. [[CrossRef](#)]
44. Zhang, J.; Chan, A.B.; Lau, E.Y.Y.; Hsiao, J.H. Individuals with insomnia misrecognize angry faces as fearful faces while missing the eyes: An eye-tracking study. *Sleep* **2019**, *42*, zsy220. [[CrossRef](#)] [[PubMed](#)]
45. Zheng, Y.; Ye, X.; Hsiao, J.H. Does adding video and subtitles to an audio lesson facilitate its comprehension? *Learn. Instr.* **2021**, *2021*, 101542. [[CrossRef](#)]
46. Coviello, E.; Chan, A.B.; Lanckriet, G.R. Clustering hidden Markov models with variational HEM. *J. Mach. Learn. Res.* **2014**, *15*, 697–747.
47. Rao, J.N.; Scott, A.J. The analysis of categorical data from complex sample surveys: Chi-squared tests for goodness of fit and independence in two-way tables. *J. Am. Stat. Assoc.* **1981**, *76*, 221–230. [[CrossRef](#)]
48. Nasreddine, Z.S.; Phillips, N.A.; Bédirian, V.; Charbonneau, S.; Whitehead, V.; Collin, I.; Cummings, J.L.; Chertkow, H. The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* **2005**, *53*, 695–699. [[CrossRef](#)] [[PubMed](#)]
49. Chan, C.Y.; Wong, J.; Chan, A.B.; Lee, T.M.; Hsiao, J.H. Analytic eye movement patterns in face recognition are associated with better performance and more top-down control of visual attention: An fMRI study. In Proceedings of the 38th Annual Conference of the Cognitive Science Society, Philadelphia, PA, USA, 10–13 August 2016.
50. Debetencourt, M.T.; Norman, K.A.; Turk-Browne, N.B. Forgetting from lapses of sustained attention. *Psychon. Bull. Rev.* **2018**, *25*, 605–611. [[CrossRef](#)] [[PubMed](#)]
51. Head, J.; Helton, W.S. Perceptual decoupling or motor decoupling? *Conscious. Cogn.* **2013**, *22*, 913–919. [[CrossRef](#)]
52. Seli, P. The attention-lapse and motor decoupling accounts of SART performance are not mutually exclusive. *Conscious. Cogn.* **2016**, *41*, 189–198. [[CrossRef](#)]
53. Murray, S.; Krasich, K.; Schooler, J.W.; Seli, P. What’s in a task? Complications in the study of the task-unrelated-thought variety of mind wandering. *Perspect. Psychol. Sci.* **2020**, *15*, 572–588. [[CrossRef](#)]
54. Smallwood, J.; Schooler, J.W. The science of mind wandering: Empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* **2015**, *66*, 487–518. [[CrossRef](#)]
55. Robison, M.K.; Miller, A.L.; Unsworth, N. Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behav. Res. Methods* **2019**, *51*, 398–408. [[CrossRef](#)]
56. Faber, M.; Bixler, R.; D’Mello, S.K. An automated behavioral measure of mind wandering during computerized reading. *Behav. Res. Methods* **2018**, *50*, 134–150. [[CrossRef](#)]
57. Johnson, K.A.; White, M.; Wong, P.S.; Murrphy, C. Aspects of attention and inhibitory control are associated with on-task classroom behaviour and behavioural assessments, by both teachers and parents, in children with high and low symptoms of ADHD. *Child Neuropsychol.* **2020**, *26*, 219–241. [[CrossRef](#)]
58. Pashler, H. Dual-task interference in simple tasks: Data and theory. *Psychol. Bull.* **1994**, *116*, 220–244. [[CrossRef](#)] [[PubMed](#)]
59. Faber, M.; Krasich, K.; Bixler, R.E.; Brockmole, J.R.; D’Mello, S.K. The eye–mind wandering link: Identifying gaze indices of mind wandering across tasks. *J. Exp. Psychol. Hum. Percept. Perform.* **2020**, *46*, 1201–1221. [[CrossRef](#)] [[PubMed](#)]
60. Pepin, G.; Fort, A.; Jallais, C.; Moreau, F.; Ndiaye, D.; Navarro, J.; Gabaude, C. Impact of mind-wandering on visual information processing while driving: An electrophysiological study. *Appl. Cogn. Psychol.* **2020**, *35*, 508–516. [[CrossRef](#)]
61. Kopp, K.; D’Mello, S. The impact of modality on mind wandering during comprehension. *Appl. Cogn. Psychol.* **2016**, *30*, 29–3040. [[CrossRef](#)]
62. Adadi, A.; Berrada, M. Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
63. Hsiao, J.H.; Ngai, H.H.T.; Qiu, L.; Yang, Y.; Cao, C.C. Roadmap of designing cognitive metrics for explainable artificial intelligence (XAI). *arXiv* **2021**, arXiv:2108.01737.

- 
64. Hsiao, J.H.; An, J.H.; Chan, A.B. The role of eye movement consistency in learning to recognise faces: Computational and experimental examinations. In Proceedings of the 42nd Annual Conference of the Cognitive Science Society, Virtual Meeting, 29 July–1 August 2020.
  65. Bates, D.; Maechler, M.; Bolker, B.; Walker, S. Linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]
  66. Chen, Y.-C.; Yeh, S.-L.; Huang, T.-R.; Chang, Y.-L.; Goh, J.O.; Fu, L.-C. Social robots for evaluating attention state in older adults. *Sensors* **2021**, *21*, 7142. [[CrossRef](#)]
  67. Lee, H.-H.; Tu, Y.-C.; Yeh, S.-L. In search of blue-light effects on cognitive control. *Sci. Rep.* **2021**, *11*, 15505. [[CrossRef](#)]