Prior EEG marks focused and mind-wandering mental states across trials

Chie Nakatani[1,*], Hannah Bernhard[1,2,3], & Cees van Leeuwen[1, 4]

Affiliations:
1: Brain and Cognition Research Unit, KU Leuven, Tiensestraat 102 - Box 3711, 3000 Leuven, Belgium

2: Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands (Current affiliation of H. Bernhard)

3: Centre for Integrative Neuroscience, Maastricht University, Maastricht, The Netherlands (Current affiliation of H. Bernhard)

4: Center for Cognitive Science, TU Kaiserslautern, Kaiserslautern, Germany

*Corresponding author: Chie Nakatani Email: chie.nakatani@kuleuven.be

Authors' Email: chie.nakatani@kuleuven.be, hannah.bernhard@maastrichtuniversity.nl,

cees.vanleeuwen@kuleuven.be

Key words: Spontaneous mental process, EEG alpha amplitude, cognitive control, default mode network, dynamical modeling

Funding: This work was supported by an Odysseus grant from The Flemish Organization for Science PEP-C3738-G.0003.12 to CvL

Running head: Predicting mind wandering

Abstract

Mind wandering is a mental process involving spontaneously generated thought irrelevant to the current task or goal. Irrelevant thought in itself, however, does not necessarily lead to mind wandering, as this could be prevented by cognitive control functions. The strength of the control varies over time, as does the rate of spontaneous thought generation. Their combined dynamics generates a complex transition pattern among mental states, e.g., from task-focusing via multi-tasking to mind wandering. We investigated neural substrates of these mental dynamics. Participants reported mind-wandering and task-focused states during a tone counting task, while their electroencephalogram (EEG) was recorded. Amplitudes of EEG alpha-band activity and auditory evoked potentials were computed as measures of default-mode network (DMN) and cognitive control network (CCN) activity, respectively. Effects of the measures on task performance and self-reported number of thoughts were estimated, allowing for effect latencies of up to 20min. Alpha amplitude effects were found to change in direction as the mental state shifts between task focusing and mind wandering; before a reported task-focused state, the amplitude effect on task performance was negative, while it was positive before a reported mind wandering. The fist effect occurred 4-6min, and the second effect 6-10min before a report. The results suggest that DMN slowly but dynamically influences cognitive control, effectuating transitions between task focused and mind wandering behavior.

Introduction

The wakeful mind continuously generates thoughts, often including ones that are irrelevant to the current situation and may trigger episodes of *mind wandering* (Antrobus, 1968; Giambra, 1995; Klinger & Cox, 1987; Smallwood & Schooler, 2006). The rate at which such thoughts occur varies over time (Campbell, 1960; Nakatani, Ganschow, & van Leeuwen, 2019). The more frequently thoughts are generated, the more likely we are to drift into mind wandering (Ganschow, 2017).

Drifting into mind wandering is typically countered by a set of mental functions that support focused, intentional, and goal-directed behavior, such as attention, vigilance, and executive control. These functions are referred to collectively as *cognitive control* (Cohen, Dunbar, & McClelland, 1990; Funes, Lupiáñez, & Humphreys, 2010; Otto, Skatova, Madlon-Kay, & Daw, 2015; Ridderinkhof, 2002). As long as cognitive control keeps the mental focus on thoughts relevant to the current task and goal, mind wandering could be prevented.

However, just like the rate of spontaneous thought generation, the strength of control fluctuates over time (Cheyne, Carriere, & Smilek, 2006; Weissman, Roberts, Visscher, & Woldorff, 2006). Mind wandering is most likely to occur when cognitive control is weak while the number of spontaneous thoughts is high. The temporal variability of both functions matters; when cognitive control is strong and spontaneous thoughts are scarce, the mental focus is most likely to be on thoughts relevant to the current task and goal.

Mind wandering and focused states are just two out of several mental states that are elicited by the combined dynamics of spontaneous thought generation and cognitive control. For example, an experience sampling study reported transitions from a focused to a mind wandering state via an intermediate mental state, in which the number of spontaneous thoughts and the strength of cognitive control are both high (Nakatani et al., 2019). Such a state has been linked to multi-tasking behavior and 'divergent thinking' (Fink, Graif, & Neubauer, 2009; Jaarsveld & Van Leeuwen, 2005) which, together with mind wandering, belong to the category of 'spontaneous mental processes' (Christoff, Irving, Fox, Spreng, & Andrews-Hanna, 2016).

Neural substrates of spontaneous mental states have been investigated by correlating neural signals and subjective reports (Fink et al., 2009; Mason et al., 2007). However, neural correlates of individual states

are not informative about how transitions from one mental state to another occur. To understand these transitions, dynamic neural effects on cognitive control strength and thought generation rate need to be considered. We studied such effects using EEG. Our results suggest that dynamic changes in alpha-band activity several minutes in advance affect whether participants subsequently find themselves in a task-focused or mind wandering state.

Neural correlates of cognitive control and spontaneous thought generation and possible effects of their activity on mental state transitions. Cognitive control and spontaneous thought generation have their own distinct neural correlates. Brain functional localization studies suggest that the neural correlates of cognitive control include the dorsal attention network, which encompasses the dorsolateral prefrontal cortex and the intraparietal sulcus (Corbetta, Kincade, Ollinger, McAvoy, & Shulman, 2000; Hopfinger, Buonocore, & Mangun, 2000), the salience network, which includes the insula and anterior cingulate cortex (Menon & Uddin, 2010; Seeley et al., 2007), and the sensory-motor systems. As a collective, these networks could be referred to as the cognitive control network (CCN) (Cole & Schneider, 2007). CCN activity is typically higher in task- than in rest conditions, a property known as task-positivity.

Spontaneous thought generation, by contrast, has been related to the default mode network (DMN). This network includes prefrontal cortex, posterior cingulate cortex, as well as precuneus and memory-related regions such as the parahippocampal and middle temporal gyri (Raichle, 2015; Raichle et al., 2001). DMN regions typically show lower activity in task conditions than at rest, a property known as task-negativity. Task-negative activity was initially reported in the ratio of oxygen consumption to blood flow in PET (Raichle et al., 2001), and later in BOLD-fMRI (Greicius, Krasnow, Reiss, & Menon, 2003), intracranial EEG (Jerbi et al., 2010), scalp EEG (Laufs et al., 2003; Mantini, Perrucci, Del Gratta, Romani, & Corbetta, 2007) and MEG (Brookes et al., 2011). Of the DMN regions, medial prefrontal cortex, precuneus, posterior cingulate cortex and insula have been related to spontaneous thought generation (Mason et al., 2007).

Task-negativity, however, is not uniform across DMN. Certain DMN regions show task-positivity rather than task-negativity when the task requires mental focus on 'internally constructed information'

(Buckner & Dinicola, 2019), e.g., recalling autobiographical events. The dorsomedial prefrontal cortex, temporal-parietal junction, lateral temporal cortex, and temporal pole showed higher BOLD levels in an autobiographic memory task than in rest conditions. This effect was further increased when the task demanded detailed examination of autobiographic memory (Andrews-Hanna, Saxe, & Yarkoni, 2014). The reversal from task-negativity to positivity could be highly relevant to a transition into mind wandering, during which spontaneously generated thoughts are being prioritized over thoughts related to external events and tasks.

Let us therefore consider how this type of reversal could influence transitions to mind wandering. To begin with, previous studies indicate that CCN facilitates cognitive control while DMN facilitates spontaneous thought generation (Fox et al., 2005; Fransson, 2005; Weissman et al., 2006). Fox et al. (2005) and Fransson (2005) report overall negative correlations between CCN and DMN activity. Thus, between CCN and spontaneous thought generation, and between DMN and cognitive control, inhibitory effects might traditionally be expected (Figure 1a). These neural effects would oscillate relative strength of cognitive control and spontaneous thought generation in counterphase. An episode in which cognitive control is strong and spontaneous thoughts generation is low would correspond to a focused state, while the opposite would correspond to mind wandering. Between the two, there is hardly a period which can be considered as another state.

The oscillating dynamics, however, changes considerably if the DMN effect on cognitive control changes from task-negative to positive as the mental focus shifts from external events to spontaneously generated thoughts (Figure 1b). The task-positive effect maintains the task performance facilitating cognitive control. As a result, during the shift from a task-focused to mind wandering states, an intermediate state would occur in which cognitive control and spontaneous thought generation both are high. Such a behavioral state was observed in Nakatani et al. (2019).

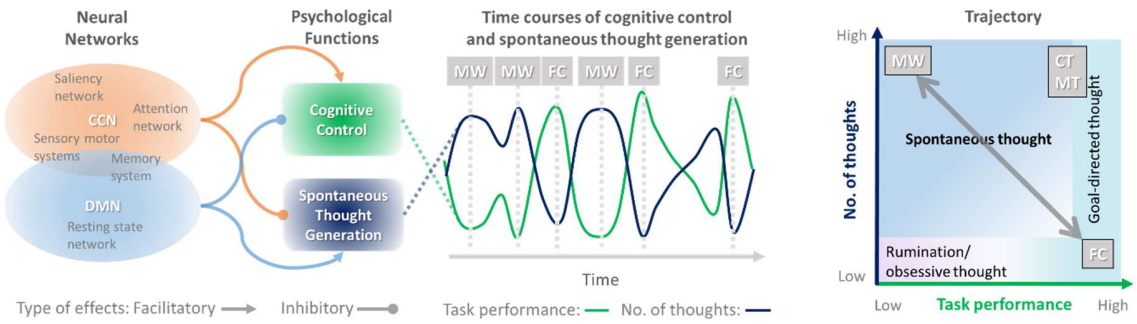
Scope of the current study. We investigated the likelihood of a reversal in DMN effect during the transition from task focused to mind wandering states. Based on previous findings (Andrews-Hanna et al.,

2014; Buckner & Dinicola, 2019; Nakatani et al., 2019), the effect on cognitive control is expected to reverse from task-negative to positive during transitions to mind wandering. We recorded these transitions in subjective reports, while test participants were counting tones presented in a slow sequence. The percentage of correct tone count was used as a measure of the strength of cognitive control, while subjective report of the number of thoughts was used as a measure of the rate of spontaneous thought generation (Nakatani et al., 2019). Together with the behavioral and subjective measures, EEG was recorded during the task. EEG was chosen because of its transferability to non-laboratory studies. Amplitude of temporal alpha-band (8-12Hz) activity was used as a measure of DMN activity (Kam, Rahnuma, Park, & Hart, 2022; Lehtelä, Salmelin, & Hari, 1997; Tiihonen et al., 1991; Weisz, Hartmann, Müller, Lorenz, & Obleser, 2011), while amplitude of the auditory evoked potential (AEP) to the tone stimuli was chosen as a measure of CCN activity (Picton & Hillyard, 1974; Picton, Hillyard, Krausz, & Galambos, 1974).

Effects of the alpha-band activity and AEP on the two psychological measures of spontaneous thought generation and cognitive control were estimated using a multivariate regression model with a moving window. This method was applied to estimate neural effects in a wide-range of latencies. In previous studies, reported effects had a short latency, typically in the order of milliseconds or seconds from a time point of a subjective report of mind wandering (Kirschner, Kam, Handy, & Ward, 2012). Short-latency effects are quasi-instantaneous. At best, they mark the current mental state, but provides little information about the history of transitions that led to this state (Nakatani et al., 2019), while long-latency effects, e.g., those with minutes-long latency, could play a significant role in such complex transitions. We, therefore, applied the method to estimates neural effects systematically over the latencies between 0 and 20 min.

Results showed that the effect of the alpha amplitude as an indicator of DMN activity on the task performance switched from task-negative to positive between task-focused and mind-wandering conditions, in accordance with previous fMRI studies (Andrews-Hanna et al., 2014; Buckner & Dinicola, 2019). However, estimated latency of the effect was much longer than that in the previous studies: 4-10 min. The results therefore suggest that DMN exults slow effects over the transition to mind wandering.

a)



b)

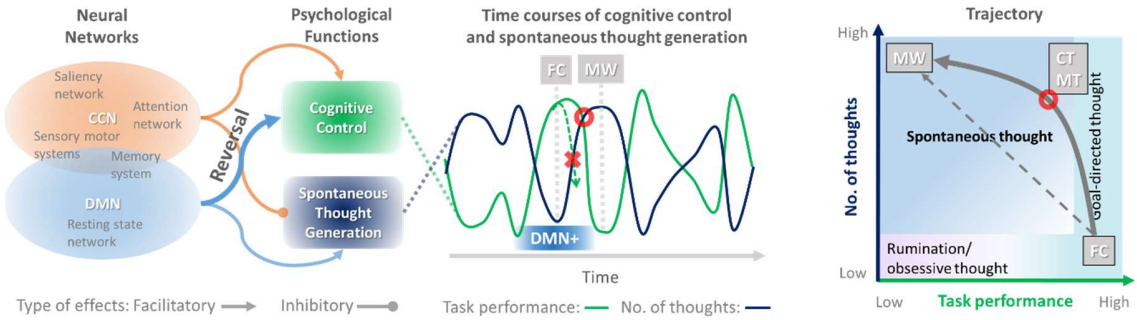


Figure 1. Direct- and non-direct transitions between task-focused and mind-wandering states

Functional neural networks collectively known as cognitive control network (CCN) facilitate cognitive control, while suppressing spontaneous thought generation. In contrast, networks, collectively known as the default mode network (DMN), facilitate spontaneous thought generation, while suppressing cognitive control (a, left panel).

Due to the opposite, facilitatory and inhibitory neural effects, the strength of cognitive control and the rate of spontaneous thought generation oscillate over time. Accuracy of primary-task performance and self-reported number of spontaneous thoughts oscillate in counterphase (a, middle panel). When task performance is high and number of thoughts is low, the mental state is most likely to be subjectively reported as task-focused (FC), while in case of the opposite, a mind-wandering state (MW) is likely to be reported. During the alternations between FC and MW, no other distinctive states, such as multi-tasking (MT), creative thinking (CT), or rumination, are visited (a, right panel).

When the DMN effect on cognitive control is allowed to change from inhibitory to facilitatory (b, left panel), the transition pattern between FC and MW changes as well. The facilitatory, thus task-positive, DMN effect (b, middle panel, DMN+) maintains task performance at a high level despite the large number of thoughts being generated (red circle). This enables a transition via an intermediate state linked to MT and CT behavior (b, right panel).

Methods

Overview

Participants performed a task of counting tones up to a target number (between 20 and 24) pre-specified at the start of each trial. Each time after reaching the target, they retrospectively rated their mental states during the trial on a scale, ranging from fully task focused to fully mind wandering (%MW), and provided a subjective report of the number of thoughts (NoTs) that had occurred during the trial (Figure 2, top). NoTs were used as measure of spontaneous thought generation (Nakatani et al., 2019), while performance of the tone count, namely, the percentage of correct count (%CC) was used as behavioral measure of cognitive control. EEG was recorded continuously during the task. As the neural measure of CCN activity, amplitude of the auditory evoked potential (AEP) to the tone stimuli was used. AEP amplitude correlates positively with auditory attention, for which the CCN is held responsible (Picton & Hillyard, 1974; Picton et al., 1974). As neural measure of DMN activity, alpha-band (8-12Hz) activity amplitude was used. In some studies, the amplitude was increased in the resting period compared to visual-, motor-, and auditory task conditions (Lehtelä et al., 1997; Tiihonen et al., 1991; Weisz et al., 2011). More recent studies also reported the opposite, lower amplitude for the resting state (Kam et al., 2022). Amplitude, therefore, is sensitive to the differences between rest- and task conditions irrespective of direction, allowing its usage as an EEG measure of DMN activity. Alpha-band activity does not overlap with the main band of the AEP which is situated in the theta band (4-8Hz, Bernhard, 2018). Thus, alpha-band activity was suitable to further distinguish DMN-related and CCN-related activity. (See **Computation of AEP amplitude** and **Computation of alpha-band amplitude**)

The five single-trial measures (%MW, NoTs, %CC, AEP, and alpha amplitude) were recorded as time series during the task. Task-focused and mind-wandering states alternated in irregular intervals. In case of a short interval, neural effects for task focusing and those for mind wandering could overlap in time. To untangle these effects, time series of trial data were averaged from the time point of a report of mind-wandering and that of a task-focused states, separately (Figure 2, middle). The event time-locked averaging reduced components which are not time-locked to the corresponding subjective report, similarly to the way

ERPs are extracted from averaged signals (Nakatani et al., 2019). Details of the averaging procedure are reported in **Preprocessing of subjective reports and behavioral data**, and **Event time-locked averaging of AEP and alpha**.

Effects of the EEG measures, AEP and Alpha, on the psychological measures, %CC and NoTs, were estimated for the task-focused and mind wandering conditions separately using multi-variate multiple regression models. The models were designed to identify the neural effects with a latency between 0 and 20min in 2-min. steps. To test the statistical reliability of effects with a long latency, we introduced auto-regressive terms to the models. The terms served as extra error terms: If the long-latency neural effects are reasonably strong relative to the auto-regressive ones, the neural effects could be considered meaningful (Granger, 1969). Details of the model are reported in **Models and main data analysis**.

Based on previous studies (Andrews-Hanna et al., 2014; Buckner & Dinicola, 2019) , we expect the effect of alpha amplitude on %CC to be task-negative in the task-focused, while task-positive in the mind-wandering conditions. Previous studies did not include neural effects with latency longer than 1-min. We considered the effect latency between 0 and 20 min, because we consider the possibility of an intermediate state in the transition from task-focused to mind-wandering states.

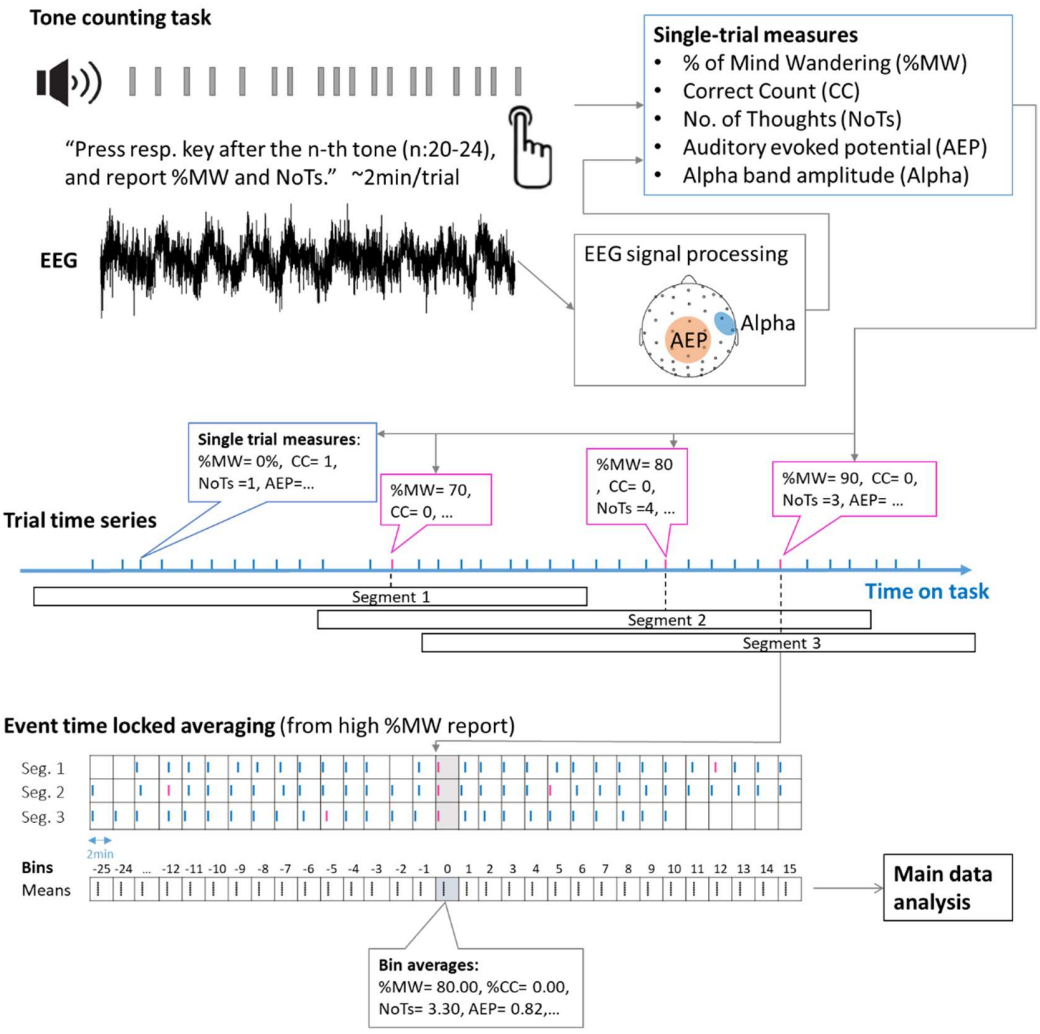


Figure 2. Tone counting task and event time-locked averaging

Tone counting task: Participants were asked to count n (20-24) tones and press a key when the count was reached. After the key press, they reported the percentage of time they spent on mind wandering (%MW) and the number of thoughts (NoTs) during the trial. EEG was recorded during the task. Amplitude of alpha activity and auditory evoked response (AEP) were computed from the EEG. Each trial thus yields 5 measures, Correctness of the count (CC), NoTs, %MW, Alpha, and AEP amplitude. On NoTs, %CC, Alpha, and AEP amplitude, event-time locked averaging was performed from the moments of mind-wandering (%MW≧70) and task-focused (%MW < 30) reports, respectively. Segmentation is illustrated for mind-wandering reports and was performed likewise for task-focused reports. Segmented time series were aligned and averaged. The averaged time series were used in the main data analysis.

Ethical approval and Participants

The study was approved by The KU Leuven Medical Ethics committee (protocol number: S56128). Twenty healthy volunteers (average 23.00 years, 6 male) participated in exchange for monetary compensation (10€/hour). They were recruited via the online participant recruitment system at KU Leuven. Prior to participation, they were given information about the study and gave oral and written informed consent. All data were collected prior to the onset of the 2020 COVID19 pandemic.

Procedure

Experiments were conducted in single sessions with individual participants, in an EEG-laboratory of the University Hospital of Leuven. After electrodes were attached, the participant was seated in a dimly lit room in front of speakers and a monitor. Baseline EEG was recorded (eyes open/relax 1 min, and eyes closed/relax, 1 min). This was followed by an auditory oddball task for the estimation of evoked response parameters, and next by the tone counting task. EEG was recorded during both tasks. Afterwards, baseline EEG measurement was performed again. Then the electrodes were removed. The participant were debriefed before leaving. The experimental took about 3 hours in total.

Tasks

To elicit a sizeable incidence rate of mind wandering during an approx. 2-hour period, a *tone counting* task was used (Figure 2, top). The task consisted of 37 trials lasting approx. 90s each, in which a sequence of brief tones, 100ms in 440Hz frequency, was presented at 70dB via the speakers located at right and left sides of a monitor in front of the participants. The tones were presented with a stimulus onset asynchrony (SOA) between 2000 and 2500ms (pseudo-randomized with 100ms steps). At the beginning of each trial, a number was presented in the center of the display. The number indicated how many tones had to be counted before a response. The number ranged between 20 and 24 (pseudo-randomized over trials). Participants put their right index finger on Key “5” of the ten-key pad and closed their eyes before the series of tones started. When the target number was reached, participants responded by pressing the key. Then they opened their eyes and reported the percentage of time spent on task irrelevant thoughts in 0 to 100% in 10% steps. An eleven-point scale (0% --10%--....-90%--100%) was shown on the display and participants

reported the percentage (%MW) using the ten-key pad. Next, participants reported on the keypad the number of thoughts (NoTs) that had occurred in their minds during the trial on a ten-point scale. The scale ranged from 1: only task-related thought to 10: task-related and 9 or more task-irrelevant thoughts. Participants were instructed to evaluate the number of thoughts in a consistent manner over trials. For instance, they were instructed to report "1" when they had only task-related thoughts; when they had one task-irrelevant thought, for instance about having lunch after the experiment, they reported "2". They were told that items included in the thoughts (e.g., a sandwich, an apple, and a bottle of soda) were not counted separately. If two different thoughts occurred at different times during the trial (e.g., thoughts on lunch, then later during the trial thinking about doing laundry), they reported "3." And so on. The instruction was the same as that in Nakatani et al., (2019). Participants had the opportunity to take a break at marked sections of the task. EEG was recorded throughout the task.

An *Auditory Oddball Task* was applied prior to the tone-counting task to specify the relevant parameters of the auditory evoked potentials (AEP), namely, scalp distribution, latency, and main frequency band, for subsequent use in the tone counting task. Participants placed their right index finger on the "5" key of the ten-key pad and closed their eyes. Then, a series of 110 tones was presented at 70 dB via the speakers, with an SOA pseudo-randomized between 2000 and 2500ms in 100ms steps. Of these tones, 101 were standard (440Hz, 100ms duration), physically identical to the ones used in the tone counting task, and 5 were odd-ball tones (510Hz, 100ms duration), which were randomly interspersed. Participants were asked to respond to the oddballs by pressing the key as fast as possible.

Both tasks were implemented using PsychoPy Version 1.82.01 (Peirce, 2009) on Windows XP.

EEG recording

EEG was recorded using the Patientguard EEG recording system (ANT Neuro, Enschede, The Netherlands). Of 64 channels in total, 59 were EEG channels and the remaining ones were used for recording the electrooculogram (EOG, vertical and horizontal) and recording reference (right mastoid, which was later re-referenced to average reference). The placement of the EEG electrodes was according to the

manufacturer-defined whole head system rather than the 10-10 system. The signal was sampled at 2048Hz. Electrode impedance was 20 kΩ or less (for the majority of electrodes, the impedance was 5 kΩ or less).

Preprocessing of subjective reports and behavioral data

Time series of NoTs and %CC were segmented for event time-locked averaging, relative to the time of a high- or low %MW report. Thresholds for high and low %MW, 70% and 30%, respectively, were decided based on the grand-average histogram of %MW (Appendix Fig. A1). From the time of the report (Time0), the NoTs and %CC series were segmented between -50min and +32min, allowing overlap between segments. The extension to the past enables estimating autoregressive effects. On average, 9.11 and 6.53 segments per participant were obtained for high- and low %MW reports, respectively. In both types of reports, segments were aligned for averaging from Time0. However, trials do not necessarily align in time (except for those including Time0 of a given high- and low %MW reports), since the inter-trial interval varied according to the number of tones per trial and to individual differences in report timing. Thus, we set bins of 2-min intervals (Figure 2, Middle); the data points between Time0 and 1min 59.99s were averaged and considered as the data for Bin0. Similarly, the next interval was between 2min and 3min 59.99s, and the averages were data for Bin1 and so on. After averaging, the time intervals were uniform 2min. For each participant, interval series were computed for NoTs and %CC. The grand averages over segments are shown in Figure 3 of the Results section. The peaks in %MW at Bin0 reflect the segmentation criteria. Beyond Bin0, variation in %MW appears to be reduced over data bins. The result illustrates the averaging out of the non-time locked fluctuations in mind-wandering due to the segmentation procedure. This means that effects observed in the data will relate to the mind wandering in Bin0, rather than to some earlier or later mind-wandering events.

Preprocessing of EEG data

EEG data were pre-processed using Brain Vision Analyzer (version 2.1.2, Brain Products GmbH, Germany) and in-house Python scripts. Data were down-sampled from 2048 to 1024Hz, re-referenced to the average of all EEG channels, and filtered between 0.32 and 45.00Hz using a 2nd order zero phase shift Butterworth filter. Independent component analysis (ICA, Infomax algorithm) was applied to the filtered

data for artifact removal. EOG channels were included in the ICA analysis. Ocular, cardiac, and muscular artifact components were visually identified and excluded from the inverse ICA. After the cleaning, two channels (one for 2 participants) still showing poor signal quality were interpolated using adjacent channels.

Computation of AEP amplitude

AEP has been reported as a complex of evoked components to auditory stimuli, with AEP amplitude reflecting selective attention and various task-related factors (Picton & Hillyard, 1974; Picton et al., 1974). Typically, the peak amplitude appears approx. 200-500ms from the onset of a pure tone stimulus, scalp distribution of the peak is central, and the peak frequency is situated in the theta-band (4-8Hz). To test the timing, scalp distribution, and frequency of the AEP peak to the current stimulus (400Hz pure tone), we conducted time-frequency analysis using the EEG data of the odd-ball task, recorded from the same group of participants. The EEG data were segmented from the onset of standard (440Hz) tone, between -500 and 2000ms. Approx.100 segments were taken per participant. To obtain scalograms of instantaneous amplitude between 1 and 40Hz in logarithmic steps, the complex Morlet wavelet (no. of cycles = 5) was applied to each of the segments. The scalograms were normalized and averaged per participant. To the group data, t-tests were applied to test the significance of the amplitude against 0 at each data point in the scalogram. Type-I error was corrected using the false discovery rate (FDR) method (Benjamini & Hochberg, 1995). FDR was applied to the p-values of all data points over the 59 EEG electrodes. A more extended description of the procedure is available elsewhere (Bernhard, 2018). We found that amplitude in the frequency band between 4 and 8Hz, in time between 130 and 430ms from the stimulus onset, and in 7 central electrodes (3Z, 4Z, 5Z, 4L, 5L, 4R, 5R in the WaveGuard Duke64 system, cf. Appendix Fig. A2) showed p(FDR-corrected) < 0.05. The frequency, latency, and scalp distribution matched those in previous AEP studies (Picton & Hillyard, 1974; Picton et al., 1974), so these parameters were employed here to extract the AEP values used as GDP measures in the subsequent tone counting task.

EEG from the tone counting task was band-passed between 4 and 8Hz using in-house Python code for band-pass filtering, which combines a Butterworth filter (scipy.signal.butter, 12dB) and a zero-phase shift filter (scipy.signal.filtfilt). Instantaneous amplitude was computed using the complex-valued analytic

signal from the Hilbert transform of the band-passed signal. The amplitude signal was segmented between -500 and 1000ms from a tone onset. Prior to averaging the segments, each segment was normalized using mean and standard deviation of the pre-stimulus period. Per trial, 20-24 segments were averaged. Based on the peak timing and scalp distribution parameters, amplitude between 130 and 430ms of the 7 central electrodes were chosen and averaged per trial per participant, and used as the AEP amplitude measure.

Computation of alpha-band amplitude

Resting state alpha-band amplitude is commonly understood as a measure of DMN activity (Chen, Feng, Zhao, Yin, & Wang, 2008; Compton, Gearinger, & Wild, 2019; Gruberger, Simon, Levkovitz, Zangen, & Hendler, 2011; Marino, Arcara, Porcaro, & Mantini, 2019). From the auditory resting state to task conditions, a decrease in alpha band amplitude is known to occur over bilateral temporal electrodes (Lehtelä et al., 1997; Tiihonen et al., 1991; Weisz et al., 2011). Task-negativity of DMN may be a factor in this reduction of alpha amplitude, suggesting the use of bilateral alpha amplitude as a measure of DMN activity. We tested the difference in alpha-band amplitude between pre-task resting (60s) and the tone-counting task (the 1st trial, ~ 80s) conditions in each of the 59 electrodes. However, the difference in the temporal electrodes was not statistically reliable. This may be due to the presence of visual resting-state alpha activity during the tone counting task, which was performed with eyes closed. Posterior visual resting-state alpha activity spreading to the temporal electrodes may have masked the auditory task effect. Because a reduction in alpha amplitude failed to occur in our data, we applied an alpha-based measure of DMN activity more specifically tuned to mind wandering states.

Recent mind wandering studies reported both decreases and increases in alpha-band amplitude from mind wandering to task focused states (Kam et al., 2022 for a review). Based on this observation, we considered variability in alpha amplitude over time in either direction as indicative of sensitivity to mind wandering. The relevant alpha activity can be separated by event time-locked averaging. When the time series of alpha amplitude is averaged from the time point of %MW report, the amplitude relevant for the particular level of mind wandering would remain, while the rest, including visual-alpha, will be cancelled out. Therefore, if a channel signal contains alpha activity relevant to mind wandering, the temporal pattern

of the average amplitude will not be flat. In other words, variability in the average signal can be used as selection criterion for alpha amplitude sensitive to mind wandering induced by the task. The temporal variability criterion allows positive and negative effects of mind wandering on amplitude.

We started signal processing by band passing the EEG between 8-12Hz using the above mentioned in-house band-pass filter for each available electrode (59ch). To the band-passed signals, Hilbert transform was applied to compute instantaneous amplitude. From the amplitude, we cut 2000ms periods, starting from the onsets of the 20-24 tone-stimuli within a trial. We averaged the amplitude of these periods at trial level to obtain time series of the single trial amplitudes. Next, the time series were segmented, aligned, and binned from a %MW report in the same way as for %CC and NoTs. In our main analysis, we are interested in segments belonging to high and low %MW (≧ 70 and < 30), respectively. To avoid circular analysis in the current validation, the trial time series was segmented from intermediate-level %MW report, 40≦%MW<60. On average, 10.58 segments were obtained per participant. Peaks and troughs in the average indicate that alpha amplitude changed, either in positive or negative direction, time-locked to the mid-level %MW report.

The variability was tested at each electrode, fitting a mixed effect model, Alpha ~ Time (in data bins, fixed effect factor) + participants (random effect factor) to the mid-level %MW segment, using the R package, lme4 (Bates, Mächler, Bolker, & Walker, 2014). The effect of Time was estimated using the chi-squared (χ^2) statistic of ANOVA between the model and its null model: Alpha ~ participants. The test was applied to all EEG electrodes. Type-I error over the univariate tests was corrected over the electrodes using the FDR method. Result showed a sizable effect of Time in two adjacent right temporal electrodes (3RB and 3RC; $\chi^2(1) = 11.23$, p(FDR-corrected) = 0.03, $\eta^2 = 0.43$, power = 0.38 and $\chi^2(1) = 12.33$, p(FDR-corrected) = 0.02, $\eta^2 = 0.44$, power = 0.34, respectively. Average waveforms are shown in Appendix Fig A3. The location of these two electrodes agrees with the temporal locations observed in previous auditory resting-state alpha studies (Lehtelä et al., 1997; Tiihonen et al., 1991). Resting-state activity has been considered as a DMN measure (Weisz et al., 2011). Moreover, the current procedure takes into account that both reductions and increases of the alpha amplitude occur in mind wandering studies. Thus, we consider

the average alpha amplitude from the two electrodes as a DMN measure, sensitive to the level of mind wandering induced by the current auditory task.

The alpha amplitude in each trial was averaged over the 2 channels to obtain a single measure of DMN-related activity. From hereon, this measure is referred to simply as *Alpha*. The Alpha measure was computed per trial, then event time-locked averaging was performed from the time point of high- and low %MW report, as explained in the following section.

Event time-locked averaging of AEP and Alpha

Event time-locked averaging was applied to the trial series of AEP and Alpha (Figure2, middle). In each measure, the time series was segmented from the high- and low %MW report, between -50 and +32min. The procedure of the averaging was same as that for NoTs and %CC. The averaging generated interval series between Bins-25 to +15 from Bin0, for the high- and low %MW reports separately.

Grand average AEP and Alpha are shown in Figure 3 of the Results section. Temporally local features, peaks and troughs, are observable in the figure. However, analyzing such features in isolation may bias our understanding. Their selection is arbitrary and extrinsic to the dynamics of the time series. Therefore, rather than pre-selected temporal features, we included entire segments in the main data analysis.

Models and main data analysis

We assumed that the rate of spontaneous thought generation and the strength of cognitive control are jointly modulated by DMN and CCN activity. For these, respectively, AEP and Alpha are proposed as EEG-derived predictor variables, while NoTs and %CC are the behavioral dependent variables. The neural predictors are given a time index which shifts the predictors to the past relative to the dependent variables to estimate neural effect with different latencies.

A set of models were built from an autoregression model which predicts the current state of NoTs and %CC using their own past states. If past states of the neural variables could improve the prediction of the current states beyond the autoregressive predictors, these neural variables are considered as stochastically causal predictors (Granger, 1969; Pearl, 2000). The following is step-by-step explanation of

our approach, starting from the autoregression model. For NoTs, the autoregression model can be written as

(Eq.1)

$$NoTs_t = \sum_{i=1}^{N} a_i * NoTs_{t-i} + \varepsilon_t$$

NoTs at a given time point *t*, $NoTs_t$, is predicted by the weighted sum of own past states $NoTs_{t-i}$, where *i* is the index for time regression and a_i is the weight. The ε_t is the error term. The past states were summed over *N*. To simplify the model, we temporarily set *N=1*. The simplified model is

(Eq.2)

$$NoTs_t \approx a_1 * NoTs_{t-1} + \varepsilon_t$$

To the autoregression model, two neural predictors, $Alpha_{t-1}$ and AEP_{t-1}, were added as follows.

(Eq.3)

$$NoTs_t \approx a_1 * NoTs_{t-1} + b_1 * Alpha_{t-1} + c_1 * AEP_{t-1} + \varepsilon_t$$

To the 3-predictor model, two more neural predictors, $Alpha_t$ and AEP_t, were added. The predictors estimate neural effects without time difference from the response variable.

(Eq.4)

$$NoTs_t \approx a_1 * NoTs_{t-1} + (b_1 * Alpha_{t-1} + c_1 * AEP_{t-1})$$
$$+ (b_0 * Alpha_t + c_0 * AEP_t) + d_0 * Participants + \varepsilon_t$$

Parentheses are added for readability. A categorical variable $Participants$ was added to account for individual differences. The same formulation was applied to predict %CC at time *t*, $\%CC_t$.

(Eq.5)

$$\%CC_t \approx a_1 * \%CC_{t-1} + (b_1 * Alpha_{t-1} + c_1 * AEP_{t-1})$$
$$+ (b_0 * Alpha_t + c_0 * AEP_t) + d_0 * Participants + \varepsilon_t$$

NoTs and %CC are two parallel data series in the tone counting task, which could covary. To account for the covariance, we chose a multivariate method to estimate the contribution of each predictor. The model for the estimation may be written as pseudo R code.

(Eq.6)

$$[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-1}) + (Alpha_{t-1} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$$

The multivariate linear regression method estimates the contribution of the autoregression, past-neural, present-neural predictors and the individual difference term to the prediction of the present state of NoTs and %CC, taking into account possible correlations among the variables. Reduced models can be obtained from this model (Eq.6) to evaluate the stochastic causality of Alpha and AEP effects; a model without $Alpha_{t-1}$ and another without AEP_{t-1}. If the reduced models predict the response variables worse than the full model did, the contribution of the omitted neural predictor is not negligible. I.e., the comparison gives the basis for attributing stochastic causality to the effect.

The above models set the time regression as *t-1*, which means the most recent past state represents the effect of all past states. This would be appropriate if we could assume that the most recent past state has always the strongest effect. However, as discussed in the Introduction, we assumed that the effect timing could vary depend on the task-focused and mind-wandering states. Thus, the order of regression was parametrized using indices *i, j, k,* and *l*.

(Eq.7)

$$[NoTs_t, \%CC_t] \sim \left(NoTs_{t-i} + \%CC_{t-j}\right) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$$

The parameters vary systematically between 1 and 10 (given the data bin size of 2min, this corresponds to times between 2 and 20min. The range was decided by the length of the data).

The full combination of the 4 parameters yields a number of 10^4 models to fit. The number of models was brought down in two steps. In the first step, *i* and *j* were searched between 1 and 10 for the highest autoregressive effects, while *k* and *l* were fixed to *k=1* and *l=1*. The larger autoregression terms were desirable since they make the criterion for stochastic causality more strict. The resulting 10 by 10 = 100 models were fit to the high- and low %MW segments, separately. Upon fitting, a data window was applied; the window for response variables, $NoTs_t$ and $\%CC_t$, was ±15 bins from Bin0 (which starts from *t*, the moment of the %MW report) based on the results of subjective report and task measure in our previous study (Nakatani et al., 2019). The same data window was applied to the present neural predictors. For the autoregression and past neural state predictors, the size of the data window was the same, however, the window was shifted bin-by-bin, up to 10bins (20min) to the past, as indicated by *i* and *j*. Of the 100 combinations of *i* and *j*, those with the highest autoregressive effects were selected using adjusted-R^2. In case of tie, the smallest *i* and *j* were chosen e.g., if (*i=1, j*=2) and (*i=2, j*=2) gave the highest adjusted-R^2, (*i=1, j*=2).

In the second step, *k* and *l* were systematically varied between 1 and 10, while the autoregression terms were fixed with $i=i^*$ and $j=j^*$, where i^* and j^* were those selected in the first step. In this manner, the number of full models were reduced from 10^4 to 10^2. The full model with i^* and j^* in pseudo R code is

(Eq.8)

$$[NoTs_t, \%CC_t] \sim \left(NoTs_{t-i*} + \%CC_{t-j*}\right) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$$

The models were fit to the high- and low %MW segment separately. To test the stochastic causality, 2 reduced models, one without the past Alpha and the other without the past AEP term, were also fit to the data. The reduced models are:

(Eq.9)

$$[NoTs_t, \%CC_t] \sim \left(NoTs_{t-i*} + \%CC_{t-j*}\right) + AEP_{t-l} + (Alpha_t + AEP_t) + Participants$$

(Eq.10)

$$[NoTs_t, \%CC_t] \sim \left(NoTs_{t-i*} + \%CC_{t-j*}\right) + Alpha_{t-k} + (Alpha_t + AEP_t) + Participants$$

The Pillai test was used to estimate statistical significance of the difference between the fits of the full and reduced models. The 100 comparisons for 2 reduced models yielded 200 tests. The R implementation of Pillai test provides the estimated F-statistic with corresponding probability. We corrected for Type-I errors over the 200 tests by using the FDR method. An arbitrary threshold ($p<0.01$) was applied to the corrected probability, as criterion for the attribution of stochastic causality to the neural effects. The indices *k* and *l* reveal the timing of the effects, e.g., a stochastically causal effect of $Alpha_{t-k}$ at *k=4* on $NoTs_t$ suggests that (given the 2min interval of data bin) the neural activity between 6-8min *before* the report of the high (or low) %MW had stochastically causal effect on the number of spontaneous thoughts 6-8 min later.

Previous studies did not report effects with a latency longer than 1 min. (Andrews-Hanna et al., 2014; Buckner & Dinicola, 2019). Moreover, they did not consider the causality of the predictors. Predicter correspond to the previous findings is $Alpha_t$: The effect on on $\%CC_t$ is expected to be task-negative in the task-focused, while positive in the mind-wandering conditions.

Results

Visual inspection of event time-locked average data

Event time-locked averages of the five measures AEP, Alpha, NoTs, %CC, and %MW, are available for visual inspection in Figure 3. Most show linear trends and local patterns, while %MW was essentially flat, except for the peaks at Bin0, which are, according to their condition, focused: %MW < 30; mind-wandering: %MW ≧ 70. Elsewhere %MW appears homogeneously distributed. Averaging has reduced non time-locked effects of %MW in the same way as averaging increases the signal-to-noise ratio in ERP. Thus, peaks and troughs in the other variables, AEP, Alpha, NoTs, and %CC, can be analyzed as time-locked to the mind-wandering report at Bin0.

AEP amplitude was higher in low- than in high %MW conditions most of the time. The difference, however, was reduced around the moment of subjective report, in Bin0. Alpha amplitude was higher in the high- than in low %MW conditions, mostly before the report. The %CC decreased with time. The task performance measure was high around the subjective report of task-focus (%MW <30 in Bin0), while it was low around the report of mind wandering (%MW ≧70 in Bin0). In contrast, while NoTs increased with time overall, there was a dip around the period task focus was reported. For all measures, preliminary analyses indicated a sizable conditions-by-time interaction and some of the bin-wise differences were reliable. However, they would be hard to interpret in isolation. An analysis based on such local differences and features relies only on part of the variance in the data and contributes little to our understanding of the dynamics. Instead, we proceed to the main analysis, which takes all of the variance in the data segments into account.

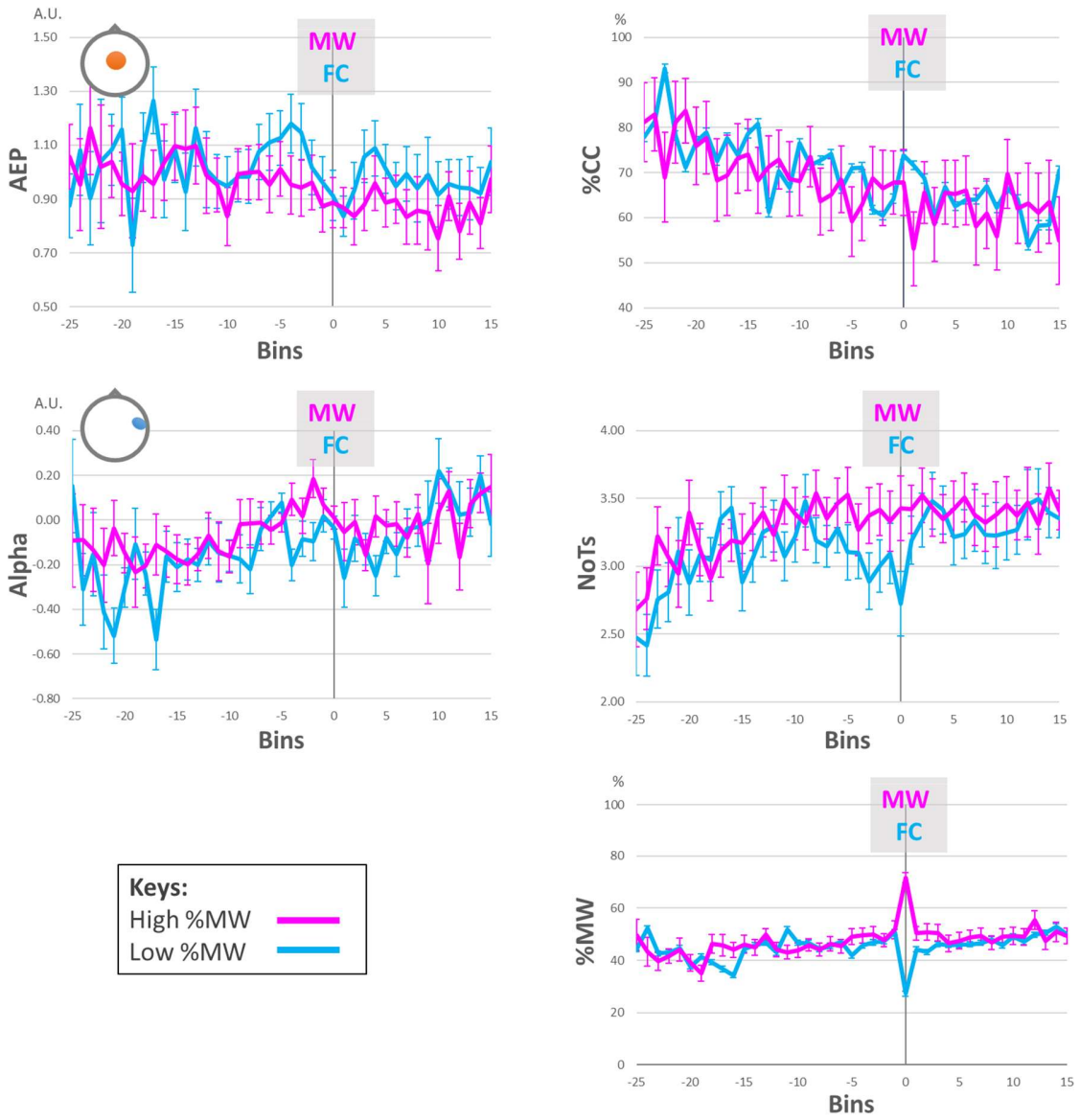


Figure 3. Event time-locked averages of AEP, Alpha, %CC, NoTs, and %MW

The horizontal axis represents time in data bins. Each bin interval equals 2 min. Bin0 starts from the point of subjective report of mind wandering (MW) in the high %MW condition (pink), and from that of task focus report (FC) in the low %MW condition (cyan). Error bars indicate SE. Head icons in the figures for AEP and Alpha indicate electrode location for the EEG measures.

Prior to the model fit, multi-variate outliers were removed using Mahalanobis distance among NoTs, %CC, Alpha and AEP data; datapoints with Mahalanobis distance > 20 were excluded from the model fit. Linear trends were removed to make the data quasi-stationary. This is to avoid bias in model estimation due to the trend. The result of the linear trend was reported in the section, *Notes on linear trends.* The detrended high- and low %MW data segments were used in the step-wise model fit.

Results of the Main Analysis

High %MW segments

The first step of model fitting indicated $i = 2$ and $j = 5$ for the highest autoregression effect (See *Notes on autoregressive effects*). For the second step, $i*$ and $j*$ were set as 2 and 5, respectively, and 100 models were generated over the combinations of k and l. The full- and corresponding reduced models were fit to the high %MW segments data. The difference between the full- and reduced models was tested using the Pillai test. Type-I error was corrected over the 200 tests using FDR method. Table1 shows the corrected probability of the estimated-F statistic of the Pillai score. With an arbitrary threshold of $p<0.01$, stochastic causality was suggested for past neural predictor $Alpha_{t-k}$ at k = 4 and 5 (Table 1a), while no causality was suggested for AEP_{t-l} (Table 1b). Timing of the effect at k = 4 and 5 correspond to 6-8 and 8-10min before the high %MW report, respectively.

Table 2 list the estimated coefficients for the full model with *k=4 and 5*. (For both cases, the index for AEP_{t-l} was *l=1.* With the choice, no crucial information was missed since the effect of AEP_{t-l} changed little over *l.*) The coefficient for the past neural effect $Alpha_{t-4}$ which was positive as a predictor for $\%CC_t$; coefficient = 8.16, the 95% Confidence Interval [2.5%, 97.5%] = [4.22, 12.11], t-value = 4.07. The effect was sizable compared with that of the autoregression term, $\%CC_{t-5}$; coefficient = 0.21, 95%CI = [0.12, 0.31], t-value = 4.33. By contrast, the effects of the present neural predictors $Alpha_t$ and AEP_t to $\%CC_t$ were negligible. For the prediction of $NoTs_t$, no neural effect was noteworthy, while the autoregression effect was considerable; coefficient of $NoTs_{t-2}$ = 0.24, 95%CI = [0.15, 0.33], t = 5.29. Results of *k=5* (Table 2b) showed the same pattern as *k=4*; The effect of $Alpha_{t-5}$ to $\%CC_t$ was positive, and also the auto regression effects were positive.

The above results show that Alpha activity 6-10min before the high %MW report had a stochastically causal and task-*positive* effect on the %CC. The rest of neural effects did not show notable contribution to predict the %CC and NoTs.

Table 1. Effect of past neural states in High %MW condition.
a) Model comparison tests the stochastic causality of past neural Alpha states, b) tests the stochastic causality of past neural AEP states.
Pillai test was applied to test the difference between full and reduced models. Probability of estimated F-statistic, DF = (1, 462) of Pillai score was shown after FDR correction.

a) $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$
vs. $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + AEP_{t-l} + (Alpha_t + AEP_t) + Participants$

		AEP $_l$									
		l=1	l=2	l=3	l=4	l=5	l=6	l=7	l=8	l=9	l=10
Alpha $_k$	k=1	0.0649	0.0649	0.0649	0.0649	0.0649	0.0649	0.0649	0.0649	0.0651	0.0649
	k=2	0.0335	0.0339	0.0335	0.0337	0.0335	0.0326	0.0337	0.0326	0.0335	0.0335
	k=3	0.2282	0.2282	0.2282	0.2282	0.2300	0.2282	0.2282	0.2282	0.2301	0.2282
	k=4	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0009	0.0005	0.0005	0.0005
	k=5	0.0025	0.0025	0.0025	0.0025	0.0025	0.0025	0.0031	0.0027	0.0025	0.0025
	k=6	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2277	0.2282
	k=7	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282	0.2282
	k=8	0.2318	0.2676	0.2282	0.2301	0.2301	0.2402	0.2402	0.2634	0.2301	0.2282
	k=9	0.5072	0.5071	0.5102	0.5071	0.5105	0.5071	0.5197	0.5071	0.5071	0.5071
	k=10	0.2301	0.2301	0.2301	0.2301	0.2372	0.2301	0.2634	0.2301	0.2402	0.2301

p<0.01 p<0.001

b) $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$
vs. $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + Alpha_{t-k} + (Alpha_t + AEP_t) + Participants$

		AEP $_l$									
		l=1	l=2	l=3	l=4	l=5	l=6	l=7	l=8	l=9	l=10
Alpha $_k$	k=1	0.5298	0.0649	0.5071	0.3540	0.9296	0.0649	0.3275	0.5071	0.5735	0.5071
	k=2	0.5551	0.0649	0.5072	0.4219	0.9709	0.0649	0.3479	0.4721	0.5210	0.5071
	k=3	0.5626	0.0649	0.5071	0.3741	0.9709	0.0649	0.2634	0.5071	0.5565	0.5071
	k=4	0.5564	0.0649	0.5071	0.3640	0.9935	0.0649	0.5071	0.4869	0.5071	0.5071
	k=5	0.5626	0.0649	0.5071	0.3741	0.9945	0.0649	0.4219	0.5197	0.5565	0.5071
	k=6	0.5626	0.0649	0.5071	0.3741	0.9629	0.0651	0.3453	0.5071	0.5071	0.5071
	k=7	0.5298	0.0649	0.5071	0.3595	0.9629	0.0649	0.3479	0.5071	0.5071	0.5071
	k=8	0.6141	0.0655	0.5071	0.3759	0.9629	0.0764	0.3595	0.5490	0.5298	0.5071
	k=9	0.5629	0.0649	0.5072	0.3670	0.9629	0.0649	0.3312	0.5071	0.5071	0.5071
	k=10	0.5626	0.0649	0.5071	0.3741	0.9709	0.0651	0.3622	0.5071	0.5551	0.5071

Table 2 Estimated coefficients of the full model for High %MW segments, separately for $Alpha_{t-4}$ (a) and $Alpha_{t-5}$ (b).

a) $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + (Alpha_{t-4} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC** $_t$	Intercept	-4.39	-7.06	-1.72	-3.23	0.001
	NoTs $_{t-2}$	-0.37	-3.81	3.07	-0.21	0.834
	%CC $_{t-5}$	0.21	0.12	0.31	4.33	0.000
	Alpha $_{t-4}$	8.17	4.22	12.11	4.07	0.000
	AEP $_{t-1}$	0.93	-5.01	6.87	0.31	0.760
	Alpha $_t$	-0.44	-4.16	3.28	-0.23	0.815
	AEP $_t$	4.78	-1.07	10.64	1.61	0.109
	Participants	0.39	0.17	0.61	3.46	0.001
Response var. = **NoTs** $_t$	Intercept	-0.02	-0.09	0.05	-0.48	0.631
	NoTs $_{t-2}$	0.24	0.15	0.33	5.29	0.000
	%CC $_{t-5}$	0.00	0.00	0.00	0.54	0.588
	Alpha $_{t-4}$	0.09	-0.01	0.19	1.71	0.088
	AEP $_{t-1}$	-0.09	-0.24	0.06	-1.17	0.244
	Alpha $_t$	0.07	-0.03	0.16	1.39	0.164
	AEP $_t$	0.11	-0.04	0.27	1.49	0.137
	Participants	0.00	0.00	0.01	1.38	0.169

b) $[NoTs_t, \%CC_t] \sim (NoTs_{t-2} + \%CC_{t-5}) + (Alpha_{t-5} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC $_t$**	Intercept	-4.48	-7.16	-1.80	-3.28	0.001
	NoTs $_{t-2}$	-0.60	-4.06	2.86	-0.34	0.732
	%CC $_{t-5}$	0.20	0.11	0.30	4.06	0.000
	Alpha $_{t-5}$	6.92	3.00	10.85	3.46	0.001
	AEP $_{t-1}$	2.35	-3.63	8.32	0.77	0.440
	Alpha $_t$	0.46	-3.22	4.14	0.25	0.807
	AEP $_t$	3.48	-2.40	9.37	1.16	0.246
	Participants	0.40	0.18	0.62	3.55	0.000

Response var. = **NoTs $_t$**	Intercept	-0.02	-0.09	0.05	-0.50	0.619
	NoTs $_{t-2}$	0.24	0.15	0.33	5.23	0.000
	%CC $_{t-5}$	0.00	0.00	0.00	0.40	0.688
	Alpha $_{t-5}$	0.09	-0.01	0.19	1.76	0.079
	AEP $_{t-1}$	-0.07	-0.23	0.08	-0.95	0.343
	Alpha $_t$	0.08	-0.02	0.17	1.59	0.113
	AEP $_t$	0.10	-0.05	0.25	1.29	0.198
	Participants	0.00	0.00	0.01	1.41	0.159

Low %MW segments

Results of the first-stage model fit suggested $i = 1$ and $j = 9$ for the highest autoregression effect. Thus, i^* and j^* of the model in the second stage was set as 1 and 9, respectively. With the selected autoregression terms, 100 models were generated over the combination of k and l between 1 and 10. The full- and reduced models were fit to the data in the low %MW segment. The Pillai test was applied in the same way for the analysis of high %MW segment data, to assess the stochastic causality of the neural effects. As shown in Table 3, the test results suggested the causality for $Alpha_{t-k}$ at k = 3, while no causality was suggested for AEP_{t-l}. The corresponding timing for $Alpha_{t-3}$ was 4-6min before the low %MW report.

Table 4 lists the estimated coefficients for the predictors at k = 3. Similar to Table 2, index l for AEP_{t-l} was set to *1* to simplify the presentation of the results. The coefficient for the past neural effect term $Alpha_{t-3}$ for $\%CC_t$ was large; coefficient = -10.04, 95% CI = [-14.6, -5.43], t-value = -4.29, relative to that of autoregression term, $\%CC_{-9}$, coefficient = 0.15, 95%CI = [0.07, 0.24], t-value = 3.51. Thus, the past neural effect could be qualified for stochastic causality similar to that in the high %MW segment, although the sign was *negative*. Moreover, in the low %MW segments, the effect of present Alpha was noticeable; coefficient of $Alpha_t$ = 7.38, 95% CI = [2.77, 12.00], t-value = 3.15.

In contrast, $Alpha_{t-3}$ predicted little of $NoTs_t$; coefficient = 0.10, 95%CI = [0.00, 0.21], t-value = 1.89, relative to that of the autoregressive term $NoTs_{-1}$ coefficient = 0.12, 95%CI = [0.03, 0.21], t-value = 2.57. Besides, the effect of present Alpha was noticeable; coefficient of $Alpha_t$ = -0.17, 95% CI = [-0.28, -0.07], t-value = -3.22.

The above results show that Alpha activity 4-6min before the low %MW report had a stochastically causal and task-*negative* effect on %CC. Moreover, the results show that the Alpha activity correlated without delay (in the current time scale), positively with %CC, while negatively with NoTs in the low %MW segment.

Table 3. Effect of past neural states in Low %MW condition.
a) Model comparison tests the stochastic causality of past neural Alpha states, b) tests the stochastic causality of past neural AEP states.
Pillai test was applied to test the difference between full and reduced models. Probability of estimated F-statistic, DF = (1, 395) of Pillai score was shown after FDR correction.

a) $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-9}) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$
vs. $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-9}) + AEP_{t-l} + (Alpha_t + AEP_t) + Participants$

		AEP $_l$									
		l=1	l=2	l=3	l=4	l=5	l=6	l=7	l=8	l=9	l=10
Alpha $_k$	k=1	0.9552	0.9552	0.9552	0.9629	0.9552	0.9552	0.9552	0.9552	0.9935	0.9552
	k=2	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935
	k=3	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014	0.0014
	k=4	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935
	k=5	0.0952	0.0952	0.0988	0.0952	0.0952	0.0952	0.0952	0.0952	0.0952	0.0952
	k=6	0.3069	0.3222	0.3222	0.2485	0.3038	0.3069	0.3069	0.3222	0.2898	0.3069
	k=7	0.5537	0.5579	0.5719	0.6140	0.5683	0.5579	0.5579	0.5286	0.5579	0.5579
	k=8	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935	0.9935
	k=9	0.0636	0.0453	0.0636	0.0636	0.0636	0.0636	0.0636	0.0928	0.0636	0.0636
	k=10	0.7600	0.7526	0.8904	0.7585	0.7585	0.7492	0.7585	0.7600	0.7492	0.7585

p<0.01

b) $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-9}) + (Alpha_{t-k} + AEP_{t-l}) + (Alpha_t + AEP_t) + Participants$
vs. $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-9}) + Alpha_{t-k} + (Alpha_t + AEP_t) + Participants$

		AEP $_l$									
		l=1	l=2	l=3	l=4	l=5	l=6	l=7	l=8	l=9	l=10
Alpha$_p$	k=1	0.9935	0.7492	0.6698	0.7492	0.9935	0.7585	0.9935	0.4371	0.1102	0.9935
	k=2	0.9935	0.7492	0.6698	0.7492	0.9935	0.7585	0.9935	0.4371	0.0952	0.9935
	k=3	0.9935	0.7492	0.6698	0.7492	0.9935	0.7600	0.9935	0.4435	0.1077	0.9935
	k=4	0.9935	0.7492	0.6698	0.7492	0.9935	0.7585	0.9935	0.4371	0.0952	0.9935
	k=5	0.9935	0.7492	0.7492	0.7492	0.9935	0.7803	0.9935	0.4535	0.0952	0.9935
	k=6	0.9935	0.7585	0.7296	0.6835	0.9935	0.7585	0.9935	0.4562	0.0952	0.9935
	k=7	0.9935	0.7492	0.6914	0.7585	0.9935	0.7585	0.9935	0.4099	0.0952	0.9935
	k=8	0.9935	0.7492	0.6698	0.7492	0.9935	0.7585	0.9935	0.4371	0.0952	0.9935
	k=9	0.9935	0.5083	0.6698	0.7526	0.9552	0.7492	0.9935	0.5683	0.0952	0.9935
	k=10	0.9935	0.7492	0.7492	0.7492	0.9935	0.7492	0.9935	0.4392	0.0952	0.9935

Table 4. Estimated coefficients of the full model for Low %MW condition for $Alpha_{t-3}$.
$[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-9}) + (Alpha_{t-3} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC** $_t$	Intercept	-6.54	-10.27	-2.82	-3.45	0.0006
	NoTs $_{t-1}$	4.83	0.81	8.86	2.36	0.0188
	%CC $_{t-9}$	0.15	0.07	0.24	3.51	0.0005
	Alpha $_{t-3}$	-10.04	-14.64	-5.43	-4.29	0.0000
	AEP $_{t-1}$	1.85	-4.59	8.30	0.57	0.5721
	Alpha $_t$	7.38	2.77	12.00	3.15	0.0018
	AEP $_t$	-1.01	-7.61	5.58	-0.30	0.7628
	Participants	0.57	0.26	0.88	3.65	0.0003

Response var. = **NoTs** $_t$	(Intercept)	-0.07	-0.15	0.02	-1.50	0.1343
	NoTs $_{t-1}$	0.12	0.03	0.21	2.57	0.0107
	%CC $_{t-9}$	0.00	0.00	0.00	-0.27	0.7887
	Alpha $_{t-3}$	0.10	0.00	0.21	1.89	0.0592
	AEP $_{t-1}$	0.03	-0.12	0.18	0.37	0.7105
	Alpha $_t$	-0.17	-0.28	-0.07	-3.22	0.0014
	AEP $_t$	0.04	-0.11	0.19	0.51	0.6115
	Participants	0.01	0.00	0.01	2.06	0.0401

Notes on autoregressive effects

The indices for autoregression terms, $i*$ and $j*$, were selected to make the criterion for stochastic causality as strict as possible, and to reduce the number of full models. We chose *i* with the highest adjusted-R^2 for NoTs, and *j* with the highest adjusted-R^2 for %CC. The selected $i*$ and $j*$ were 2 and 5 for the high-, and 1 and 9 for the low %MW segments. The timing of the effects, in particular those for the autoregression effect of %CC, were further in the past than might have been expected. As a reality check, we also fit full models with autoregression terms with *i*= 1 and *j*=1 to the data. The results showed that the contribution of the terms were slightly less than those with $i*$ and $j*$; e.g., $\%CC_{t-1}$ had coefficient= 0.11, 95%CI=[0.01, 0.21], t=2.17 (prob.=0.031), in comparison with $\%CC_{t-9}$ coefficient = 0.15, 95%CI=[0.07, 0.24], t=3.51 (prob.=0.0005) for the low %MW segment (See Appendix 4 for more results). The observations suggest that the $i*$ and $j*$ gave a stricter baseline for causality. Thus, we conclude the $i*$ and $j*$ chosen are reasonable for the purpose in the current study.

Notes on linear trend

Prior to the model fit, the linear trend was removed to make the data segment quasi-stationary. However, the linear trend is useful to check if the %CC and NoTs were negatively correlated during the task. We expected the %CC to decrease, while NoTs to increase with time, thus, a negative and positive linear trend for %CC and NoTs, respectively.

The linear trend was estimated applying a mixed-effect model, Y ~ Time (fixed variable) + participants (random variable) to $\%CC_t$ and $NoTs_t$ in of the low- and high %MW segment, separately. In the low %MW segment, the linear trend was negative for %CC; Linear coef.= -23.27, the 95%CI= [-32.20, -14.34], t= -4.97, and positive for NoTs; Linear coef.= 0.55, the 95%CI= [0.31, 0.78], t= 4.42. In the high %MW segment, the linear coefficient for $\%CC_t$was again negative; Linear coef.= -15.79, the 95%CI= [-23.66, -7.84], t= -3.80. For $NoTs_t$, the positive trend was not strong; Linear coef.= 0.18, the 95%CI= [-0.03, 0.39], t= 1.66. Overall, the results suggest that the %CC and NoTs negatively correlated during the task. The weak linear trend in the high %MW segment might suggests that the NoTs were near their ceiling;

the number of thoughts varied around 3.40 which is near 4, the limit in working memory capacity according to Luck and Vogel (1997). Moreover, this is above the 90% level of the standardized number of thoughts observed in our previous study (Nakatani et al., 2019). In contrast, the number varied around 2.72 in the low %MW segment.

Moreover, we checked the linear trend of AEP_t and $Alpha_t$. In the low %MW segment, the AEP amplitude decreased with time; Linear coef.= -0.28, the 95%CI= [-0.43, -0.14], t= -3.75, while the Alpha amplitude increased; Linear coef.= 0.42, the 95%CI= [0.21 0.64], t= 3.75. Similarly, in the high %MW segment, AEP amplitude decreased; Linear coef.= -0.37, the 95%CI= [-0.50, -0.24], t= -5.53, while Alpha amplitude increased with time; Linear coef.= 0.26, the 95%CI= [0.07, 0.46], t= 2.59.

The above results are in agreement with our assumption; the CCN and DMN activity overall have antagonistic effects to the cognitive control and spontaneous thought generation, but the antagonism is disrupted by the positive effect of $Alpha_{t-k}$, effective with a delay of 4-10 min (corresponding to *k=3,4 and 5*) on cognitive control.

Discussion

A good proportion of our thoughts does not have immediate relevance to our current task or goal (Antrobus, 1968; Giambra, 1995; Klinger & Cox, 1987). When the proportion is high, chances are high that irrelevant thoughts instead of task-relevant ones receive mental focus (Ganschow, 2017). Such mind wandering could be prevented by the mental functions for goal directed behavior known as cognitive control. The strength of control changes over time. Relative to the rate of the thought generation at the time, various mental states could arise, e.g., task-focused state where the control is strong and the generation rate is low, mind wandering where the control is weak and the rate is high, multi-tasking or creative thinking where the control and the rate are both high, and obsessive thought or rumination where both are low (Christoff et al., 2016; Jaarsveld & Van Leeuwen, 2005; Nakatani et al., 2019). These states can be visited in a sequence, e.g., from a task-focused, via multi-tasking, to mind-wandering (Nakatani et al., 2019). Instead of identifying neural correlates for each state, we investigated dynamic neural effects which enable transitions among these states.

Previous studies suggest that two functional neural networks, Default Mode Network (DMN) and the Cognitive Control Networks (CCN), jointly modulate the psychological functions during the complex transition to mind wandering (Christoff et al., 2016). Overall, DMN and CCN have opposite effect on the spontaneous thought generation and cognitive control, which could enable an alteration between task-focused and mind-wandering states. In addition to these overall effects, a specific effect has been suggested for DMN when internally generated information, such as autobiographical memory, is receiving mental focus: the DMN effect changes task-negative to positive (Andrews-Hanna et al., 2014). The reversal could also happen as the mental state shifts to mind wandering. Such a reversal would be resolved as the mental state comes back to a task-focused state. The reversal and the recovery from it could therefore affect the time course of cognitive control. The perturbed function could implement various spontaneous mental states, including that reported in a previous study (Nakatani et al., 2019).

The current study investigated the effect of the reversal in the role of DMN activity on the time course of mental transition between task-focused and mind wandering states. Transitions were induced using

a tone counting task during which EEG, task performance, and subjective reports were recorded. From the time series data, Alpha and AEP amplitude, percentage of correct tone counting performance (%CC), self-reported numbers of thoughts (NoTs), were computed separately from the time points of self-report of mind wandering and task-focused, separately. Events were time-locked to obtain neural effects specific to a mind-wandering, and those to a task-focused state.

Results showed that the effect of Alpha on %CC changed from task-negative to positive between the task-focused and mind-wandering conditions. This was consistent with that in previous fMRI studies which suggested that the switch from task-negative to positive in DMN reflects a shift in mental focus from external task-related events to internally generated information (Buckner & Dinicola, 2019). Thus, the current EEG results may reflect the shift from external stimuli of the task to spontaneously generated thoughts.

Moreover, the current results suggested that Alpha activity several minutes before the mind-wandering or task-focused report had the strongest effect. Notably, the latencies were longer than observed in previous studies. In previous EEG studies, data segments of seconds before the subjective report of mind wandering or the response to the task were analyzed, i.e., the segments 0-12s before the report of mind wandering (Kirschner et al., 2012). FMRI studies have a larger time scale, because of the delay in the hemodynamic function (e.g., the default setting of the most widely used hemodynamic function in SPM is 6s). Yet, BOLD signal analyses were limited to within 1min from the reference point (e.g., Andrews-Hanna et al., 2014, a meta analysis). These studies did not estimate short and long-latency effects separately. Thus, their effects could include the long-latency ones we observed.

The estimated latency was slightly longer in the mind-wandering (6-10min) than in the task-focused (4-6min) conditions. At this moment, we have no convincing explanation for this difference. At least, the current result reports the relevance of past activity at the scale of several minutes, and therefore suggests that at least a part of DMN activity is involved in not immediate modulation of cognitive control.

Neural activities subsequent to the alpha activity eventually leads to a subjective report of mind wandering. During the several minutes, meta-cognitive processes for an awareness of mind wandering

(Smallwood & Schooler, 2006) also take place. However, neural dynamics for conscious awareness have been discussed in the time scales of milliseconds and seconds (Blumenfeld, 2023; Weissman et al., 2006). To discuss the dynamics together with those of cognitive control and spontaneous thought generation, the dynamics need to be further specified in the time scale of minutes.

A subjective experience of mind wandering develops, most likely, in the time scale of minutes. As we have shown, relevant psychological functions and neural activity are also in action in this time scale. Thus, we consider that minute is a reasonable time scale to build a cohesive understanding over various elements relevant to mind wandering. Nevertheless, investigation in this time scale has rarely been done. Situated in-between the time scales investigated in previous studies; seconds and shorter with neural measures (Braboszcz & Delorme, 2011; Ellamil et al., 2016; Kirschner et al., 2012; Kucyi, Esterman, Riley, & Valera, 2016), and hours and longer others with behavioral and subjective report measurements (Bennike, Wieghorst, & Kirk, 2017; Nakatani et al., 2019; Smith, Mills, Paxton, & Christoff, 2018), the current study fills the gap in multi-timescale investigation on mind wandering and related spontaneous mental processes.

In contrast to Alpha, AEP showed no stochastic causal effect on any of our measures. It is generally assumed that AEP amplitude is related to the level of attention to the incoming stimulus (Picton & Hillyard, 1974; Picton et al., 1974). As a stimulus evoked potential, AEP could be less sensitive to sustained attention than ongoing activity. For example, Docree and colleagues (Dockree, Kelly, Foxe, Reilly, & Robertson, 2007) reported that visual evoked potential (VEP) amplitude was less sensitive to sustained attention than visual alpha-band amplitude. In the current dataset, however, the sensitivity between the AEP amplitude (full band) and the auditory alpha-band power did not differ (Bernhard, 2018). Whereas level of attention is critical when processing is difficult, this is unlikely to be the case for the tone counting task. Thus, the current result may be explained better by the level of attention required by the tone counting task. It was explicitly chosen to be easy and monotonous, in order to elicit mind wandering. At least under such conditions, any attentional gain offered by the CCN may have limited influence on the mental shift from task-focused to mind wandering states.

Cognitive control in our experiment was modulated, not by CCN, but by DMN activity. This might appear to contradict the consensus in the literature, which links cognitive control to areas belonging to the CCN, such as the dorsal attention network, the salience network, and sensory-motor systems (Aben, Verguts, & Van den Bussche, 2017; Cohen et al., 1990; Funes et al., 2010; Ridderinkhof, 2002). However, there is no contradiction if we consider that the workload of the task was low and thus no strong contribution from CCN was required.

No strong evidence was found in our data for a stochastically causal neural effect on spontaneous thought generation. The only noticeable result was an effect of $Alpha_t$ in the low %MW segment. This effect was instantaneous, and thus a correlate rather than a cause. The current results therefore provide little information about which neural activity modulates the rate of spontaneous thought generation.

By correlating neural data directly with subjective reports of mind wandering, the wider context of spontaneous mental process, in which mind wandering is situated, would be lost. The context has been discussed in a conceptual level in previous studies, e.g., a 'conceptual space relating different types of thoughts' in which mind wandering and creative thinking are placed side-by-side (Christoff et al., 2016). The conceptual space helps us to grasp mental dynamics intuitively. However, for an empirical study of mental dynamics, the space needs to be specified further. We specified the space based on two psychological functions: spontaneous thought generation and cognitive control. Measures of the functions, the number of thoughts and task performance, were found weakly correlated to a sufficient extent to be treated as the basis of the state space in which a transition to mind wandering could take a complex trajectory, e.g., visiting 'multi-tasking' state on the way (Nakatani et al., 2019). In the current study, neural candidates for the mental dynamics were scrutinized: The overall opposite DMN and CCN effects (Fox et al., 2005; Fransson, 2005; Weissman et al., 2006) and the specific DMN effect which reverses depending on the direction of mental focus (Andrews-Hanna et al., 2014; Andrews-Hanna, Smallwood, & Spreng, 2014; Buckner & Dinicola, 2019). Results suggest the both effects contribute the dynamic change of the psychological functions. In particular, the specific DMN effect on cognitive control modulate the control

independent from CCN effect. Thus, it could also play a role in transitions to other states, such as obsessive thought. Neural effects on transitions to various spontaneous mental processes can be studied in a unified manner in our approach.

DMN and CCN have distinctive subnetworks. However, the effect of each subnetwork was not considered in the current study to avoid over-complication of models. An fMRI study reported contributions of various subsystems to mind wandering (Kucyi et al., 2016). BOLD/fMRI signal from 7 subsystems (3 DMN subsystems: Core DMN, dorsomedial prefrontal cortex, and middle temporal lobule networks and 4 CCN ones: salience, dorsal attention, frontoparietal control, and sensorimotor networks) were correlated with the level of mind wandering (which participants rated in response to probes presented during a sustained visual attention task) and variability of response times (RT) to trials in the 30s intervals before the probes. In all DMN subsystems, the correlation between the BOLD signal and the level of mind wandering was positive, while that between the signal and the RT variance was negative (except for the medial temporal subsystem, which fell short of statistical significance). The latter result indicates that RTs were more stable when the DMN subsystems were highly activated. It, therefore, parallels our observed task-positive DMN effect. Kucyi et al (2016) found no correlation between the BOLD signal of the CCN subsystems and the level of mind wandering. The authors attribute this to the unreliability of self-report. A positive correlation was found between the signal in the salience and dorsal attention networks and RT variability, which is in accordance with response variability studies (Esterman, Noonan, Rosenberg, & DeGutis, 2013). These results suggest that, as far as mind wandering is concerned, the respective subsystems of DMN operate in consort with each other, thereby validating the collective label over the subsystems. Disagreement with the results for subregion-based analysis (Andrews-Hanna et al., 2014; Andrews-Hanna et al., 2014; Buckner & Dinicola, 2019) needs to be resolved in future studies. At this moment, we keep the collective term 'DMN' for a simpler conceptualization of the neural effect.

The current DMN measure, event time-locked average of right temporal alpha-band amplitude, was newly developed for this study. The channels for the alpha activity were selected based on the variability of alpha amplitude averaged from the mid-level %MW report. The current results provide some vindication

for our procedure. The results, however, neither inform us where the source of the selected alpha-band activity is, nor how the selected signal is related to the auditory resting alpha rhythm, except that it corresponds to it in frequency band and rough scalp distribution (Lehtelä et al., 1997; Tiihonen et al., 1991; Weisz et al., 2011). Further investigation is necessary to establish the relationship between our measure and standard auditory resting-state alpha-band activity.

Of the signal processing methods, event time-locked averaging increased signal-to-noise ratio of the time series measure effectively. For example, to investigate the dynamics of mental states on the time scale of minutes, the level of mind wandering and the number of thoughts were reported every ~2min, after each trial of tone counting. Repeated reporting could affect the dynamics itself. This is a problem faced by all experimental studies in which some form of subjective report is required. Most likely in our study, the repeated probing kept the participants from engaging more fully in mind wandering, as it made the task less monotonous. The overall level of reported mind wandering was approx. 50% on average, suggesting the probing was not too frequent to interfere completely with mind wandering behavior. Neither does it seem likely that the repeated probing has produced the observed effect of Alpha on mind wandering. Between cause and effect, there is a random scattering of multiple probes that, as shown in Figure 3, has no temporally specific effect on later mind wandering behavior, whereas the Alpha effect on mind wandering is temporally specific.

Data acquisition and processing methods in the current study were kept relatively simple. The methods chosen allow including neural data collection in studies of mind wandering outside of the laboratory. EEG recording using a mobile system suffices for measurement, as an EEG signal was able to detect neural effects in accordance with those reported in previous fMRI studies. A small number of channels may be sufficient to obtain the relevant signals, e.g., one channel in the central and another in the right temporal scalp location. The current results specified the timing of possible marker signal, 4-10 minutes before the subjective report of %MW. The signal could be combined with bandpass filtering, which can be performed (quasi) in real-time, to develop a reliable neural marker for the prediction of mind wandering in various applications.

The current model concerns 4 effects from 2 neural networks on 2 psychological functions. The model could, in principle, be applied to investigate neural effects on transitions to other spontaneous mental states, such as creative thinking, multi-tasking, rumination or obsessive thought. For a larger array of mental states, the current state space of the model could be expanded. For instance, the dimension of wakefulness could be added to include lucid dreaming. This, however, would increase the search space for the parameters considerably. Heuristic search (e.g., through stepwise regression) would become necessary. With careful adjustment, the current approach could bring a unified understanding over various spontaneous mental states and relevant neural dynamics, beyond piecemeal neural correlates obtained for each of the states.

Open Practice Statement

All data are at the project page on Open Science Framework (https://osf.io/5ucvs/). They will be made publicly available after publication of the manuscript.

References

Aben, B., Verguts, T., & Van den Bussche, E. (2017). Beyond trial-by-trial adaptation: A quantification of the time scale of cognitive control. *Journal of Experimental Psychology: Human Perception and Performance, 43*(3), 509-517. doi:10.1037/xhp0000324

Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2014). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *Neuroimage, 91*, 324-335. doi:10.1016/j.neuroimage.2014.01.032

Andrews-Hanna, J. R., Smallwood, J., & Spreng, R. N. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the new York Academy of Sciences, 1316*(1), 29-52.

Antrobus, J. S. (1968). Information theory and stimulus-independent thought. *British Journal of Psychology, 59*(4), 423-430.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological), 57*(1), 289-300.

Bennike, I. H., Wieghorst, A., & Kirk, U. (2017). Online-based Mindfulness Training Reduces Behavioral Markers of Mind Wandering. *Journal of Cognitive Enhancement, 1*(2), 172-181. doi:10.1007/s41465-017-0020-9

Bernhard, H. (2018). *Tracking the mind: measuring the cortical dynamics during mind wandering in an auditory counting task. .* (Research Master). Maastricht University, Maastricht, The Netherlands.

Blumenfeld, H. (2023). Brain Mechanisms of Conscious Awareness: Detect, Pulse, Switch, and Wave. *Neuroscientist, 29*(1), 9-18. doi:10.1177/10738584211049378

Braboszcz, C., & Delorme, A. (2011). Lost in thoughts: neural markers of low alertness during mind wandering. *Neuroimage, 54*(4), 3040-3047.

Brookes, M. J., Hale, J. R., Zumer, J. M., Stevenson, C. M., Francis, S. T., Barnes, G. R., . . . Nagarajan, S. S. (2011). Measuring functional connectivity using MEG: methodology and comparison with fcMRI. *Neuroimage, 56*(3), 1082-1104.

Buckner, R. L., & Dinicola, L. M. (2019). The brain's default network: updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience, 20*(10), 593-608. doi:10.1038/s41583-019-0212-7

Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review, 67*(6), 380.

Chen, A. C., Feng, W., Zhao, H., Yin, Y., & Wang, P. (2008). EEG default mode network in the human brain: spectral regional field powers. *Neuroimage, 41*(2), 561-574.

Cheyne, J. A., Carriere, J. S., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. *Consciousness and cognition, 15*(3), 578-592.

Christoff, K., Irving, Z. C., Fox, K. C., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: a dynamic framework. *Nature Reviews Neuroscience, 17*(11), 718-731.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*(3), 332-361. doi:10.1037/0033-295x.97.3.332

Cole, M. W., & Schneider, W. (2007). The cognitive control network: integrated cortical regions with dissociable functions. *Neuroimage, 37*(1), 343-360.

Compton, R. J., Gearinger, D., & Wild, H. (2019). The wandering mind oscillates: EEG alpha power is enhanced during moments of mind-wandering. *Cognitive, Affective, & Behavioral Neuroscience, 19*(5), 1184-1191. doi:10.3758/s13415-019-00745-9

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., & Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nature Neuroscience, 3*(3), 292-297. doi:10.1038/73009

Dockree, P. M., Kelly, S. P., Foxe, J. J., Reilly, R. B., & Robertson, I. H. (2007). Optimal sustained attention is linked to the spectral content of background EEG activity: Greater ongoing tonic alpha (~ 10 Hz) power supports successful phasic goal activation. *European Journal of Neuroscience, 25*(3), 900-907.

Ellamil, M., Fox, K. C., Dixon, M. L., Pritchard, S., Todd, R. M., Thompson, E., & Christoff, K. (2016). Dynamics of neural recruitment surrounding the spontaneous arising of thoughts in experienced mindfulness practitioners. *Neuroimage, 136*, 186-196. doi:10.1016/j.neuroimage.2016.04.034

Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral cortex, 23*(11), 2712-2723.

Fink, A., Graif, B., & Neubauer, A. C. (2009). Brain correlates underlying creative thinking: EEG alpha activity in professional vs. novice dancers. *Neuroimage, 46*(3), 854-862. doi:10.1016/j.neuroimage.2009.02.036

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., & Raichle, M. E. (2005). From The Cover: The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences, 102*(27), 9673-9678. doi:10.1073/pnas.0504136102

Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping, 26*(1), 15-29. doi:10.1002/hbm.20113

Funes, M. J., Lupiáñez, J., & Humphreys, G. (2010). Sustained vs. Transient cognitive control: Evidence of a behavioral dissociation. *Cognition, 114*(3), 338-347. doi:10.1016/j.cognition.2009.10.007

Ganschow, B. (2017). *Mind wandering at and under capacity: an on-line daily life assessment of mental activity.* . (Master). KU Leuven, Leuven, Belgium.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and cognition, 4*(1), 1-21.

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica, 37*(3), 424-438. doi:10.2307/1912791

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences, 100*(1), 253-258. doi:10.1073/pnas.0135058100

Gruberger, M., Simon, E., Levkovitz, Y., Zangen, A., & Hendler, T. (2011). Towards a Neuroscience of Mind-Wandering. *Frontiers in human neuroscience, 5*(56). doi:10.3389/fnhum.2011.00056

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience, 3*(3), 284-291. doi:10.1038/72999

Jaarsveld, S., & Van Leeuwen, C. (2005). Sketches from a Design Process: Creative Cognition Inferred From Intermediate Products. *Cognitive science, 29*(1), 79-101. doi:10.1207/s15516709cog2901_4

Jerbi, K., Vidal, J. R., Ossandon, T., Dalal, S. S., Jung, J., Hoffmann, D., . . . Lachaux, J.-P. (2010). Exploring the electrophysiological correlates of the default-mode network with intracerebral EEG. *Frontiers in Systems Neuroscience*. doi:10.3389/fnsys.2010.00027

Kam, J., Rahnuma, T., Park, Y., & Hart, C. (2022). Electrophysiological markers of mind wandering: A systematic review. *Neuroimage*, 119372.

Kirschner, A., Kam, J. W. Y., Handy, T. C., & Ward, L. M. (2012). Differential synchronization in default and task-specific networks of the human brain. *Frontiers in human neuroscience, 6*, 139.

Klinger, E., & Cox, W. M. (1987). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality, 7*(2), 105-128.

Kucyi, A., Esterman, M., Riley, C. S., & Valera, E. M. (2016). Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proceedings of the National Academy of Sciences, 113*(48), 13899-13904. doi:10.1073/pnas.1611743113

Laufs, H., Krakow, K., Sterzer, P., Eger, E., Beyerle, A., Salek-Haddadi, A., & Kleinschmidt, A. (2003). Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest. *Proceedings of the National Academy of Sciences, 100*(19), 11053-11058. doi:10.1073/pnas.1831638100

Lehtelä, L., Salmelin, R., & Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neuroscience Letters, 222*(2), 111-114. doi:10.1016/s0304-3940(97)13361-4

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279-281. doi:10.1038/36846

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., & Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proceedings of the National Academy of Sciences, 104*(32), 13170-13175. doi:10.1073/pnas.0700668104

Marino, M., Arcara, G., Porcaro, C., & Mantini, D. (2019). Hemodynamic Correlates of Electrophysiological Activity in the Default Mode Network. *Frontiers in Neuroscience, 13*(1060). doi:10.3389/fnins.2019.01060

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science, 315*(5810), 393-395.

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function, 214*(5-6), 655-667.

Nakatani, C., Ganschow, B., & van Leeuwen, C. (2019). Long-term dynamics of mind wandering: ultradian rhythms in thought generation. *Neuroscience of Consciousness, 2019*(1). doi:10.1093/nc/niz007

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2015). Cognitive Control Predicts Use of Model-based Reinforcement Learning. *Journal of cognitive neuroscience, 27*(2), 319-333. doi:10.1162/jocn_a_00709

Pearl, J. (2000). Causality: Models, reasoning and inference cambridge university press. *Cambridge, MA, USA, 9*, 10-11.

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II: Effects of attention. *Electroencephalography and clinical neurophysiology, 36*, 191-200.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography and clinical neurophysiology, 36*, 179-190.

Raichle, M. E. (2015). The Brain's Default Mode Network. *Annual Review of Neuroscience, 38*(1), 433-447. doi:10.1146/annurev-neuro-071013-014030

Raichle, M. E., Macleod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences, 98*(2), 676-682. doi:10.1073/pnas.98.2.676

Ridderinkhof, R. (2002). Micro- and macro-adjustments of task set: activation and suppression in conflict tasks. *Psychological Research, 66*(4), 312-323. doi:10.1007/s00426-002-0104-7

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *Journal of Neuroscience, 27*(9), 2349-2356. doi:10.1523/jneurosci.5587-06.2007

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958. doi:10.1037/0033-2909.132.6.946

Smith, G. K., Mills, C., Paxton, A., & Christoff, K. (2018). Mind-wandering rates fluctuate across the day: evidence from an experience-sampling study. *Cogn Res Princ Implic, 3*(1), 54. doi:10.1186/s41235-018-0141-4

Tiihonen, J., Hari, R., Kajola, M., Karhu, J., Ahlfors, S., & Tissari, S. (1991). Magnetoencephalographic 10-Hz rhythm from the human auditory cortex. *Neuroscience Letters, 129*(2), 303-305. doi:10.1016/0304-3940(91)90486-d

Weissman, D. H., Roberts, K. C., Visscher, K. M., & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. *Nature Neuroscience, 9*(7), 971-978. doi:10.1038/nn1727

Weisz, N., Hartmann, T., Müller, N., Lorenz, I., & Obleser, J. (2011). Alpha rhythms in audition: cognitive and clinical perspectives. *Front Psychol, 2*, 73. doi:10.3389/fpsyg.2011.00073

Appendix

1. Mind wandering percentage (%MW) histogram

Figure A1 provides the %MW histogram over all participants. Low-, intermediate-, and high %MW trials are marked in yellow (%MW<30), gray (40≦%MW<60), and purple (%MW≧70), respectively

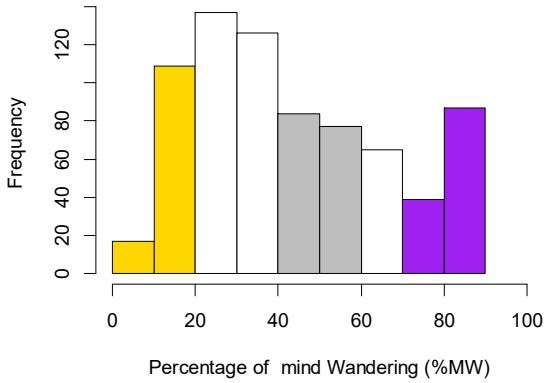


Fig A1. Histogram of the reported mind wandering percentage (%MW)

2. Electrodes for auditory evoked potential (AEP) and Alpha amplitudes

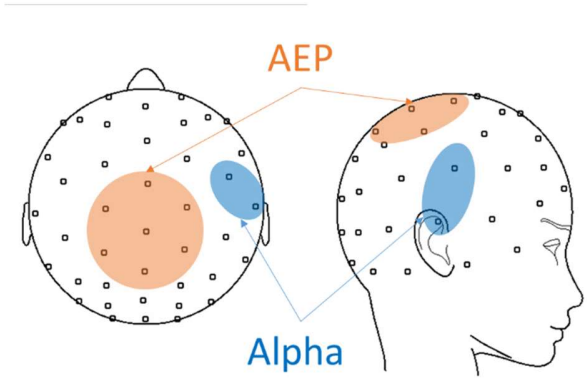


Figure A2. Electrodes for auditory evoked potential (AEP) and Alpha amplitudes
Electrodes used to compute the AEP and Alpha amplitude are indicated. Note that the electrode placement is not the international 10-10, but a manufacturer-defined system.

3. Alpha amplitude averaged from mid-level %MW

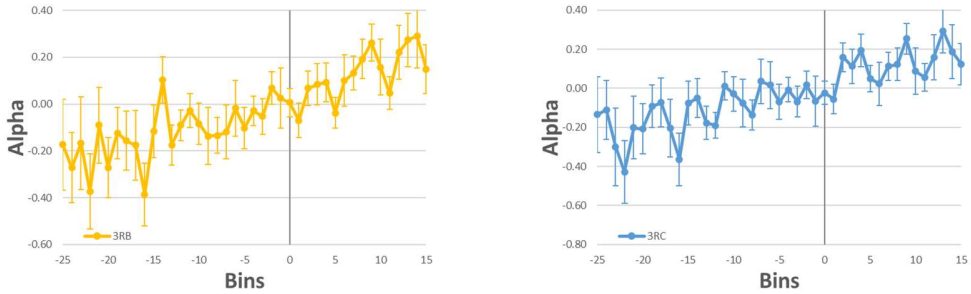


Figure A3. Average waveform of two right temporal electrodes, 3RB (left) and 3RC (right) Bin0 starts from the mid-level %MW report (40≦%MW<60). Error bars indicate SE.

4. Auto regression effects at r=1, s=1

The effect of autoregression at *i=1, j=1* was estimated to compare that with pre-selected i^* and l^* (See main text Notes on autoregression effect). In the same way as that in models for Tables 2, 3, and 5 in main text, the index for AEP was set as *l=1*.

a) Estimated coefficients of the full model with (*i=1, j=1, k=4, l=1*) for high %MW segments

model: $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-1}) + (Alpha_{t-4} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC$_t$**	(Intercept)	-4.39	-7.07	-1.70	-3.21	0.0014
	NoTs $_{t-1}$	2.02	-1.48	5.52	1.14	0.2571
	%CC $_{t-1}$	0.18	0.09	0.27	3.88	0.0001
	Alpha $_{t-4}$	8.07	4.11	12.04	4.00	0.0001
	AEP $_{t-1}$	2.46	-3.43	8.35	0.82	0.4124
	Alpha $_t$	-0.72	-4.45	3.00	-0.38	0.7031
	AEP $_t$	3.90	-2.02	9.83	1.30	0.1960
	Participants	0.36	0.14	0.59	3.16	0.0017

Response var. = **NoTs $_t$**	(Intercept)	0.00	-0.07	0.07	-0.11	0.9120
	NoTs $_{t-1}$	0.19	0.10	0.28	4.06	0.0001
	%CC $_{t-1}$	0.00	0.00	0.01	2.44	0.0150
	Alpha $_{t-4}$	0.07	-0.04	0.17	1.28	0.2020
	AEP $_{t-1}$	-0.12	-0.27	0.03	-1.54	0.1240
	Alpha $_t$	0.08	-0.02	0.18	1.64	0.1030
	AEP $_t$	0.11	-0.04	0.26	1.41	0.1600
	Participants	0.00	0.00	0.01	0.98	0.3280

b) Estimated coefficients of the full model with (*i=1, j=1, k=5, l=1*) for high %MW segments

model: $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-1}) + (Alpha_{t-5} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC$_t$**	(Intercept)	-4.50	-7.20	-1.80	-3.28	0.0011
	NoTs $_{t-1}$	1.75	-1.78	5.29	0.98	0.3293
	%CC $_{t-1}$	0.17	0.08	0.26	3.56	0.0004
	Alpha $_{t-5}$	6.74	2.75	10.72	3.32	0.0010
	AEP $_{t-1}$	3.82	-2.09	9.74	1.27	0.2047
	Alpha $_t$	0.19	-3.50	3.87	0.10	0.9212
	AEP $_t$	2.68	-3.27	8.63	0.88	0.3769
	Participants	0.38	0.15	0.60	3.28	0.0011

Response var. = **NoTs $_t$**	(Intercept)	0.00	-0.07	0.06	-0.13	0.8939
	NoTs $_{t-1}$	0.18	0.09	0.28	3.98	0.0001
	%CC $_{t-1}$	0.00	0.00	0.01	2.30	0.0217
	Alpha $_{t-5}$	0.06	-0.04	0.16	1.19	0.2362
	AEP $_{t-1}$	-0.11	-0.26	0.04	-1.39	0.1656
	Alpha $_t$	0.09	-0.01	0.18	1.80	0.0718
	AEP $_t$	0.10	-0.05	0.25	1.28	0.2030
	Participants	0.00	0.00	0.01	1.02	0.3088

c) Estimated coefficients of the full model with (*i=1, j=1, k=3, l=1*) for Low %MW segments

model: $[NoTs_t, \%CC_t] \sim (NoTs_{t-1} + \%CC_{t-1}) + (Alpha_{t-3} + AEP_{t-1}) + (Alpha_t + AEP_t) + Participants$

	Predictors	Coefficients	Confidence interval 2.50%	97.50%	t-value	Pr(>|t|)
Response var. = **%CC$_t$**	(Intercept)	-6.09	-9.89	-2.30	-3.16	0.0017
	NoTs $_{t-1}$	6.46	2.36	10.55	3.10	0.0021
	%CC $_{t-1}$	0.11	0.01	0.21	2.17	0.0310
	Alpha $_{t-3}$	-10.48	-15.12	-5.84	-4.44	0.0000
	AEP $_{t-1}$	1.53	-4.98	8.04	0.46	0.6449
	Alpha $_t$	7.30	2.64	11.96	3.08	0.0022
	AEP $_t$	-2.23	-8.89	4.43	-0.66	0.5110
	Participants	0.52	0.21	0.83	3.26	0.0012

Response var. = **NoTs $_t$**	(Intercept)	-0.05	-0.14	0.03	-1.21	0.2285
	NoTs $_{t-1}$	0.14	0.04	0.23	2.88	0.0042
	%CC $_{t-1}$	0.00	0.00	0.00	2.02	0.0440
	Alpha $_{t-3}$	0.10	0.00	0.21	1.92	0.0558
	AEP $_{t-1}$	0.02	-0.13	0.17	0.29	0.7758
	Alpha $_t$	-0.17	-0.27	-0.06	-3.13	0.0019
	AEP $_t$	0.03	-0.12	0.18	0.37	0.7103
	Participants	0.01	0.00	0.01	1.75	0.0804