

Neural Dynamics of Spontaneous Thought: An Electroencephalographic Study

Manesh Girn¹(✉), Caitlin Mills¹, Eric Laycock¹, Melissa Ellamil²,
Lawrence Ward¹, and Kalina Christoff¹

¹ University of British Columbia, Vancouver, Canada
maneshg@alumni.ubc.ca

² Max Planck Institute for Human Cognitive and Brain Sciences,
Leipzig, Germany

Abstract. Spontaneous thinking is a ubiquitous aspect of our mental life and has increasingly become a hot topic of research in cognitive neuroscience. To date, functional neuroimaging studies of spontaneous thought have revealed general brain recruitment centered on a combination of default mode network and executive regions. Despite recent findings about general brain recruitment, very little is known about how these regions are recruited dynamically over time. The current research addresses this gap in the literature by using EEG to investigate the fine-grained temporal dynamics of brain activity underlying spontaneous thoughts. We employed the first-person reports of experienced meditators to index the onset of spontaneous thoughts, and examined brain electrical activity preceding indications of spontaneous thought onset. An independent component analysis-based source localization procedure recovered sources very similar to those previously found with fMRI (Ellamil et al. in *NeuroImage* 136:186–196, 2016). In addition, phase synchrony analyses revealed a temporal trajectory that begins with default network midline and salience network connectivity, followed by the incorporation of language and executive regions during the period from thought generation to appraisal.

Keywords: Spontaneous thought · Neural dynamics · Electroencephalography · Default mode network · Frontoparietal control network · Independent-component analysis

1 Introduction

The human brain has a remarkable propensity to spontaneously generate mental content that captures our attention. Memories and projections into the future arise in our awareness unbidden, a previously contemplated problem may suddenly return to cognizance, and so on. Such mental phenomena fall under the rubric of ‘spontaneous thought,’ which can be understood as unintended mental phenomena that occur without conscious generation, and which are relatively free-flowing with a lack of strong deliberate control [1]. Types of spontaneous thought include mental states such as dreaming, mind wandering, and creativity [1]. Psychological research has linked spontaneous thought to various both beneficial (creativity, prospective planning, etc.

[2]) and detrimental (disruptions and reductions in memory encoding [3, 4]) processes and outcomes [5, 6]. However, much less is understood about the neural markers of spontaneous thought, particularly with respect to its initial generation and onset – a focus of the current research.

To date, functional neuroimaging studies of spontaneous thought have revealed general brain recruitment centered on a combination of default mode network (DMN) and frontoparietal control network (FPCN) regions [1, 7]. The DMN, which encompasses medial temporal memory regions, is thought to supply the general content of thought [8], while the FPCN likely pertains to processes such as attentional appraisal, thought evaluation and elaboration, as well as maintenance in working memory [9, 10]. This general network recruitment has been further fractionated into distinct sub-components and specific brain regions, each of which putatively has a distinct functional contribution to spontaneous thought [8, 11].

Moreover, in line with the highly dynamic and interactive nature of the brain, these various sub-components are very unlikely to be uniformly recruited; rather, they are likely to be differentially integrated into varying dynamic patterns contingent on both the content of the thought, as well as the current temporal stage of the thought process [11–13]. Thus, it is reasonable to expect that different stages of neural activity may map onto different stages in the subjective experience of spontaneous thought: thought generation, attentional appraisal, evaluation, and elaboration.

We currently know very little about the neural dynamics underlying the temporal components of spontaneous thought. A particular challenge has been to isolate the neural activity corresponding to the initial generation of a spontaneous thought and its subsequent attentional appraisal. The reason for this is evident in the methodological difficulty of experimentally isolating the onset of an elusive mental phenomenon such as a spontaneous thought. By definition, a spontaneous thought cannot be induced by external means; it must rise of its own accord – thus differentiating it from the stimulus-based paradigms that constitute the bulk of cognitive neuroscience research. Moreover, there is currently no known objective neural or behavioral index for the moment of spontaneous thought onset. Studies that hope to investigate this phenomenon, therefore, must employ first-person measures. Yet, first-person measures are also imperfect given individuals' typically poor ability to indicate the onset of mental content, often only realizing their engagement in a mind-wandering episode some time after the fact [5, 14].

One study by Ellamil and colleagues found a way to address these challenges: using trained meditators to investigate the temporal dynamics of spontaneous thought [11]. Indeed, experienced meditators are suggested to have a refined introspective capacity, and may as such constitute a uniquely suitable population to investigate subtler aspects of first-person experience. This contention is encapsulated by the 'neurophenomenology' research program [15], which promotes understanding the neural basis of first-person experience by relating rigorous first-person reports from individuals with introspective training to neural measures [16].

Meditation is generally conceived as a form of attentional training [17], and mindfulness meditation in particular has an explicit goal of increasing one's ability to maintain awareness of the rise and fall of spontaneous mental phenomena [18]. Meditators/mindfulness practitioners have been shown to display greater awareness of

subtle emotional and interoceptive feelings [19, 20] and have shown increased meta-cognitive accuracy on memory judgments [21]. In addition, meditators have shown improved ability on a number of perceptual tasks, presumably as a result of increased attentional stability and efficiency [22, 23]. Performance improvements also include, for example, increased perceptual sensitivity and vigilance on a visual threshold line-length discrimination task [24] and improved detection of the second target during attentional blink tasks [25]. As such, experienced meditators appear to be an ideal population to investigate subtler aspects of spontaneous thought, as their attentional capacities may enable them to recognize the onset of thoughts with greater accuracy.

Ellamil and colleagues took advantage of the heightened metacognitive abilities of meditators and employed a procedure that had long-term meditators self-report the onset of spontaneous thoughts while they were maintaining a meditative focus. They characterized brain region recruitment in blocks of 2 s prior to and after indications of spontaneous thought onset [11]. This study found activations spanning both the DMN and FPCN, with a distinct temporal trajectory that originated in the medial temporal lobes as well as lateral and posteromedial parietal cortex (2 s preceding onset), spread to a number of regions including medial prefrontal cortex (onset), and later spread to executive regions such as dorsolateral prefrontal cortex and anterior cingulate cortex (2 s following onset) [11]. Notably, however, this study was greatly limited by the relatively poor temporal resolution of fMRI. Neural activity contains significant variability at the sub-second level, and as such a large amount of information is lost in the use of 2-second blocks. Here we aim to extend the findings of the Ellamil study by using EEG to address the temporal limitations of fMRI.

1.1 Current Study and Research Questions

The present study employed high-density electroencephalography (EEG) to overcome the temporal limitations inherent in fMRI. High-density EEG is useful in this context as it allows the measurement of brain electrical activity at the millisecond time scale while also offering reasonably accurate source localization. Notably, at least one previous study has successfully used EEG source localization to recover putative spontaneous thought-related regions [26]. We adopted a similar experimental paradigm to Ellamil et al. [11], capitalizing on the improved first-person reports afforded by using experienced meditators as participants.

Our primary goals were (1) to determine whether brain regions activated during spontaneous thought converge across our EEG analyses and the previous fMRI findings [11], and (2) to investigate the temporal progression of interregional functional connectivity (information sharing) between these regions. In line with our first goal, we hypothesized that we would recover sources of neural electrical activity that correspond to regions previously implicated in spontaneous thought, including DMN regions such as the posterior cingulate cortex, medial prefrontal cortex, and medial temporal lobes, and executive regions such as dorsolateral prefrontal cortex and dorsal anterior cingulate [1, 7]. Moreover, we expected a particular temporal trajectory of activations spanning these regions. In line with the results of Ellamil and colleagues' study [11],

we hypothesized that spontaneous-thought related activity will first begin in the medial temporal lobe (MTL), and then spread to medial and lateral parietal regions, followed by other default-mode network regions such as the medial prefrontal cortex (MPFC). Additionally, executive regions, such as the dorsolateral prefrontal cortex (DLPFC), are expected to only show activity immediately prior to responses, and not during the early onset stage. This expectation is consistent with the DLPFC's role in the volitional top-down deployment of cognitive control [27], implying that executive regions may not play a role early on in the spontaneous thought generation process.

In line with our second goal, we made a number predictions afforded by the greater degree of temporal specificity in comparison to the previous fMRI results [11]. Importantly, we analyze brain activity at much smaller time intervals, up to four times smaller than the previous fMRI study. For example, Ellamil and colleagues found a peak in both medial temporal lobe (MTL) and posterior cingulate cortex (PCC) activity at -2 s relative to thought onset indication, followed by medial prefrontal cortex activation at 0 s (thought onset) [11]. Using EEG, we were able increase the temporal specificity from 2 s temporal windows to 500 ms. We expect connectivity to begin between the MTL and PCC, immediately followed by PCC-MPFC and PCC-lateral parietal connectivity. This is consistent with a recent intracranial EEG study that found that MTL-retrosplenial cortex (a region ventrally adjacent to the PCC) phase locking peaked at around ~ 200 ms following cued autobiographical memory retrieval, with retrosplenial cortex activity peaking ~ 300 ms later [28]. We expect this to then be followed by connectivity with the MPFC in virtue of dense MPFC-PCC interconnectivity and frequent coactivation [29], and connectivity with lateral parietal regions in virtue of previous electrophysiological work indicating PCC and lateral parietal coactivation during episodic memory retrieval [30].

2 Methods

2.1 Participants and Design

A total of 23 participants took part in the study (13 females; mean age = 31.4 years old, $SD = 5.3$). 3 participants had to be excluded either because of too few thought reports (>20), or because of technical issues with the data acquisition software. Participants were experienced meditators recruited from the Greater Vancouver area. Minimum meditation experience cut-off was 500 total hours within the past 2 years (mean meditation experience = 726.6 h, $SD = 374.4$ h). Participants received \$10 an hour as compensation.

We employed a within-subjects yoked-control design (Fig. 1), similar to the one used in Ellamil et al. [11]. Participants engaged in two 10-minute conditions while sitting in front of a computer monitor: a thought condition (monitoring and reporting spontaneous thoughts) and a word condition (monitoring and reporting words that appeared onscreen, yoked to the timing of the thought condition). The two conditions were completed back-to-back twice, totaling 20 min each (40 min total).

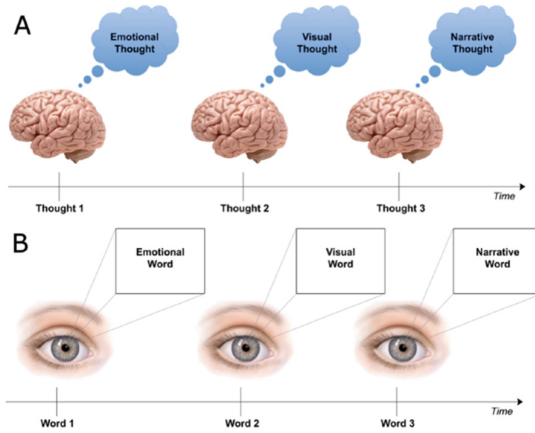


Fig. 1. A visual representation of the two experimental conditions. Adapted from Ellamil et al. [11] with permission.

2.2 Materials

The task and stimuli were implemented and presented using E-Prime 2.0 (Psychology Software Tools, Sharpsburg, PA, USA). The words and fixation cross appeared as gray font on a black background.

Word Task Stimuli. Words in the ‘verbal’ list consisted of 30 nouns (e.g., work, money, family, goals, health) selected from the Edinburgh Associative Thesaurus (EAT) [31]. The specific words were chosen based on their association with the typical concerns that constitute the bulk of spontaneous thought content [32]. These included home and household matters; employment and finance; partner, family, and relatives; friends and acquaintances; love, intimacy, and sexual matters; self-changes; education and training; health and medical matters; spiritual matters; and hobbies, pastimes, and recreation. The ‘visual’ list consisted of 30 nouns (e.g., mountain, beach, rain, sun, pet) selected from the Medical Research Council (MRC) Psycholinguistics Database [33] which had imageability, concreteness, and familiarity ratings of 500–700 (on scales of 100 = very low to 700 = very high). The ‘somatic’ list consisted of 30 nouns and adjectives selected from the EAT that were associated with various body sensations (e.g., warmth, tickle, vibration, pressure, pain), whereas the ‘affective’ list consisted of 30 adjectives associated with various emotions (e.g., happiness, sadness, anger, disgust, fear, surprise) [31]. Across all types, each word contained 3–10 letters and 1–3 syllables.

2.3 Procedure

Thought Condition. Participants were instructed to focus their attention on an aspect of their breathing (sensations on the nostrils or the rise and fall of the abdomen) throughout both conditions. They were told to rest their eyes on a black computer

screen with a white fixation cross in the center. Participants reported when they detected the onset of a spontaneous thought by pressing ‘I’ on the keyboard, then immediately indicated the type of thought it was with an additional button press (I = verbal, J = visual, K = somatic, L = affective). Verbal thoughts were defined as thoughts that represent an internal narrative and/or that were embedded in terms of language/inner speech. Visual thoughts were defined as any form of mental imagery or symbols. Somatic thoughts were defined as thoughts directly related to bodily sensations. Affective thoughts were defined as thoughts directly related to emotions. These thought types were chosen to correspond the categories of thought frequently identified during mindfulness training [34].

Word Condition. For the word condition, participants were also instructed to attend to their breathing and rest their eyes on a black computer screen with a white fixation cross in the center. Rather than reporting on thoughts, however, participants were instructed to press a button when a word was presented on the computer screen. Word presentation was yoked to the thought condition, matched precisely for timing and type of thought reported in the preceding thought condition via a real-time algorithm. Participants were explicitly instructed to read and interpret the word prior to responding with a button press, and each word stayed onscreen until the first button press. The first button press was always followed by an asterisk (*) and the second by two asterisks (**), in order to indicate that the responses were received. Participants were explicitly told not to report thoughts during the word condition.

3 EEG Acquisition and Analysis

3.1 EEG Recording and Data Pre-processing

EEG was recorded using 60 electrodes on a standard electrode cap (International 10–10 System). The reference electrode was over the right mastoid and electrode AFz served as the ground. Eye movements were recorded by 4 peri-ocular electrodes. Electrode impedances were below 10 k Ω (input impedance of the amplifier was 2 g Ω).

Prior to analysis, all signals were re-referenced to an average reference and down-sampled to 250 Hz. Signals were also digitally filtered using the EEGLAB toolkit in MATLAB [35], to only contain activity within the 1–50 Hz range. The continuous data were then separated into 10-second time-bins (epochs) time locked to the thought/word onset button presses (5.5 s prior and 4.5 s after). Distinct sets of epochs were delineated for each type of thought (verbal, visual, somatic, affective) and for the word category in general – for a total of 5 conditions, which were used for comparisons. Epoching was performed in order to remove task-irrelevant inter-trial activity.

3.2 EEG Source Localization

In order to recover the sources of neural activity responsible for the EEG signals, we ran independent component analysis (ICA) on the EEG data. ICA is a method of blind

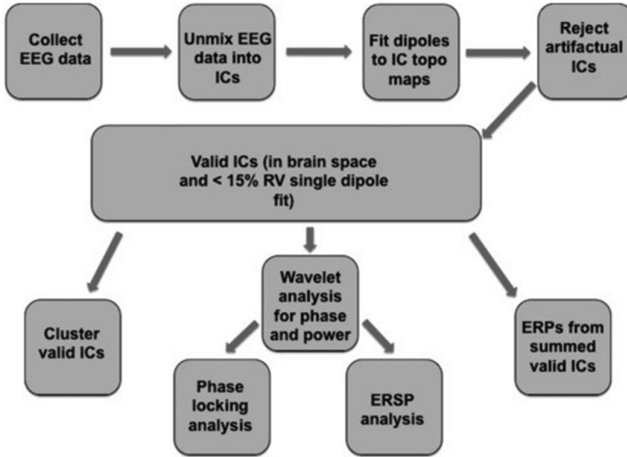


Fig. 2. EEG data analysis pipeline. Event-related spectral perturbations (ERSPs) and event-related potentials (ERPs) were excluded from discussion in the present paper.

source separation, which separates the EEG data into apparent neural sources without a priori constraints. Essentially, ICA takes into account the activity across all 60 channels of the raw EEG data, and organizes it into a new set of 60 new channels (independent components; ICs) that are free of the volume conduction that characterizes the raw scalp EEG [36]. ICs represent non-Gaussian neural sources that are maximally independent of each other, as defined by sharing minimal mutual information. Specifically, the procedure involves iteratively solving a neural network for an IC \times channel ‘unmixing’ matrix, U , according to $I = UX$, where X is the electrode channel \times time data matrix and I is the IC \times time matrix of independent component activations. We computed ICA by using the *runica* function in EEGLAB, which implements extended infomax ICA [36].

Once IC’s were derived from the data of all 23 participants (a total of 1380 ICs), we began the process of pruning and rejecting artifactual and unreliable sources. First, we localized the neural sources of the ICs using the *dipfit* algorithm in EEGLab. Electrode locations were co-registered to the Montreal Neurological Institute (MNI) average brain, which enabled the production of Talairach brain space coordinates for the IC dipoles. All ICs with dipoles localized outside of the brain were rejected as artifacts. We then examined the spectral distributions for all of the included ICs, and rejected the ICs that significantly deviated from a $1/f$ distribution; which would be indicative of non-neural sources of activity such as eye movements or muscle twitches [37].

Clusters were then created using EEGLAB’s k-means clustering algorithm, which minimizes intra-cluster distances while maximizing the inter-cluster distances, based on each IC’s location in Talairach brain space. Large values of k relative to the number of ICs yield many small, highly localized, clusters with only a few ICs per cluster, whereas small values of k yield a few large, diffuse, non-localized, clusters with many ICs per cluster. We investigated a range of k -values from 10 to 20, and ultimately settled on a k -value of 13: thirteen clusters composed from a total of 289 valid ICs

(ranging from 11 to 17 ICs contributed by each participant). Importantly, other k -values within this range produced highly similar results.

In addition, we compared these results to the clusters generated by a seed-based cluster analysis. This analysis generated clusters based on their Euclidean distance from a priori ROIs (distance <3 cm), as defined based on the Talairach coordinates of the peak activations from Ellamil et al. [11]. We compared the clusters across these two analyses in order to better determine sources of convergence and divergence between our results and those of the original fMRI study.

We then selected the six IC clusters that best overlapped between the two cluster analyses, based on their putative relationship with spontaneous thought-relevant processing. Although several of the unselected clusters are also of some interest, their ICs were not analyzed further in the interest of our focus on spontaneous thought and of minimizing statistical error. We pruned these six clusters to contain only the most representative IC from each participant, which consisted in choosing the IC in closest proximity to the centroid and with the best fit to a single dipole. In a few ambiguous cases, we also examined scalp maps and event-related spectral perturbations to determine the inclusion of a particular IC over another.

It is important to emphasize here that dipole localization and clustering is for interpretational purposes only, and has no bearing on the subsequent spectral and connectivity analyses. These subsequent analyses are computed solely on the basis of the statistically derived IC activations themselves.

In order to decompose the broadband signals into their component frequencies, we applied wavelet analyses on all of the included ICs. Specifically, a Morlet wavelet analysis on each IC time series yielded wavelet coefficients of the sinusoidal oscillations between 1 and 50 Hz, from which phase at each time-frequency point was calculated to be used in computing phase synchrony analyses.

3.3 Phase Synchrony Analysis

In order to assess functional connectivity (an index of information sharing) between regions, we conducted phase synchrony analyses. To do this, we computed phase-locking values (PLVs) between IC pairs of interest, each of which were localized to a specific brain region. PLVs were computed via the following formula [35]:

$$PLV_{1,2}(f, t) = \frac{1}{N} \sum_{k=1}^N \frac{W_{1,k}(f, t) W_{2,k}^*(f, t)}{|W_{1,k}(f, t) W_{2,k}(f, t)|}$$

where $W_{i,k}(f, t)$ are the wavelet coefficients for each time point, t , and frequency, f , for each IC, i , and $k = 1$ to N is the index of epochs. The PLVs as computed by this equation represent the degree to which the phase differences between signals at a specific oscillatory frequency are constant across trials. PLVs can range from 0 to 1, where 0 indicates a total absence of phase locking, and 1 indicates that the phase difference between two ICs at any given time point remains constant across all trials. Due to neural noise, only $0 < PLV < 1$ is expected from any time series of brain activity.

We first identified the ICs from participants that were common to both clusters and then computed PLVs across all cluster pairs, across all conditions (verbal thought, somatic thought, affective thought, visual thought, word condition), and all frequency bands. We then ran t-tests comparing the PLVs between each thought condition and the word condition. This gave us a PLV/time two-dimensional matrix for each comparison (a total of four), for each region pair. We then looked at each frequency band separately (theta, alpha, beta, gamma). Ultimately, for the scope of this paper we decided to focus on the alpha frequency band for subsequent analyses. The reasoning for this was two-fold. First, previous work has indicated that alpha processing plays a critical role in attentional processes [38–40], and also in internally-oriented processing [38, 41]. Second, and in line with the abovementioned findings, activity in the alpha band displayed the greatest amount of significant activity around the events of interest.

Looking at the alpha frequency band, we calculated the maximum t value at each time point (i.e., across its component frequencies; 8–12 Hz). This was to determine the most significant point in this frequency band at a given point in time; allowing us to collapse across the band to one t value per time point. We then determined whether each of these maximum t values was statistically significant at $p < 0.01$. This gave us a binary output: 0 if not statistically significant, 1 if statistically significant.

Next, we did a chi-squared analysis as a means to measure synchrony over time in 500 ms time-bins. This analysis technique and time-bin length was chosen in order to overcome limitations associated with the number of trials we received. We looked at time bins of 500 ms, from 4 s pre-button press to the button press. Chi Squared tests were used to determine whether the proportion of significant time points (the number of 1 s from the t value analysis) within a given time bin was significantly greater at $p < 0.01$ in one condition relative to a comparison. We compared each individual thought condition relative to the word condition, and the word condition relative to the thought condition as a whole. As an additional check, the condition in question also required at least one instance of a full cycle (133 ms for alpha) of consecutive significant time points within the time bin to be considered significant. This was to control for spurious inconsistent time points of significance within a time bin. Ultimately, these tests indicate whether one condition is exhibiting higher connectivity between a given region pair, during a given time bin.

4 Results

4.1 Behavioral Results

Before addressing our main research questions, we first assessed the frequency and types of thoughts reported by the meditators. Participants reported a mean of 46 thoughts (2.3 thoughts per minute). Verbal thoughts were the most common (40%), followed by somatic (37%), visual (16%), and affective thoughts (7%). As mentioned in the methods, the words were yoked to the reported thoughts in the preceding thought condition and thus had the same frequency and distribution. The trials for each individual thought type were collapsed (into four groups), and all of the words types were collapsed into one group. In subsequent analyses, we focused on the verbal thought

type due to their high frequency (most common) and comparability with the word condition, where participants were asked to interpret the meaning of a word.

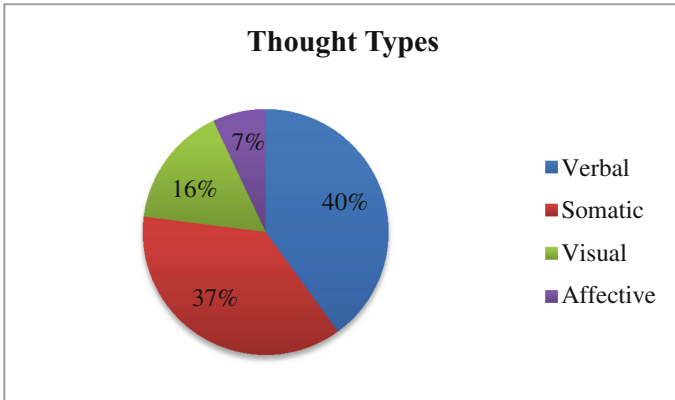


Fig. 3. Thought type frequencies.

4.2 EEG Sources Converge with Neuroimaging Data

Next, we determined whether the sources of neural activity recovered from the EEG data spatially aligned with the regions implicated by previous fMRI investigations of spontaneous thought. Specifically, we assessed whether prominent clusters localized using EEG converged with regions from Ellamil et al. [11]. In line with our hypotheses, a number of the clusters that were recovered through the ICA, dipole fitting, and cluster analysis exhibited high spatial correspondence to regions previously identified with fMRI using the same paradigm (see Table 1 and Fig. 2) [11]. The scalp maps of the ICs indicated single dipole sources (Fig. 3), which further suggests that these locations represent compact cortical generators [42]. Convergence between our EEG analyses and previously reported hemodynamic results strongly suggests that we have indeed identified reliable regions of spontaneous thought-related activity.

Table 1. Talaraich coordinates for the centroid of recovered neural source clusters.

Region	Talarach coordinates (centroid)				
	L/R/M	BA	x	y	z
Medial frontal gyrus (MPFC)	M	10	-3	47	25
Dorsal anterior cingulate	M	32	-8	2	36
Middle frontal gyrus (DLPFC)	L	9	-30	37	31
Insula	R	13	43	-12	16
Posterior cingulate cortex	M	31	-19	45	21
Superior temporal gyrus	L	22	-62	-12	1

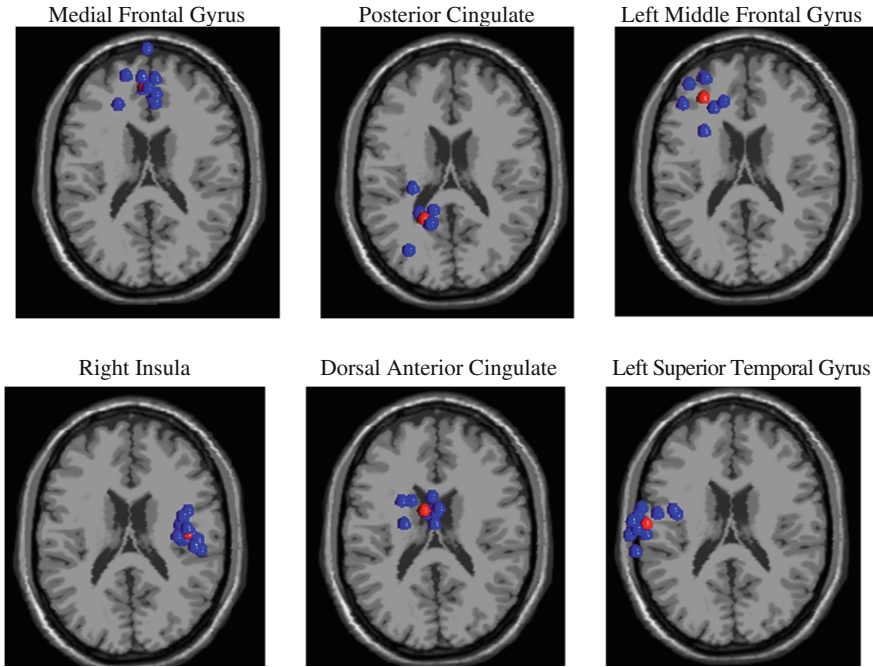


Fig. 4. Cluster locations on a Talarach brain. Blue circles represent independent components, and the red circle is the centroid. (Color figure online)

However, one notable absence was evident: our analyses did not localize a medial temporal lobe (MTL) cluster. On the one hand, this is surprising given that MTL activity prior to thought onset was a primary finding in the previous fMRI study of the same paradigm [11]. MTL regions are consistently recruited by spontaneous thoughts, and are associated with a number of relevant mnemonic processes [1, 8, 43, 44]. On the other hand, the absence of MTL activity makes sense given the relatively poor ability of EEG to detect signals from deeper brain regions. EEG detects the scalp distribution of the aggregate electrical charge of sets of neurons (dipoles). Due to the conductivity of neural tissue, a charge generated by a deeper structure will spatially spread on the way to the surface, and will therefore have weaker detectability (and localizability) at the scalp [45]. As such, due to the location of MTL regions deep below the cortical surface, it would require a very strong and consistent activation for them to be detected by EEG. This kind of activity may not have been feasible with the current paradigm and/or with the number of trials we observed. Future studies should further explore whether EEG can detect MTL structures in relation to spontaneous thought.

4.3 Functional Connectivity over Time

Our second research question was addressed using phase synchrony analyses, which allowed us to observe changes in interregional functional connectivity (i.e., information

sharing) over time. We were specifically interested in contrasting the significant activity that occurred in small windows prior thought to onset. This was done using two key comparisons: (1) all words vs. all thoughts, and (2) verbal thoughts vs. all words.

We focused on connectivity between the IC clusters that exhibited high spatial overlap with regions previously detected with fMRI. As mentioned in Sect. 3.4, we investigated four time-bins of 500 ms leading up to thought onset to give us finer-grained temporal resolution compared to the previous fMRI investigation. Our chi-squared analysis (see Sect. 2) gave us connectivity matrices, which were then projected onto three-dimensional brains (see Figs. 5 and 6 for results projected on an MNI brain). In Figs. 5 and 6, lines are drawn between two regions when the comparison condition (i.e. Fig. 5: all words, Fig. 6: verbal thoughts) significantly differs from the control condition (i.e. Fig. 5: all thoughts, Fig. 6: all words). More specifically, lines indicate there was (1) a significantly greater proportion of time points of significant alpha-band synchrony between those two regions and (2) at least one instance of a full alpha cycle's worth of consecutive significant time-points.

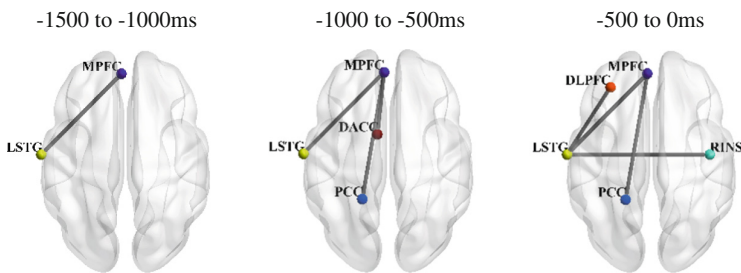


Fig. 5. Greater alpha synchrony for all words relative to all thoughts.

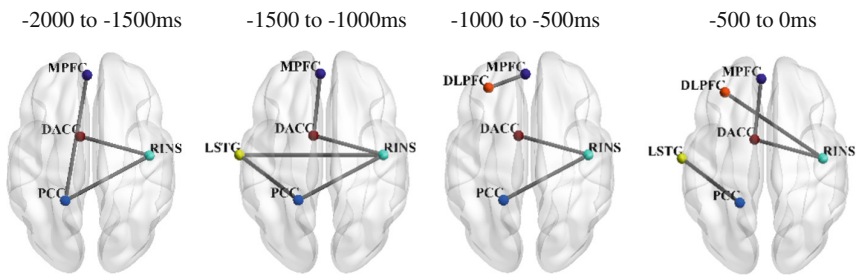


Fig. 6. Greater alpha synchrony for verbal thoughts relative to all words.

Comparing all words vs. all thoughts ultimately served as a check of methodological validity. Based on previous studies of word reading and evaluation [10, 46], we should see consistent LSTG connectivity with DMN midline and executive region involvement. For this comparison (Fig. 5), we used a time window of -1.5 s to the button press. This window was chosen due to the fact that individuals indicated word onset at a mean of 1128 ms following word presentation. Indeed, LSTG connectivity was consistent with our predictions, providing evidence for our methodological approach.

The verbal thoughts vs. all words comparison is the key analysis interest. A time window of -2 s relative to the button press was used for the verbal thoughts vs. all words comparison. We chose this window because there was very sparse significant connectivity prior to this window of time; consistent with participants focusing on their breath during this time. In line with our general hypotheses, we observed a progression of DMN connectivity followed by greater executive region recruitment closer to thought onset. Also, we interestingly found consistent significant connectivity between putative ‘salience network’ [47] network regions.

5 Discussion

5.1 Overview

Following up on an fMRI study that employed the same behavioral paradigm [11], we used EEG to investigate the neural dynamics underlying the generation and attentional appraisal of spontaneous thoughts. Specifically, we characterized the temporal dynamics of connectivity between the associated brain regions in the 2 s leading up to indications of thought onset. Notably, and as hypothesized, ICA and dipole fitting revealed sources of neural electrical activity that were highly convergent with activation peaks found in the original fMRI study (Fig. 4 and Table 1) [11]. This study is the second to use this EEG source localization procedure to identify putative spontaneous thought-related neural activity [26]. With these regions of spatial correspondence recovered, we then conducted functional connectivity analyses. In particular, we examined the temporal dynamics of interregional alpha-band phase synchronization, a frequency band for which oscillations have been associated with attention [38–40] and internally-directed processing [38, 41]. These analyses revealed a general trajectory that begins with DMN midline and salience network connectivity, followed by the subsequent incorporation of language and executive regions (Fig. 5), which was generally in line with our hypotheses.

5.2 Major Findings

We recovered six sources of electrical brain activity that have general spatial correspondence to spontaneous-thought related regions found by the previous fMRI study (Fig. 2), which included DMN regions (MPFC and PCC) [8], salience network regions (DACC and RINS) [47], a language processing region (LSTG) [48], and an executive region (DLPFC) [27]. As mentioned, we analyzed changes in significant connectivity between these regions (in pairs) in 500 ms time-bins, from -2 s to indications of thought onset.

Although we did not recover any MTL activity as hypothesized, we observed connectivity between the MPFC and PCC early in the thought generation process (-2000 to -1500 ms). As mentioned, the PCC is highly structurally and functionally connected with the MTL [29], and was coactivated with the MTL in the original fMRI study [11]. It has also been found to activate in response to a variety of episodic memory-based tasks [44]. This offers the interpretation that the PCC may serve as a

(likely temporally delayed) proxy of MTL activity. Further support for this comes from the recent intracranial EEG finding that MTL-retrosplenial cortex phase locking occurred early during autobiographical memory retrieval [28].

The MPFC-PCC connectivity may represent the initial evaluative processing and affective valuation of the spontaneously arising mental content [8]. The concurrent connectivity between the RINS and DACC may indicate initial bottom-up salience processing of this mental content [49]. Interestingly, the RINS also indicated significant connectivity with the PCC. Although causal directionality cannot be inferred based on the current analyses, it is intriguing to consider whether the RINS could be engaging its putative network switching role in this instance by recruiting the PCC to switch from external to internal processing in response to emerging mental content [49].

The second following time-bin, -1500 to -1000 ms, features LSTG connectivity with the RINS and PCC. We interpret the LSTG-PCC coupling as possibly representing a form of linguistic encoding of thought content, consistent with the LSTG's role in language processing [48]. Underscoring this interpretation, comparisons with other thought types (not presently discussed) revealed that LSTG-PCC coupling only occurred in relation to verbal thoughts. The LSTG-RINS connectivity during this time-bin may be indicative of the RINS signaling the need for LSTG thought encoding, as a result of salience tagging [49].

The time-bin from -1000 to -500 ms uniquely includes DLPFC-MPFC connectivity, which may represent evaluative processes occurring on the emergent mental content; consistent with past work implicating these regions in thought evaluation [10].

The final 500 ms immediately preceding the thought onset-button presses, may pertain to the conscious attentional/emotional appraisal of the thought culminating in the decision to report via button press. The RINS-DLPFC connectivity found here very likely indicates salience network recruitment of the executive network in the initiation of the required behavioral response [27, 49].

5.3 Limitations and Methodological Considerations

There are a number of notable limitations to the present study. Firstly, although the recovered sources had high spatial correspondence to the results of an fMRI studying using the same paradigm, it is impossible to be certain that they represent the same activity. Moreover, since each measure has its basis in a different indicator of brain activity (i.e., electrical vs. hemodynamic), it is unclear whether exact spatial overlap is to be expected or required for the validation of the EEG results. Our use of EEG source localization combined with a measure of the temporal progression of interregional synchrony is, to our knowledge, the first of its kind as applied to spontaneous thought. As such, this study was as much a proof-of-principle methodological evaluation as it was a study of spontaneous thought. The correspondence and interpretability of the results in relation to previous research is suggestive, but further research employing a similar set of analyses is required to validate the approach.

Another limitation corresponds to the use of meditator's self-reports to index the onset of spontaneous thought. Although we specifically only recruited individuals with a moderate to high amount of meditation experience (range: ~ 500 – 2000 h), it was not

possible to evaluate the accuracy with which they were able to report the emergence of a spontaneous thought. Additionally, it was also not possible to objectively evaluate their meditation expertise and validate the hours of experience that they claimed to possess. We were also limited in our analyses due to the relatively low amount of trials. This precluded our ability to characterize differential patterns of activity for each individual thought type, may have also contributed to our inability to detect MTL regions, and limited us to 500 ms time-bins.

6 Concluding Remarks

Our results are generally consistent with those of Ellamil et al. [11] and additionally move beyond them to provide a more fine-grained temporal analysis of the neural dynamics underlying spontaneous thought. The electrophysiological dynamics revealed by EEG appear to share similar sources to the regions found in the original study, but the sequence of connectivity revealed at the millisecond level indicate that there may be important specificities that are overlooked by fMRI. These results underscore the neural heterogeneity of different temporal stages of spontaneous thought, and the need for additional studies in characterizing their dynamics. Our results also suggest that EEG spatial localization combined with phase synchrony analysis may constitute a useful approach to study the neural dynamics of spontaneous thought. Overall, the present study, in addition to Ellamil et al. [11], suggests that the neural correlates pertaining to the temporally distinct processes of thought generation, crystallization, and appraisal can be differentiated, and future research is needed to further fractionate the complex dynamic process of spontaneous thought.

References

1. Christoff, K., Irving, Z.C., Fox, K.C., Spreng, R.N., Andrews-Hanna, J.R.: Mind-wandering as spontaneous thought: a dynamic framework. *Nat. Rev. Neurosci.* **17**, 718–731 (2016)
2. Baird, B., Smallwood, J., Schooler, J.W.: Back to the future: autobiographical planning and the functionality of mind-wandering. *Conscious. Cogn.* **20**, 1604–1611 (2011)
3. Unsworth, N., McMillan, B.D.: Mind wandering and reading comprehension: examining the roles of working memory capacity, interest, motivation, and topic experience. *J. Exp. Psychol. Learn. Mem. Cogn.* **39**, 832 (2013)
4. Mrazek, M.D., Smallwood, J., Franklin, M.S., Chin, J.M., Baird, B., Schooler, J.W.: The role of mind-wandering in measurements of general aptitude. *J. Exp. Psychol. Gen.* **141**, 788 (2012)
5. Smallwood, J., Schooler, J.W.: The science of mind wandering: empirically navigating the stream of consciousness. *Annu. Rev. Psychol.* **66**, 487–518 (2015)
6. Mooneyham, B.W., Schooler, J.W.: The costs and benefits of mind-wandering: a review. *Can. J. Exp. Psychol.* **67**, 11–18 (2013)
7. Fox, K.C.R., Spreng, R.N., Ellamil, M., Andrews-Hanna, J.R., Christoff, K.: The wandering brain: meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *NeuroImage* **111**, 611–621 (2015)

8. Andrews-Hanna, J.R., Smallwood, J., Spreng, R.N.: The default network and self-generated thought: component processes and dynamic control. *Ann. New York Acad. Sci.* **1316**, 29–52 (2014)
9. Vincent, J.L., Kahn, I., Snyder, A.Z., Raichle, M.E., Buckner, R.L.: Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.* **100**, 3328–3342 (2008)
10. Ellamil, M., Dobson, C., Beeman, M., Christoff, K.: Evaluative and generative modes of thought during the creative process. *NeuroImage* **59**, 1783–1794 (2012)
11. Ellamil, M., Fox, K.C., Dixon, M.L., Pritchard, S., Todd, R.M., Thompson, E., Christoff, K.: Dynamics of neural recruitment surrounding the spontaneous arising of thoughts in experienced mindfulness practitioners. *NeuroImage* **136**, 186–196 (2016)
12. Zabelina, D.L., Andrews-Hanna, J.R.: Dynamic network interactions supporting internally-oriented cognition. *Curr. Opin. Neurobiol.* **40**, 86–93 (2016)
13. Dixon, M.L., Andrews-Hanna, J.R., Spreng, R.N., Irving, Z.C., Mills, C., Girn, M., Christoff, K.: Interactions between the default network and dorsal attention network vary across default subsystems, time, and cognitive states. *NeuroImage* **147**, 632–649 (2016)
14. Nisbett, R.E., Wilson, T.D.: Telling more than we can know: verbal reports on mental processes. *Psychol. Rev.* **84**, 231 (1977)
15. Varela, F.J.: Neurophenomenology: a methodological remedy for the hard problem. *J. Conscious. Stud.* **3**, 330–349 (1996)
16. Lutz, A., Thompson, E.: Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *J. Conscious. Stud.* **10**, 31–52 (2003)
17. Lutz, A., Slagter, H.A., Dunne, J.D., Davidson, R.J.: Attention regulation and monitoring in meditation. *Trends Cogn. Sci.* **12**, 163–169 (2008)
18. Tang, Y.-Y., Hölzel, B.K., Posner, M.I.: The neuroscience of mindfulness meditation. *Nat. Rev. Neurosci.* **16**, 213–225 (2015)
19. Nielsen, L., Kaszniak, A.W.: Awareness of subtle emotional feelings: a comparison of long-term meditators and nonmeditators. *Emotion* **6**, 392 (2006)
20. Fox, K.C.R., Zakarauskas, P., Dixon, M.L., Ellamil, M., Thompson, E., Christoff, K.: Meditation experience predicts introspective accuracy. *PLoS ONE* **7**, e45370 (2012)
21. Baird, B., Mrazek, M.D., Phillips, D.T., Schooler, J.W.: Domain-specific enhancement of metacognitive ability following meditation training. *J. Exp. Psychol. Gen.* **143**, 1972 (2014)
22. Kozasa, E.H., Sato, J.R., Lacerda, S.S., Barreiros, M.A., Radvany, J., Russell, T.A., Sanches, L.G., Mello, L.E., Amaro Jr., E.: Meditation training increases brain efficiency in an attention task. *NeuroImage* **59**, 745–749 (2012)
23. Lutz, A., Slagter, H.A., Rawlings, N.B., Francis, A.D., Greischar, L.L., Davidson, R.J.: Mental training enhances attentional stability: neural and behavioral evidence. *J. Neurosci.* **29**, 13418–13427 (2009)
24. MacLean, K.A., Ferrer, E., Aichele, S.R., Bridwell, D.A., Zanesco, A.P., Jacobs, T.L., King, B.G., Rosenberg, E.L., Sahdra, B.K., Shaver, P.R.: Intensive meditation training improves perceptual discrimination and sustained attention. *Psychol. Sci.* **21**, 829–839 (2010)
25. Slagter, H.A., Lutz, A., Greischar, L.L., Francis, A.D., Nieuwenhuis, S., Davis, J.M., Davidson, R.J.: Mental training affects distribution of limited brain resources. *PLoS Biol.* **5**, e138 (2007)
26. Kirschner, A., Kam, J.W.Y., Handy, T.C., Ward, L.M.: Differential synchronization in default and task-specific networks of the human brain. *Front. Hum. Neurosci.* **6**, 139 (2012)
27. Niendam, T.A., Laird, A.R., Ray, K.L., Dean, Y.M., Glahn, D.C., Carter, C.S.: Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* **12**, 241–268 (2012)

28. Foster, B.L., Kaveh, A., Dastjerdi, M., Miller, K.J., Parvizi, J.: Human retrosplenial cortex displays transient theta phase locking with medial temporal cortex prior to activation during autobiographical memory retrieval. *J. Neurosci.* **33**, 10439–10446 (2013)
29. Greicius, M.D., Supekar, K., Menon, V., Dougherty, R.F.: Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* **19**, 72–78 (2008)
30. Foster, B.L., Rangarajan, V., Shirer, W.R., Parvizi, J.: Intrinsic and task-dependent coupling of neuronal population activity in human parietal cortex. *Neuron* **86**, 578–590 (2015)
31. Kiss, G.R., Armstrong, C., Milroy, R., Piper, J.: An associative thesaurus of English and its computer analysis. *Comput. Literary Stud.* 153–165 (1973)
32. Klinger, E., Cox, W.M.: Dimensions of thought flow in everyday life. *Imagination Cogn. Pers.* **7**, 105–128 (1987)
33. Wilson, M., Psycholinguistic, M.R.C.: Database: machine-usable dictionary, version 2.00. *Behav. Res. Methods, Instrum. Comput.* **20**, 6–10 (1988)
34. Sayadaw, M., Maung, T.N.: Fundamentals of Vipassana Meditation (2002)
35. Delorme, A., Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004)
36. Bell, A.J., Sejnowski, T.J.: An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995)
37. Viola, F.C., Thorne, J., Edmonds, B., Schneider, T., Eichele, T., Debener, S.: Semi-automatic identification of independent components representing EEG artifact. *Clin. Neurophysiol.* **120**, 868–877 (2009)
38. Palva, S., Palva, J.M.: Functional roles of alpha-band phase synchronization in local and large-scale cortical networks. *Front. Psychol.* **2**, 204 (2011)
39. Green, J.J., McDonald, J.J.: Electrical neuroimaging reveals timing of attentional control activity in human brain. *PLoS Biol.* **6**, e81 (2008)
40. Doesburg, S.M., Bedo, N., Ward, L.M.: Top-down alpha oscillatory network interactions during visuospatial attention orienting. *NeuroImage* **132**, 512–519 (2016)
41. Cooper, N.R., Burgess, A.P., Croft, R.J., Gruzelier, J.H.: Investigating evoked and induced electroencephalogram activity in task-related alpha power increases during an internally directed attention task. *NeuroReport* **17**, 205–208 (2006)
42. Delorme, A., Palmer, J., Onton, J., Oostenveld, R., Makeig, S.: Independent EEG sources are dipolar. *PLoS ONE* **7**, e30135 (2012)
43. Squire, L.R., Stark, C.E., Clark, R.E.: The medial temporal lobe. *Annu. Rev. Neurosci.* **27**, 279–306 (2004)
44. Spreng, R.N., Mar, R.A., Kim, A.S.: The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* **21**, 489–510 (2009)
45. Luck, S.J.: An Introduction to the Event-Related Potential Technique. MIT Press, Cambridge (2014)
46. Bedo, N., Ribary, U., Ward, L.M.: Fast dynamics of cortical functional and effective connectivity during word reading. *PLoS ONE* **9**, e88940 (2014)
47. Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A. L., Greicius, M.D.: Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007)
48. Jobard, G., Crivello, F., Tzourio-Mazoyer, N.: Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. *NeuroImage* **20**, 693–712 (2003)
49. Menon, V., Uddin, L.Q.: Saliency, switching, attention and control: a network model of insula function. *Brain Struct. Funct.* **214**, 655–667 (2010)