UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

When the brain takes a break: A model-based analysis of mind wandering

Mittner, M.; Boekel, W.; Tucker, A.M.; Turner, B.M.; Heathcote, A.; Forstmann, B.U.

DOI
10.1523/JNEUROSCI.2062-14.2014

Publication date
2014

Document Version
Final published version

Published in
The Journal of Neuroscience

Link to publication

Citation for published version (APA):
Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., & Forstmann, B. U. (2014). When the brain takes a break: A model-based analysis of mind wandering. *The Journal of Neuroscience*, *34*(49), 16286-16295. https://doi.org/10.1523/JNEUROSCI.2062-14.2014

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Download date:16 Oct 2026

Behavioral/Cognitive

When the Brain Takes a Break: A Model-Based Analysis of Mind Wandering

Matthias Mittner,[1] Wouter Boekel,[2] Adrienne M. Tucker,[2] Brandon M. Turner,[4] Andrew Heathcote,[3] and Birte U. Forstmann[2]

[1]Department of Psychology, University of Tromsø, 9037 Tromsø, Norway, [2]Cognitive Science Center Amsterdam, 1018 VZ Amsterdam, The Netherlands, [3]Newcastle Cognition Laboratory, School of Psychology, University of Newcastle, 2308 Newcastle, Australia, and [4]Stanford University Center for Mind, Brain and Computation, Department of Psychology, Stanford University, Stanford, California 94305

Mind wandering is an ubiquitous phenomenon in everyday life. In the cognitive neurosciences, mind wandering has been associated with several distinct neural processes, most notably increased activity in the default mode network (DMN), suppressed activity within the anti-correlated (task-positive) network (ACN), and changes in neuromodulation. By using an integrative multimodal approach combining machine-learning techniques with modeling of latent cognitive processes, we show that mind wandering in humans is characterized by inefficiencies in executive control (task-monitoring) processes. This failure is predicted by a single-trial signature of (co)activations in the DMN, ACN, and neuromodulation, and accompanied by a decreased rate of evidence accumulation and response thresholds in the cognitive model.

Key words: **cognitive modeling; fMRI; mind wandering; model-based neuroscience; MVPA; resting state**

Introduction

Mind wandering, or shifts in attention from outward, stimulus-based processing to inward, introspective cognition, is ubiquitous: recent research estimates the frequency of mind wandering in everyday life to be ~50%, independent of current activity (Killingsworth and Gilbert, 2010). Understanding the cognitive mechanisms involved in mind wandering is critical for understanding and avoiding the potentially disastrous impacts off-task cognition can have, e.g., on performance in driving (He et al., 2011; Yanko and Spalek, 2014) and aviation (Wiegmann et al., 2005). In recent years, researchers have recognized the importance of investigating cognitive mechanisms underlying mind wandering, resulting in a surge of studies probing the phenomenon from different perspectives, most notably within the fields of neuroimaging and cognitive psychology (Schooler et al., 2011; Christoff, 2012; Callard et al., 2013). As a result, several more or less independent strands of research have localized correlates of mind wandering in activity in specific brain regions (Mason et al., 2007; Christoff et al., 2009), periodic activity fluctuations (O'Connell et al., 2009), and behaviors (Cheyne et al., 2009; Bastian and Sackur, 2013).

Activity in the default-mode network (DMN; Raichle et al., 2001), a network of nodes comprising most notably the posterior cingulate cortex (PCC), and the medial prefrontal cortex (mPFC), has been related to self-reported mind wandering (Christoff et al., 2009), behavioral errors (Eichele et al., 2008), and attentional lapses (Weissman et al., 2006). A different set of brain regions comprising predominantly lateral prefrontal and parietal areas are commonly activated during demanding tasks (Fox et al., 2005). Activity in this anti-correlated network (ACN) is negatively correlated with activity in the DMN. It has been shown that internetwork (i.e., between DMN and ACN) dynamic functional connectivity (i.e., time-dependent correlations between brain areas; Hutchison et al., 2013b) is related to vigilant attention (Thompson et al., 2013), arguably a theoretical construct inversely related to mind wandering. In addition, stronger functional connectivity between DMN and ACN has been negatively correlated across subjects with each subjects variability in reaction time (RT; Kelly et al., 2008). In another relevant strand of research, norepinephrine released by the locus ceruleus (LC) in the brainstem has shown to be a key mechanism in sustained attention detailed in the influential adaptive gain theory (Aston-Jones and Cohen, 2005; Nieuwenhuis et al., 2010).

Despite these insights into the neural basis of mind wandering, its impact on cognitive control processes is still an open question (Smallwood and Schooler, 2006; Kane et al., 2007; Watkins, 2008; McVay and Kane, 2009, 2010). In this study, we combine measures of neural activity to predict episodes of mind wandering on a single-trial level. Taking a model-based neuroscience approach (Forstmann et al., 2011) by using predictions derived from the neural data with computational models of cog-

Received May 21, 2014; revised Oct. 6, 2014; accepted Oct. 17, 2014.

Author contributions: M.M., A.M.T., and B.U.F. designed research; W.B. and A.M.T. performed research; M.M., B.M.T., and A.H. analyzed data; M.M., W.B., A.M.T., B.M.T., A.H., and B.U.F. wrote the paper.

M.M. and B.U.F. were supported by a Vidi grant from the Netherlands Organization for Scientific Research and a European Research Council grant. Data were collected under an internal University of Amsterdam grant awarded to A.M.T., A.H. was supported by ARC discovery projects (DP120102907 and DP110100234) and B.M.T. by an NIH award (F32GM103288). We thank E. J. Wagenmakers and Jérôme Sackur for insightful discussions.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr Matthias Mittner, Department of Psychology University of Tromsø, Huginbakken 32, 9037 Tromsø, Norway. E-mail: matthias.mittner@uit.no.

DOI:10.1523/JNEUROSCI.2062-14.2014

Copyright © 2014 the authors 0270-6474/14/3416286-10$15.00/0

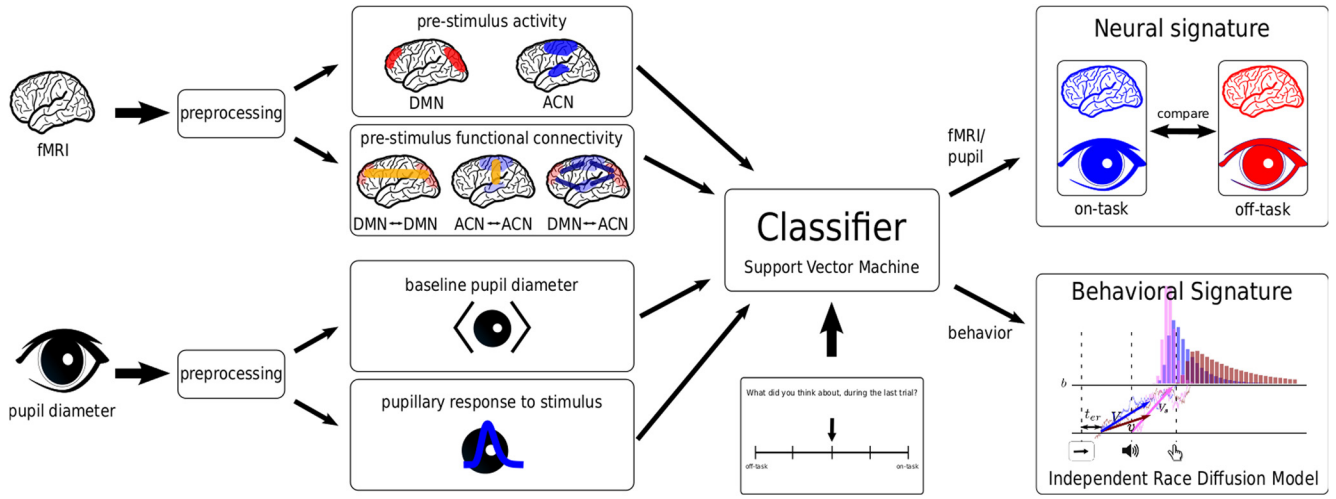

Figure 1. Summary of the analysis procedure. Data from both fMRI and pupil-diameter recordings are preprocessed and theoretically derived features are extracted and fed into a classification algorithm (SVM). Self-reported mind wandering scores are used as labels and the SVM is trained to predict them. After training, all trials are classified and the neural and behavioral signature of off-task behavior is analyzed.

nition (Fig. 1), we can identify and measure a goal-monitoring and a response-inhibition component of executive control and evaluate the impact of mind wandering on these component processes.

Materials and Methods

Overview. An overview over the experiment and data-analysis procedure is given in Figure 1. We recorded fMRI data and pupil dilation from subjects performing the stop-signal task. After preprocessing these data, we extracted features that were indicated by previous research to be affected by mind wandering. These features included prestimulus activity and functional connectivity in and between two large-scale brain networks, the DMN and ACN. We also extracted prestimulus baseline pupil diameter and the pupillary response elicited by each stimulus. Mind wandering was operationalized using introspective thought-probes (Smallwood and Schooler, 2006) and we trained a support-vector machine (Schölkopf and Smola, 2002) to classify each individual trial into either on-or off-task based on subjects' responses to these probes. As a result of this classification, each trial was assigned a probability reflecting the likelihood that the individual was on- or off-task. This classification was instrumental in enabling us to analyze both neural and cognitive processes underlying on- and off-task cognition because we could analyze the data from every single trial rather than being restricted to those trials that were accompanied by a thought-probe. Based on the single-trial labels derived from the classifier, we analyzed the neural and behavioral differences between on- and off-task states using a cognitive process model that is often used for analyzing data from a stop-signal task. The stop-signal task is frequently used to operationalize response-inhibition (Logan and Cowan, 1984; Aron and Poldrack, 2006; Chikazoe et al., 2009; Forstmann et al., 2012; Jahfari et al., 2012) and requires the participant to occasionally withhold responses in a choice task (the go task) when presented with a stop signal. We use a cognitive model of the stop-signal task with racing stop and go processes (Logan et al., 2014) to quantify two executive functions. The first is a goal-monitoring process that balances between making fast go responses and the necessity to stop some responses, which is reflected in model parameters describing the go process. This process comprises both an attentional component, captured by rate parameters in the cognitive model, and a strategic component reflected in changes in the response threshold parameter. The second, response inhibition, is the fast inhibitory process necessary to interrupt a response, which is described by stopping-related parameters. Thus, the stop-signal task enables us to distil from behavior model parameters that index both goal-monitoring and response-inhibition processes that are hallmarks of executive control (Verbruggen and Logan, 2008).

Subjects. The study was approved by the ethics committee of the Psychology Department of the University of Amsterdam. A group of 20 neurologically healthy subjects were recruited (8 male; age range, 22–35 years; mean age, 24.0 years). The data from a subgroup of 16 subjects was acquired in the control condition in the larger setup of a sleep-deprivation study implying that subjects sleep and nicotine/caffeine consumption was controlled before testing. The additional four subjects were acquired to increase confidence in the results and subjects were required to abstain from coffee and nicotine 24 h before testing. Because of a technical error in the eye-tracker, pupil-diameter was only acquired for 19 of 20 subjects. This subject's data were therefore excluded when training the classifier but were included in all other analyses.

Behavioral paradigm. All subjects completed a standard stop-signal paradigm (Lappin and Eriksen, 1966; Logan and Cowan, 1984): in each trial, subjects were presented with a fixation cross (500 ms) followed by an arrow pointing to the left or to the right presented for 400 ms. Subjects were instructed to respond as quickly and accurately as possible with the index finger of their left and right hand corresponding to the orientation of the arrow. In one-third of the trials, an additional auditory stimulus (2.2 kHz) was presented at a variable delay (stop-stimulus delay; SSD) relative to the onset of the visual stimulus. Subjects were instructed to try to withhold their response once the stop-signal was perceived. The SSD was adapted according to a staircase procedure, such that the probability of successfully inhibiting a response approached chance level (Levitt, 1971). The delay between the arrow and the stop signal was varied separately for left and right hand responses. After each stop trial, the probability of responding given all earlier stop-trials was calculated. Probabilities <0.5 resulted in an decrease of the SSD on the next stop-trial, and probabilities >0.5 resulted in a increase of the SSD on the next stop-trial in steps of 50 ms. The SSD value for stop-signal trials was initially set at 190 ms and had a possible range from 40 to 790 ms. In total, 284 go-trials and 142 stop-trials were presented. In addition, there were 24 null trials in which only a fixation-cross was presented. The null trials were included to counteract continuous building up of the BOLD-response over successive trials. There were 40 thought-probes, which consisted of a blank screen with the question "Where was your attention in the previous trial?" and subjects responded on a five-point Likert scale ranging from task-independent to task-centered. Subjects adjusted an arrow pointing to one of the five responses using their left and right index fingers. Thought probes were pseudorandomly intermixed in the trial sequence but always presented after go trials to avoid disruption of mind-wandering episodes by the stop-signal.

MRI image acquisition. Imaging data were acquired on a 3T Philips Achieva scanner using a 32-channel head coil. For each subject, a T1 anatomical scan was acquired [T1 multishot turbo-field echo, 220 trans-

Table 1. Specification of DMN and ACN regions based on a seed in the PCC

Region	Peak coordinates			N_{vox}	Correlation (*r*)
	MNI_x	MNI_y	MNI_z		
DMN					
PCC	−1	−55	26	5790	0.86*
mPFC	−0	60	5	2373	0.45
lIPC	−45	−68	30	1957	0.62
rIPL	48	−66	27	1082	0.55
ACN					
rIPS	59	−35	41	2907	−0.17
lIPS	−60	−36	40	2298	−0.16
SMA	7	−1	55	2150	−0.12
rIL	48	6	−1	2114	−0.17
lIL	−48	1	−2	1822	−0.15
rDLPFC	38	42	15	488	−0.12
lDLPFC	−37	38	23	276	−0.13

*The high correlation for the PCC is partly determined by its role as the seed region.

verse slices of 1 mm, with a resolution of 1 mm^3; repetition time (TR) = 8.2 ms, TE = 3.8 ms]. Functional images were acquired in transverse orientation using a single shot EPI sequence with 37 3 mm slices with an in-plane resolution of 3 × 3 mm (field-of-view = 122 × 240 × 240 mm, TR = 2000 ms, TE = 27.63 ms, flip-angle = 76.1°, voxel size = 3 × 3 × 3 mm, slice gap 0.3 mm).

fMRI analysis: preprocessing. All analyses were conducted using FSL (Jenkinson et al., 2012) and custom Python (Van Rossum and Drake, 2011) scripts combined using the NiPype framework (Gorgolewski et al., 2011). The analyses were performed in individual functional space where not mentioned otherwise. Raw EPI data were slice-time corrected, realigned to correct for effects of motion over time, and spatially smoothed with a 6 mm full-width half-maximum Gaussian kernel. The data were high-pass filtered using a cutoff frequency of 1/50 Hz to correct for baseline drifts in the signal.

Because we wanted to investigate task-unrelated rather than task-related activity, we preprocessed the fMRI data using a standard resting-state analysis approach. We constructed a residual general linear model (GLM) including the following nuisance regressors: motion direction and amplitude (6 variables), mean time courses of white matter, and CSF voxels extracted using FMRIB's Automated Segmentation Tool (FAST; Zhang et al., 2001), a regressor coding for the percentage of time the eyes of the subject were closed during acquisition of each volume and task-regressors (stimulus onset convolved with a standard HRF). The average global signal was not used in the nuisance model because it can result in spurious negative correlations (Murphy et al., 2009). All subsequent analysis where performed on the normalized residuals obtained by subtracting the model obtained via ordinary least-square linear regression from the preprocessed data.

Data-driven ROI specification. To select regions-of-interest (ROIs) within the DMN and the ACN, we used a mask of the PCC (van Maanen et al., 2011) as a seed-region to derive a spatial correlation map. The PCC-mask was transformed into individual space and Pearson-correlation coefficients were calculated between the mean time course in the mask and all other voxels in the brain, yielding one correlation map for each subject. After projection into MNI space, these correlation maps were averaged after applying Fisher's Z-transform. The resulting cross-subject correlation map was thresholded at the 95th percentile to yield the 5% voxels with highest correlation with the PCC seed to extract the nodes of the DMN. Voxels within the PCC that were correlated with the seed region were included. Similarly, the ACN was extracted by thresholding the inverted map at the 95th percentile. The resulting maps were automatically segmented into spatially separated clusters (Table 1). Four prominent nodes of the DMN were extracted including the PCC/precuneus, mPFC, and the bilateral inferior parietal lobules (IPL) all of which have previously been reported as being part of the DMN (Andrews-Hanna et al., 2010). Seven nodes of the ACN (Fox et al., 2005) were extracted comprising bilateral intraparietal sulci (IPS), supplemental motor area (SMA), bilateral dorsolateral prefrontal cortex (DLPFC), and both insula lobules (IL). The group ROIs were projected into individual space and a 3 × 3 × 3 cube centered on the individual peak-correlation voxel within the ROI was used to extract a mean time-series for each ROI and individual. Note that, even though a PCC seed was used to extract a ROI in the same region, the remaining analysis focusing on (co)activations in and between ROIs is not contingent on how the ROIs were defined.

fMRI feature extraction. We derived measures of activity as well as mutual connectivity from all of the extracted nodes that were subsequently used as features to train a classification algorithm. For each ROI, mean activity in the TR before the onset of a stimulus in each trial was extracted yielding four features for the DMN and seven features for the ACN. In addition, we calculated sliding-window correlations (Hutchison et al., 2013a) with a window size of 40 s (this value was chosen based on previous work showing that correlations can be reliably estimated on 30–60 s of data; Shirer et al., 2012) between the mean time courses of each node and every other node, yielding another six features for DMN↔ DMN, 21 features for ACN↔ CAN, and 28 features for DMN↔ ACN correlations.

Pupil diameter data acquisition/analysis. Several previous studies have established a correlation between pupil diameter and mind wandering (Smallwood et al., 2011, 2012; Franklin et al., 2013). The theoretical interpretation that is usually used to explain this effect is based on previously reported correlations between pupil diameter and activity of LC neurons (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005). Note, however that even though several studies have used pupil diameter as a direct measure of LC activity (Gilzenrat et al., 2010; Jepma et al., 2010; Nieuwenhuis et al., 2010; Jepma and Nieuwenhuis, 2011; de Gee et al., 2014), there does not yet exist published empirical evidence for this claim. Interpreting pupil diameter as a proxy for activity of LC neurons is attractive because it allows the application of the adaptive-gain theory: LC neurons are the main source of norepinephrine in the brain and the adaptive gain theory asserts that norepinephrine tunes the neural gain in cortical cell populations that can result in participants being more or less susceptible to competition among different cognitive goals (the exploration–exploitation tradeoff; Jepma and Nieuwenhuis, 2011).

Pupil diameter (PD) was recorded using an Eyelink II system operating at a sampling rate of 1000 Hz. Points in time in which no signal was available (blinks) were removed (along with signal transients 200 ms before and after each blink) and linearly interpolated before analysis. As a measure of baseline PD, the signal in the time window (−1,0) preceding stimulus presentation was averaged. To determine the transient response of the pupil, the PD time series was modeled using a GLM with stimulus onset regressors (de Gee et al., 2014). The onset of each trial was convolved with the response function of the pupil as measured by (Hoeks and Levelt, 1993) and used as regressors in the GLM. The β weights corresponding to each trial were used as a measure for the pupillary response.

Classifier. Classification was performed by a nonlinear support vector machine (SVM; Schölkopf and Smola, 2002) with Gaussian radial basis functions. This nonlinear, supervised classifier finds an optimal decision boundary in feature space that allows a classification of trials into on- and off-task trials as accurately as possible. The 11 prestimulus ROI activities, 55 between-ROI correlations, baseline PD, and pupillary response were normalized and used for training the SVM. As target labels, we used a dichotomized version of the subjects' responses to the thought probes. To account for individual variability in introspective certainty as to what was considered on or off task, we used the per-subject minimum and maximum score as signifying off- and on-task cognition during the preceding trial (the other probes were ignored; median number of usable trials was 22 of 40). The tuning parameters of the SVM (soft-margin parameter C and kernel-width parameter γ) were optimized by grid-search using the area under the receiver-operating characteristic curve (AUC) criterion with a leave-one-out cross-validation approach across subjects. This means that, for all possible combinations, we trained the SVM classifier on all subjects except one and predicted the behavior for the subject whose data were not included in the training of the classifier. The final cross-validation score was averaged over all possible permutations. Importantly, the classifier was therefore trained and evaluated on completely independent datasets.

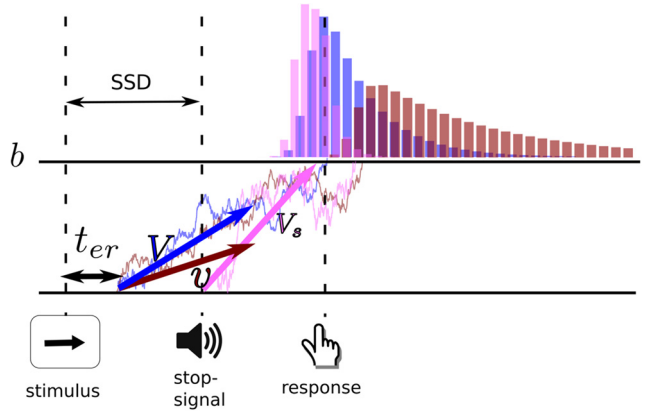

Figure 2. An illustration of the main parameters of the independent race diffusion model: nondecision time t_{er}, boundary b, correct (blue), incorrect (brown), and stop (pink) drift-rate (V, v, V_s). The stop-signal accumulator is delayed by the SSD. The first of the three accumulators that hits the boundary wins the race and the corresponding action is executed.

After obtaining the optimal parameters for the SVM, we calculated noise-perturbation scores as implemented in PyMVPA (Hanke et al., 2009) for each feature. This score is a rough estimate of the relative importance of each feature for the classification performance. The noise-perturbation sensitivity measure was calculated by adding random perturbations individually to each feature and calculating its impact on the cross-validated predictive score. If the classifier is on average sensitive to perturbations to a feature, this feature is regarded as being more important for overall classification performance. In addition, we performed recursive feature-elimination by successively dropping the least informative feature and choosing the feature set that produced optimal classification performance. This was done because dropping noninformative features can significantly improve performance of the classifier. In addition, this procedure enabled us to evaluate whether all the feature groups we extracted from the brain and pupil data were indeed yielding independent information that could help classification. To evaluating the information contained in the labels we performed a random permutation test by generating $N = 20{,}000$ random permutations of the assignment of the labels to the trials and recalculating the performance of the classifier. The result clearly indicated that classification performance on the actual labels was superior to that on random labels ($p < 0.0001$). Finally, we trained the optimal SVM on the complete dataset and derived probabilities for each single trial to be either on or off task.

Analysis of behavioral data. To studying behavioral correlates of mind wandering, we used an independent race diffusion model (Logan et al., 2014), which describes decision-making as a race between independent stochastic accumulators. The distribution of a single accumulator is described by the shifted Wald-distribution parameterized by the time for nondecision processes (including stimulus encoding time, response production time and in the case of the stop accumulator the SSD) t_{er}, drift-rate v, and boundary b (Matzke and Wagenmakers, 2009). We modeled the stop-signal paradigm as a race between three accumulators, one for correct decisions, one for incorrect decisions, and one for stopping the response with, respectively, drift-rates V, v, and V_s (Fig. 2*a*). The response associated with the first accumulator that hits boundary b is executed (correct, error, or response-stop). In addition, each accumulator has a nondecision time parameter t_{er}.

The classifier described in the previous section yields, in addition to a classification of each trial t, a probability of having correctly classified it, i.e., whether the subject was on or off task during this trial ($p_{\text{on},t}$ and $p_{\text{off},t} = 1 - p_{\text{on},t}$). Taking this uncertainty into account, we modeled the likelihood for each trial's data D_t as a mixture of the densities f for on- and off-task state

$$f(D_t|\theta) = p_{\text{on},t} f(D_t|\theta, \text{on} - \text{task}) + p_{\text{off},t} f(D_t|\theta, \text{off} - \text{task}).$$

This approach allows to compensate for the noise created by misclassifications.

To extract parameter estimates at the group level, we modeled the behavioral data across subjects in a hierarchical Bayesian framework. All log-transformed parameters θ on the subject level were modeled as being distributed according to a normal distribution with group-level mean μ_{θ_i} and standard-deviation σ_{θ_i}:

$$\log(\theta_i) \sim \text{Normal}(\mu_{\theta_i}, \sigma_{\theta_i})$$

for $\theta = (\theta_1, \ldots, \theta_n)$ where n is the number of model parameters on the subject level. We assigned mildly informative priors to the group-level parameters, as follows:

$$\mu_{\theta_i} \sim \text{Normal}(1.5, 0.8), \qquad \forall \theta_i \in \{\theta\} \setminus \{t_{er}\}$$

$$\mu_{t_{er}} \sim \text{Normal}(0.75, 0.5)$$

$$\sigma_{\theta_i} \sim \text{Inverse Gamma}(4, 10), \qquad \forall \theta_i,$$

that allowed the parameter estimates to vary across a large number of parameter values while constraining them to be in a plausible range (Gelman and Shalizi, 2013; Gelman et al., 2013).

Eight different models implementing all possible combinations of free parameters between on- and off-task trials were fitted and compared, testing for the most likely parameter configuration. We used the deviance information criterion (DIC; Spiegelhalter et al., 2002) which is a generalization of Akaikes information criterion to hierarchical models for model selection.

For the eight models, we sampled from the posterior distribution of the parameters given the model using a blocked differential evolution Markov-chain Monte-Carlo algorithm with migration step (turned off after half of the burn-in period) described fully by Turner et al. (2013). This nonstandard sampler was necessary because of the high intrinsic correlations between the parameter values of the race model, which is very well handled by the differential-evolution algorithm. We used 24 concurrent chains, a burn-in period of 5000 samples per chain and sampled another 5000 samples resulting in $24 \times 5000 = 120{,}000$ samples per variable. The tuning parameters of the differential evolution algorithm was set to $\gamma = 2.38/(2n)$, where n was the number of parameters and $b = 0.001$. Posterior predictions for the best-fitting model were generated by randomly sampling 12,000 parameter settings from the posterior distributions each of which was used to sample 462 trials (one-third stop-trials). The plots in Figure 6*c* show averages over parameter settings were each estimate was calculated across the 462 trials.

Results

Neural data can reliably predict mind wandering

As shown in Figure 1, we extracted features of interest both from the fMRI and the pupil-dilation recordings and fed them into a SVM classifier. The choice of these features was motivated by previous findings from the literature (Aston-Jones and Cohen, 2005; Christoff et al., 2009; Thompson et al., 2013) and can be categorized in seven distinct groups: prestimulus activity in several ROIs in the DMN and ACN (Table 1), respectively, prestimulus connectivity within and between the networks (DMN↔DMN, ACN↔ACN, DMN↔ACN) and pupil diameter (baseline PD and pupil response). The classifier was trained using each subject's response to randomly interspersed thought probes presented to sample the subjects' current state of attention (on- vs off-task). We trained and evaluated the classifier using a between-subject cross-validation approach, such that the classifier was trained and tested on completely independent datasets (see Materials and Methods).

With the best parameter settings ($C = 6.39$, $\gamma = 0.027$) and an optimal set of 28 features, the SVM achieved across-validation median AUC of 0.75. The cross-validation median accuracy of this classifier was 79.7%, implying that we can expect to correctly

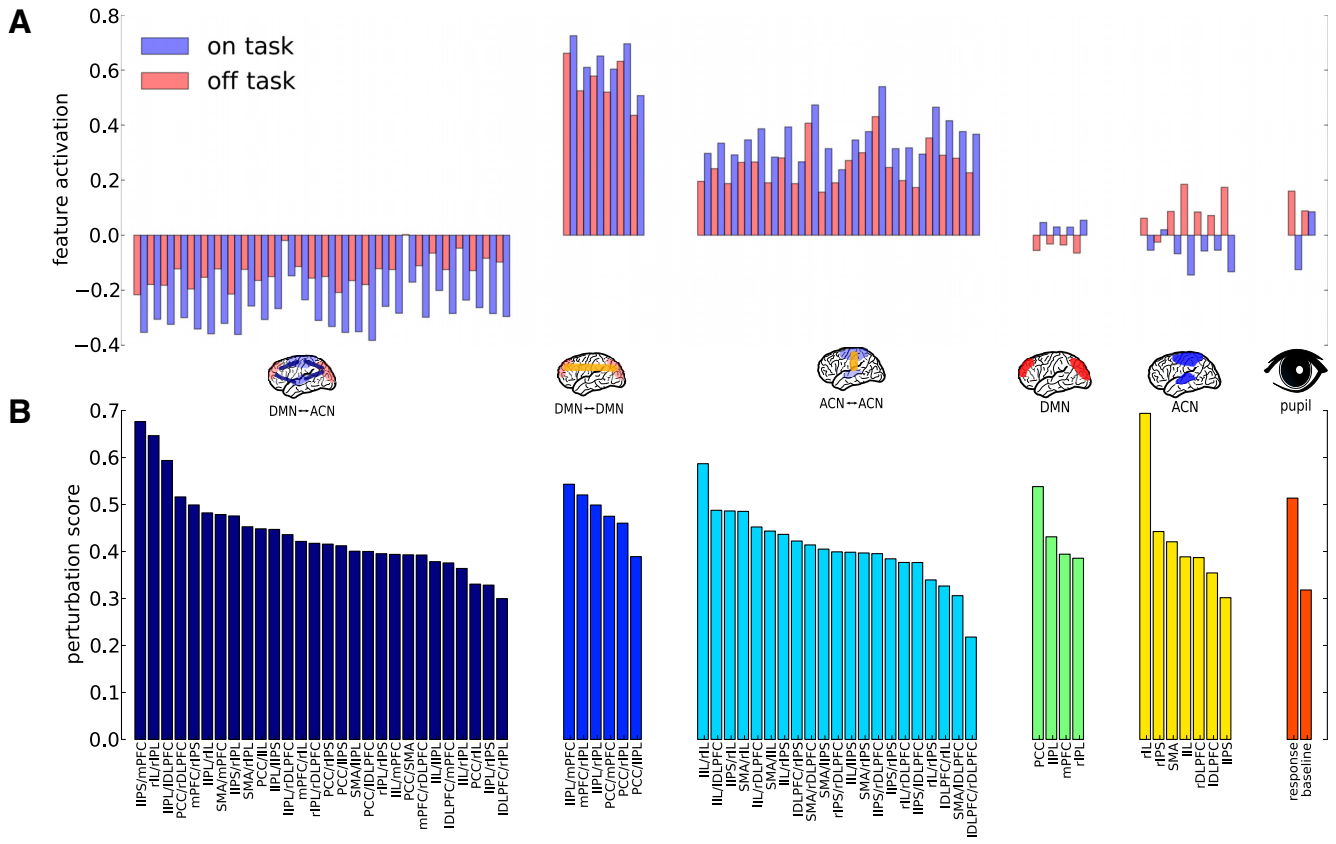

Figure 3. Analysis of feature activations and importance for classification performance. ***A***, Feature activation averaged over subjects and trials classified as on and off task. ***B***, Noise-perturbation score (measure of importance) for each feature. Features are sorted by group (network connectivity, network activity, and pupil) and perturbation score within each group.

classify four of five trials as either on or off task. Results of a random-permutation test confirmed classification performance (see Materials and Methods). The feature-selection procedure showed that features from all groups (DMN/ACN activity, DMN/ACN correlation, pupil) were necessary for optimal predictive performance. This implies that each set of features carried unique information about responses to thought probes.

We also conducted a noise-perturbation analysis that estimates the importance of each particular feature for classification performance (Fig. 3*b*). In each group of features, there were a few ROIs carrying most of the information. Activity in the PCC and lIPL was most important in the DMN, and the most informative ACN node was the left insula. The latter is reflected in the importance of bilateral correlations between left and right insula and cross-network correlations between right insula and rIPL. The mPFC's role is mainly reflected in the importance of between (IPS/mPFC) and within-network correlations (mPFC/IPL). Finally, the pupillary response elicited by the stimulus appears to be more important for classification than the baseline pupil diameter.

The signature of the classifier replicates previous findings

To highlight which properties of the neural data were used by the classifier, we calculated the mean score for each feature in trials classified as on- versus off-task (Fig. 3*a*). The plot shows that this direct contrast carries much of the classification-relevant information across the various ROIs within each group, even though additional, specific information may be contained in linear or nonlinear interactions used by the full classifier. The direct contrasts have the advantage that they allows us average across the features (Fig. 4). The resulting main effects agree with previous work (Mason et al., 2007; Christoff et al., 2009; Stawarczyk et al., 2011). In trials classified as on-task, DMN activity was below baseline but it was above baseline in off-task trials ($t_{(19)} = -4.25$, $p = 0.003$). The opposite pattern was observed for ACN activity ($t_{(19)} = 8.07$, $p < 0.0001$). Absolute synchronicity within and between networks was higher in off-task trials (DMN↔ DMN: $t_{(19)} = -4.80$, $p = 0.00087$, ACN↔ ACN: $t_{(19)} = -7.27$, $p < 0.0001$, DMN↔ ACN: $t_{(19)} = -0.8.88$, $p < 0.0001$). In addition, the pupillary response was reduced in off-task trials ($t_{(19)} = 9.80$, $p < 0.0066$) and so was baseline PD ($t_{(19)} = 3.91$, $p < 0.0001$). Note that p values were Bonferroni-corrected for multiple comparisons across all reported t tests.

Efficiency of goal monitoring is reduced in mind wandering

We used an independent race diffusion model (Logan et al., 2014) to studying the behavioral signature of mind wandering. The model describes decision-making as a race between independent, stochastic accumulators (Fig. 2) striving to reach their respective boundaries. As an addition to the usual setup used in race models of the go task (e.g., the linear ballistic accumulator model; Brown and Heathcote, 2008), the stop-signal task requires an extra accumulator starting at the onset of the stop-signal and representing accumulating evidence for stopping the response. All of the parameters of this model have intuitive interpretations with respect to the underlying cognitive processes: drift rates reflect the efficiency of the process, an increased threshold can be interpreted as reflecting response caution and the nondecision parameter estimates the time used for stimulus encoding and response execution. To account for uncertainty in the predictions

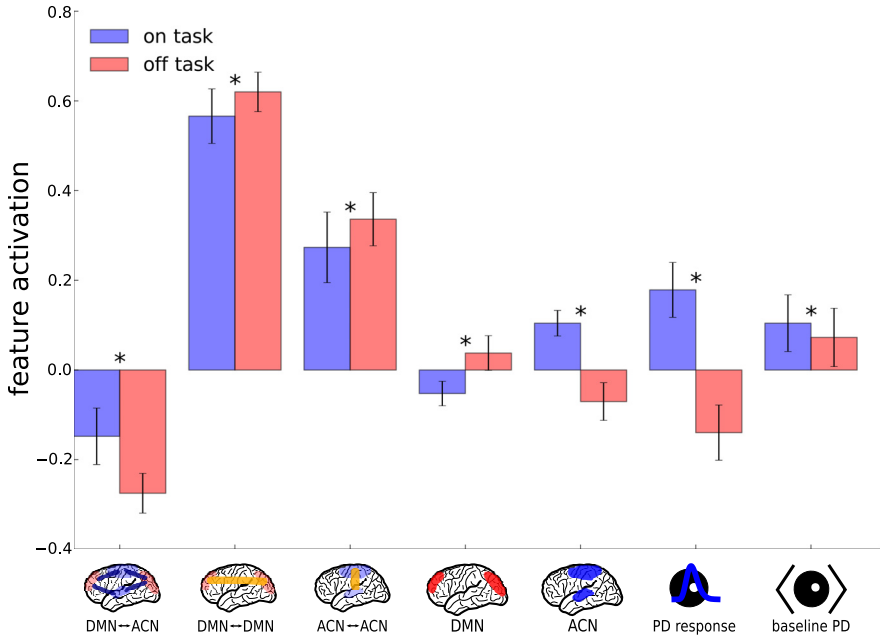

Figure 4. Comparison of feature activations between trials classified as on or off task. Error bars are confidence intervals (Morey, 2008), asterisks indicate significance at the $p < 0.01$ level after adjusting for multiple comparisons.

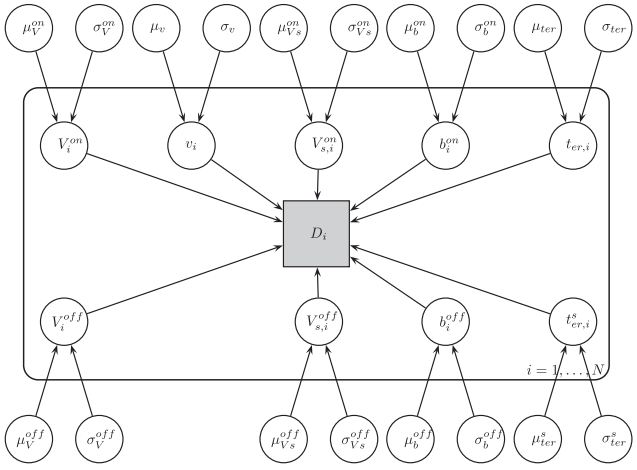

Figure 5. The best-fitting hierarchical Bayesian model: individual parameters V, V_s, and b (indexed by subject i) are free to vary between on- and off-task. Individual data D_i is modeled according to the independent race diffusion model illustrated in Figure 2 and each of the log-transformed, subject-level parameters is normally distributed at the group-level with parameters μ_p, σ_p for each parameter p (see Materials and Methods).

derived from the classifier, we extended the model to account probabilistically for classification errors (see Materials and Methods). To make inference both on the group-and the subject-level, we modeled the data in a hierarchical Bayesian framework (Gelman et al., 2013; Fig. 5).

To determine which parameters, and ultimately cognitive processes, were impacted by mind wandering, we used a Bayesian model-selection approach. All tested models had the following parameters: drift rate of the correct and incorrect response (V and v), as well as drift rate for stopping (V_s); response threshold (b), which was assumed to be the same for all accumulators, and separate nondecision times for go and stop accumulators (t_{er} and t_{ers}, respectively). The models differed with respect to which of these parameters were allowed to vary between trials classified as on or off task. We restricted the model selection procedure to the drift rates V and V_s, as well as the threshold b because of the theoretical assumptions that are reflected in these parameters, while no effect is predicted for the nondecision times. The drift rate for incorrect decisions v was not varied because it had a minor impact due to the low number of error trials and was difficult to estimate on subsets of the data. Among the eight competing models varying the free parameters between on- and off-task conditions, we found that a model allowing both drift rates (V and V_s) and decision threshold b to vary gave the best account of our data in terms of the DIC (Table 2). Differences in DIC larger than 10 can be considered strong (Pratte and Rouder, 2012). In our analysis, the difference from the best model to the next best was 154 DIC units, which clearly indicates that all three parameters are adjusted during mind wandering.

The posterior distributions of the mean group-level parameters are displayed in Figure 6*a*,*b*. We can summarize the main findings in terms of odds-ratios, i.e., how much more likely the effect is compared with no effect. Because 73% of the group-level distribution of the difference between V for trials with and without mind wandering is above zero, the correct drift-rate parameter V is 2.74 [i.e., 0.73/(1−0.73)] times more likely to be increased in on-task trials versus off-task trials across individuals. Similarly, the boundary is 2.73 times more likely to be increased for on-task relative off-task trials (73% mass above zero). Although the stop-drift rate V_s was increased on on-task trials, the effect was weaker, with a factor of only 1.5 times (60% of the posterior mass >zero). We can interpret these effects as mainly reflecting a reduction of the efficiency of monitoring the balance between go and stop task. The reduced drift rate is indicative of reduced attention during mind wandering, whereas the lowered boundary points to a strategic adaptation to compensate this effect. The result of combined reduction drift rate and response threshold was that go errors were more frequent and go responding was more variable (mainly reflected in the tails of the distribution, i.e., an overrepresentation of short and long RTs) during mind wandering, whereas efficiency of the stop-process was not significantly impaired (Fig. 6*c*).

Discussion

In a multimodal classification study, we showed that episodes of mind wandering can be classified with high accuracy on the single-trial level using prestimulus (co)activations of brain structures belonging to the DMN and ACN. An inclusion of pupillary measures further improves predictive performance. Our results indicate that network activity and transient network correlations, as well as baseline PD and the pupillary response, provide unique sources of information regarding the current attentional state of the subject. Furthermore, a model-based analysis of behavior associated with on- and off-task state as identified by the classifier show that both the rate of evidence accumulation of the go process and the response threshold are reduced during mind wandering. Because the efficiency of the inhibitory process necessary for stopping the

response was not impaired, a nonspecific effect was not supported. Rather, our results indicate a specific effect on executive processes controlling goal monitoring, which become less efficient during mind wandering in the stop-signal task. The dissociation between effects on go and stop processes rates is consistent with Logan et al.'s (2014) finding that these processes do not share attentional capacity as operationalized by drift-rate parameters.

By modeling the behavioral data from the stop-signal task, we were able to disentangle the impact of mind wandering on different executive control processes involved in processing the stop-signal task, namely goal monitoring and response inhibition. The first is mainly reflected in adjusting the cognition of the go task and managing the tradeoff between fast going and stopping, the latter is specific to processing the stop signal. Our model-based analysis of the cognitive processes underlying mind wandering in this task shows a subtle but specific pattern of results. Response inhibition is largely unimpaired, with inefficiencies in the go process reflected in lowered drift rates and response thresholds during mind wandering. The combination of lower drift rates for the correct response accumulator and lower thresholds is reflected in more impulsive behavior; in the sense that more errors and more behavioral variability are observed. This finding is in correspondence with previous work which found increased variability during mind wandering (Stawarczyk et al., 2011; Bastian and Sackur, 2013). However, our results give a further indication of which underlying processes are responsible for these findings.

Table 2. Model selection criteria for the eight tested models

Model	Free parameters*			DIC
	V	V_s	b	
1	–	–	–	−8917
2	×	–	–	−9751
3	–	×	–	−9733
4	–	–	×	−9698
5	×	×	–	−9788
6	–	×	×	−9763
7	×	–	×	−9867
8	×	×	×	−10021

*Parameters that were free to vary between on- and off-task trials are marked by an ×; fixed parameters are marked by a –.

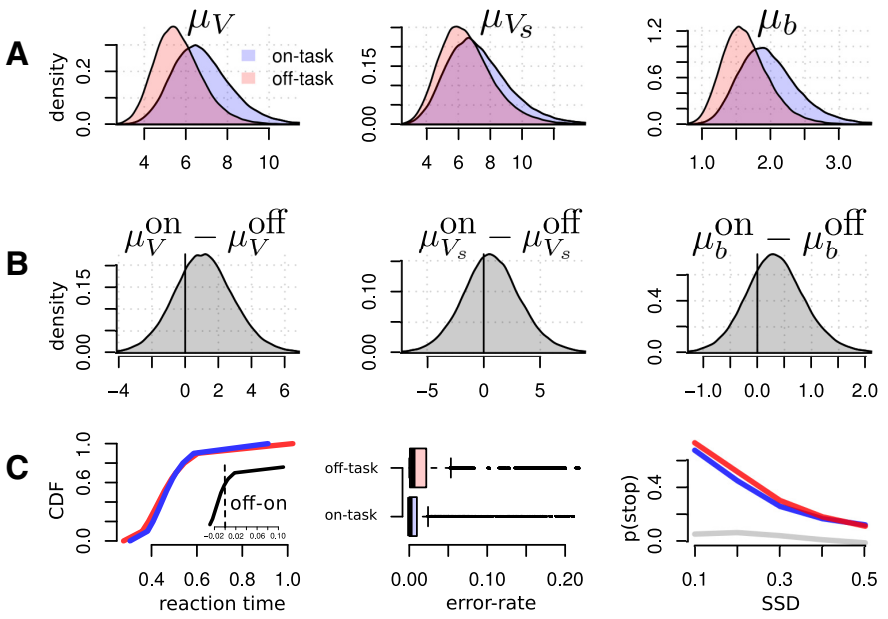

Figure 6. Posterior distribution of the group-level means and posterior predictions. ***A***, Posterior distributions of group-level mean parameter for V, V_s, and b overlaid for on- versus off-task. ***B***, The distribution of differences between on- versus off-task and (***C***) posterior predictions for RT, error-rate and probability of stopping (left to right). Blue signifies on-task and red codes for off-task parameters.

The posterior distributions estimated in our Bayesian analysis were relatively broad implying that there is strong interindividual variation of cognitive and behavioral consequences of mind wandering. This is unsurprising when considering the unspecific nature of the thought-probes as used in this study, which measure a potentially complex phenomenon on a single scale. It is reasonable to assume that our subjects engaged in different cognitive processes during the episodes they classified as "off-task." For example, while some might have been planning their evening activities, more motivated subjects could have used the time to think about how to improve their performance on the current task. Although both kinds of mentation may engage associative brain regions, their impact on behavior is potentially different. To shed more light on this interindividual variability, it might be beneficial to use more specific introspective measures (Stawarczyk et al., 2011) allowing a refined analysis.

In addition to investigating the processes involved in mind wandering on a psychological level, our method also shed light on the neural origins of these processes. Dissecting the trained classifier, we found that activations in the DMN were consistently higher and activations in ACN regions consistently lower in trials that were classified as off task. This finding is in correspondence with earlier work using different paradigms (Weissman et al., 2006; Christoff et al., 2009; Stawarczyk et al., 2011), and has been interpreted in terms of an emergence of internal mentation during mind wandering (Andrews-Hanna, 2012). Extending these results, our study considered also functional connectivity in mind wandering (Kucyi and Davis, 2014). We found that the functional connections within and between the investigated networks were stronger during mind wandering. Even though this might at first sight be counterintuitive (assuming that higher interconnectivity means higher efficiency of information processing), it is consistent with recent theoretical and empirical work (Eldar et al., 2013). Based on simulations of artificial neural networks, as well as empirical data from whole-brain connectivity analyses, these authors showed that an increase in neural-gain as caused by the release of norepinephrine heightens clustering and functional connectivity of brain networks. Because high norepinephrine levels are associated with an exploratory state (Aston-Jones and Cohen, 2005) in which competing internal goals have a higher chance of becoming active (i.e., mind wandering can be initiated), higher functional connectivity should be observed during periods of mind wandering. A prediction from this viewpoint is that these modulations should be observable

across all cortical areas because of the widespread connectivity of the norepinephrine system (Eldar et al., 2013). Our results are in line with this interpretation since within-network correlations were more positive and between-network correlations were more negative during mind wandering. Note that, even though the main effects were in the same direction for all functional couplings and may therefore seem to be redundant, unique information about mind wandering was carried by within-DMN, within-ACN, and between-network correlations. The specific dynamic functional connectivity between individual brain regions does therefore seem to be modulated by more than neural gain.

The theoretical framework by Eldar et al. (2013) coupling functional connectivity and norepinephrine is also in line with our finding that the transient pupillary response evoked by the presentation of the stimulus was reduced during mind wandering, if one accepts the putative link between norepinephric neuromodulation and pupil diameter. The pupillary response, when interpreted as a measure of the phasic LC-bursts initiated by target-processing (Aston-Jones et al., 1994), is strongest during periods of optimal task-performance (Aston-Jones and Cohen, 2005). The relationship between these phasic bursts and tonic levels of LC activity is of the Yerkes-Dodson type: both low and high tonic activity suppresses the transient LC response, which is strongest at intermediate tonic LC levels. Therefore, the increase in functional connectivity and the reduction of the pupillary response are well explained by this theory. However, interpreting the baseline PD as reflecting tonic LC levels, we are faced with the puzzling finding that baseline PD was lower in off- rather than on-task trials in our experiment. From the adaptive-gain theory, we would have expected that baseline PD would be increased rather than reduced. Even though usually a negative correlation between baseline PD and pupillary response is observed, there are exceptions to this rule (de Gee et al., 2014, SI). Further research targeting this specific effect is necessary to investigate this seemingly contradictory pattern of results.

The vast majority of previous studies investigating mind wandering have used a go/no-go task, the sustained attention to response task (SART). Because of its highly repetitive and undemanding nature, this task has proven well suited to provoking episodes of mind wandering. In addition, there is good evidence that particular behavioral patterns can be observed in this task when a person is not concentrating: increased error rates, anticipatory responses, response omissions (Cheyne et al., 2009), RT variability, and an overrepresentation of long RTs (Bastian and Sackur, 2013). Despite the obvious benefits of using such a simple task, more complex tasks are required to investigate how mind wandering impacts higher cognitive processes. Converging results from many studies indicate that mind wandering occupies a substantial amount of time whenever humans engage in cognitive processing both in the laboratory (Smallwood and Schooler, 2006) and the real world (Killingsworth and Gilbert, 2010). Indeed it has been stated that "[…] every laboratory study is at least partially a study of mind wandering" (Smallwood and Schooler, 2006). The main reason why research so far has been mainly restricted to the SART (Reichle et al., 2010; Uzzaman and Joordens, 2011; with the exception of studies investigating mind wandering during reading, Schad et al., 2012) is that it is problematic to detect periods of mind wandering without that detection affecting behavior. Because most research relies on introspective thought-probes, a large number of those needs be interspersed throughout the experiment to increase statistical power. The single-trial based approach we presented in the current study of the stop-signal task overcomes this problem and is extensible to other experimental paradigms. By relating neural measures to occasionally interspersed thought-probes using classification algorithms, all trials in the experiment can be assigned a probability of belonging to an episode of mind wandering. In combination with suitable analysis methods, this offers a powerful new approach for future studies in this research field.

References

Andrews-Hanna JR (2012) The brain's default network and its adaptive role in internal mentation. Neuroscientist 18:251–270. CrossRef Medline

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic fractionation of the brain's default network. Neuron 65:550–562. CrossRef Medline

Aron AR, Poldrack RA (2006) Cortical and subcortical contributions to stop signal response inhibition: role of the subthalamic nucleus. J Neurosci 26:2424–2433. CrossRef Medline

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450. CrossRef Medline

Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. J Neurosci 14:4467–4480. Medline

Bastian M, Sackur J (2013) Mind-wandering at the fingertips: automatic parsing of subjective states based on response time variability. Front Psychol 4:573. CrossRef Medline

Brown SD, Heathcote A (2008) The simplest complete model of choice response time: linear ballistic accumulation. Cognit Psychol 57:153–178. CrossRef Medline

Callard F, Smallwood J, Golchert J, Margulies DS (2013) The era of the wandering mind? Twenty-first century research on self-generated mental activity. Front Psychol 4:891. CrossRef Medline

Cheyne JA, Solman GJ, Carriere JSA, Smilek D (2009) Anatomy of an error: a bidirectional state model of task engagement/disengagement and attention-related errors. Cognition 111:98–113. CrossRef Medline

Chikazoe J, Jimura K, Hirose S, Yamashita K, Miyashita Y, Konishi S (2009) Preparation to inhibit a response complements response inhibition during performance of a stop-signal task. J Neurosci 29:15870–15877. CrossRef Medline

Christoff K (2012) Undirected thought: neural determinants and correlates. Brain Res 1428:51–59. CrossRef Medline

Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW (2009) Experience sampling during fmri reveals default network and executive system contributions to mind wandering. Proc Natl Acad Sci U S A 106:8719–8724. CrossRef Medline

de Gee JW, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. Proc Natl Acad Sci U S A 111:E618–625. CrossRef Medline

Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, von Cramon DY, Ullsperger M (2008) Prediction of human errors by maladaptive changes in event-related brain networks. Proc Natl Acad Sci U S A 105:6173–6178. CrossRef Medline

Eldar E, Cohen JD, Niv Y (2013) The effects of neural gain on attention and learning. Nat Neurosci 16:1146–1153. CrossRef Medline

Forstmann BU, Wagenmakers EJ, Eichele T, Brown S, Serences JT (2011) Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? Trends Cogn Sci 15:272–279. CrossRef Medline

Forstmann BU, Keuken MC, Jahfari S, Bazin PL, Neumann J, Schäfer A, Anwander A, Turner R (2012) Cortico-subthalamic white matter tract strength predicts interindividual efficacy in stopping a motor response. Neuroimage 60:370–375. CrossRef Medline

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A 102:9673–9678. CrossRef Medline

Franklin MS, Broadway JM, Mrazek MD, Smallwood J, Schooler JW (2013) Window to the wandering mind: pupillometry of spontaneous thought while reading. Q J Exp Psychol 66:2289–2294. CrossRef Medline

Gelman A, Shalizi CR (2013) Philosophy and the practice of bayesian statistics. Br J Math Stat Psychol 66:8–38. CrossRef Medline

Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D (2013) Bayesian Data Analysis, Third Edition. Boca Raton, FL: Chapman and Hall/CRC.

Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD (2010) Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cogn Affect Behav Neurosci 10:252–269. CrossRef Medline

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011) Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 5:13. CrossRef Medline

Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby JV, Pollmann S (2009) Pymvpa: a python toolbox for multivariate pattern analysis of fMRI data. Neuroinformatics 7:37–53. CrossRef Medline

He J, Becic E, Lee YC, McCarley JS (2011) Mind wandering behind the wheel: performance and oculomotor correlates. Hum Factors 53:13–21. CrossRef Medline

Hoeks B, Levelt WJ (1993) Pupillary dilation as a measure of attention: a quantitative system analysis. Behav Res Meth Instr Comput 25:16–26. CrossRef

Hutchison RM, Womelsdorf T, Gati JS, Everling S, Menon RS (2013a) Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques. Hum Brain Mapp 34:2154–2177. CrossRef Medline

Hutchison RM, Womelsdorf T, Allen EA, Bandettini PA, Calhoun VD, Corbetta M, Della Penna S, Duyn JH, Glover GH, Gonzalez-Castillo J, Handwerker DA, Keilholz S, Kiviniemi V, Leopold DA, de Pasquale F, Sporns O, Walter M, Chang C (2013b) Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80:360–378. CrossRef Medline

Jahfari S, Verbruggen F, Frank MJ, Waldorp LJ, Colzato L, Ridderinkhof KR, Forstmann BU (2012) How preparation changes the need for top-down control of the basal ganglia when inhibiting premature actions. J Neurosci 32:10870–10878. CrossRef Medline

Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM (2012) Fsl. Neuroimage 62:782–790. CrossRef Medline

Jepma M, Nieuwenhuis S (2011) Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. J Cogn Neurosci 23:1587–1596. CrossRef Medline

Jepma M, Te Beek ET, Wagenmakers EJ, van Gerven JM, Nieuwenhuis S (2010) The role of the noradrenergic system in the exploration-exploitation trade-off: a psychopharmacological study. Front Hum Neurosci 4:170. CrossRef Medline

Kane MJ, Brown LH, McVay JC, Silvia PJ, Myin-Germeys I, Kwapil TR (2007) For whom the mind wanders, and when: an experience-sampling study of working memory and executive control in daily life. Psychol Sci 18:614–621. CrossRef Medline

Kelly AM, Uddin LQ, Biswal BB, Castellanos FX, Milham MP (2008) Competition between functional brain networks mediates behavioral variability. Neuroimage 39:527–537. CrossRef Medline

Killingsworth MA, Gilbert DT (2010) A wandering mind is an unhappy mind. Science 330:932. CrossRef Medline

Kucyi A, Davis KD (2014) Dynamic functional connectivity of the default mode network tracks daydreaming. Neuroimage 100:471–480. CrossRef Medline

Lappin JS, Eriksen CW (1966) Use of a delayed signal to stop a visual reaction-time response. J Exp Psychol 72:805. CrossRef

Levitt H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467. CrossRef Medline

Logan GD, Cowan WB (1984) On the ability to inhibit thought and action: a theory of an act of control. Psychol Rev 91:295. CrossRef

Logan GD, Van Zandt T, Verbruggen F, Wagenmakers EJ (2014) On the ability to inhibit thought and action: general and special theories of an act of control. Psychol Rev 121:66–95. CrossRef Medline

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN (2007) Wandering minds: the default network and stimulus-independent thought. Science 315:393–395. CrossRef Medline

Matzke D, Wagenmakers EJ (2009) Psychological interpretation of the ex-gaussian and shifted wald parameters: a diffusion model analysis. Psychon Bull Rev 16:798–817. CrossRef Medline

McVay JC, Kane MJ (2009) Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. J Exp Psychol Learn Mem Cogn 35:196–204. CrossRef Medline

McVay JC, Kane MJ (2010) Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). Psychol Bull 136:188–197; discussion 198–207. CrossRef Medline

Morey RD (2008) Confidence intervals from normalized data: a correction to Cousineau (2005). Tutor Quant Methods Psychol 4:61–64.

Murphy K, Birn RM, Handwerker DA, Jones TB, Bandettini PA (2009) The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? Neuroimage 44:893–905. CrossRef Medline

Nieuwenhuis S, De Geus EJ, Aston-Jones G (2010) The anatomical and functional relationship between the p3 and autonomic components of the orienting response. Psychophysiology 48:162–175. CrossRef Medline

O'Connell RG, Dockree PM, Robertson IH, Bellgrove MA, Foxe JJ, Kelly SP (2009) Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. J Neurosci 29:8604–8611. CrossRef Medline

Pratte MS, Rouder JN (2012) Assessing the dissociability of recollection and familiarity in recognition memory. J Exp Psychol Learn Mem Cogn 38:15911607. CrossRef Medline

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proc Natl Acad Sci U S A 98:676–682. CrossRef Medline

Rajkowski J, Kubiak P, Aston-Jones G (1993) Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. Soc Neurosci Abstr 18:538.

Reichle ED, Reineberg AE, Schooler JW (2010) Eye movements during mindless reading. Psychol Sci 21:1300–1310. CrossRef Medline

Schad DJ, Nuthmann A, Engbert R (2012) Your mind wanders weakly, your mind wanders deeply: objective measures reveal mindless reading at different levels. Cognition 125:179–194. CrossRef Medline

Schölkopf B, Smola A (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond adaptive computation and machine learning. Cambridge, MA: MIT.

Schooler JW, Smallwood J, Christoff K, Handy TC, Reichle ED, Sayette MA (2011) Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn Sci 15:319–326. CrossRef Medline

Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012) Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22:158–165. CrossRef Medline

Smallwood J, Schooler JW (2006) The restless mind. Psychol Bull 132:946–958. CrossRef Medline

Smallwood J, Brown KS, Tipper C, Giesbrecht B, Franklin MS, Mrazek MD, Carlson JM, Schooler JW (2011) Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. PLoS One 6:e18298. CrossRef Medline

Smallwood J, Brown KS, Baird B, Mrazek MD, Franklin MS, Schooler JW (2012) Insulation for daydreams: a role for tonic norepinephrine in the facilitation of internally guided thought. PLoS One 7:e33706. CrossRef Medline

Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A (2002) Bayesian measures of model complexity and fit. J Roy Stat Soc B Met 64:583–639. CrossRef

Stawarczyk D, Majerus S, Maquet P, D'Argembeau A (2011) Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. PLoS One 6:e16997. CrossRef Medline

Thompson GJ, Magnuson ME, Merritt MD, Schwarb H, Pan WJ, McKinley A, Tripp LD, Schumacher EH, Keilholz SD (2013) Short-time windows of correlation between large-scale functional brain networks predict vigilance intraindividually and interindividually. Hum Brain Mapp 34:3280–3298. CrossRef Medline

Turner BM, Sederberg PB, Brown SD, Steyvers M (2013) A method for efficiently sampling from distributions with correlated dimensions. Psychol Methods 18:368–384. CrossRef Medline

Uzzaman S, Joordens S (2011) The eyes know what you are thinking: eye

movements as an objective measure of mind wandering. Conscious Cogn 20:1882–1886. CrossRef Medline

van Maanen L, Brown SD, Eichele T, Wagenmakers EJ, Ho T, Serences J, Forstmann BU (2011) Neural correlates of trial-to-trial fluctuations in response caution. J Neurosci 31:17488–17495. CrossRef Medline

Van Rossum G, Drake F (2011) The python language reference manual. Bristol, UK: Network Theory.

Verbruggen F, Logan GD (2008) Response inhibition in the stop-signal paradigm. Trends Cogn Sci 12:418–424. CrossRef Medline

Watkins ER (2008) Constructive and unconstructive repetitive thought. Psychol Bull 134:163–206. CrossRef Medline

Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. Nat Neurosci 9:971–978. CrossRef Medline

Wiegmann D, Faaborg T, Boquet A, Detwiler C, Holcomb K, Shappell S (2005) Human error and general aviation accidents: a comprehensive, fine-grained analysis using HFACS. Office of Aerospace Medicine. Washington, DC: DOT/FAA/AM-05/24.

Yanko MR, Spalek TM (2014) Driving with the wandering mind the effect that mind-wandering has on driving performance. Hum Factors 56:260–269. CrossRef Medline

Zhang Y, Brady M, Smith S (2001) Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 20:45–57. CrossRef Medline