NeuroImage 114 (2015) 249–256

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Full Length Articles

Predicting moment-to-moment attentional state

Monica D. Rosenberg [a,*], Emily S. Finn [b], R. Todd Constable [b,c,d], Marvin M. Chun [a,b,e]

[a] Department of Psychology, Yale University, USA
[b] Interdepartmental Neuroscience Program, Yale University, USA
[c] Department of Diagnostic Radiology, Yale University School of Medicine, USA
[d] Department of Neurosurgery, Yale University School of Medicine, USA
[e] Department of Neurobiology, Yale University School of Medicine, USA

ARTICLE INFO

Article history:
Received 29 October 2014
Accepted 14 March 2015
Available online 20 March 2015

Keywords:
Sustained attention
Attentional fluctuations
Attentional states
fMRI
Multi-voxel pattern analysis

ABSTRACT

Although fluctuations in sustained attention are ubiquitous, most psychological experiments treat them as noise, averaging performance over many trials. The current study uses multi-voxel pattern analysis (MVPA) to decode whether, on each trial of a cognitive task, participants are in an optimal or suboptimal attentional state. During fMRI, participants performed *n*-back tasks, composed of central face images overlaid on distractor scenes, with low, perceptual, and working memory load. Instructions were to respond to novel faces and withhold response to rare repeats. To index attentional state, reaction time variability was calculated at each correct response. Participants' 50% least variable trials were labeled optimal, or "in the zone," and their 50% most erratic trials were labeled suboptimal, or "out of the zone." Support vector machine classifiers trained on activity in the default mode network (DMN), dorsal attention network (DAN), and task-relevant fusiform face area (FFA) distinguished in-the-zone and out-of-the-zone trials in all tasks. Consistent with evidence that distractors are processed when central task load is low, parahippocampal place area (PPA) classifiers were only successful in the low load task. Classification in anatomical regions across the brain revealed widespread coding of attentional state. In contrast to these robust pattern analyses, univariate signal in DMN, DAN, FFA, and PPA did not distinguish states, suggesting a nuanced relationship to sustained attention. In sum, MVPA can be used to decode trial-by-trial attentional state throughout much of cortex, helping to characterize how attention network fluctuations correlate with performance variability.

Introduction

Maintaining attention to task is nearly always critical for successful performance (Chun et al., 2011), but our best efforts often fail to prevent mind wandering or distraction. Despite the ubiquity of attention lapses—which can lead to performance errors (Cheyne et al., 2006; Robertson et al., 1997) and even catastrophic accidents (Hudock and Duchon, 1988; Edkins and Pollock, 1997)—they frequently go undetected by individuals lacking meta-awareness (Schooler et al., 2011) and experiments averaging performance across many trials.

Attempting to characterize intrinsic attention fluctuations, Esterman et al. (2013, 2014) defined distinct states of attention based on behavioral response variability: an optimal "in-the-zone" state marked by consistent responding, and an error-prone "out-of-the-zone" state marked by erratic responding. These states mapped onto brain activity in somewhat surprising ways: Being in the zone was associated with increased default mode network (DMN) activity, typically implicated in mind-wandering (Christoff et al., 2009) and task-unrelated thought (Buckner et al., 2008; Weissman et al., 2006). In contrast, out-of-the-zone performance relied on dorsal attention network (DAN) activity, thought to subserve externally focused attention (Corbetta and Shulman, 2002) and associated with decreased distractibility (Leber, 2010) and error rates (Padilla et al., 2006). Findings linking DMN activity to better and DAN activity to worse performance are not without precedent, however: DMN activity has been associated with practice (Mason et al., 2007) and better target detection, and DAN activity with worse target detection (Sadaghiani et al., 2009). Thus, although attention fluctuates between optimal and suboptimal states characterized by distinct neural activity, the precise roles of attention networks remain unclear.

High-level visual areas are also likely impacted by fluctuating attention, and are thus good candidate regions from which to decode attentional state. For example, the parahippocampal place area (PPA) processes distractor scenes only when the perceptual load of a central task is low (Yi et al., 2004), but this effect is modulated by attentional state, such that PPA processes distractor scenes during in-the-zone, but not out-of-the-zone, performance (Esterman et al., 2014).

Here, we use multi-voxel pattern analysis (MVPA) of fMRI data to predict whether participants are in the zone or zoning out. Participants performed low load (1-back), working memory load (2-back), and perceptual load (degraded 1-back) tasks with central face and distractor

* Corresponding author at: Department of Psychology, Yale University, 2 Hillhouse Avenue, New Haven, CT 06511, USA.
E-mail address: monica.rosenberg@yale.edu (M.D. Rosenberg).

http://dx.doi.org/10.1016/j.neuroimage.2015.03.032
1053-8119/

scene stimuli. We hypothesized that DMN and DAN activity would predict attentional state in all tasks. Consistent with perceptual load theory (Lavie, 2005; Yi et al., 2004), we hypothesized that activity in the fusiform face area (FFA; selective to central faces) would distinguish state in all tasks, while patterns in PPA (selective to distractor scenes) would distinguish state in the low and working memory load tasks only. Moment-to-moment attentional state classification has broad applications, from monitoring performance in psychological studies to preventing real-world failures of attention and vigilance.

Materials and methods

Participants

Twenty-two participants (ten females, ages 21–33 years, mean age = 25.3 years) were recruited from Yale University and the surrounding community. All participants gave written informed consent in compliance with procedures approved by the Yale University Human Subjects Committee and were paid for their participation. Participants were right handed and had normal or corrected-to-normal vision.

Paradigm and stimuli

Participants performed three continuous *n*-back tasks composed of grayscale face photographs centrally overlaid on grayscale scene photographs during fMRI (see Fig. 1). Faces were cropped to show the eyes, nose, and mouth and were sized to 132 × 132 pixels; scenes were 440 × 400 pixels. A border of width 5 pixels surrounded the faces. On a back-projected display that the subject viewed with a mirror mounted on the head coil of the MRI system, faces subtended approximately 3° × 3° and background scenes subtended approximately 10° × 10° of visual angle.

On each trial, a face–scene composite appeared on the screen for 1 s followed by a 1-s mask (a phase-scrambled face overlaid on a phase-scrambled scene). A fast event-related design with a predictable inter-trial interval was employed to maximally tax sustained attention and avoid beneficial effects of jitter on performance (Wodka et al., 2009; Ryan et al., 2010; Lee et al., 2012). Similar non-jittered designs have been used in previous studies of attention fluctuations (Smith et al., 2006; Suskauer et al., 2008; Chikazoe et al., 2009; Solanto et al., 2009; Christoff et al., 2009; Esterman et al., 2013, 2014). A univariate analysis comparing activity evoked by error vs. correct trials further supported the validity of this design (see Supplementary Material). Participants were instructed to attend to faces and ignore background scenes.

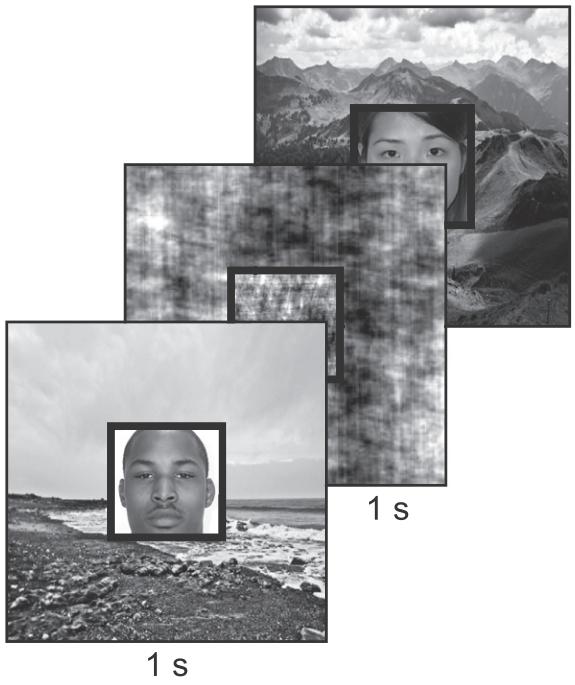

Fig. 1. Trials consisted of face–scene composite images presented for 1 s followed by 1 s masks. Participants were instructed to attend to faces while ignoring background scenes, and to respond to non-repeated faces and withhold response to repeats (1-back for the low load and perceptual load tasks; 2-back for the working memory load task).

Task runs consisted of 252 trials divided evenly into three blocks. During 1-back, or low load, task blocks, participants were instructed to respond via button press to every face that was different than the previous (non-targets; ~90.5%), and to withhold response to repeated faces (targets; ~9.5%). Response accuracy was emphasized without reference to speed. In perceptual load blocks, faces were degraded by adding 20% salt-and-pepper noise and instructions remained the same. During working memory load blocks, faces were not degraded and participants were instructed to respond when faces were different than the face presented two trials back. Non-target faces were only shown once per run, and task type was indicated by the color of a border around faces such that a blue or orange border indicated that the subjects were to perform the 1-back or 2-back task, respectively. Across participants, color mappings were counterbalanced, and task order was pseudorandomized using a Latin square design.

A face/scene region of interest (ROI) localizer in which scenes and faces alternated every minute was also administered. Participants were instructed to indicate via button press whether a face was male or female and whether a scene was indoor or outdoor.

Procedure

Before scanning, participants practiced each *n*-back task for 1 min. In the MRI scanner, an anatomical magnetization prepared rapid gradient echo (MPRAGE) volume scan was acquired, followed by a 6-min resting-state (blood oxygenation level dependent) BOLD fMRI scan and three 8.4-min runs of continuous *n*-back tasks during BOLD imaging. Following task runs, another 6-min resting scan and a 6-min face/scene localizer scan were collected. Due to excessive motion (defined a priori as >2 mm translation or >3° rotation over the course of a run) or sleepiness, one task run from each of four participants was excluded from analysis, and one rest run was excluded from each of two.

Imaging parameters

FMRI data acquisition was performed on a 3T Siemens Trio TIM system equipped with a 32-channel head coil at the Yale Magnetic Resonance Research Center. Functional runs included 504 (task) or 363 (rest and localizer) whole-brain volumes acquired using a multiband echo-planar imaging sequence with the following parameters: repetition time (TR) = 1000 ms, echo time (TE) = 30 ms, flip angle = 62°, acquisition matrix = 84 × 84, in-plane resolution = 2.5 mm^2, 51 axial-oblique slices parallel to the ac–pc line, slice thickness = 2.5, multiband 3, acceleration factor = 2. MPRAGE parameters were as follows: TR = 2530 ms, TE = 3.32, flip angle = 7°, acquisition matrix = 256 × 256, in-plane resolution = 1.0 mm^2, slice thickness = 1.0 mm, 176 sagittal slices.

Behavioral analysis

For each task and each subject, sensitivity (d′) was calculated as a measure of overall performance, and RT coefficient of variation (standard deviation divided by mean correct trial RT) was calculated as a measure of intraindividual response variability (IIV). IIV has been linked to performance on attention and executive control tasks in healthy adult populations (Bellgrove et al., 2004; Kelly et al., 2008;

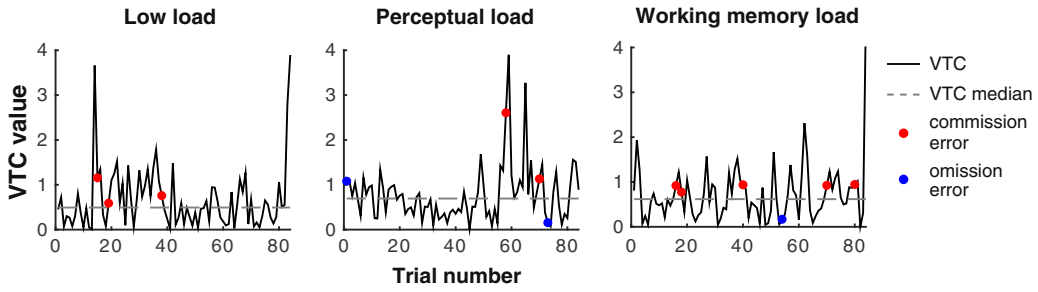

Fig. 2. Shown are example VTCs from one block of each task. Trials with VTC values below the median are considered in-the-zone trials, and trials with VTC values above the median are out-of-the-zone trials.

Prado et al., 2011; Rosenberg et al., 2013; Prado and Weissman, 2011), as well as in populations with attention deficit hyperactivity disorder (ADHD; for review see Castellanos et al., 2006), schizophrenia (Schwartz et al., 1989; Kaiser et al., 2008), dementia (Hultsch et al., 2000), and traumatic brain injury and frontal lobe lesions (Stuss et al., 1989, 1994, 2003). Current theories suggest that IIV reflects the efficiency with which cognitive resources are deployed and is more sensitive to changes in cognitive function than mean performance (MacDonald et al., 2006, 2009; Kelly et al., 2008).

A continuous measure of RT variability—a variance time course (VTC; Esterman et al., 2013; Esterman et al., 2014; Rosenberg et al., 2013)—was also computed for each subject and task (see Fig. 2 for example VTCs). To create VTCs, each correct commission (correct response to a non-target trial) was assigned a value corresponding to the absolute z-score of its RT.[1] To control for potential drifts in RT over time, RT mean and standard deviation were calculated using correct commission RTs from the current task block only. Values for correct omissions (correct response inhibitions to target trials), omission errors (failures to press to non-target trials), and commission errors (incorrect presses to target trials) were interpolated linearly from the two surrounding correct commission RTs. "In-the-zone" trials were those below the VTC median (the 50% most stable trials) while "out-of-the-zone" trials were those above (the 50% most erratic trials).

FMRI analysis

General methods

FMRI data were analyzed using Analysis of Functional NeuroImages (AFNI; Cox and Hyde, 1997) and custom Matlab scripts (Mathworks). Preprocessing included skull-stripping, de-obliquing, motion correction with a 6-parameter, rigid body, least-squares alignment procedure, spatial smoothing to a 6-mm FWHM Gaussian kernel, automated coregistration and normalization of anatomical and functional volumes to MNI space, and scaling of functional dataset values to percent signal change. Data for classification analyses were preprocessed identically but no spatial smoothing was applied. Individual participant data were analyzed with a stepwise regression procedure described in more detail in RT Variability in Materials and methods. To obtain group-level statistical maps, regression coefficients were combined across participants and evaluated with voxel-wise *t*-tests. These statistical maps were thresholded at $p < 0.05$ cluster corrected (individual-voxel intensity threshold of $p < 0.01$, cluster size of 70 contiguous voxels), where minimum voxel extent was determined using Monte Carlo simulations estimating the probability of noise-only clusters after accounting for the smoothness of the data (Forman et al., 1995).

[1] Although comparing RTs to the median rather than the mean is often preferred due to the rightward skew of RTs in many cognitive tasks, here we used mean removal to most closely replicate previous work finding relationships between VTCs, behavior, and neural activity (Rosenberg et al., 2013; Esterman et al., 2013, 2014). The choice made little practical difference: VTCs calculated with mean and median removal were highly correlated, $r = 0.98$, $p < 0.001$.

ROI definition

Face/scene localizer data were used to define participant-specific PPA and FFA. To define right and left PPA, spherical ROIs with a 4-mm radius were centered on activation peaks within the parahippocampal gyrus from the thresholded scene > face contrast (see Table 1). The right and the left FFA were defined in an identical manner on peaks from the thresholded face > scene contrast (Table 1).

To define DMN and DAN, each participant's preprocessed resting-state data were concatenated across runs and submitted to independent component analysis using FSL's MELODIC software (www.fmrib.ox.ac.uk/fsl/melodic/html). Components best representing the DMN and the DAN were selected based on visual inspection (e.g., Esterman et al., 2013). To define the four most robust DMN regions (posterior cingulate cortex, ventromedial prefrontal cortex and bilateral lateral parietal cortex), the peak 200 contiguous voxels were extracted in each ROI (Esterman et al., 2013). An identical procedure was employed to extract four DAN regions: bilateral inferior parietal sulci and frontal eye fields (Table 1).

RT variability

To isolate the effect of RT variability—a continuous measure of attentional state—on BOLD signal, a stagewise regression procedure (Esterman et al., 2013, 2014) that controlled for the effects of trial type and RT was employed. First, for each participant for each task, a general linear model (GLM) accounted for signal associated with the evoked response for correct omissions, omission errors and commission errors, and included nuisance regressors for signal mean, linear drift, six realignment parameters, and mean signal from ROIs centered in each subject's deep white matter and lateral ventricle cerebrospinal fluid.

A second-stage GLM, performed on the residuals of the first-stage model, modeled effects of RT by implementing amplitude-modulated regression using the RT of each correct commission convolved with a 1-parameter gamma variate hemodynamic response function. The residuals of this regression were submitted to the analyses described in the Pattern classification and Univariate ROI analysis sections below.

To find voxels whose activity fluctuated with RT variability independent of raw RT, a third-stage GLM implemented amplitude-modulated

Table 1
Mean peak coordinates of participant-specific regions of interest. Coordinates (*x*, *y*, *z*) are in MNI stereotaxic space.

Region	Left hemisphere coordinate			Right hemisphere coordinate		
	x	*y*	*z*	*x*	*y*	*z*
Posterior cingulate cortex				0	−58	24
Ventromedial prefrontal cortex	−7	54	7			
Lateral parietal cortex	−46	−71	25	45	−66	25
Frontal eye field	−32	1	49	32	0	50
Inferior parietal sulcus	−25	66	48	26	63	48
Fusiform face area (FFA)	−40	−48	−17	41	−49	−18
Parahippocampal place area (PPA)	−24	−42	−8	25	−42	−9

regression on residuals of the second-stage model using VTC values (see Behavioral analysis in Materials and methods).

Pattern classification

MVPA was performed to predict in- and out-of-the-zone periods of performance. For each participant, linear support vector machines (SVMs; http://www.csie.ntu.edu.tw/~cjlin/libsvm) were trained to discriminate voxel-wise patterns of activity associated with in- and out-of-the-zone correct commissions. For each task, a linear support vector machine was trained on activation, averaged 4–6 after the onset of each correct commission in order to account for hemodynamic delay, in each of four independently defined ROIs (FFA, PPA, DMN, and DAN). Training data were the residuals of the second-stage GLM (see RT Variability in Materials and methods), which accounted for error-evoked and RT-specific activity, scaled to range from zero to one. Test data were scaled based on scaling parameters from the training data. Classifiers were tested with a 10-fold cross-validation procedure that was repeated 1000 times to confirm that results were not due to binning idiosyncrasies. For each round of cross-validation, data were divided randomly into ten bins. A classifier was trained on data from nine of the bins and tested on the excluded data. This was repeated until each bin had been left out once, and accuracy was averaged across these ten iterations. Reported classifier accuracy is the average of this procedure repeated 1000 times.

Chance-level performance was determined for each individual participant and task by shuffling condition labels and running the classification procedure 1000 times ($M = 50.16\%$; $SD = 0.59\%$). Paired t-tests were conducted to determine if classification accuracy differed from chance.

In addition to the a priori ROIs, classifiers were trained in each of the 48 bilateral cortical ROIs in the Harvard–Oxford anatomical atlas to determine if patterns of activity elsewhere in the brain distinguished attentional state. Procedures in this analysis were identical to those described previously, except, due to computational constraints, the 10-fold cross-validation procedure was performed 10 times (as opposed to 1000 times in a priori ROIs) and chance accuracy was assumed to be 50% in all regions.

Univariate ROI analysis

To determine whether overall activity, as opposed to multivariate patterns, distinguished attentional state, mean activity was submitted to a 4 (ROI: FFA, PPA, DMN, DAN) × 3 (task: low load, perceptual load, working memory load) × 2 (attentional state: in and out of the zone) repeated measures ANOVA. Mean activity was calculated by averaging the residuals of the second-stage regression (see RT variability in Materials and methods) 4–6 s after the onset of every correct commission. Residuals—rather than percent signal change—were used in this analysis in order to look for activity differences due to attentional state above and beyond those evoked by rare targets and errors.

Results

Behavioral analyses

Sensitivity

Sensitivity (d′) differed by task, $F_2 = 35.58$, $p < 0.001$. As predicted, participants showed greater sensitivity in the low load task ($M = 2.62$, $SD = 0.849$) than in the perceptual load task ($M = 2.16$, $SD = 0.692$), $t_{21} = 3.43$, $p = 0.0025$, and in the working memory load task ($M = 1.61$, $SD = 0.689$), $t_{21} = 7.92$, $p < 0.001$. Sensitivity was significantly greater in the perceptual load than in the working memory load task, $t_{21} = 5.74$, $p < 0.001$.

RT and intraindividual RT variability

Mean RT was 561 ms for the low load task, 585 ms for the perceptual load task, and 597 ms for the working memory load task (*SDs* = 81, 74, and 89 ms, respectively). RT differed as a function of task, $F_2 = 4.80$, $p = 0.013$, such that RTs were longer in the perceptual load, $t_{21} = 2.71$, $p = 0.013$, and working memory tasks, $t_{21} = 2.44$, $p = 0.023$, than in the low load task. Mean RT variability (coefficient of variation) did not differ across tasks, $F_2 = 0.123$, $p = 0.885$ (low load task $M = 0.269$, $SD = 0.130$; perceptual load task $M = 0.279$, $SD = 0.176$; working memory load task $M = 0.278$, $SD = 0.126$).

Because trial-to-trial RT variability was used to assign trials to in-the-zone and out-of-the-zone states, we confirmed that more variable responses were related to worse performance. Across subjects, RT variability was negatively correlated with d′ in the low load (Pearson's $r = -0.85$, $p < 0.001$), perceptual load ($r = -0.66$, $p < 0.001$) and working memory load ($r = -0.65$, $p = 0.001$) tasks. In contrast, mean RT did not relate to d′ in the low load (Pearson's $r = -0.12$, $p = 0.6$) or working memory load ($r = -0.05$, $p = 0.8$) tasks. In the perceptual load task, there was a trend such that subjects with faster RTs had higher d′ values ($r = -0.41$, $p = 0.063$). Thus subjects who responded more erratically—not just more slowly—performed worse in all three *n*-back tasks.

RT variability, but not raw RT, was also related to error rates within subjects. To examine the effects of RT variability on performance, we compared accuracy between subjects' 50% least variable trials (those closest to the mean; "stable") and 50% most variable trials (those furthest from the mean, whether fast or slow; "erratic"). This split was defined using VTCs smoothed with a Gaussian smoothing kernel of seven trials full width at half maximum, incorporating information from the surrounding 15 trials (Rosenberg et al., 2013; Esterman et al., 2013, 2014). Subjects showed lower error rates during periods of stable compared to erratic responding in the perceptual load task (in-the-zone vs. out-of-the-zone error rate: *Ms* = 8.4% vs. 10.4%, *SDs* = 9.8% vs. 10.4%; $t_{21} = 2.19$, $p = 0.040$) and in the working memory load task (*Ms* = 12.6% vs. 14.6%, *SDs* = 14.7% vs. 14.5%; $t_{21} = 2.38$, $p = 0.027$). During the low load task, error rates were numerically but not significantly lower during stable than erratic periods (*Ms* = 7.3% vs. 8.5%, *SDs* = 10.1% vs. 9.4%; $t_{21} = 1.59$, $p = 0.13$). Further, although d′ values between stable and erratic periods did not significantly differ, consistent with predictions, d′ was numerically higher during stable than erratic periods in the low load task (*Ms* = 2.74 vs. 2.48, *SDs* = 1.25 vs. 0.81; $t_{21} = 1.14$, $p = 0.27$), perceptual load task (*Ms* = 2.29 vs. 2.03, *SDs* = 0.88 vs. 0.74; $t_{21} = 1.47$, $p = 0.16$), and working memory load task (*Ms* = 1.74 vs. 1.50, *SDs* = 0.79 vs. 0.75; $t_{21} = 1.69$, $p = 0.11$).

To explore potential effects of raw RT on performance, we also compared accuracy between subjects' 50% fastest and 50% slowest trials defined using vectors of correct-trial raw RTs (smoothed identically to the VTCs described above). For each task block, error rates for each half of the data were calculated as the percentage of omission and commission errors occurring in that half. We did not observe any differences in d′ between the fastest and slowest 50% of trials (low load: *Ms* = 2.47 vs. 2.68, *SDs* = 0.89 vs. 1.07; perceptual load *Ms* = 2.07 vs. 2.19, *SDs* = 0.81 vs. 0.69; working memory load *Ms* = 1.73 vs. 1.45, *SDs* = 1.01 vs. 0.90; all $|t_{21}| < 1.61$, *p* values > 0.12), or in error rates between the fastest and slowest trials (low load: *Ms* = 7.8% vs. 8.0%, *SDs* = 9.5% vs. 10.1%; perceptual load *Ms* = 8.6% vs. 10.1%, *SDs* = 8.5% vs. 11.8%; working memory load *Ms* = 13.2% vs. 13.4%, *SDs* = 14.5% vs. 13.5%; all $|t_{21}| < 1.37$, *p* values > 0.18). Although these null results should be interpreted with caution, it is interesting to note that error rates and d′ also did not differ between the fastest and slowest 25% of trials (all $|t_{21}| < 1.60$, *p* values > 0.12), challenging notions that, in the current tasks, fast responding represents increased attention to task while slow responses represent attention lapses. Thus, at least in the perceptual and working memory load tasks, participants made more errors when they responded more erratically but not more slowly (or quickly). These results justify using absolute RT variability over raw RT as an index of attention, in which the fastest and slowest trials are grouped together and treated as the out-of-the-zone state (Rosenberg et al., 2013; Esterman et al., 2013, 2014).

FMRI analyses

Whole brain correlates of attention fluctuations

To identify regions whose activity fluctuates along with sustained attention, whole-brain multiple regression was performed for each subject using the VTC as a regressor (see RT Variability in Materials and methods). In the low load task, regions of the DAN were positively correlated with variability, including bilateral intraparietal sulci, frontal eye fields, dorsolateral prefrontal cortex, and presupplementary motor area. In the perceptual load task, bilateral frontal eye fields and left lateral parietal cortex positively correlated with RT variability. Finally, in the working memory load task, frontal eye fields, inferior parietal lobules, and medial frontal gyrus fluctuated with variability (Fig. 3; Table 2). Thus, in all three tasks, regions traditionally associated with the on-task state (e.g., Fox et al., 2005) correlated with erratic, out-of-the-zone performance.

One possible explanation for the positive correlation between VTCs and DAN activity is that periods of high variability are driven by exaggerated pre-error speeding and post-error slowing, requiring attention to engage more strongly after an error (Chun and Wolfe, 1996; Gehring and Knight, 2000; Botvinick et al., 2004). Several analyses, however, suggest that high variability here is not fully explained by error-related changes in RT. First, perhaps because our task was so easy, commission errors were relatively rare: Participants made 2.9 per task block on average. Second, there was no evidence for post-error slowing in any of the current tasks: RTs directly following commission errors were not different from RTs of correct commissions that did not immediately precede or follow an error (p values > 0.12). Further post-hoc analyses are available in the supplement.

Table 2
Peak coordinates of regions significantly correlated with RT variability across subjects. Coordinates (x, y, z) are in MNI stereotaxic space.

	Left hemisphere coordinate				Right hemisphere coordinate			
	x	y	z	Cluster size	x	y	z	Cluster size
Low load task					58	22	−7	15,724
	−10	−66	−57	104	10	−68	−57	581
	−40	−83	−20	476				
	−55	10	46	348				
	−60	−58	−17	239	58	−63	−22	196
					30	−56	−17	175
					0	−86	5	157
					10	5	11	126
	−5	−21	−7	116				
	−43	−46	−52	83				
	−43	−66	−52	73				
Perceptual load task					0	20	38	396
	−58	20	−5	234	60	20	−5	302
	−8	−73	63	302				
					65	−56	33	219
					18	15	73	194
					38	35	43	106
					48	15	43	92
					45	7	26	80
Working memory load task	−70	−28	21	499	70	−33	31	257
	−60	15	6	176	58	20	−5	200
					3	7	56	107

Distinguishing attentional state: MVPA

In DMN, DAN, and FFA, classifiers predicted whether correct commission trials were in the zone or out of the zone with above-chance accuracy in all three tasks. In the PPA, classifiers achieved above-chance accuracy in the low load task, but not in the perceptual or working memory load tasks (see Fig. 4).

Within the 48 bilateral anatomical ROIs of the Harvard–Oxford atlas, classification analyses reached significance (exceeding a Bonferroni-corrected p-value of 1.04×10^{-3}) in 27 regions in the low load task (Fig. 5), one region in the perceptual load task, and two regions in the working memory load task (Table 3).

Distinguishing attentional state: univariate analyses

To determine whether univariate activity differed as a function of attentional state, residuals of the second-stage regression (see RT variability in Materials and methods) were submitted to a 4 (ROI: FFA, PPA, DMN, DAN) × 3 (task: low load, perceptual load, working memory load) × 2 (attentional state: in and out of the zone) repeated measures ANOVA. There were no significant main effects of ROI, $F_{(3,63)} = 1.28$, $p = 0.29$; task, $F_{(2,42)} = 0.09$, $p = 0.91$; or Zone, $F_{(1,21)} = 2.04$, $p = 0.17$ on overall activity; and no two- or three-way interactions (p-values > 0.18).

Although we did not observe a main effect of attentional state on ROI activity, post-hoc t-tests, prompted by the finding that DAN activity correlates with RT variability, revealed higher DAN activity during out-of-the-zone than in-the-zone epochs in the low load task ($t_{21} = 2.58$, $p = 0.0174$), but not in the perceptual ($t_{21} = 1.23$, $p = 0.23$) or working memory load tasks ($t_{21} = 0.70$, $p = 0.49$). The effect of attentional state on DAN activity did not survive Bonferroni correction for multiple comparisons (p value = 0.0166), but does suggest that, in accordance with past (Esterman et al., 2014) and current univariate results, increased DAN activity may underlie out-of-the-zone performance in certain tasks.

Together these analyses demonstrate that overall activity in stimulus-specific ROIs and attention networks of interest is not sufficient to consistently distinguish attentional state. Rather, there are complex relationships, reflected in patterns of activity, between brain activity and attentional states.

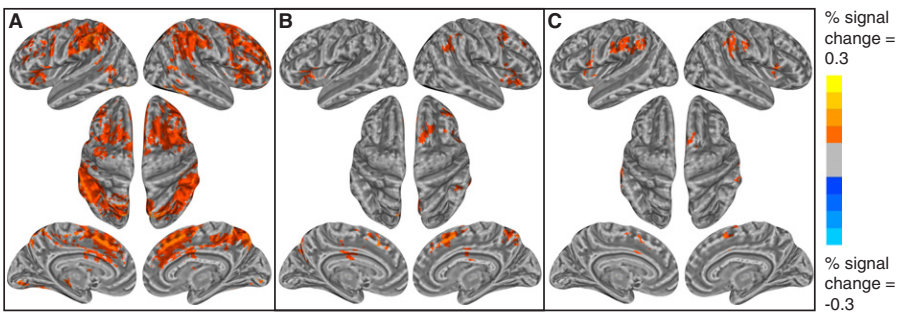

Fig. 3. Regions correlated with RT variability in (A) the low load task, (B) the perceptual load task, and (C) the working memory load task. Each inflated brain shows volume-based group level results projected onto the cortical surface. Maps have a corrected significance level of $\alpha = 0.05$ (individual-voxel intensity threshold of $p < 0.01$; cluster size of 70); colors represent regressor coefficients (here, equivalent to percent signal change because functional datasets were scaled to percent signal change during preprocessing). Orange shades represent positive correlations with RT variability and blue shades represent negative correlations with variability.

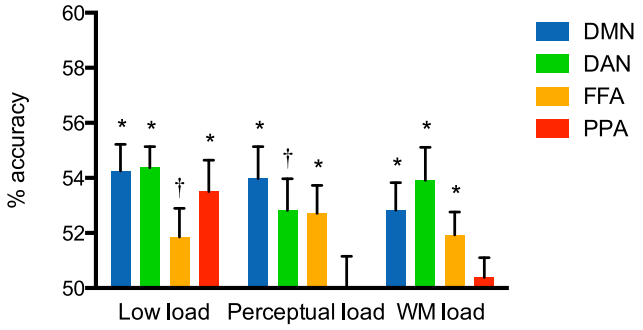

Fig. 4. Percent classification accuracy is averaged across subjects. In each ROI, chance level performance was determined for individual participants and tasks by shuffling condition labels and running the classification procedure 1000 times. Degrees of freedom for all t-tests = 21. * $p < 0.05$ (Bonferroni-corrected for 4 regions); † $p < 0.05$ (uncorrected). Error bars represent standard error.

Discussion

The present study used patterns of fMRI activity in large-scale brain networks and high-level visual areas to predict participants' attentional state at each trial of three *n*-back tasks. Classifiers trained in regions of the DMN and DAN, where activity has been extensively linked to the focus of attention (e.g., Buckner et al., 2008; Spreng et al., 2013), successfully predicted whether, at each trial, a participant was in the zone or out of the zone in all tasks. Classifiers trained in FFA, where activity reflects processing of the task-relevant face stimuli, also distinguished in-the-zone and out-of-the-zone performance in all tasks. However, classifiers trained in PPA, where activity corresponds to task-irrelevant distractor processing, were only successful in the easiest low load task. Of note, a univariate approach confirmed that overall activity in these regions did not consistently distinguish attentional states; although there was a trend such that DAN activity was higher during out-of-the-zone than in-the-zone epochs in the low load task, patterns of DAN activity reliably distinguished state in all tasks. Patterns of activity that discriminate attentional state may be widespread across the cortex, given that classification was successful in the majority of anatomical ROIs in the low load task. Together these results confirm that MVPA of BOLD signal can be used to predict fluctuating attentional state during task performance.

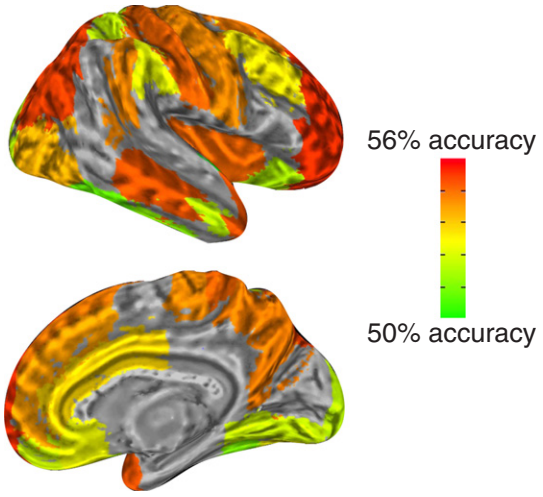

Fig. 5. Harvard–Oxford regions in which classification during the low load task exceeded chance ($p < 0.05$; Bonferroni-corrected for 48 regions) are colored according to average accuracy across subjects. Because ROIs are bilateral, only right hemisphere results are shown.

The current results support recent characterizations of the DMN and DAN as playing complex roles in attentional performance (e.g., Spreng et al., 2013; Prado and Weissman, 2011). Specifically, they challenge strict notions of the DMN as task-negative and the DAN as task-positive, and

Table 3
In-the-zone vs. out-of-the-zone classification accuracy in cortical ROIs from the Harvard–Oxford Atlas. Percent classification accuracy is averaged across subjects in each of the 48 Harvard–Oxford ROIs. Chance was considered 50% in all regions. * $p < 0.05$ (Bonferroni-corrected for 48 regions); † $p < 0.01$ (uncorrected).

	Low load task		Perceptual load task		Working memory load task	
ROI	Mean (SE)		Mean (SE)		Mean (SE)	
Angular gyrus	54.14 (1.12)	†	53.34 (1.23)		52.98 (1.13)	
Central opercular cortex	54.21 (1.23)	†	53.12 (1.14)		53.35 (1.33)	
Cingulate gyrus anterior	54.74 (1.15)	*	53.18 (1.30)		53.20 (1.25)	
Cingulate gyrus posterior	54.05 (1.07)	†	52.70 (1.28)		53.37 (0.93)	†
Cuneal cortex	53.19 (1.12)	†	52.32 (1.12)		51.24 (0.91)	
Frontal medial cortex	54.16 (0.86)	*	52.74 (1.17)		53.16 (1.15)	
Frontal operculum cortex	53.63 (1.04)	†	52.54 (0.96)		52.37 (0.88)	
Frontal orbital cortex	53.97 (0.94)	*	53.89 (1.32)	†	53.36 (1.36)	
Frontal pole	55.75 (0.96)	*	54.49 (1.41)	†	53.80 (1.43)	
Heschl's gyrus	52.15 (0.62)	†	53.09 (1.06)	†	51.27 (0.93)	
Inferior frontal gyrus pars opercularis	53.14 (0.97)	†	52.72 (1.13)		52.82 (1.37)	
Inferior frontal gyrus pars triangularis	53.74 (1.00)	†	53.04 (1.11)		52.06 (1.01)	
Inferior temporal gyrus anterior	51.39 (0.59)		52.08 (1.11)		51.41 (1.07)	
Inferior temporal gyrus posterior	53.72 (0.91)	*	52.60 (0.94)		54.22 (1.33)	†
Inferior temporal gyrus temporooccipital	53.47 (0.83)	*	52.45 (0.99)		54.21 (0.99)	*
Insular cortex	55.43 (1.04)	*	53.40 (1.09)	†	54.13 (1.23)	†
Intracalcarine cortex	52.68 (1.16)		52.43 (0.94)		51.61 (0.90)	
Juxtapositional lobule cortex	52.74 (1.12)		54.63 (1.14)	*	53.09 (1.09)	
Lateral occipital cortex inferior	54.95 (0.86)	*	53.27 (1.47)		52.96 (1.32)	
Lateral occipital cortex superior	55.63 (1.17)	*	53.57 (1.33)		53.74 (1.38)	
Lingual gyrus	54.15 (1.02)	*	53.59 (1.29)		53.31 (1.25)	
Middle frontal gyrus	54.56 (1.01)	*	52.54 (1.50)		54.61 (1.63)	
Middle temporal gyrus anterior	54.02 (0.88)	*	52.40 (0.98)		52.09 (0.87)	
Middle temporal gyrus posterior	55.52 (0.96)	*	53.15 (1.29)		53.72 (1.05)	†
Middle temporal gyrus temporooccipital	54.29 (1.42)	†	52.29 (1.11)		53.75 (1.25)	†
Occipital fusiform gyrus	54.46 (1.17)	*	53.54 (1.12)	†	53.17 (0.90)	†
Occipital pole	54.03 (0.89)	*	54.37 (1.27)	†	54.03 (1.32)	†
Paracingulate gyrus	55.11 (1.17)	*	53.69 (1.22)	†	53.76 (1.39)	
Parahippocampal gyrus anterior	53.51 (1.14)	†	51.76 (1.24)		50.80 (1.23)	
Parahippocampal gyrus posterior	52.46 (0.94)		51.99 (0.80)		51.40 (0.73)	
Parietal operculum cortex	53.34 (0.88)	†	52.22 (1.14)		52.33 (0.96)	
Planum polare	52.93 (0.72)	*	52.03 (0.96)		51.55 (0.86)	
Planum temporale	52.10 (0.79)		51.08 (1.22)		53.06 (1.15)	
Postcentral gyrus	55.44 (1.01)	*	54.25 (1.39)	†	53.84 (1.41)	
Precentral gyrus	55.20 (1.09)	*	54.64 (1.52)	†	54.28 (1.23)	†
Precuneous cortex	55.30 (1.04)	*	53.33 (1.24)		54.41 (1.06)	*
Subcallosal cortex	54.37 (0.99)	*	52.86 (1.16)		52.38 (1.14)	
Superior frontal gyrus	55.29 (0.90)	*	54.58 (1.21)	†	53.40 (1.15)	†
Superior parietal lobule	54.08 (0.87)	*	51.80 (1.34)		54.40 (1.32)	†
Superior temporal gyrus anterior	52.91 (0.95)	†	52.59 (0.80)	†	52.70 (1.12)	
Superior temporal gyrus posterior	53.73 (0.99)	†	51.78 (1.02)		52.43 (0.87)	
Supracalcarine cortex	52.67 (1.04)		52.44 (1.18)		50.89 (1.00)	
Supramarginal gyrus anterior	54.33 (1.02)	*	53.76 (1.33)	†	53.70 (1.41)	
Supramarginal gyrus posterior	55.23 (0.95)	*	53.14 (1.31)		53.31 (1.26)	
Temporal fusiform cortex anterior	52.33 (0.89)		52.09 (1.05)		52.04 (0.94)	
Temporal fusiform cortex posterior	54.21 (1.12)	†	52.28 (1.18)		50.89 (0.90)	
Temporal occipital fusiform cortex	53.62 (0.84)	*	52.50 (1.04)		52.84 (1.44)	
Temporal pole	55.53 (0.93)	*	54.39 (1.18)	†	53.27 (1.23)	

suggest that their roles in attention are task-dependent (Spreng et al., 2010; Esterman et al., 2013). A strict interpretation of DMN as task-negative and DMN and task-positive (e.g., Fox et al., 2005)—that is, that DMN activity is detrimental to successful performance and DAN is necessary for it—might predict overall activity differences across attentional states. However, we did not observe such differences: DMN activity was not higher during out-of-the-zone states and DAN activity was not higher during in-the-zone states. Rather, DAN activity was positively correlated with erratic performance. Although this finding may be counterintuitive at first, one possible explanation is that during performance of a task requiring sustained effort, the over-engagement of external attention may lead to poor performance (Ling and Carrasco, 2006; Esterman et al., 2014). Instead, moderate levels of activity in DMN and DAN may support optimal performance (Esterman et al., 2014). The finding that multivariate patterns of activity in DMN and DAN predicted attentional state, and thus, that both networks are involved in on- and off-task performance, supports this hypothesis. Another possibility is that periods of erratic performance reflect the frequent disengagement and re-engagement of external attention and thus the DAN. While clearly observed in other behavioral contexts (Chun and Wolfe, 1996; Gehring and Knight, 2000; Botvinick et al., 2004), this explanation appears less applicable for our study given the lack of post-error slowing. In sum, although more work is necessary to determine the precise roles of these networks in tasks with varying attentional demands, our results compel roles for DMN and DAN as more nuanced than simply off- and on-task.

In high-level visual areas, decoding results reflect task demands. Patterns of activity in FFA, selective to face stimuli of the central task, carried state-relevant information regardless of task load. Interestingly, although one may have predicted that the suboptimal attentional state would be accompanied by decreased FFA activity—caused, for instance, by attention to distractors at the expense of the central task or even a simple failure to look at the stimuli—the finding that univariate FFA activity did not vary by attentional state suggests that these periods of performance are instead characterized by subtler patterns of activity. In contrast, patterns of PPA activity predicted attentional state only in the low load task, when resources were available to process both central task and distractor stimuli (e.g., Yi et al., 2004). These results are partially supportive of Lavie's load theory (Lavie, 2005): Under conditions of high perceptual load, one would expect less distractor processing. However, under conditions of high working memory load, load theory predicts greater distractor processing, which should have resulted in above-chance classification in the PPA during the current working memory load task (Yi et al., 2004). We did not see this pattern of results. Potentially with more participants or task trials, or if performance on the working memory load task were matched to that on the perceptual load task (as in Yi et al., 2004), decoding in distractor-selective regions during a working memory load task would reach significance.

Patterns of activity discriminating in- and out-of-the-zone states of attention are not confined to a priori regions of interest, at least in the easiest low load task. In fact, classification accuracy exceeded a Bonferroni-corrected p-value in 27 of the 48 anatomical ROIs in this task, and, when accuracy was above chance at $p < 0.05$, average classification accuracy across tasks was not numerically higher in the a priori than in the anatomical ROIs (53.26% vs. 53.55%). One reason for the lack of spatial specificity of these patterns could be that the anatomical ROIs overlapped with our attention networks of interest. A broader explanation for the findings, however, is that wide networks of brain regions are involved in various states of sustained attention, a finding that parallels ubiquitous coding of reward in decision making tasks (Vickery et al., 2011). Functional connectivity analyses of these data support this view, revealing that networks comprised of hundreds of connections spanning cortical and subcortical regions predict attentional performance (Finn et al., 2013). Thus, although attention-relevant and stimulus-specific regions are reasonable places to begin looking for information about attentional state in the brain, future work should take a whole-brain network level approach to the representation of these states.

Limitations of our study can motivate future work. To begin, it is likely that sustained attention fluctuates along a continuum of task engagement rather than falling into two distinct states. Although in the future it will be valuable to consider a continuous metric of sustained attention, for the purposes of our proof-of-concept trial-by-trial prediction, and to ensure sufficient statistical power given three tasks, it was useful to binarize attentional state as in and out of the zone. In addition, the small number of correct commissions in each task block (at maximum, 76) and small number of runs (three) prompted the use of a 10-fold cross-validation (between five and ten folds has been considered optimal [e.g., Arlot and Celisse, 2010]) instead of a leave-one-run-out approach that would have more stringently controlled for noise overlap between training and test data sets (e.g., Coutanche and Thompson-Schill, 2012). Future work utilizing more trials per run, or more runs, and a leave-one-run-out cross-validation approach would be useful in replicating the present findings.

Because attention dysregulation is a symptom of many psychiatric disorders, including ADHD (Barkley, 1997), schizophrenia (Braff, 1993; Luck and Gold, 2008), and post-traumatic stress disorder (Vasterling et al., 1998), decoding attentional state could be especially useful in clinical contexts. For example, individuals could be trained via neurofeedback to maintain an optimal attentional state (e.g., deBettencourt et al., 2015), or alerted to drifting attention in order to prevent errors during task performance. Although online in vs. out of the zone classification is yet untested, future work could explore its efficacy in attention training.

Conclusions

The ability to predict moment-to-moment attentional state from patterns of fMRI activity is of both theoretical and practical significance. Rather than relying on condition averages to evaluate attention to task or distractor processing, investigators can use MVPA of activity in large-scale networks and stimulus-selective regions to track the attentional states of experimental participants over time. Importantly, future work may be able to use online monitoring of attentional state to alert participants and experimenters to mid-task zone-outs, reducing the frequency of performance errors.

Funding sources

This work was supported by the Yale FAS MRI Program funded by the Office of the Provost and the Department of Psychology. M.D.R. and E.S.F. are each supported by a National Science Foundation Graduate Research Fellowship.

Conflict of interest

The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.03.032.

References

Arlot, S., Celisse, A., 2010. A survey of cross-validation procedures for model selection. Stat. Surveys 4, 40–79.
Barkley, R.A., 1997. Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. Psychol. Bull. 121 (1), 65–94.
Bellgrove, M.A., Hester, R., Garavan, H., 2004. The functional neuroanatomical correlates of response variability: evidence from a response inhibition task. Neuropsychologia 42 (14), 1910–1916.
Botvinick, M.M., Cohen, J.D., Carter, C.S., 2004. Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. 8 (12), 539–546.

Braff, D.L., 1993. Information processing and attention dysfunctions in schizophrenia. Schizophr. Bull. 19 (2), 233–259.
Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L., 2008. The brain's default network: anatomy, function, and relevance to disease. Ann. N. Y. Acad. Sci. 1124, 1–38.
Castellanos, F.X., Sonuga-Barke, E.J., Milham, M.P., Tannock, R., 2006. Characterizing cognition in ADHD: beyond executive dysfunction. Trends Cogn. Sci. 10 (3), 117–123.
Cheyne, J.A., Carriere, J.S.A., Smilek, D., 2006. Absent-mindedness: lapses of conscious awareness and everyday cognitive failures. Conscious. Cogn. 15 (3), 578–592.
Chikazoe, J., Jimura, K., Asari, T., Yamashita, K., Morimoto, H., Hirose, S., Miyashita, Y., Konishi, S., 2009. Functional dissociation in right inferior frontal cortex during performance of a go/no-go task. Cereb. Cortex 19 (1), 146–152.
Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., Schooler, J.W., 2009. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. Proc. Natl. Acad. Sci. U. S. A. 106 (21), 8719–8724.
Chun, M.M., Wolfe, J.M., 1996. Just say no: how are visual searches terminated when there is no target present? Cogn. Psychol. 30 (1), 39–78.
Chun, M.M., Golomb, J.D., Turk-Browne, N.B., 2011. A taxonomy of external and internal attention. Annu. Rev. Psychol. 62, 73–101.
Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat. Rev. Neurosci. 3 (3), 201–215.
Coutanche, M.N., Thompson-Schill, S.L., 2012. The advantage of brief fMRI acquisition runs for multi-voxel pattern detection across runs. NeuroImage 61, 1113–1119.
Cox, R.W., Hyde, J.S., 1997. Software tools for analysis and visualization of fMRI data. NMR Biomed. 10 (4–5), 171–178.
deBettencourt, M.T.Cohen J.D.Lee R.F.Norman K.A.Turk-Browne N.B., 2015. Closed-loop training of attention with real-time brain imaging. Nat Neurosci. 18, 470–475.
Edkins, G.D., Pollock, C.M., 1997. The influence of sustained attention on railway accidents. Accid. Anal. Prev. 29 (4), 533–539.
Esterman, M., Noonan, S.K., Rosenberg, M., DeGutis, J., 2013. In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. Cereb. Cortex 23 (11), 2712–2723.
Esterman, M., Rosenberg, M.D., Noonan, S., 2014. Intrinsic fluctuations in sustained attention and distractor processing. J. Neurosci. 34 (5), 1724–1730.
Finn, E.S., Rosenberg, M.D., Shen, X., Chun, M.M., Constable, R.T., 2013. Predicting Attention and Performance Across Varying Task Loads From Complex Networks During Task and at Rest. Presented at. Society for Neuroscience, San Diego, CA.
Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn. Reson. Med. 33 (5), 636–647.
Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., Raichle, M.E., 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc. Natl. Acad. Sci. U. S. A. 102 (27), 9673–9678.
Gehring, W.J., Knight, R.T., 2000. Prefrontal–cingulate interactions in action monitoring. Nat. Neurosci. 3 (5), 516–520.
Hudock, S.D., Duchon, J.C., 1988. A safety risk evaluation of vigilance tasks in the U.S. surface mining industry. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 32(15), pp. 990–994.
Hultsch, D.F., MacDonald, S.W.S., Hunter, M.A., Levy-Bencheton, J., Strauss, E., 2000. Intraindividual variability in cognitive performance in the elderly: comparison of adults with mild dementia, adults with arthritis, and healthy adults. Neuropsychology 14, 588–598.
Kaiser, S., Roth, A., Rentrop, M., Friederich, H.C., Bender, S., Weisbrod, M., 2008. Intra-individual reaction time variability in schizophrenia, depression and borderline personality disorder. Brain Cogn. 66, 73–82.
Kelly, A.M., Uddin, L.Q., Biswal, B.B., Castellanos, F.X., Milham, M.P., 2008. Competition between functional brain networks mediates behavioral variability. NeuroImage 39 (1), 527–537.
Lavie, N., 2005. Distracted and confused?: selective attention under load. Trends Cogn. Sci. 9 (2), 75–82.
Leber, A.B., 2010. Neural predictors of within-subject fluctuations in attentional control. J. Neurosci. 30 (34), 11458–11465.
Lee, R.W.Y., Jacobson, L.A., Pritchard, A.E., Ryan, M.S., Yu, Q., Denckla, M.B., Mostofsky, S., Mahone, E.M., 2012. Jitter reduces response-time variability in ADHD: an ex-Gaussian analysis. J. Atten. Disord.
Ling, S., Carrasco, M., 2006. When sustained attention impairs perception. Nat. Neurosci. 9 (10), 1243–1245.
Luck, S.J., Gold, J.M., 2008. The construct of attention in schizophrenia. Biol. Psychiatry 64 (1), 34–39.
MacDonald, S.W.S., Nyberg, L., Bäckman, L., 2006. Intra-individual variability in behavior: links to brain structure, neurotransmission, and neuronal activity. Trends Neurosci. 29, 474–480.
MacDonald, S.W.S., Li, S.C., Bäckman, L., 2009. Neural underpinnings of within-person variability in cognitive functioning. Psychol. Aging 24 (4), 792–808.
Mason, M.F., Norton, M.I., Van Horn, J.D., Wegner, D.M., Grafton, S.T., Macrae, C.N., 2007. Wandering minds: the default network and stimulus-independent thought. Science 315 (5810), 393–395.
Padilla, M.L., Wood, R.A., Hale, L.A., Knight, R.T., 2006. Lapses in a prefrontal-extrastriate preparatory attention network predict mistakes. J. Cogn. Neurosci. 18 (9), 1477–1487.
Prado, J., Weissman, D.H., 2011. Heightened interactions between a key default-mode region and a key task-positive region are linked to suboptimal current performance but to enhanced future performance. NeuroImage 56, 2276–2282.
Prado, J., Carp, J., Weissman, D.H., 2011. Variations of response time in a selective attention task are linked to variations of functional connectivity in the attentional network. NeuroImage 54, 541–549.
Robertson, I.H., Manly, T., Andrade, J., Baddeley, B.T., Yiend, J., 1997. 'Oops!': performance correlates of everyday attention failures in traumatic brain injured and normal subjects. Neuropsychologia 35 (6), 747–758.
Rosenberg, M., Noonan, S., Degutis, J., Esterman, M., 2013. Sustaining visual attention in the face of distraction: a novel gradual-onset continuous performance task. Atten. Percept. Psychophys. 75 (3), 426–439.
Ryan, M., Martin, R., Denckla, M.B., Mostofsky, S.H., Mahone, E.M., 2010. Interstimulus jitter facilitates response control in children with ADHD. J. Int. Neuropsychol. Soc. 16 (2), 388–393.
Sadaghiani, S., Hesselmann, G., Kleinschmidt, A., 2009. Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. J. Neurosci. 29 (42), 13410–13417.
Schooler, J.W., Smallwood, J., Christoff, K., Handy, T.C., Reichle, E.D., Sayette, M.A., 2011. Meta-awareness, perceptual decoupling and the wandering mind. Trends Cogn. Sci. 15 (7), 319–326.
Schwartz, F., Carr, A.C., Munich, R.L., Glauber, S., Lesser, B., Murray, J., 1989. Reaction time impairment in schizophrenia and affective illness: the role of attention. Biol. Psychiatry 25 (5), 540–548.
Smith, R., Keramatian, K., Smallwood, J., Schooler, J., Luus, B., Christoff, K., 2006. Mind-wandering with and without awareness: an fMRI study of spontaneous thought processes. In: Sun, R. (Ed.), Proceedings of the 28th Annual Conference of the Cognitive Science Society. Erlbaum, Vancouver, pp. 804–809.
Solanto, M.V., Schulz, K.P., Fan, J., Tang, C.Y., Newcorn, J.H., 2009. Event-related FMRI of inhibitory control in the predominantly inattentive and combined subtypes of ADHD. J. Neuroimaging 19 (3), 205–212.
Spreng, R.N., Stevens, D.W., Chamberlain, J.P., Gilmore, A.W., Schacter, D.L., 2010. Default network activity, coupled with the frontoparietal network, supports goal-directed cognition. NeuroImage 53 (1), 303–317.
Spreng, R.N., Sepulcre, J., Turner, G.R., Stevens, W.D., Schacter, D.L., 2013. Intrinsic architecture underlying the relations among the default, dorsal attention, and frontoparietal control networks of the human brain. J. Cogn. Neurosci. 25, 74–86.
Stuss, D.T., Stethem, L.L., Hugenholtz, H., Picton, T., Pivik, J., Richard, M.T., 1989. Reaction time after head injury: fatigue, divided and focused attention, and consistency of performance. J. Neurol. Neurosurg. Psychiatry 52 (6), 742–748.
Stuss, D.T., Pogue, J., Buckle, L., Bondar, J., 1994. Characterization of stability of performance in patients with traumatic brain injury: variability and consistency on reaction time tests. Neuropsychology 8 (3), 316–324.
Stuss, D.T., Murphy, K.J., Binns, M.A., Alexander, M.P., 2003. Staying on the job: the frontal lobes control individual performance variability. Brain 126 (11), 2363–2380.
Suskauer, S.J., Simmonds, D.J., Caffo, B.S., Denckla, M.B., Pekar, J.J., Mostofsky, S.H., 2008. fMRI of intrasubject variability in ADHD: anomalous premotor activity with prefrontal compensation. J. Am. Acad. Child Adolesc. Psychiatry 47 (10), 1141–1150.
Vasterling, J.J., Brailey, K., Constans, J.I., Sutker, P.B., 1998. Attention and memory dysfunction in posttraumatic stress disorder. Neuropsychology 12 (1), 125–133.
Vickery, T.J., Chun, M.M., Lee, D., 2011. Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron 72 (1), 166–177.
Weissman, D.H., Roberts, K.C., Visscher, K.M., Woldorff, M.G., 2006. The neural bases of momentary lapses in attention. Nat. Neurosci. 9 (7), 971–978.
Wodka, E.L., Simmonds, D.J., Mahone, E.M., Mostofsky, S.H., 2009. Moderate variability in stimulus presentation improves motor response control. J. Clin. Exp. Neuropsychol. 31 (4), 483–488.
Yi, D.J., Woodman, G.F., Widders, D., Marois, R., Chun, M.M., 2004. Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. Nat. Neurosci. 7 (9), 992–996.