NeuroImage xxx (2015) xxx–xxx

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 xxxx

*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.

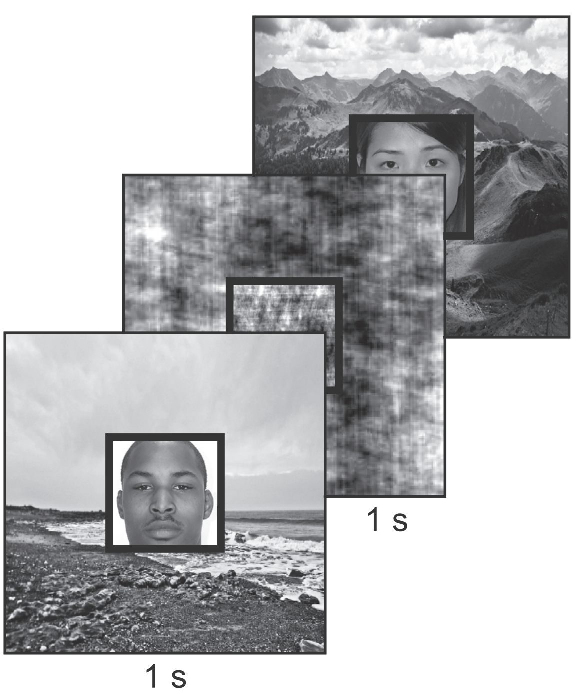


**Fig. 1.** Trials consisted of face–scene composite images presented for 1 s followed by 1 s masks. Participants were instructed to attend to faces while ignoring background scenes, and to respond to non-repeated faces and withhold response to repeats (1-back for the low load and perceptual load tasks;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