

Intrinsic Fluctuations in Sustained Attention and Distractor Processing

Michael Esterman,^{1,2} Monica D. Rosenberg,^{1,3} and Sarah K. Noonan^{1,4}

¹Veterans Administration Boston Healthcare System, Neuroimaging Center for Veterans, Boston, Massachusetts 02130, ²Boston University School of Medicine, Department of Psychiatry, Boston, Massachusetts 02118, ³Yale University, Department of Psychology, New Haven, Connecticut 06520, and ⁴Veterans Administration Puget Sound Health Care System, Seattle, Washington 98108

Although sustaining a moderate level of attention is critical in daily life, evidence suggests that attention is not deployed consistently, but rather fluctuates from moment to moment between optimal and suboptimal states. To better characterize these states in humans, the present study uses a gradual-onset continuous performance task with irrelevant background distractors to explore the relationship among behavioral fluctuations, brain activity, and, in particular, the processing of visual distractors. Using fMRI, we found that reaction time variability, a continuous measure of attentional instability, was positively correlated with activity in task-positive networks and negatively correlated with activity in the task-negative default mode network. We also observed greater processing of distractor images during more stable and less error prone “in the zone” epochs compared with suboptimal “out of the zone” epochs of the task. Overall, the data suggest that optimal states of attention are accomplished with more efficient and potentially less effortful recruitment of task-relevant resources, freeing remaining resources to process task irrelevant features of the environment.

Key words: attentional fluctuations; attentional states; distraction; fMRI; repetition attenuation; sustained attention

Introduction

Sustaining an appropriate level of attention is critical for completing most tasks in daily life, but our ability to do so is far from stable. Attention can vary from optimal levels to extremes of either under-engagement (e.g., mindlessly daydreaming while driving) or overengagement with the task at hand (e.g., overthinking the service in tennis). In a previous fMRI study, we demonstrated that attention fluctuates between at least two distinct states during sustained performance: (1) a behaviorally stable, less error-prone state associated with moderate activity in brain regions of the task-negative network (TNN), a state that we have termed “in the zone,” and (2) a behaviorally unstable, error-prone state in which performance relies more heavily on perceptual/attentional task-positive network (TPN) regions, which we have termed “out of the zone” (Esterman et al., 2013). This added to a growing body of evidence linking activity in these networks to intrinsic fluctuations in performance (Weissman et al., 2006; Hahn et al., 2007; Christoff et al., 2009; Sadaghiani et al., 2009).

Despite progress in identifying the neural underpinnings of different attentional states, relatively little is known about the

cognitive processes that distinguish them. We reasoned that an fMRI study of distractor processing during sustained performance would be well suited to adjudicate between several alternative characterizations. One possibility is that, during periods of successful performance, spatial attention is relatively focused such that perceptual resources are narrowly directed toward task-relevant stimuli (McMains and Somers, 2005). This account predicts filtering of irrelevant distractors and thus reduced distractor processing during in the zone relative to out of the zone performance. Alternatively, successful performance could reflect more efficient, economical recruitment of attentional and/or perceptual resources as a central task is performed with ease (Rosenberg et al., 2013). In this scenario, processing resources would be available for distractor processing when perceptual demands of the central task are low (Yi et al., 2004; Lavie et al., 2004), predicting greater distractor processing in the zone than out of the zone. A third possibility is that fluctuations in performance reflect changes in nonperceptual, higher-level processes such as response control, working memory, or task-set maintenance. This account would predict little relationship between performance and distractor processing. That is, distractor processing would not differ between the in the zone and out of the zone periods.

The present study explored these alternatives by examining neural indicators of distractor processing across in the zone and out of the zone periods during performance of a sustained attention task (the gradual onset continuous performance task; gradCPT). Subjects performed the gradCPT with a central face task and irrelevant background distractor scenes. We examined the relationships between attention states and overall cortical scene-network activity known to reflect voluntary

Received June 23, 2013; revised Dec. 17, 2013; accepted Dec. 19, 2013.

Author contributions: M.E. and M.R. designed research; M.E. and M.R. performed research; S.K.N. contributed unpublished reagents/analytic tools; M.E., M.R., and S.K.N. analyzed data; M.E. and M.R. wrote the paper.

This work was supported by the Translational Research Center for TBI and Stress Disorders (TRACTS), a VA Rehabilitation Research and Development Traumatic Brain Injury Center of Excellence (B6796-C).

The authors declare no competing financial interests.

Correspondence should be addressed to Michael Esterman, PhD, Boston Attention and Learning Laboratory, VA Boston Healthcare System (182 JP), 150 South Huntington Avenue, Boston, MA 02130. E-mail: esterman@bu.edu.

DOI:10.1523/JNEUROSCI.2658-13.2014

Copyright © 2014 the authors 0270-6474/14/341724-07\$15.00/0

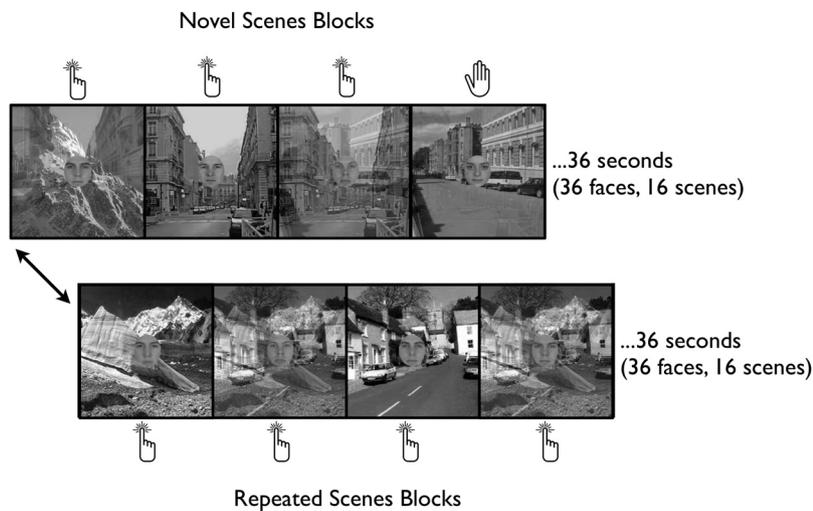


Figure 1. The gradCPT. Faces gradually transition from one to the next every second. Participants are instructed to respond to male faces and withhold response to a target female face. Irrelevant background scenes gradually transition every 2250 ms. In novel-scene blocks, all scenes are unique (16 novel scenes). In repeating-scenes blocks, 2 scenes alternate (8 times each). Scenes from a single block are never repeated in the experiment.

attention and goal relevance (Esterman and Yantis, 2010). Further, we manipulated scene repetition to examine more subtle changes in depth of processing of distractors via repetition attenuation (Yi et al., 2004).

Materials and Methods

Participants

Sixteen participants (5 males, ages 18–24 years) performed the gradCPT (Esterman et al., 2013; Rosenberg et al., 2013) during fMRI. Participants were right-handed, had normal or corrected-to-normal vision, and reported no history of major illness, head trauma, or neurological/psychiatric disorder. The study was approved by the VA Boston Healthcare System Institutional Review Board and all participants gave written informed consent. One participant was excluded for high error rate (>80%, 4 SDs above the mean) and one was excluded for excessive motion during fMRI.

Paradigm and stimuli

Participants performed the gradCPT, comprised of grayscale face photographs (MIT Face Database; Russell, 2009, with permission from Richard Russell) centrally overlaid on grayscale scenes (SUN Database; Xiao et al., 2010) during fMRI (Fig. 1). Faces consisted of one female and 10 males cropped to show the eyes, nose, and mouth. Viewed through an MR-compatible goggle system (VisuaStim Digital; Resonance Technology), faces subtended a diameter of 4.8° of visual angle and background scenes subtended 13° in height and width.

On each trial, a face photograph gradually transitioned from one to the next using linear pixel-by-pixel interpolation. Transitions took 800 ms and faces paused for 200 ms when fully cohered. Faces were presented randomly (90% male, 10% target female) without allowing identical faces to repeat on consecutive trials (repeats were banned because there was only one target female, making the overall target probability 9%).

Background scenes transitioned out of sync with faces at a rate of 2250 ms/image. Blocks of repeated and novel scenes alternated every 36 s. Repeated-scenes blocks consisted of two alternating scene images such that each pair repeated eight times (16 scenes per block). Novel-scenes blocks consisted of 16 unique scene images. Scenes never appeared in multiple blocks. This repetition procedure of alternating identical images was based on the paradigm developed by Yi et al. (2004) in which background scenes were used with a central face task. Given our parallel questions regarding attentional state's influence on repetition attenuation, we used this study as a model.

Participants were instructed to press a button in response to each male face and to withhold response to the target female face. Accuracy was emphasized without reference to speed.

Procedure

Before scanning, participants were given a 1–2 min practice with the gradCPT. In the MRI scanner, participants performed three 8.67 min runs of the gradCPT (8.4 min of continuous task with an additional 16 s of fixation). An anatomical magnetization prepared rapid gradient echo (MPRAGE) was also acquired.

Imaging parameters

Scanning was performed on a 3T Siemens MAGNETOM Trio system with a 12-channel head coil at the VA Boston Neuroimaging Center for Veterans. Functional runs included 260 whole-brain volumes acquired using an echoplanar imaging sequence with the following parameters: TR = 2000 ms, TE = 30 ms, flip angle = 90°, acquisition matrix = 64 × 64, in-plane resolution = 3.0 mm², 33 oblique slices, slice thickness = 3, 0.75 mm gap. MPRAGE parameters were as follows: TE = 3.32, TR = 2530 ms, flip angle = 7°, acquisition matrix = 256 × 256, in-plane resolution = 1.0 mm², 176 sagittal slices, slice thickness = 1.0 mm.

Behavioral analysis

Reaction time. Reaction times (RTs) were calculated relative to the beginning of each face transition such that an RT between 800 and 1000 ms indicated a response that occurred when the face on the current trial was 100% cohered and not yet mixed with the following image. An RT shorter than 800 ms indicated that the current face was still transitioning from the previous one and an RT longer than 1000 ms indicated that the face was transitioning to the next. On rare trials with highly deviant RTs (before 70% coherence of the current face and after 40% coherence of the following face) or multiple responses, an iterative algorithm that maximized correct responses was used. First, the algorithm assigned unambiguous correct responses (presses to male faces that occurred after 70% coherence of the current face and before 40% coherence of the following). Second, the remaining ambiguous presses (<5% of trials) were assigned to an adjacent trial if one of the two had no response or to the closest trial if both had no response (unless one was the target female face, in which case participants were given the benefit of the doubt that they had correctly omitted). Finally, if multiple presses could be assigned to a trial, the fastest response was selected. Slight variations to the algorithm yielded highly similar results because the vast majority of button presses showed a 1-to-1 correspondence with the presented faces.

Accuracy. Trials on which participants correctly inhibited response to the target female face were considered correct omissions or “correct” trials. Trials on which participants erroneously responded to the female face were considered commission errors or “lapses.” Omission errors, or failures to respond to male faces, occurred rarely (average of 6.2% across participants) and thus were excluded from fMRI analyses. Trials on which participants correctly responded to male faces (the majority) were considered correct commissions or “baseline” trials.

RT variability. To examine trial-to-trial fluctuations in attentional stability, we computed a variance time course (VTC) from the ~450 baseline trials in each run (Esterman et al., 2013; Rosenberg et al., 2013). Each trial was assigned a value corresponding to the absolute deviance of its RT from the mean RT of the run quartile in which it occurred (to control for possible low-frequency drifts in RT). Values for correct trials, lapses, and omission errors were interpolated linearly by averaging the two neighboring baseline trial RTs. The VTC was smoothed using a Gaussian ker-

nel integrating information from the surrounding 20 trials (20 s) with a 9 s full width at half maximum (FWHM).

In each participant, smoothed VTC values were used to assign trials to low-variability or high-variability epochs via median split (in the zone and out of the zone epochs). In the zone epochs represent periods in which RTs were close to the mean, whereas out of the zone epochs represent periods in which RTs were highly deviant from the mean.

fMRI analysis

General methods. fMRI data were analyzed using Analysis of Functional NeuroImages (AFNI; Cox and Hyde, 1997) and custom scripts in MATLAB (The MathWorks). Preprocessing included slice-time correction; motion correction using a 6-parameter, rigid body, least-squares alignment procedure; spatial smoothing to an 8 mm FWHM Gaussian kernel; automated coregistration and normalization of anatomical and functional volumes to Talairach space; and scaling of functional dataset values to percent signal change (relative to the mean of the run). Data from individual participants were analyzed with linear multiple regression (see RT variability: variance time course, below). Regression coefficients for effects of interest were evaluated using voxelwise group-level t tests. All resulting statistical maps were corrected for multiple comparisons using voxel-cluster Monte Carlo-type α simulations (Forman et al., 1995), resulting in a corrected significance level of $\alpha = 0.05$ (individual-voxel intensity threshold of $p < 0.01$, cluster size of 54 contiguous voxels).

Functional regions of interest. An independent dataset (Esterman et al., 2013) was used to define the cortical scene network (Fig. 3A). A blocked design-scene-selective localizer that used identical scene stimuli was analyzed to contrast 30 s scene and face blocks. To define right and left scene-selective regions, spherical regions of interest (ROIs; 6 mm radius, 33 voxels) were centered on group-level peaks from the thresholded scene > face contrast within the parahippocampal gyrus (PPA) ($-22, -44, -7$) and ($25, -44, -7$), retrosplenial cortex (RSC) ($-16, -56, 11$) and ($17, -53, 11$), and transverse occipital sulcus/occipital place area (OPA) ($-38, -80, 23$) and ($37, -83, 23$).

RT variability: variance time course. A stagewise regression procedure was used to isolate blood oxygenation level-dependent (BOLD) effects due to RT variability. For each participant, a first-stage general linear model (GLM) accounted for signal variance associated with the mean evoked response for each trial type, as well as each trial's RT. Additional nuisance regressors included terms for signal mean, linear drift, six realignment parameters, and mean signal from spherical ROIs centered in deep white matter and lateral ventricle CSF. The RT variability analysis was performed on the residuals of this first-stage model and implemented with amplitude-modulated regression using an unsmoothed VTC convolved with a one-parameter gamma variate hemodynamic response function.

Repetition attenuation. To investigate the interaction between repetition attenuation and attentional state, percent signal change values in the independently defined scene selective regions were submitted to a 2 (block type: repeated vs novel) \times 2 (attentional state: in the zone vs out of the zone) \times 3 (ROI: PPA, OPA, RSC) \times 2 (hemisphere: left vs right) ANOVA. Specifically, the percent signal change values were extracted from each ROI 4–6 s after each baseline trial and sorted by whether the background scene was repeated or novel and whether the trial occurred in or out of the zone. Note that this analysis and all subsequent analyses employed the same stagewise procedure outlined above (VTC) and are thus conducted on the residuals of the first-stage GLM.

Lapse precursors. To evaluate lapse precursors in the cortical scene network, activity during the TR preceding appearance of the target female face was estimated separately for lapse and correct trials in scene-selective regions (Esterman et al., 2013). Again, to isolate spontaneous signal fluctuations, these analyses were conducted on the residuals of a first-stage GLM (see RT variability: variance time course, above). Linear time interpolation was conducted to estimate the BOLD response at each image transition (every second), ensuring that interpolated responses were only estimated from the nearest TRs. BOLD signal values were averaged across a pretrial window spanning 1 TR (2 s, 2 trials) before target appearance.

To determine whether lapse precursors differed across blocks and attentional states, BOLD signal averaged in this window was submitted to a 2 (trial type: lapse vs correct trial) \times 2 (attentional state: in the zone vs out of the zone) \times 3 (ROI: PPA, OPA, RSC) \times 2 (hemisphere: left vs right) ANOVA.

Results

Accuracy

On average, participants made lapses (incorrect presses) on 28.2% of female face trials (range = 12.6–53.4%, SD = 13.8%) and omission errors (failures to press to male faces) to 6.2% of male faces (range = 0.6–15.9%, SD = 4.6%). Mean RT on baseline trials was 849 ms (SD = 95.5 ms).

RT variability

Participants made significantly more lapses ($t_{13} = 3.57, p = 0.003$) and omission errors ($t_{13} = 5.55, p < 0.001$) when out of the zone than when in the zone (lapses: means = 33.2% vs 23.5%; SDs = 14.8% vs 13.6%; omission errors: means = 8.1% vs 4.2%; SDs = 5.6% vs 3.8%). Baseline RTs did not differ across zone epoch, $t_{13} = 0.85, p = 0.41$ (out of the zone vs in the zone: means = 852 vs 846 ms, SDs = 87 vs 105 ms).

Extrinsic versus intrinsic fluctuations

Although the 10 male faces were presented repeatedly and randomly to promote familiarity and potentially alleviate any stimulus-specific effects, there remained the possibility that objective differences between the faces partially accounted for the behavioral fluctuations. To address this, we examined the frequency at which each of the 10 male faces occurred in and out of the zone. On average, in a single run, each stimulus appeared on 7.7 trials (range 7.1–8.3) during in the zone and out of the zone epochs. *Post hoc* comparisons revealed that none of the faces occurred more often in the zone than out of the zone. Nevertheless, we cannot establish definitively that the fluctuations in variability are purely intrinsic and not in some way associated with objective differences in the task difficulty, such as higher-order relationships of stimulus order or transitions.

fMRI: correlates of RT variability

To determine the neural correlates of trial-to-trial fluctuations in RT variability, whole-brain multiple regression was performed for each subject using the VTC as a regressor (see Materials and Methods). Several classic TPN regions in the dorsal attention and salience networks, including bilateral intraparietal sulci, frontal eye fields, dorsolateral prefrontal cortex, insular/opercular cortex, and presupplementary motor area, showed higher activity levels during moments of relative variability (Fig. 2, Table 1), suggesting greater recruitment of these areas during periods of instability and poorer performance (out of the zone). In contrast, several classic TNN regions in the default mode network (DMN; ventromedial prefrontal cortex, posterior cingulate, and left lateral parietal cortex), as well as the left putamen, showed higher activity during periods of relative stability and more accurate performance (in the zone).

fMRI: repetition attenuation

As expected, there was a main effect of repetition across the scene-selective ROIs, $F_{(1,13)} = 10.02, p = 0.01$, such that these regions exhibited more activity during novel than repeated background blocks. A block type \times attentional state interaction was observed ($F_{(1,13)} = 5.40, p = 0.037$) such that there was a greater

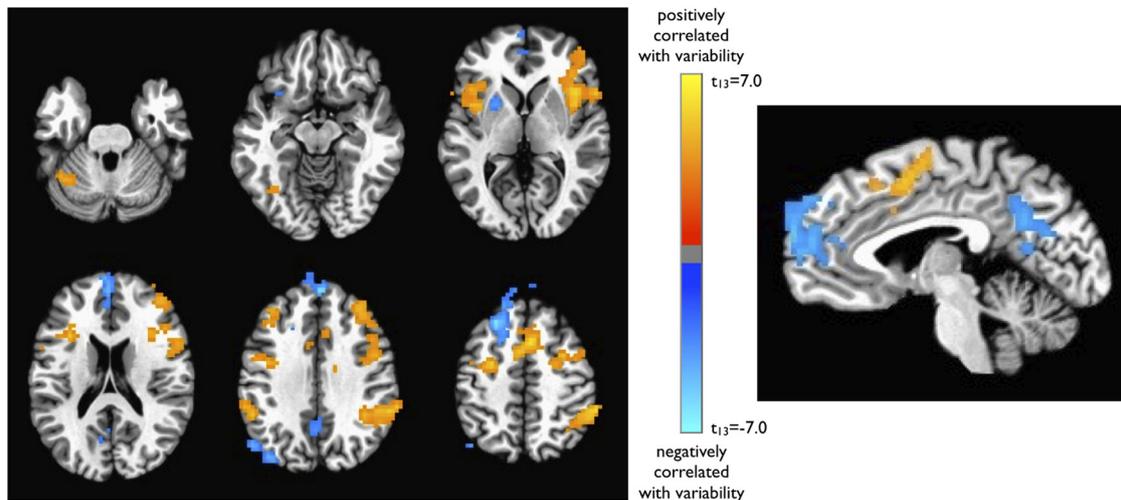


Figure 2. Reaction time variability: VTC-BOLD signal correlation. VTC represents trial-to-trial fluctuations in response variability. Regions in blue are negatively correlated with the VTC, such that they are associated with relative stability of reaction times (lower variability; in the zone). Regions in orange are positively correlated with the VTC and thus are associated with relative instability of RTs (higher variability; out of the zone). Maps are displayed after correction for multiple comparisons (corrected $p < 0.05$; nominal $p < 0.01$, cluster size > 54 voxels).

Table 1. Regions with significant variance time course-BOLD signal correlations (Figure 2)

Region	Talairach coordinates		
Positive correlation with variability			
R inferior /middle frontal gyrus	41	26	15
R inferior parietal/IPS	56	-41	48
B superior frontal/SMA	5	11	48
L inferior frontal gyrus	-29	23	12
L cerebellum	-29	-62	-43
L precentral/middle frontal gyrus	-26	-5	48
L middle frontal gyrus	-32	32	39
L inferior parietal cortex	-50	-44	33
L cerebellum/fusiform gyrus	-35	-53	-25
Negative correlation with variability			
B superior/medial frontal gyrus	5	53	36
L angular gyrus	-47	-65	45
B posterior cingulate	-5	-56	30
L putamen	-23	2	-4

Talairach coordinates indicate the center of mass of each cluster. $p < 0.05$ corrected threshold (nominal $p < 0.01$, minimum cluster size of 54 contiguous voxels). R, right; L, left; B, bilateral.

repetition effect (difference in percent signal change between novel and repeated blocks) during in the zone compared with out of the zone epochs (Fig. 3B). Greater repetition attenuation during in the zone epochs suggests that these periods are characterized by fewer demands on attention by, and/or lower perceptual load of, the central face task. Therefore, more perceptual resources are available for the background scenes. This block type \times attentional state interaction did not further interact with ROI (PPA, RSC, OPA); therefore, the effect was approximately comparable across regions. ANOVAs for each ROI separately revealed that the interaction term was significant in the OPA ($p < 0.05$) and marginally significant in the RSC and PPA (p values < 0.1). There was a marginally significant higher-order interaction between block type \times attentional state \times hemisphere, $F_{(1,13)} = 3.56, p = 0.082$, which suggested that the effect of attentional state on repetition attenuation may have been greater in the left hemisphere scene network. Note that a percent signal change of zero represents the mean of the run, so negative numbers do not reflect deactivation below neural baseline.

fMRI: scene network activity by attentional state

There was no significant main effect of attentional state on scene network activity. The interaction between repetition and attentional state described above was driven by decreased activity in response to repeated scenes during in the zone epochs ($t_{13} = 2.06, p = 0.06$) rather than by increased activity in response to novel scenes ($p > 0.3$). This suggests that attentional states differ in the depth of processing that background scenes are afforded, but not in overall selective attention to the scenes.

fMRI: lapse precursors

When comparing lapse to correct trial precursors in scene-selective cortex (see Materials and Methods), there was no main effect of trial type, attentional state, ROI, or hemisphere on BOLD signal. Importantly, there was an attentional state \times trial type interaction ($F_{(1,13)} = 7.46, p = 0.017$) such that, when participants were in the zone, the BOLD signal was higher before lapse than correct trials ($p < 0.05$), but when participants were out of the zone, activity did not differ as a function of trial type (Fig. 3C). This suggests that perceptual interference from the scenes was more likely to occur in the zone, when there was overall greater depth of scene processing (see “Repetition Attenuation” section above). However, this was only a transient effect, because overall scene activity was not sustained at a higher level during in the zone epochs (see “Scene Network Activity by Attentional State” section above). This pattern of activity was observed in all individual ROIs and no higher-order interactions were significant. Again, zero percent signal change reflects the mean of the run, so negative numbers do not reflect deactivation per se.

Discussion

We investigated the nature of intrinsic fluctuations in sustained attention by examining distractor processing as a function of attentional state during a continuous performance task. We found evidence that periods of successful attentional performance were associated with greater depth of distractor processing, which was akin to states of low perceptual load. Specifically, we observed enhanced neural sensitivity to irrelevant distractors, as measured by repetition attenuation and

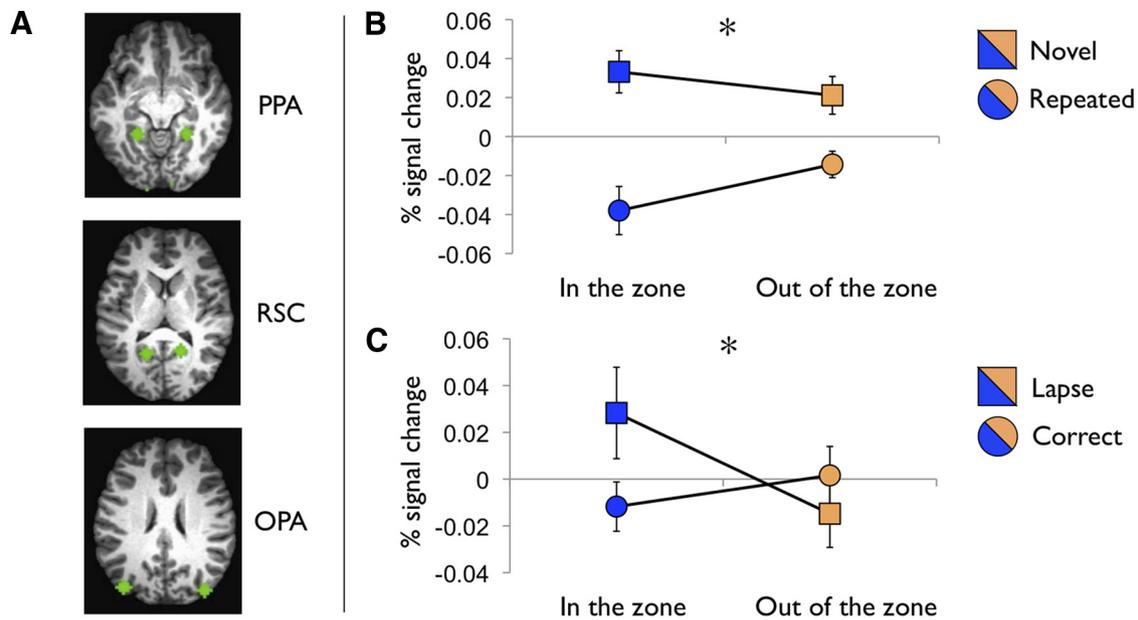


Figure 3. Effects of attentional state in cortical scene network. **A**, ROIs in the cortical scene network: PPA, RSC, and OPA. Spherical ROIs (6 mm radius, 33 voxels) derived from Esterman et al. (2013). **B**, Repetition attenuation. During in the zone (low variability) epochs, there was greater repetition attenuation (novel-repeating block) than during out of the zone epochs (high variability). *Significant interaction ($p < 0.05$) between scene condition (repeated vs novel) and attentional state (in the zone vs out of the zone). **C**, Precursors of attention lapses. Attention lapses (commission errors) were preceded by higher activity in cortical scene network, only during in the zone epochs. *Significant interaction ($p < 0.05$) between pretrial activity (commission errors vs correct omissions) and attentional state (in the zone vs out of the zone).

error precursors, during periods of superior performance (in the zone epochs).

We also observed unique neural signatures of attentional states across large-scale networks, finding that more erratic, error prone performance (i.e., being out of the zone) is associated with greater engagement of classic TPN regions, whereas relatively stable, in the zone performance is associated with greater activity in TNN/DMN regions. The finding that DMN activity is associated with increased stability replicates our prior work (Esterman et al., 2013); however, overrecruitment of TPN regions during out of the zone epochs is a result unique to the present study. One potential explanation is that the current version of the gradCPT, which includes visual distractors, is more challenging than the distractor-free gradCPT used previously and thus relies more on top-down spatial and object-based attentional control. Although speculative, it may be that some degree of activity in TPN regions is necessary to filter distracting information, but too much activity is suboptimal for sustained performance. Together, these data suggest that, in the current task, in the zone performance is accomplished by attaining a more efficient attentional state that requires less engagement of visual and attentional resources.

Perceptual load theory (Lavie et al., 2004) provides a useful framework for characterizing intrinsic fluctuations in resource demands during the gradCPT. Previous studies have demonstrated greater behavioral and neuroimaging markers of distractor processing during tasks with fewer perceptual/attentional demands because more resources are available to process task-irrelevant information (Lavie et al., 2004). In particular, Yi et al. (2004) found greater repetition attenuation to background distractors during a task with low perceptual load using fMRI. In the present study, in the zone periods are comparable to such periods of low perceptual/attentional load in that they show greater repetition attenuation than out of the

zone epochs. Errors that do occur during in the zone periods are preceded by increased activity in scene-selective regions, potentially reflecting the interference from distractors that can occur during low perceptual load. Another account of task irrelevant information processing attempts to resolve recent challenges to perceptual load theory (Tsal and Benoni, 2010), suggesting that distractor processing is dictated by the degree of competition for representation (Scalf et al., 2013). According to this account, less distractor processing will occur in conditions requiring greater top-down bias. Our results are consistent with this interpretation: during out of the zone epochs, characterized by activity in TPN regions associated with top-down bias, we observe weaker repetition attenuation to distractors. In sum, our results suggest that in the zone periods are characterized by more efficient engagement of perceptual/attentional resources, whereas out of the zone epochs reflect overengagement of these resources.

The current result that increased TPN activity is associated with poor performance challenges the strict interpretation that engagement of these regions reflects being on-task (Pardilla et al., 2006), whereas activity in TNN regions signals the opposite (Christoff et al., 2009). Instead, the role of these networks in sustained attention performance must be more nuanced (Sadaghiani et al., 2009). We propose the optimal engagement hypothesis that moderate levels of activity in TNN and TPN are most conducive to successful sustained attention performance. Specifically, less task-positive and more task-negative activity may reflect moments of efficient (and potentially less effortful) performance, whereas overengagement of TPN and/or oversuppression of TNN may undermine performance. Concurrently, too extreme a dip in TPN or spike in TNN activity may put one at risk for errors or lapses (Esterman et al., 2013). Therefore, more balanced activity profiles across networks support accurate sustained performance.

Deleterious effects of overengagement on sustained performance have been noted in previous studies. For example, findings in the attentional blink task demonstrate that overengagement of cognitive resources directed toward a first target impairs detection of second target (Olivers and Nieuwenhuis, 2005). In addition, Wieser and Keil (2011) found, using EEG, that initial overengagement of perceptual resources leads to poor sustained attention and Ling and Carrasco (2006) demonstrated behaviorally that transient attentional engagement has initial benefits but subsequent costs to sustained performance. Although speculative, it may be that overengagement is less sustainable in the gradCPT than in other cognitive tasks because it requires continuous rather than transient attention.

There are two possible explanations for why task-positive regions show less activity during in the zone than out of the zone epochs. One possibility is that, when in the zone, successful performance relies less on these networks and may be accomplished more “automatically.” Alternatively, more economical recruitment of these regions could reflect more efficient neural coding or greater reliance on their coordinated activity for successful sustained performance. Neurostimulation methods may help resolve these alternatives. A finding that temporary disruption of task-positive regions selectively impairs performance during out of the zone epochs would support the automatization hypothesis, whereas a result that disruption impairs in the zone performance would support the efficient-economical recruitment hypothesis.

One caveat to our findings is that whereas distractor processing interacted with attentional state, as indexed by lapses precursors and repetition attention, overall scene network activity did not differ between in the zone and out of the zone epochs. This suggests that participants are not consistently directing voluntary, selective attention to the irrelevant scenes during in the zone epochs. This is perhaps unsurprising because consistent attention to scenes would likely undermine performance on the central task, rather than being associated with stable and less error-prone performance that characterizes in the zone epochs. Instead, we found that the greater repetition attenuation observed during in the zone epochs was driven more by decreased activity for repeated scenes than by increased activity to novel scenes. This fits well with the proposal that “involuntary” processing of ignored stimuli is reflected in the degree of suppression of repeated items, rather than the evoked response to novel items (Lim et al., 2008). Therefore, our results suggest that repetition attenuation may index a subtler phenomenon, such as depth of processing and/or more automatic resource distribution, rather than voluntary acts of attention.

Another caveat to the current approach is the possibility that behavioral fluctuations are in part a reflection of objective differences between faces. For example, idiosyncratic differences in the difficulty of certain stimulus transitions might contribute to behavioral variability. Although no systematic differences were observed between faces presented in and out of the zone, the role of higher-order stimulus relationships cannot be excluded. In either case, our results are unlikely to be stimulus specific because similar findings were obtained in a previous study in which the gradCPT consisted of scenes.

Although the gradual nature of the gradCPT makes it conducive to studying sustained attention, other task features may limit the generalizability of the current results. First, tasks with simultaneous demands (in which only the current stimulus is task relevant, such as go/no-go tasks) may not tax sustained attention to the same degree as tasks with sequential demands

(in which previous stimuli affect responses, such as *n*-back tasks; Parasuraman, 1979). Therefore, the performance fluctuations we characterize in the current task may be driven by changes in present-moment perceptual-attentional demands, whereas in a sequential task, fluctuations may be caused by ever-changing engagement of working memory. Further, most studies of distractor processing use distractors that compete with targets, either via congruency (Eriksen and Eriksen, 1974) or bottom-up salience (Leber, 2010), and have associated TPN activity with distractor suppression. In contrast, the distractors in the present study were irrelevant both temporally (changing at a different rate than faces) and categorically. We reasoned that task-irrelevant distractors may have greater ecological validity and would better allow us to tease apart the neural contributions of target and distractor processing. Nevertheless, future work should explore how attentional states interact with distractibility when distractors have bottom-up or top-down salience. Finally, we acknowledge the possibility that despite the gradual nature of the task, distractor processing may have occurred during implicit breaks between trials. However, this interpretation would predict greater repetition attenuation or overall scene processing during periods of (or for subjects with) relatively fast versus slow RTs. *Post hoc* analyses to examine these possibilities did not reveal any such patterns, suggesting that the observed scene processing is not likely yoked to the time between trials.

In sum, the present findings illustrate that sustained attention fluctuates between in the zone periods of response stability, optimal engagement of neural resources, and relative perceptual ease and out of the zone periods of instability, overengagement, and perceptual difficulty. The hypothesized optimal engagement framework has wide-ranging implications for understanding how sustained performance can be maintained and the reasons that fluctuations in performance inevitably occur.

References

- 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](#)
- Cox RW, Hyde JS (1997) Software tools for analysis and visualization of fMRI data. *NMR Biomed* 10:171–178. [CrossRef Medline](#)
- Eriksen BA, Eriksen CW (1974) Effects of noise letters upon identification of a target letter in a nonsearch task. *Perception & Psychophysics* 16:143–149. [CrossRef](#)
- Esterman M, Yantis S (2010) Perceptual expectation evokes category-selective cortical activity. *Cereb Cortex* 20:1245–1253. [CrossRef Medline](#)
- Esterman M, Noonan SK, Rosenberg M, Degutis J (2013) In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cereb Cortex* 23:2712–2723. [CrossRef Medline](#)
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn Reson Med* 33:636–647. [CrossRef Medline](#)
- Hahn B, Ross TJ, Stein EA (2007) Cingulate activation increases dynamically with response speed under stimulus unpredictability. *Cereb Cortex* 17:1664–1671. [CrossRef Medline](#)
- Lavie N, Hirst A, de Fockert JW, Viding E (2004) Load theory of selective attention and cognitive control. *J Exp Psychol Gen* 133:339. [CrossRef Medline](#)
- Leber AB (2010) Neural predictors of within-subject fluctuations in attentional control. *J Neurosci* 30:11458–11465. [CrossRef Medline](#)
- Lim SL, Padmala S, Pessoa L (2008) Affective learning modulates spatial competition during low-load attentional conditions. *Neuropsychologia* 46:1267–1278. [CrossRef Medline](#)
- Ling S, Carrasco M (2006) Sustained and transient covert attention enhance

- the signal via different contrast response functions. *Vision Res* 46:1210–1220. [CrossRef Medline](#)
- McMains SA, Somers DC (2005) Processing efficiency of divided spatial attention mechanisms in human visual cortex. *J Neurosci* 25:9444–9448. [CrossRef Medline](#)
- Olivers CN, Nieuwenhuis S (2005) The beneficial effect of concurrent task-irrelevant mental activity on temporal attention. *Psychol Sci* 16:265–269. [CrossRef Medline](#)
- Padilla ML, Wood RA, Hale LA, Knight RT (2006) Lapses in a prefrontal-extrastriate preparatory attention network predict mistakes. *J Cogn Neurosci* 18:1477–1487. [CrossRef Medline](#)
- Parasuraman R (1979) Memory load and event rate control sensitivity decrements in sustained attention. *Science* 205:924–927. [CrossRef Medline](#)
- 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* 75:426–439. [CrossRef Medline](#)
- Russell R (2009) A sex difference in facial contrast and its exaggeration by cosmetics. *Perception* 38:1211–1219. [CrossRef Medline](#)
- Sadaghiani S, Hesselmann G, Kleinschmidt A (2009) Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *J Neurosci* 29:13410–13417. [CrossRef Medline](#)
- Scalf PE, Torralba A, Tapia E, Beck DM (2013) Competition explains limited attention and perceptual resources: implications for perceptual load and dilution theories. *Front Psychol* 4:243. [CrossRef Medline](#)
- Tsal Y, Benoni H (2010) Diluting the burden of load: perceptual load effects are simply dilution effects. *J Exp Psychol Hum Percept Perform* 36:1645–1656. [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](#)
- Wieser MJ, Keil A (2011) Temporal trade-off effects in sustained attention: dynamics in visual cortex predict the target detection performance during distraction. *J Neurosci* 31:7784–7790. [CrossRef Medline](#)
- Xiao J, Hays J, Ehinger KA, Oliva A, Torralba A (2010) SUN database: large-scale scene recognition from abbey to zoo. In: *Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition*, pp 3485–3492.
- Yi DJ, Woodman GF, Widders D, Marois R, Chun MM (2004) Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nat Neurosci* 7:992–996. [CrossRef Medline](#)