

Attentional Fluctuations Influence the Neural Fidelity and Connectivity of Stimulus Representations

David Rothlein¹, Joseph DeGutis^{1,2}, and Michael Esterman^{1,3}

Abstract

■ Attention is thought to facilitate both the representation of task-relevant features and the communication of these representations across large-scale brain networks. However, attention is not “all or none,” but rather it fluctuates between stable/accurate (in-the-zone) and variable/error-prone (out-of-the-zone) states. Here we ask how different attentional states relate to the neural processing and transmission of task-relevant information. Specifically, during in-the-zone periods: (1) Do neural representations of task stimuli have greater fidelity? (2) Is there increased communication of this stimulus information across large-scale brain networks? Finally, (3) can the influence of performance-contingent reward be differentiated from zone-based fluctuations? To address these questions, we used fMRI and representational similarity analysis during a visual sustained attention task (the gradCPT). Participants ($n = 16$) viewed a series of city or mountain scenes, respond-

ing to cities (90% of trials) and withholding to mountains (10%). Representational similarity matrices, reflecting the similarity structure of the city exemplars ($n = 10$), were computed from visual, attentional, and default mode networks. Representational fidelity (RF) and representational connectivity (RC) were quantified as the interparticipant reliability of representational similarity matrices within (RF) and across (RC) brain networks. We found that being in the zone was characterized by increased RF in visual networks and increasing RC between visual and attentional networks. Conversely, reward only increased the RC between the attentional and default mode networks. These results diverge with analogous analyses using functional connectivity, suggesting that RC and functional connectivity in tandem better characterize how different mental states modulate the flow of information throughout the brain. ■

INTRODUCTION

Maintaining attention to a single task for an extended duration requires considerable cognitive effort. Consequently, sustained attention is not a constant all-or-none phenomenon but fluctuates, leading to varying degrees of stable/accurate (in-the-zone) and variable/error-prone (out-of-the-zone) performance. During experiments that tax sustained attention, attentional fluctuations result in missed targets, incorrect responses, irregular RTs, or other task-dependent lapses. Outside the lab, these lapses can have consequences that range from relatively trivial (missing one’s exit while driving) to practical (spacing out during a lecture), to profound (car accidents). To mitigate errors resulting from lapses in attention, characterizing how the peaks of these attentional fluctuations differ from the troughs could provide clues to help optimize one’s attentional state for a particular task. Given the consequences of attentional lapses, an understanding of the cognitive and neural mechanisms that support optimal attentional states has considerable potential for

real-world benefit. Complicating this understanding, however, is the fact that fluctuations in attention have multiple causes. Some of these causes may result from intrinsic factors like sleepiness/alertness, intrusions from distracting thoughts, or oscillations in the effectiveness of perceptual processing mechanisms (Buschman & Kastner, 2015) to name a few. Conversely, external sources of motivation—like performance contingent reward—can substantially boost task performance and reduce these fluctuations (e.g., Esterman, Poole, Liu, & DeGutis, 2017; Esterman, Reagan, Liu, Turner, & DeGutis, 2014).

In the current study, we sought to better understand sustained attention by examining how attentional fluctuations impact the processing of task-relevant information. The effect of attention on stimulus processing has been well studied in the transient attention and visual search literature. Transient attention contrasts with sustained attention largely in terms of the duration of the maintained attentional state—where transient attention pertains to the effects of attention on a trial-to-trial basis and sustained attention is invoked to explain fluctuations in attention over time. However, it is unknown whether fluctuations in sustained attention correspond to stimulus processing in the same way that transient attention does. Broadly, transient attention is thought to enhance stimulus processing in at least two ways: First, it is

This paper is part of a Special Focus deriving from a symposium at the 2017 annual meeting of Cognitive Neuroscience Society, entitled, “Fluctuations in Attention and Cognition.”

¹VA Boston Healthcare System, ²Harvard Medical School, ³Boston University School of Medicine

© No rights reserved. This work was authored as part of the Contributor’s official duties as an Employee of the United States Government and is therefore a work of the United States Government. In accordance with 17 U.S.C. 105, no copyright protection is available for such works under U.S. law.

thought to boost the fidelity of task-specific information processing through the construction of attentional templates (Lee & Geng, 2017; Eimer, 2014; Desimone & Duncan, 1995). Attentional templates facilitate the selection of task-relevant features (e.g., colors, shapes, locations) while inhibiting distracting or irrelevant features (Reeder, Perini, & Peelen, 2015; Peelen & Kastner, 2014). In humans, one prominent source of evidence comes from fMRI, which has been used to show that attention and task context leads to preparatory activation in domain-relevant perceptual processing regions (Esterman & Yantis, 2010). Similarly, attention increases classification accuracy of preparatory or stimulus-evoked multivoxel patterns across the task-relevant feature dimensions in visual processing regions (Cohen & Tong, 2015; Harel, Kravitz, & Baker, 2014; Peelen & Kastner, 2011) as well as frontal and parietal regions associated with more flexible, domain-general processing (Jackson, Rich, Williams, & Woolgar, 2017; Woolgar, Hampshire, Thompson, & Duncan, 2011). Furthermore, attention warps the pattern-based representational similarity of a stimulus set to fit the context of a task (Bracci, Daniels, & Op de Beeck, 2017; Nastase et al., 2017).

Second, transient attention is thought to boost performance by facilitating the communication of stimulus information throughout the large-scale neural networks that enable flexible, domain-general processing at a limited capacity (Buschman & Kastner, 2015; Peelen & Kastner, 2014; Duncan, 2013; Dehaene & Changeux, 2011; Woolgar et al., 2011; Dehaene & Naccache, 2001). Generally localized to regions in the frontal and parietal cortices, these networks are critical for representing the rules and goals of a given task (Chiu, Esterman, Han, Rosen, & Yantis, 2011; Badre, 2008), and fMRI studies have shown that the fidelity or classification accuracy of such task-rule representations are modulated by task context (Qiao, Zhang, Chen, & Egner, 2017) and reward (Etzel, Cole, Zacks, Kay, & Braver, 2016; Wisniewski, Reverberi, Momennejad, Kahnt, & Haynes, 2015). Access to such networks from stimulus-processing regions is thought to be critical for integrating and adapting incoming stimulus information to meet the demands of a given task—likely at the expense of other cognitive processes that would be competing for these computational resources (Shenhav et al., 2017). Evidence for this largely comes from primate physiology, showing that attention boosts the synchrony of neural responses between visual cortices and the FEFs (Gregoriou, Gotts, Zhou, & Desimone, 2009), as well as across frontal and parietal cortices (Buschman & Miller, 2007). In humans, examination of the dynamic functional connectivity (FC) of electrical signals measured by magnetoencephalography across a wide array of brain regions show that FC increases as task difficulty—and corresponding attentional and memory load—increases, suggesting that attention and effort result in the communication of task-relevant information into this global workspace (Wen, Yao, Liu, & Ding, 2012; Kitzbichler, Henson, Smith, Nathan, &

Bullmore, 2011; see Parks & Madden, 2013, for a review of similar findings). However, although FC and other measures of neural synchrony provide important clues as to the nature of the information transmitted across these connected regions, the representational content encoded within these signals is often indirectly inferred from the nature of the task and not directly observed from the correlated neural signals.

The information processing framework described above—emphasizing the role of attention as a process that (1) boosts the fidelity of task-relevant representations and (2) mediates the communication of these representations across large-scale brain networks—has received less consideration in the sustained attention literature. Instead, focus has been largely dedicated to identifying brain regions associated with the fluctuations in performance during sustained attention tasks (for a comprehensive review, see Fortenbaugh, DeGutis, & Esterman, 2017). Broadly, accurate responses are associated with increased activation in “task-positive” networks. For example, in a task that requires the identification of natural scenes, accurate responses would be preceded by higher activation in task-dependent stimulus processing regions like the parahippocampal place area (PPA). In addition, increased activation is observed in the dorsal attention network (DAN), which is a network composed of bilateral FEFs and intraparietal sulci (IPS; Corbetta & Shulman, 2002), that is active across many types of tasks. Unlike the DAN, accurate performance is associated with decreased activation in task-negative networks, such as the default mode network (DMN), however, the relationship between the DMN activation and performance is complex (e.g., Kucyi, Hove, Esterman, Hutchison, & Valera, 2017; Raichle, 2015). The DMN is widely believed to be responsible for generating thoughts that distract from the task at hand, increasing the likelihood of attentional lapses. Indeed, increased activation in the DMN is associated with self-reports of mind wandering, and FC between the DMN and brain regions associated with a global workspace is associated with poorer task performance within (Kucyi et al., 2017; Kucyi, Esterman, Riley, & Valera, 2016) and across (Fortenbaugh, Rothlein, McGlinchey, DeGutis, & Esterman, 2018) individuals. Therefore, a view emerges where fluctuations in sustained attention arise, in part, when task-related representations generated in perceptual processing regions (e.g., PPA) compete with task-unrelated representations generated in the DMN for access to the flexible, but limited, computational resources of a global workspace (e.g., DAN).

This study investigated three questions regarding the relationship between fluctuations in sustained attention and the processing of stimulus information. First, similar to studies of transient attention, do fluctuations in sustained attention modulate the fidelity of task-relevant features? Second, do fluctuations in attention correspond to changes in the communication of stimulus-specific information between perceptual processing regions and the

DAN while reducing the transmission of task-unrelated information between the DAN and the DMN? Finally, do extrinsic sources of motivation, such as performance-contingent reward, modulate both the fidelity and communication of stimulus information in a similar manner to intrinsically driven fluctuations of attention (i.e., the natural fluctuations between in- vs. out-of-the-zone periods) or does reward differentially influence this stream of information processing? To examine these questions, we used fMRI to measure neural activity during the gradual onset continuous performance task (gradCPT)—a visual sustained attention task that requires participants to respond to pictures of city scenes (90% of trials) and withhold responses to rare pictures of mountain scenes (10% of trials). The frequent responses (90% of trials) enabled us to use RT variability to infer individual moment-to-moment fluctuations of attention and split the trials by whether they occurred during in-the-zone (low variability) or out-of-the-zone (high variability) epochs. These zone differences have been shown to reflect intrinsic attentional fluctuations (Esterman, Rosenberg, & Noonan, 2014; Esterman, Noonan, Rosenberg, & Degutis, 2013; Rosenberg, Noonan, Degutis, & Esterman, 2013). To test the influence of externally driven motivation or effort, half of the task blocks included performance-contingent rewards.

Examining the influence—attentional fluctuations and reward have on the flow of information processing required isolating and measuring the stimulus-specific information contained within neural signals. However, stimulus-evoked neural signals are imprecise, reflecting any number task-unrelated thoughts that just happened to coincide with the presentation of a stimulus as well nonneural events such as breathing or motion. The FC between two regions is difficult to interpret for the same reasons (Simony et al., 2016). To address these issues, we adapted representational similarity analysis (RSA) and more specifically representational connectivity analysis (RCA; Kriegeskorte, Mur, & Bandettini, 2008) to compute two measures: (1) representational fidelity (RF), which quantified the cross-participant stability of the similarity structures of the set of 10 city exemplars (representational similarity matrices [RSMs]) within a particular ROI, and (2) representational connectivity (RC), which provided an analogous measure across a pair of ROIs. Specifically, RSA was used to convert the ROI- and participant-specific stimulus-evoked multivoxel patterns into the common format of similarity space (Kriegeskorte & Kievit, 2013; Aguirre, 2007; Shepard & Chipman, 1970) that can be compared directly across participants and brain regions. Importantly, both RF and RC were assessed as the interparticipant reliability of these RSMs (i.e., how similar the RSMs were across participants), ensuring that stimulus-unrelated mental and neural events, such as mind wandering, as well as participant-specific strategies or trial order effects would only act to increase the idiosyncrasy of an RSM and consequently reduce the overall RF or RC.

Using RF and connectivity analyses, we directly assessed how intrinsic (zone-based) and extrinsic (reward-based) attentional fluctuations influenced the patterns of information processing by examining how these states altered RF and RC within and across the lateral occipital (LO), PPA, DAN, and DMN. Furthermore, to compare RC with more standard measures of connectivity, we examined how the FC for each pair of brain regions changed as a function of reward or attentional zone as well. Specifically, we had four predictions: (1) Being in the zone (optimal attentional state), analogous to transient attention, should be associated with an overall increase in the fidelity of stimulus representations. (2) Being in the zone should increase overall RC across regions of the brain, particularly between stimulus processing regions like LO or PPA and networks that support more global and flexible cognitive operations like the DAN. (3) Performance-contingent rewards should boost performance by enhancing attention and therefore should be indistinguishable from intrinsic fluctuations in terms of how they related to the RF and connectivity of stimulus information. (4) Suboptimal attentional states (being out of the zone and unrewarded blocks) should be associated with increased mind-wandering behavior and, consequently, should correspond to an increase in the FC between the DAN and DMN. Critically, increases of DAN–DMN FC during suboptimal attentional states should not correspond to analogous increases of the more specific measure of RC, as the information driving these FC increases is thought to be unrelated to the task and will therefore not increase the interparticipant stability that RC measures.

METHODS

Participants

Sixteen participants (10 men; mean age = 22 years, range = 19–29 years) performed the gradual onset continuous performance task (gradCPT) during an fMRI session. Fourteen participants were right-handed, and all were considered healthy, had normal or corrected-to-normal vision, and had no reported history of major illness, head trauma, or neurological/psychiatric disorders. All were screened to confirm no metallic implants or history of claustrophobia. Drug/medication use was not explicitly assessed. The study protocol was approved by the VA Boston Healthcare System institutional review board, and all participants gave written informed consent. The data used in this study and portions of the methods have been published (Esterman et al., 2017), but the current analyses and results reported have not been published elsewhere.

Paradigm and Stimuli

The gradCPT consisted of 20 grayscale photographs of city (10 exemplars) and mountain scenes (10 exemplars), and participants were instructed to respond via button

press to frequently occurring city scenes and withhold responses to rare mountain scenes. The stimulus images were resized to a 256×256 pixel image and then cropped to appear within a circular frame (radius ~ 126 pixels) and projected to a screen placed at the back of the scanner at a 800×600 screen resolution. In the gradCPT, the scene images were presented in a unique pseudorandom order for each participant with the following constraints: 10% of trials displayed mountain scenes, the remaining 90% were city scenes, and the same exemplar could not repeat on adjacent trials. Scene images gradually transitioned from one to the next in a linear pixel-by-pixel interpolation, with each transition occurring over 800 msec. The new image increased in clarity, whereas the old image decreased in clarity. The task instructions emphasized response accuracy without reference to speed. However, as a new image replaced the previous image every 800 msec, there was an implicit response deadline in the task.

Each 8-min task run was divided into alternating 1-min rewarded and unrewarded blocks, which were differentiated by a continuous color border (green for rewarded, blue for unrewarded). To have the background colors be more intuitive and avoid confusion, we chose “green” for money-rewarded blocks in all participants rather than counterbalancing green and blue colors. This yielded 4 min of each block type per run. During rewarded blocks, participants earned \$0.01 for correctly pressing to city scenes and \$0.10 for correctly withholding a response to mountain scenes. However, if a participant failed to press to a city scene, they would lose \$0.01, and if a participant incorrectly pressed to a mountain scene, they would lose \$0.10. During the unrewarded blocks, no money could be gained or lost. As has been shown, these reward contingencies produced reliable improvements in accuracy and RT variability in studies using the gradCPT (Esterman, Reagan, et al., 2014) and the present data set (Esterman et al., 2017).

Procedure

A MacBook Pro with MATLAB (Mathworks, Inc.) delivered stimuli to a rear-facing projector. Participants viewed the stimuli on a rear projector screen via a mirror inside of the MRI bore. Responses were collected using a fiber-optic button box. Before scanning, participants were given a 1-min practice of the task. Inside of the scanner, participants completed three to five runs of the task (13 participants completed five runs, two completed four runs, and one completed three runs). Participants were informed of their accrued reward after each run (mean = \$4.84, range = \$2.93–6.63) and were told in advance that two runs would randomly be selected for bonus payment at the end of the experiment; however, the two highest runs were actually selected as the additional bonus payment. An anatomical scan (MPRAGE sequence) and a resting-state scan (not used in this

study) were also acquired and interspersed to provide breaks between task runs, such that no more than two runs were done consecutively.

MRI Acquisition and Preprocessing

Scanning was performed on a 3T Siemens MAGNETOM Trio system equipped with a 32-channel head coil at the VA Boston Neuroimaging Research Center for Veterans. Structural volumes were acquired via an MPRAGE sequence with the following parameters: echo time = 3.32 msec, repetition time (TR) = 2530 msec, flip angle = 7° , acquisition matrix = 256×256 , in-plane resolution = 1.0 mm^2 , 176 sagittal slices, slice thickness = 1.0 mm. All structural images were processed using standard analysis of functional imaging (AFNI) pipelines (Cox, 1996).

Each gradCPT functional run included 248 whole-brain volumes acquired using an EPI sequence with the following parameters: TR = 2000 msec, echo time = 30 msec, flip angle = 90° , acquisition matrix = 64×64 , in-plane resolution = $3.0 \times 3.0 \text{ mm}^2$, 33 oblique slices aligned to the anterior and posterior commissures, slice thickness = 3 mm with a 0.75-mm gap. Following acquisition, the functional scans were processed using AFNI and custom written routines in MATLAB (Mathworks, Inc.). Preprocessing steps included slice-time correction, motion correction using a six-parameter, rigid body, least squares alignment procedure, spatial smoothing with a 6-mm full-width at half-maximum (FWHM) Gaussian kernel, automated coregistration and normalization of anatomical and functional volumes to Talairach space, and scaling of functional data set values to percent signal change. During preprocessing, automated segmentation algorithms generated three masks from the Talairached anatomical volume. These included masks covering gray matter, white matter, and cerebral spinal fluid (CSF). Average time series from the functional scan were extracted from eroded white matter and CSF masks to use as nuisance regressors. Average global signal time series from the functional scans were extracted to use as a nuisance regressor for the FC analyses.

To control for unrelated sources of variance in the task-evoked BOLD signal, residuals from initial whole-brain general linear model (GLM) analyses (using AFNI's 3dDeconvolve function) were extracted. For each GLM, the preprocessed signal from each run of the gradCPT was corrected using a set nuisance regressors—specifically, white matter and CSF signal time courses; six head-motion parameter time courses; linear, quadratic, and cubic trends (-polart 3 in AFNI); and, for the FC analyses, a global signal time course. Time points with a framewise displacement greater than 0.5 mm (along any dimension) were censored.

ROIs

Analyses were carried out across a set of ROIs selected using the NeuroSynth platform (Yarkoni, Poldrack,

Nichols, Van Essen, & Wager, 2011). NeuroSynth constructs keyword-specific activation likelihood maps from coordinates published in a database of studies. To account for the visual processing of scene stimuli, bilateral PPA and LO ROIs were selected from NeuroSynth maps¹ generated using the keyword “scenes” and “lateral occipital,” respectively. The PPA was included to account for domain-specific scene processing (Epstein & Kanwisher, 1998), whereas the LO was included to account for more domain-general object processing (Grill-Spector & Malach, 2004). Attention and DMN ROIs consisted of the maps generated by the keywords “attention” and “DMN,” respectively. The attention network clusters consisted of the regions surrounding bilateral IPS and bilateral FEFs, and these clusters closely corresponded to what is commonly referred to as the DAN, and we will adopt this terminology. The DMN consisted of the ventromedial prefrontal cortex (vmPFC), posterior cingulate cortex (PCC), and bilateral supramarginal gyri. All NeuroSynth maps were converted from Montreal Neurological Institute to Talairach space using AFNI’s 3dWarp function.

Categorizing Attentional State (In vs. Out of the Zone)

Attentional state was categorized based on continuous fluctuations in RT variability throughout a given task run. RTs to each trial were determined using an iterative algorithm that assigned button presses to individual trials. RTs were calculated relative to the beginning of each image onset. Thus, an RT of 800 msec would indicate that the current image was 100% coherent, whereas shorter RTs indicate that the current image was still in the process of transitioning in from the previous image. Following the analysis first outlined in Esterman et al. (2013), this analysis inferred instantaneous attentional state by using trial-by-trial variations in RT to calculate the variance time course (VTC). VTCs were computed for each run using the correct responses to the nontarget city scenes. First, RTs were z -transformed to normalize values within participants, and the absolute value of the z -scores was calculated so that higher values indicated greater deviations from the mean, including both very slow and very fast RTs, whereas lower values indicated RTs closer to the mean of the run. Values for trials without responses (omission errors and correct omissions to target mountain scenes) were linearly interpolated from the RTs of the two surrounding trials. A smoothed VTC was then computed using a Gaussian kernel of 20 trials (~16 sec) FWHM integrating information from the surrounding 20 trials with a weighted average. From the smoothed VTC, a median split was used to divide performance into low- and high-variability epochs. This median split was done separately for the trials associated with each city exemplar to ensure the number of repetitions for each exemplar was matched across zone. Based on

previous work, these 4-min epochs were referred to as being “in the zone” (low RT variability) and “out of the zone” (high RT variability).

RF and RC Analyses

Carrying out the RSA-based RC analyses required (1) estimating the set of ROI-specific activation patterns in response to each city exemplar split by either in-the-zone versus out-of-the-zone or rewarded versus unrewarded trials, (2) computing RSMs from the city-evoked activation patterns in each ROI (split by zone or reward), (3) constructing a set of nuisance RSMs (nRSMs) to remove uninteresting sources of variability, and (4) computing the RF and the RC for each ROI and ROI pair, respectively, and examining how this fidelity and connectivity changes as a function of zone/reward.

Estimating Exemplar Specific Activation Patterns

The residuals from the preprocessing GLMs were concatenated across runs and served as input to two GLMs that estimated the activation patterns for each of the 10 city exemplars split across rewarded and unrewarded blocks (GLM_{reward}) or in-the-zone and out-of-the-zone epochs (GLM_{zone}). The GLM_{reward} consisted of 22 regressors: one for each of the 10 city exemplars during either rewarded and unrewarded blocks (20 conditions), one for all of the omission errors (failures to respond to city scenes), and one for all of the mountain trials. The GLM_{zone} had a similar set up with 22 regressors, splitting each of the 10 city exemplars by zone instead of reward. The experimental regressors for each GLM were created by convolving a stick-based impulse corresponding to the point of maximum stimulus clarity (800 msec after trial onset) with a gamma hemodynamic response function. For each GLM type, resulting beta-maps were converted to t maps—one map for each of the 20 experimental conditions (10 city exemplars \times 2 reward/zone conditions). Only the t maps from the 20 city exemplars were used in the RSA analyses. These exemplar-specific t maps served as the activation values for the MVPA-RSA. MVPA-RSA was carried out using in house code run in MATLAB (Mathworks, Inc.). The AFNI MATLAB library was used to integrate AFNI output with MATLAB.

Observed RSMs

City-evoked activation patterns (from the exemplar-specific t maps) were used to generate ROI-based RSMs (Figure 1). The activation patterns derived from the stimulus-specific t maps across the set of voxels contained within an ROI were reshaped into a set of vectors (one for each exemplar). An RSM was constructed by computing pairwise (Pearson) correlations for each possible vector pairing and placing the correlation coefficient

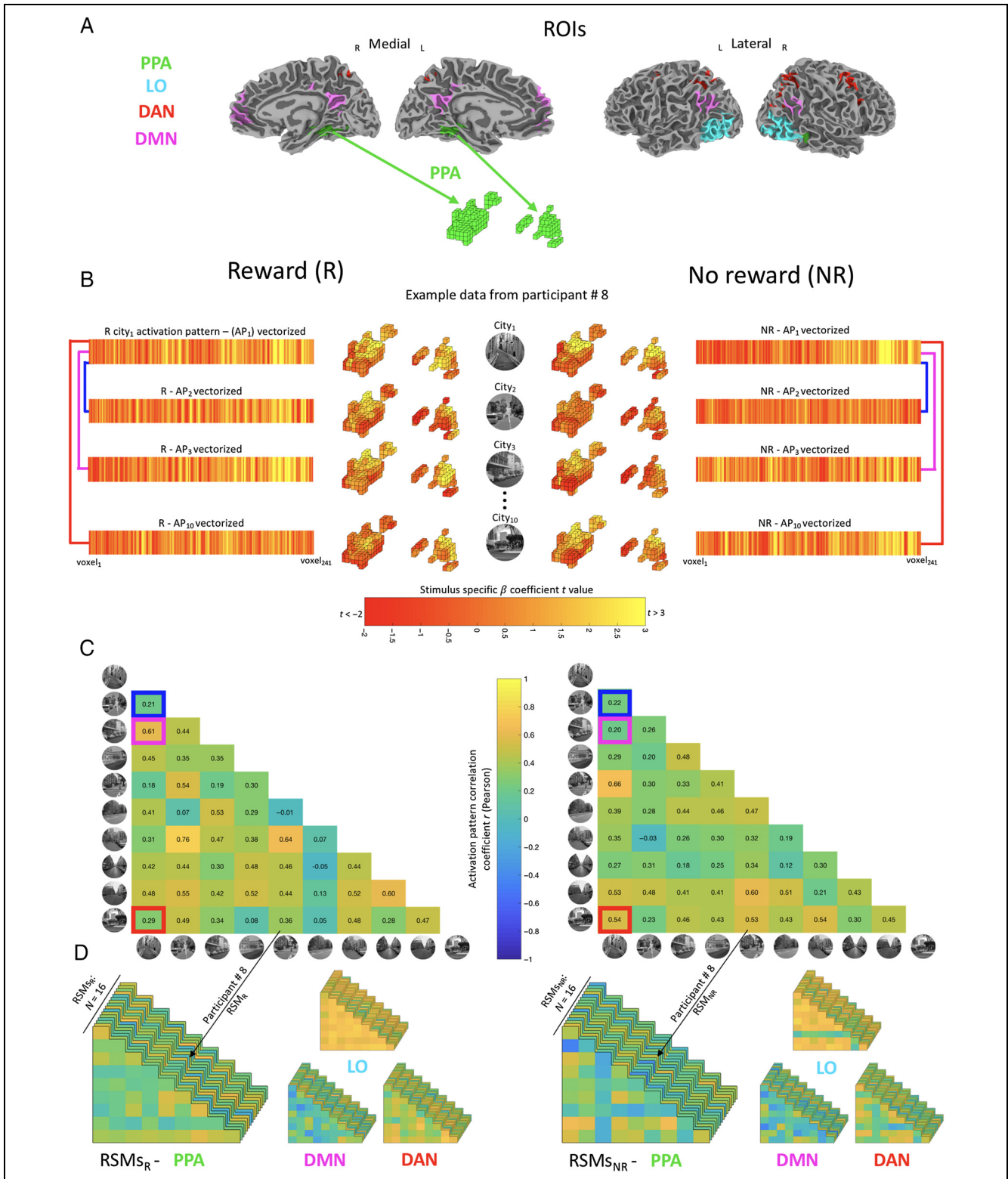


Figure 1. Constructing RSMs. (A) The four ROIs consisted of the bilateral LO cortex in cyan, the PPA in green, the DMN in magenta, and the DAN in red. From an example participant, an RSM is constructed from the activation pattern within the PPA. The exemplar-specific activation patterns are shown in B. Activation patterns (APs), composed of beta t value, were generated by modeling the BOLD response evoked for each of the 10 city exemplars split by whether they occurred during in- versus out-of-the-zone epochs or rewarded versus unrewarded blocks—the latter is displayed above. To construct an RSM as shown in C, each activation pattern was vectorized and an array of correlation coefficients was computed that quantified the pairwise similarity of the set of 10 activation patterns. Example pairs of activation patterns are connected with colored lines in B, and the corresponding r values are highlighted in the RSMs in C using boxes of the same color. Because the RSM is symmetric across the diagonal, only the unique values (excluding the diagonal of the matrix) were included. (D) For each condition, a set of 16 RSMs (one per participant) was computed from each ROI (shown: reward, no reward; not shown: full, in the zone, and out of the zone).

in the corresponding cell within the RSM. Each RSM was triangular (due to symmetry across the diagonal), and the cells in the diagonal and lower triangle were excluded. For each participant and ROI, four observed RSMs (rewarded, unrewarded, in the zone, and out of the zone) were computed.

Nuisance RSMs

Inherent to any fMRI analysis is the issue of differentiating the signal of interest from noise. Here we created two types of nRSMs that quantified the expected consequences of potentially confounding information directly onto each participant's RSMs. The nuisance similarity structures were removed from the ROI-based RSMs by running a linear regression analysis where the nRSMs (along with a constant) served as regressors for a given observed RSM and the residuals from this regression served as the RSM that was used for the RF and connectivity analyses.

The first nRSM modeled the expected impact of condition (city exemplar) differences in ROI-based mean activation levels on the RSM computed from the same ROI. The concern is that multivoxel pattern similarity methods may be encoding univariate effects within the similarity structure. Therefore, seemingly multivariate and distributed information contained in RSMs may simply be repackaging univariate activation or, more specifically, how univariate activation interacts with the variance structure across voxels (Davis et al., 2014). To account for this, we constructed a univariate activation nRSM that assigned greater similarity to pairs of city conditions that had similar mean activity levels across the set of voxels within an ROI. Specifically, at the participant level within an ROI, we calculated the mean activation for each of the 10 city exemplars. We then constructed an nRSM that indexed, for each pair of city conditions, the absolute value of the activation differences of the mean activation between the two city conditions. Finally, to ensure that larger values indicated more similarity, we multiplied these nRSMs by -1 and linearly scaled the resulting values to have a range of 0–1 where 1 indicated the means were identical and 0 indicated the means were maximally different within the set. By removing the predicted similarity structure due to differences in univariate activation across conditions, the resulting residuals were more likely to reflect multivariate and multidimensional representations.

The second nRSM addressed the concern that the similarity structure did not reflect representations of the city exemplars but instead reflected information that was incidental to, but correlated with, each city exemplar—namely RT (Todd, Nystrom, & Cohen, 2013). Specifically, because certain city exemplars elicited slower RTs on average (Rothlein et al., 2018), the RSMs could reflect these performance-based differences. Therefore, we constructed an RT-based nRSM that predicted that exemplars

that had similar RTs should have elicited similar activation patterns. Specifically, at the participant level, we computed the mean RT for each city exemplar, and in an identical manner to the activation-based nRSM, we converted the set of 10 mean RTs to an RSM with a range of 0–1, where 1 indicated the RTs were identical and 0 indicated that the RTs were maximally different within the set of 10 means. By accounting for RT-based similarity, the resulting residual RSMs were more likely to directly reflect representations of the city exemplars and not reflect information that correlated with each exemplar (e.g., the motor consequences of the RT differences or stimulus difficulty).

RF and RC

The procedures for computing RF and RC are shown in Figure 2. RF was quantified as the interparticipant similarity of the RSMs from a given ROI, and RC was quantified as the correlation between the RSMs derived from different ROIs (Kriegeskorte et al., 2008) where larger correlations indicated greater RF/RC. To ensure our measures of RF/RC uniquely measured task-related representations of exemplar-specific properties and not task-unrelated thoughts or other events that happened to coincide with the different conditions, all measures of RF and RC were quantified in a cross-validated manner. Specifically, for a given ROI pair (ROI₁ and ROI₂), we used an interparticipant reliability procedure that entailed the following: (1) selecting RSMs from half of the participants (8 out of 16) from ROI₁ (split_A) and selecting RSMs from the remaining (nonoverlapping) participants (split_B) from ROI₁ (for RF) and ROI₂ (for RC); (2) averaging together the set of eight RSMs from each split and each ROI in the pair²; (3) correlating (Pearson) the two group average RSMs (45 values in each RSM). Specifically, RF was the correlation coefficient between the group average RSM from split₁ and split₂ within ROI₁, and RC was the correlation coefficient between the group average RSM from split₁ ROI₁ and split₂ ROI₂; (4) repeating this split-half correlation procedure 1000 times and recording the correlation coefficients from each split. We quantified RF within ROI₁ and RC between ROI₁ and ROI₂ as the average of these split-half correlation coefficients. Importantly, we computed RF and RC separately for RSMs derived from rewarded blocks and unrewarded blocks (RF_{Reward}, RF_{NoReward}, RC_{Reward}, and RC_{NoReward}) and quantified the influence of reward (RF_{RewardDiff} and RC_{RewardDiff}) as RF_{Reward} – RF_{NoReward} and RC_{Reward} – RC_{NoReward}, respectively. We computed RF_{ZoneDiff} and RC_{ZoneDiff} in an analogous manner, subtracting RF_{OutZone} and RC_{OutZone} from RF_{InZone} and RC_{InZone}, respectively. Significance for each RF/RC value and the RF/RC_{Diff} values were assessed using a label-scramble permutation analysis where the above split-half procedure was repeated (10000 repetitions) after scrambling the city exemplar labels (no replacement) for each RSM.³ Null permutation distributions for RF and

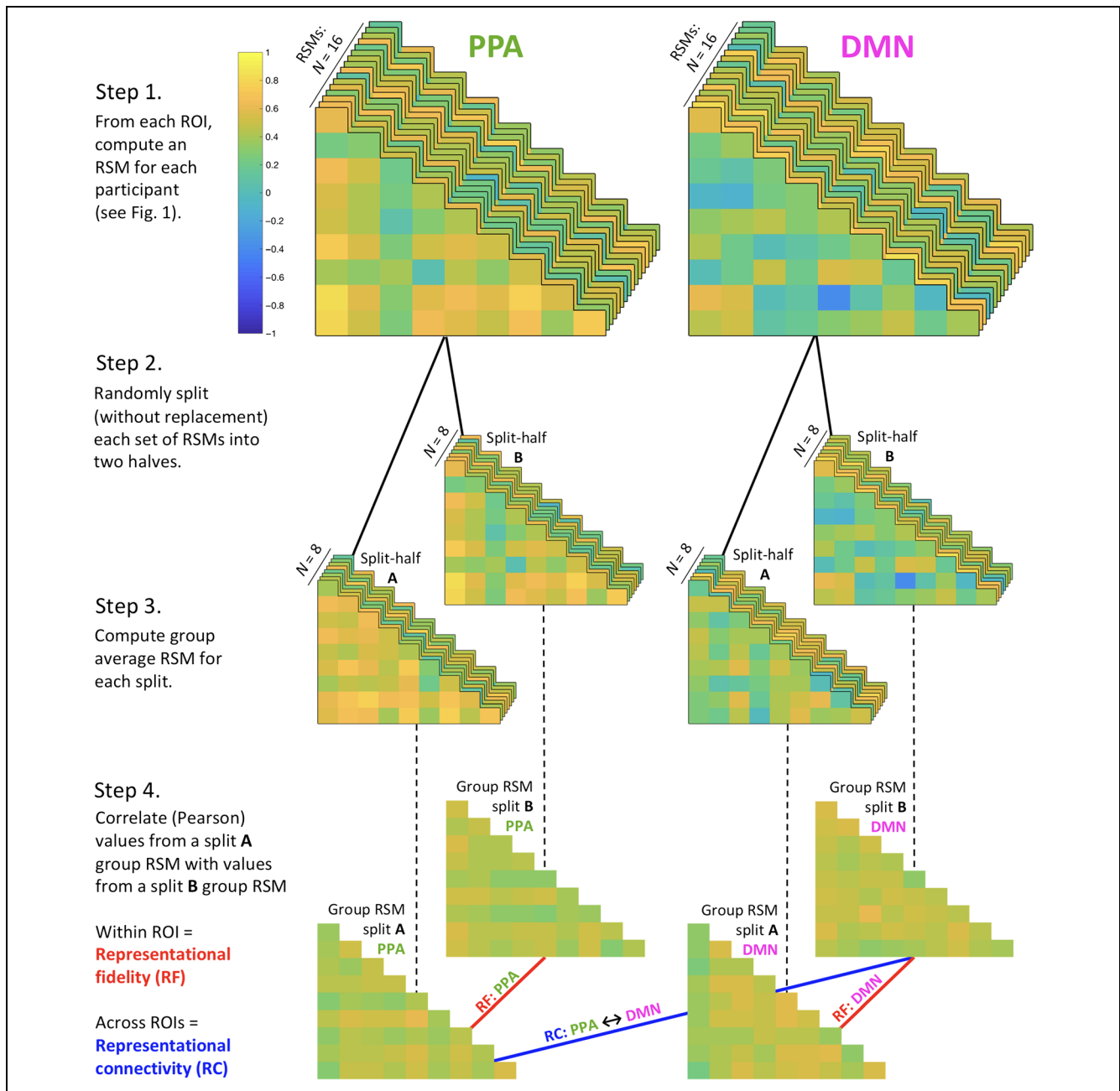


Figure 2. Computing RF and RC. Each step in this process is described in the figure and in the text. RF and RC are estimated by repeating Steps 2–4 1000 times and averaging the resulting distribution of within (RF) and across (RC) ROI correlation coefficients (1000 each).

RC values from each of the four conditions, as well as RF_{Diff} and RC_{Diff} values reflecting zone and reward differences, were formed from this procedure, and significance was assessed by the rank of the real mean values relative to their permutation distributions.

FC Analyses

The residuals from the preprocessing GLM (including the global signal regressor) from each participant were concatenated across task runs and used to compute the

functional time series for the FC analyses. Specifically, the set of time series from each voxel within the ROI were averaged together—removing the first six and the final two TRs from each run. The removed TRs corresponded to fixation periods (after adjusting for a BOLD signal delay of two TRs or 4000 msec). These ROI-based time series were then split into separate in-the-zone and out-of-the-zone or rewarded and unrewarded time series (adjusting for a two-TR hemodynamic response delay). For each participant, correlation matrices were computed from the reward (FC_{Reward}) and no-reward ($FC_{NoReward}$)

or in-the-zone (FC_{In}) and out-of-the-zone (FC_{Out}) ROI-based time series, and all resulting correlation coefficients underwent Fisher z transformations. The influence of reward on FC for a given ROI pair was quantified as $FC_{Reward} - FC_{NoReward}$ or FC_{Diff} , and the analogous FC_{Diff} value was computed for zone. One-sample t tests (average $FC_{Diff} \neq 0$) were used to evaluate the influence of reward and zone on FC.

RESULTS

We sought to explore how the RF and RC of task-relevant stimulus information was influenced by both intrinsic attentional fluctuations as well as performance-contingent rewards. We examined RF and RC within and across perceptual processing regions (bilateral LO and PPA), the DAN, and the DMN. Specifically, we predicted that being in the zone would correspond to (1) an increase in overall RF across the task-positive regions (LO, PPA, and DAN), (2) an increase in the RC from the LO and PPA to the DAN, and (3) that task-contingent reward would have a similar effect on RF and RC as being in the zone. Although not discussed in detail here, both being in the zone and reward were associated with increased accuracy (fewer commission errors and greater d'), indicating the validity of exploring information processing differences across these states. Results pertaining to the behavioral effects of zone and reward on this data set were reported in detail in Esterman et al. (2017). To start, the overall RF and RC without splitting by attentional zone or reward were computed. All results reported as significant survived a family-wise Bonferroni correction for multiple comparisons, unless otherwise noted.

Overall RF and RC

The full data set was examined for RF and RC before splitting by reward and zone (see Figure 3A). This revealed the greatest fidelity in the LO (RF = 0.53, 95% CI [0.28, 0.79], $p < .001$) with significant RF in the PPA as well (RF = 0.39, 95% CI [0.10, 0.66], $p = .002$) while no RF in the DAN (RF = 0.07, 95% CI [-0.24, 0.35], $p = .327$) or DMN (RF = 0.06, 95% CI [-0.22, 0.34], $p = .338$). As Figure 3B shows, the right LO had greater RF than the left LO (RF = 0.60 and RF = 0.48, respectively). Likewise, the right PPA had greater RF than the left (RF = 0.43 and RF = 0.28, respectively). In summary, without factoring zone or reward, the representational similarity structures computed from the LO and PPA (biased toward the right in both) were reliable across participants. This suggests that these regions represented stimulus information about each city exemplar, that the features represented were consistent across participants, and that the fidelity of these representations was the greatest in LO but also high in PPA. The presence of RF in regions known to process visual information provides an important validation

on this method. Specifically, reliable stimulus information was contained within the RSMs across participants despite short trial ISIs, gradual stimulus transitions, and the high degree of overall stimulus complexity and similarity (all photographs of city scenes).

In addition to examining the degree of representational content within an ROI (RF), we also examined the overall RC across ROIs. This revealed significant RC between the LO and PPA (RC = 0.24, 95% CI [0.05, 0.43], $p = .005$), LO and DAN (RC = 0.27, 95% CI [0.08, 0.46], $p = .006$), and PPA and DAN (RC = 0.19, 95% CI [-0.03, 0.39], $p = .033$); however, the latter RC value did not survive correction for multiple comparisons. Figure 3C and D shows the RC of all the clusters within and across the three ROIs. Although the cluster-level results were intended to be interpreted more qualitatively to show how distributed these effects were across each cluster, we report which connections survived the family-wise Bonferroni correction for multiple comparisons.

The Influence of Attentional Zone on RF

Based on the literature into the effects of transient attention on stimulus information processing, we predicted that optimal (in the zone) attentional states should be associated with increased RF within perceptual processing brain regions (LO, PPA, and DAN to a lesser extent). We split the data from each run into in-the-zone and out-of-the-zone trials where in-the-zone trials occurred when the participant's trial-to-trial RTs were most consistent and out-of-the-zone trials occurred when the RTs were most variable. We computed RF separately from in-the-zone and out-of-the-zone epochs (Figures 4A and 5A). Within the perceptual processing ROIs, we found a high degree of RF while participants were in the zone (LO: $RF_{in-zone} = 0.58$, 95% CI [0.31, 0.87], $p < .001$; PPA: $RF_{in-zone} = 0.35$, 95% CI [0.08, 0.61], $p = .003$), whereas none of the ROIs had significant RF while participants were out of the zone. Directly testing the difference in RF across zones, the LO had significantly greater $RF_{in-zone}$ versus $RF_{out-zone}$ ($RF_{zone-diff} = 0.46$, 95% CI [0.08, 0.86], $p = .012$), whereas the PPA had a marginally significant numerical trend ($RF_{zone-diff} = 0.27$, 95% CI [-0.11, 0.63], $p = .079$). We failed to find such zone differences in the DAN ($RF_{zone-diff} = 0.18$, 95% CI [-0.22, 0.56], $p = .169$) and DMN ($RF_{zone-diff} = 0.06$, 95% CI [-0.38, 0.32], $p = .522$). These results fit with our predictions, suggesting that attentional fluctuations influenced the fidelity of representations in a similar manner to the influences of transient attention. One notable exception, however, was the lack of an effect of attentional zone on the RF in the DAN.

The Influence of Attentional Zone on RC

As shown above, attentional zone influenced the RF of stimulus features within the LO and PPA. We also predicted that zone would have a substantial influence on

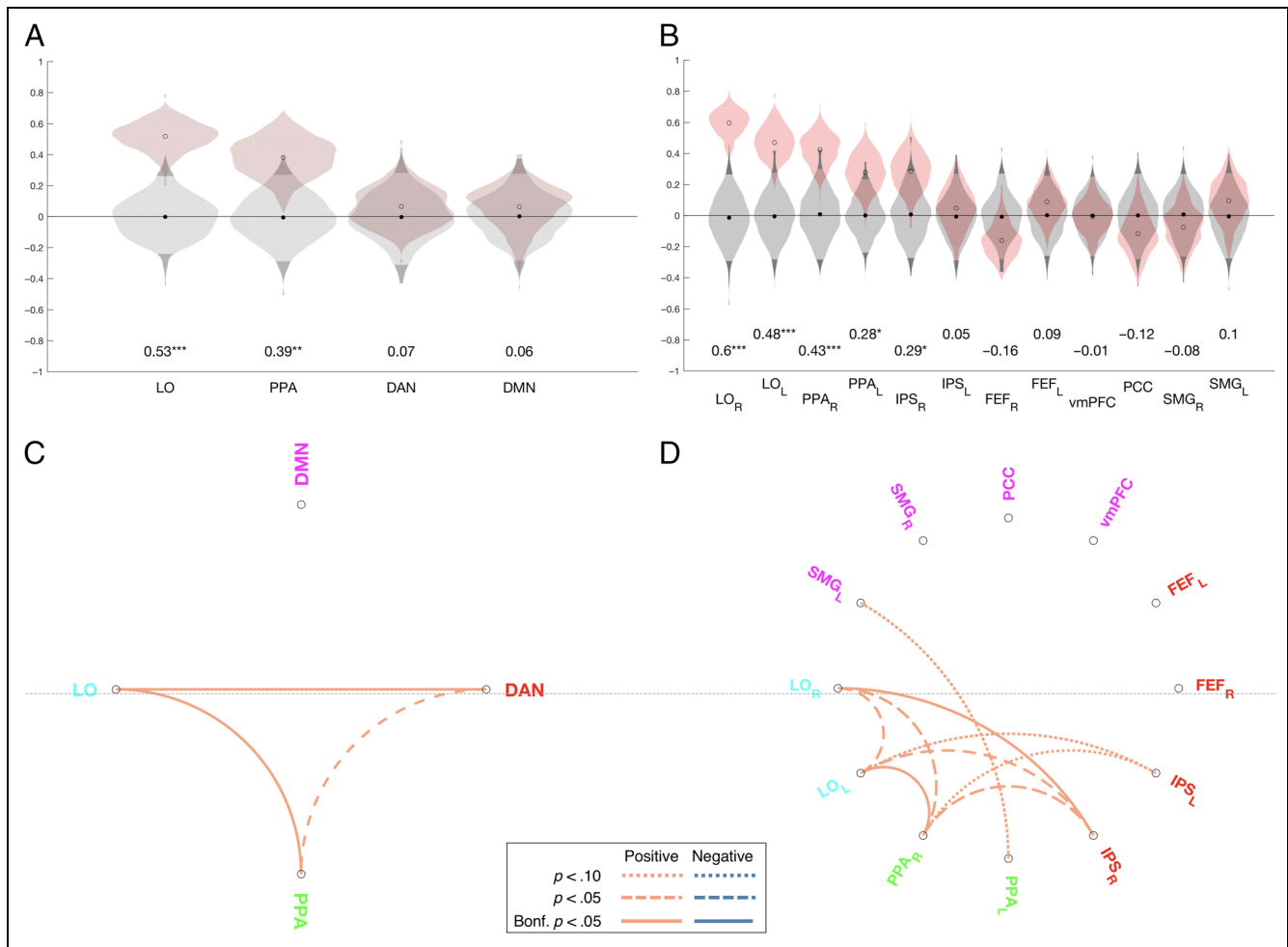


Figure 3. RF and connectivity using full data set. Violin plots of the RF are shown by ROI network in A and by ROI cluster in B. The permutation-based null distribution is depicted in gray with the top and bottom 2.5% of values in a darker shade of gray and the mean value of the null distribution plotted as a black dot. The red-toned distribution depicts the variance from the 1000 split-half reliability values—the average of which is plotted as an open circle and indicates that ROI's RF value. The RF value is printed below the plot. Circular graphs of the RC are shown by network in C and by cluster in D. Circular graphs made in MATLAB using circularGraph 2.0. * $p < .05$, ** $p < .01$, *** $p < .005$.

RC between these perceptual processing regions and more flexible, domain-general networks like the DAN. Confirming our prediction (see Figures 4B and 5B), we observed significant $RC_{in-zone}$ between the LO and DAN ($RC_{in-zone} = 0.27$, 95% CI [0.07, 0.47], $p = .005$), between the PPA and DAN ($RC_{in-zone} = 0.31$, 95% CI [0.13, 0.50], $p < .001$), as well as between the LO and PPA ($RC_{in-zone} = 0.34$, 95% CI [0.14, 0.52], $p < .001$). None of the $RC_{out-zone}$ values, however, reached significance. Directly comparing the differences ($RC_{zone-diff} = RC_{in-zone} - RC_{out-zone}$) revealed significant $RC_{zone-diff}$ between the PPA and DAN ($RC_{zone-diff} = 0.38$, 95% CI [0.11, 0.65], $p = .004$) and the LO and PPA ($RC_{zone-diff} = 0.36$, 95% CI [0.08, 0.63], $p = .002$). The $RC_{zone-diff}$ between the LO and DAN trended in the same direction ($RC_{zone-diff} = 0.27$, 95% CI [0, 0.53], $p = .021$) but failed to survive the correction for multiple comparisons. It is worth noting, however, that strong a priori predictions as well as analogous findings between the PPA and DAN boost confidence in this result.

To briefly summarize, we observed that being in the zone was associated with an increase in the fidelity of stimulus representations within perceptual processing regions (LO and marginally PPA) as well as an increase in the RC between these two regions. Furthermore, the RC between the DAN and both LO and PPA increased when participants were in the zone, suggesting that sustaining attention increases the communication of stimulus-specific information between perceptual processing regions and the more flexible and task-oriented processing regions in the DAN.

The Influence of Reward on RF and Connectivity

Being in the zone was associated with increases in RF in perceptual processing regions (LO and PPA) and increases in RC between these regions and the DAN. Here we predicted that reward influenced information processing in a similar manner to the intrinsic fluctuations

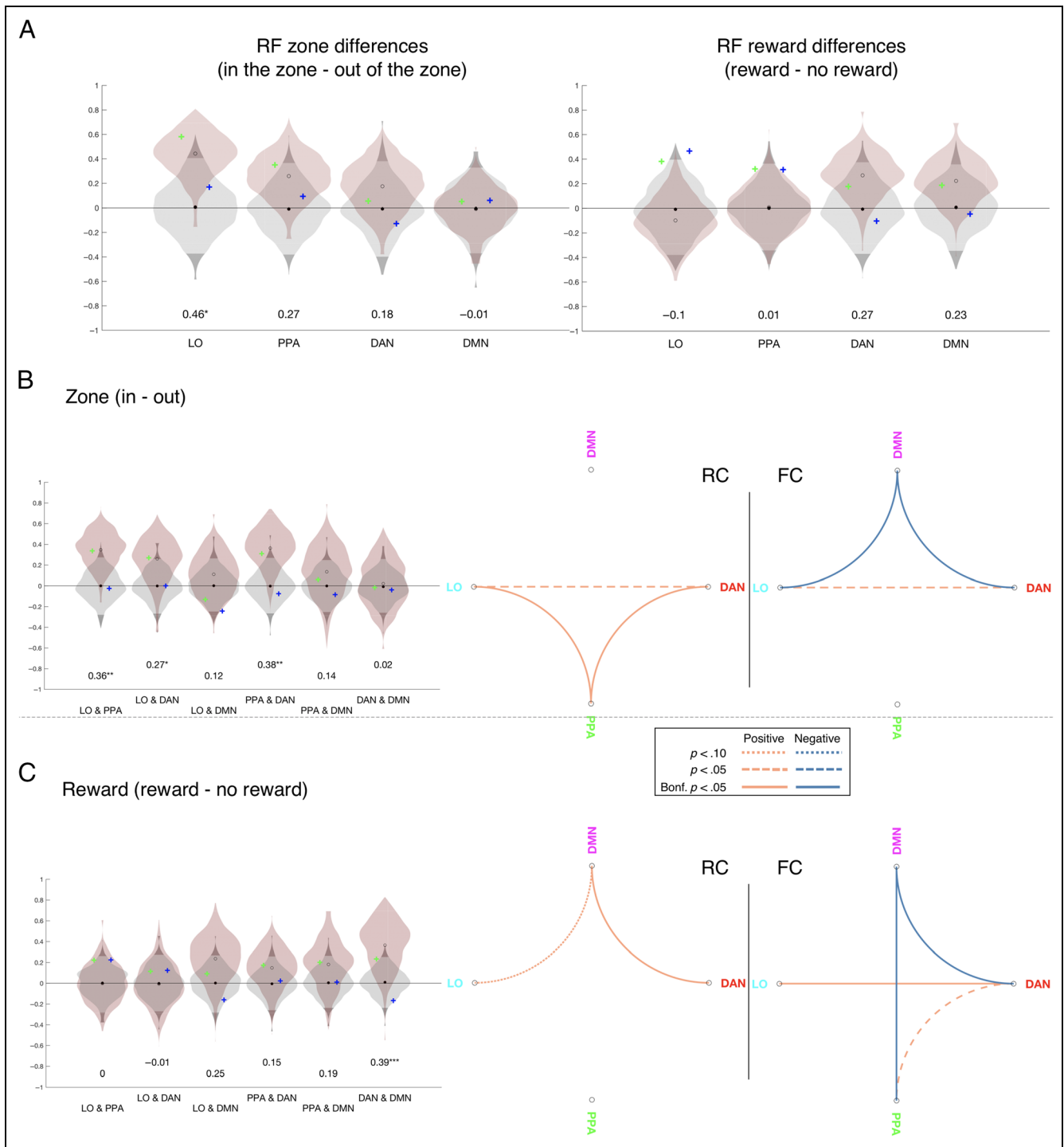


Figure 4. Network-level differences across zone and reward in RF, RC, and FC. (A) Violin plots depicting zone (left) and reward (right) differences in RF from each network ROI. The permutation-based null distribution of RF differences is depicted in gray with the top and bottom 2.5% of values in a darker shade of gray and the mean value of the null distribution plotted as a black dot. The red-toned distribution depicts the variance from the 1000 split-half reliability values—the average of which is plotted as an open circle and indicates that ROI's RF difference value. The RF difference value is also printed below the plot. The green cross on the left side of each violin plot indicates the in-the-zone/reward RF value, and the blue cross on the right indicates the out-of-the-zone/no-reward RF value. (B) A summary of the network-level results for RC and FC. The violin plots depict the RC zone difference values for each network pair and the circular graph depicts the same results. The rightmost circular plot depicts the zone difference values for FC. The same set of plots are depicted for reward differences in C. Circular graphs made in MATLAB using circularGraph 2.0. * $p < .05$, ** $p < .01$, *** $p < .005$.

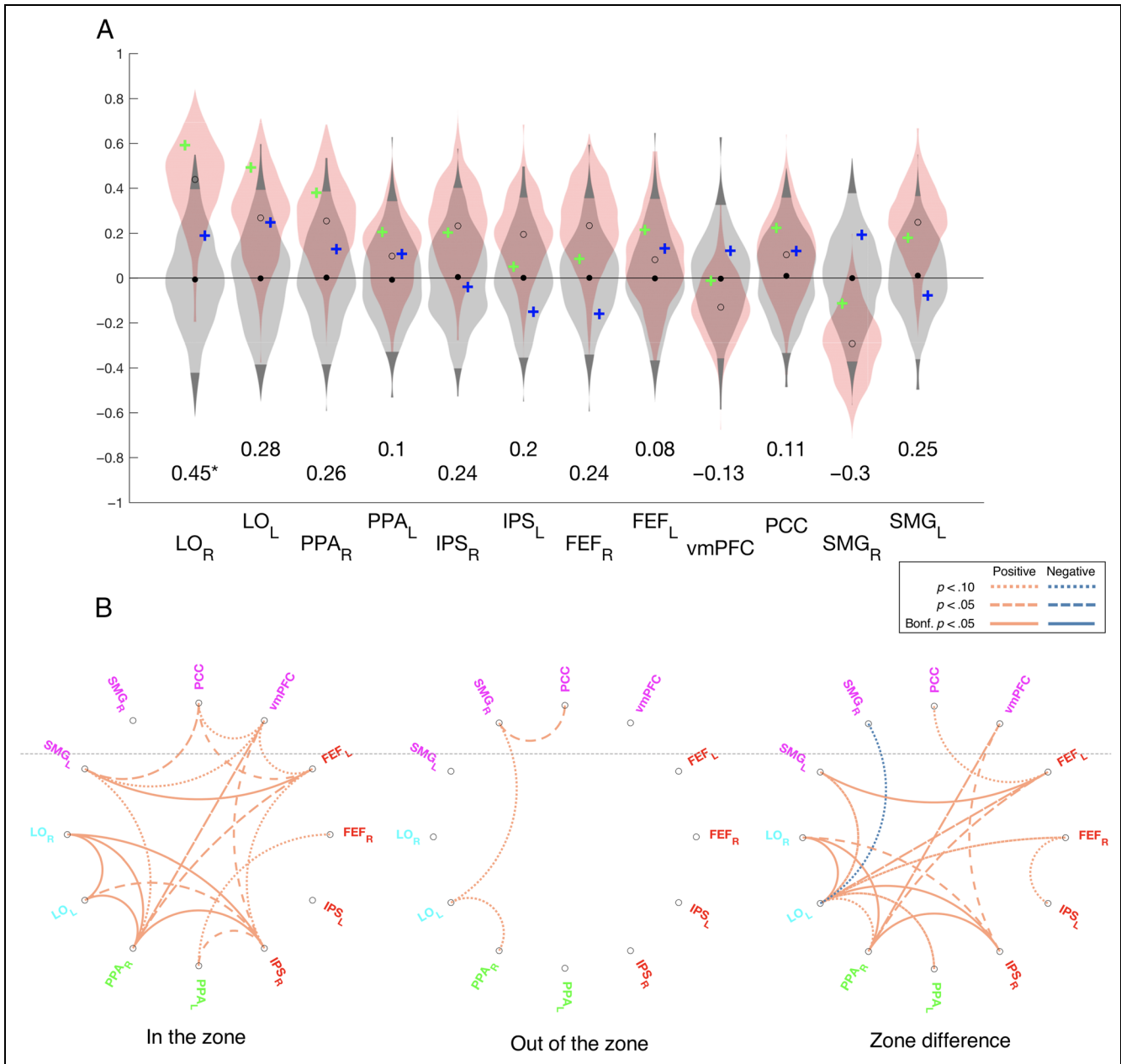


Figure 5. Cluster-level differences in attentional zone for representational fidelity (A) and representational connectivity (B). Violin plots of the RF zone differences (zone difference = in the zone – out of the zone) are shown by cluster in A. The permutation-based null distribution for zone difference is depicted in gray with the top and bottom 2.5% of values in a darker shade of gray and the mean value of the null distribution plotted as a black dot. The red-toned distribution depicts the variance from the 1000 split-half reliability values—the average of which is plotted as an open circle and indicates that ROI’s RF zone difference value. The RF difference value is also printed below the plot. The green cross on the left side of each violin plot indicates the in-the-zone RF value, and the blue cross on the right indicates the out-of-the-zone RF value. Circular graphs (B) of the RC are shown by cluster from in-the-zone trials, out-of-the-zone trials, and the zone differences. Circular graphs made in MATLAB using circularGraph 2.0. * $p < .05$.

(zone). We split the data from each run into rewarded and unrewarded blocks and computed RF and RC after this split. As Figures 4A and 6A show, the LO and PPA ROIs had significant RF_{reward} (LO: $RF_{\text{reward}} = 0.38$, 95% CI [0.10, 0.66], $p < .001$; PPA: $RF_{\text{reward}} = 0.32$, 95% CI [0.06, 0.58], $p = .004$); however, these regions also had significant RF during the unrewarded blocks (LO: $RF_{\text{no-reward}} = 0.47$, 95% CI [0.10, 0.66], $p < .001$; PPA:

$RF_{\text{no-reward}} = 0.31$, 95% CI [0.06, 0.58], $p = .004$). Directly testing the differences between RF_{reward} and $RF_{\text{no-reward}}$ revealed only nonsignificant numerical trends toward greater RF_{reward} in the DAN ($RF_{\text{reward-diff}} = 0.27$, 95% CI [-0.10, 0.65], $p = .082$).

Examining RF identified that reward failed to show the same influence as zone on RF in the LO and PPA and suggested a larger effect in the DAN. Surprisingly, as

Figures 4C and 6B show, the effect of reward further dissociated when examining RC. Unlike $RC_{in-zone}$, we observed RC_{reward} between the PPA and DMN ($RC_{reward} = 0.17$, 95% CI $[-0.01, 0.36]$, $p = .03$) as well as the DAN and DMN ($RC_{reward} = 0.17$, 95% CI $[-0.01, 0.36]$, $p = .03$), although neither survived correction for multiple comparisons. Critically, we found the $RC_{reward-diff}$ between the DAN and DMN was robust ($RC_{reward-diff} =$

0.39 , 95% CI $[0.15, 0.64]$, $p < .001$). In addition, numerical trends revealed that reward increased RC between the LO and DMN ($RC_{reward-diff} = 0.25$, 95% CI $[-0.03, 0.52]$, $p = .036$), as well as the PPA and DMN ($RC_{reward-diff} = 0.19$, 95% CI $[-0.05, 0.45]$, $p = .073$).

To briefly summarize, reward was associated with marginally significant increases in the fidelity of stimulus representations within the DAN. Interestingly, RC

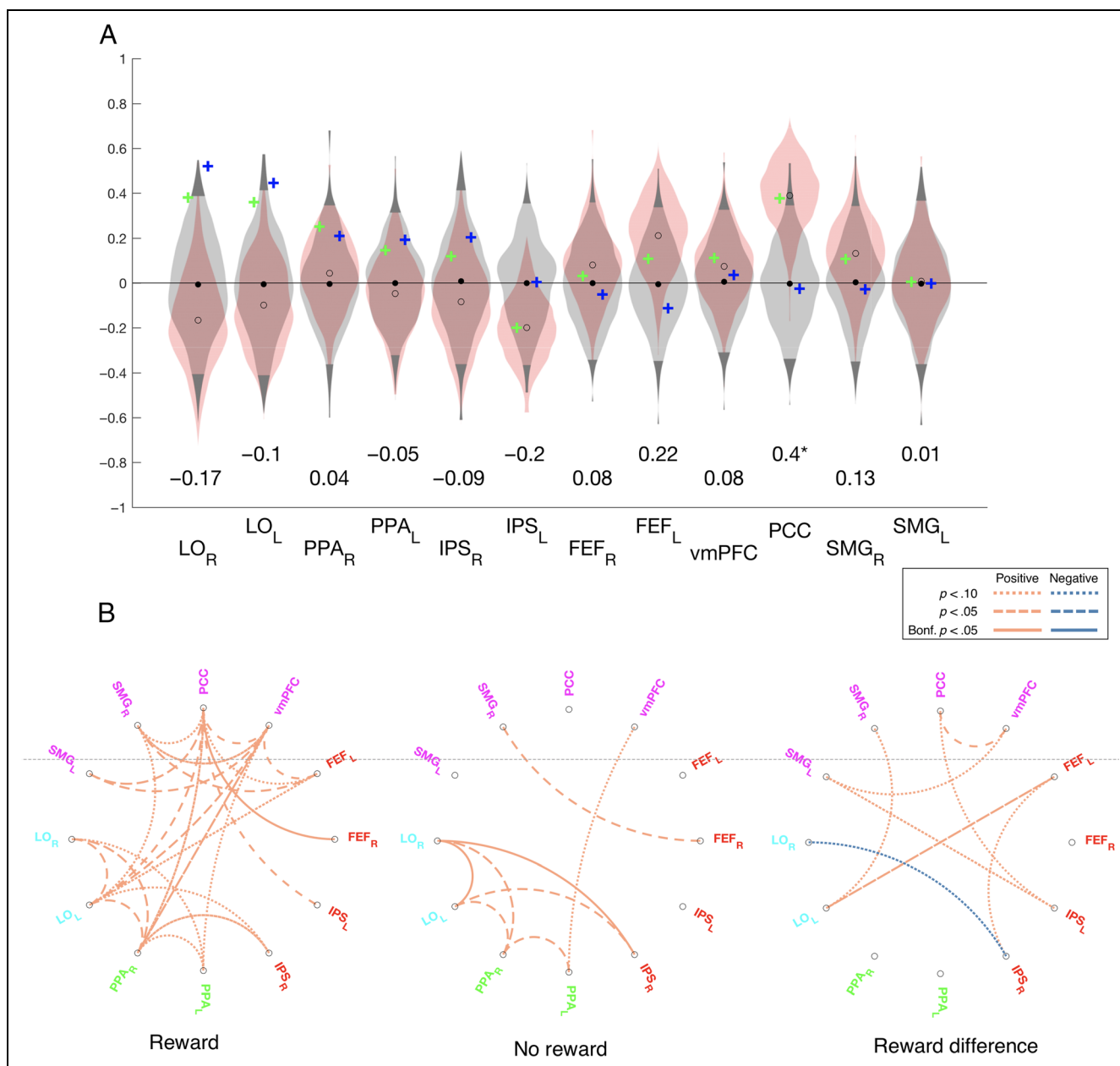


Figure 6. Cluster-level differences in reward for representational fidelity (A) and representational connectivity (B). Violin plots of the RF reward differences (reward difference = reward – no reward) are shown by cluster in A. The permutation-based null distribution for reward difference is depicted in gray with the top and bottom 2.5% of values in a darker shade of gray and the mean value of the null distribution plotted as a black dot. The red-toned distribution depicts the variance from the 1000 split-half reliability values—the average of which is plotted as an open circle and indicates that ROI’s RF reward difference value. The RF difference value is also printed below the plot. The green cross on the left side of each violin plot indicates the rewarded RF value, and the blue cross on the right indicates the no-reward RF value. Circular graphs (B) of the RC are shown by cluster from rewarded trials, no-reward trials, and the reward differences. Circular graphs made in MATLAB using circularGraph 2.0. * $p < .05$.

between the all three ROIs and the DMN increased during rewarded relative to unrewarded blocks—with smaller effects between the perceptual processing regions and the DMN but a robust effect between the DAN and DMN. This provided a striking contrast to the effects of attentional zone, which was characterized by RC increases between perceptual processing regions and from these regions to the DAN.

The Influence of Reward and Zone on FC

Examining RC across the LO, PPA, DAN, and DMN enabled us to quantify the degree of communication of stimulus-specific information between these regions. This revealed that reward and zone were associated with distinct changes in RC. However, univariate measures like FC (time series correlations) have traditionally been used to infer the communication of information between ROIs. Therefore, to compare our RC results with these more traditional measures, we examined how the FC between these regions was modulated by reward and zone. Because FC could reflect changes in the communication of task unrelated information, we predicted that the DMN—which is thought to be critical in generating task unrelated thoughts—should be sensitive to changes in the communication of this information in manner that RC would not be. Specifically, being in the zone and rewarded blocks should correspond to decreases in the FC between the DMN and the other ROIs.

To analyze the influence of zone and reward, we examined how FC changed between the four ROIs while participants were in versus out of the zone (Figure 4B) and rewarded versus unrewarded (Figure 4C). First, being in the zone and rewarded were characterized by increased FC between perceptual processing regions and the DAN with both zone and reward showing this effect for the FC between the LO and DAN (zone: $FC_{diff} = 0.06$, 95% CI [0.01, 0.10], $t(15) = 2.82$, $p = .013$; reward: $FC_{diff} = 0.07$, 95% CI [0.04, 0.10], $t(15) = 4.46$, $p < .001$) and for reward (but not zone) this effect was between the PPA and DAN as well ($FC_{diff} = 0.06$, 95% CI [0.01, 0.10], $t(15) = 2.62$, $p = .019$). Second, being in the zone and rewarded were characterized by a general decrease in connectivity between the DMN and the other three ROIs with both zone and reward showing this effect between the DAN and DMN (zone: $FC_{diff} = -0.07$, 95% CI [-0.11, -0.03], $t(15) = -3.90$, $p < .001$; reward: $FC_{diff} = -0.08$, 95% CI [-0.12, -0.04], $t(15) = -3.99$, $p < .001$). Furthermore, this effect was shown between the LO and DMN (zone: $FC_{diff} = -0.08$, 95% CI [-0.14, -0.03], $t(15) = -3.41$, $p = .004$; reward (marginal): $FC_{diff} = -0.05$, 95% CI [-0.10, 0], $t(15) = -1.96$, $p = .069$). Finally, reward (but not zone) showed this effect between the PPA and DMN ($FC_{diff} = -0.07$, 95% CI [-0.10, -0.03], $t(15) = -3.56$, $p = .003$). These results, in conjunction with the RC results, suggested that the increased FC between the perceptual processing

regions and the DAN likely reflected stimulus-specific information, whereas FC involving the DMN mostly reflected the communication of distracting, stimulus-unrelated information.

DISCUSSION

The goal of this study was to investigate how attentional fluctuations—due to intrinsic (attentional zone) or extrinsic causes (reward)—relate to both the fidelity and connectivity of exemplar-specific stimulus representations (pictures of city scenes) within and across the LO, PPA, DAN, and DMN. We had four predictions: (1) Being in the zone (optimal attentional state), analogous to studies of transient attention, should be associated with an overall increase in the fidelity of stimulus representations. (2) Being in the zone should increase overall RC across regions of the brain, particularly between visual processing regions like LO or PPA and networks that support more global and flexible cognitive operations like the DAN. (3) Performance-contingent rewards should boost performance by enhancing attention and therefore should be indistinguishable from intrinsic fluctuations in terms of the fidelity and connectivity of stimulus information. (4) The FC between the DAN and DMN should increase when participants are not rewarded or out of the zone, reflecting an increase in the communication of task-unrelated information that would conversely decrease our measures of RC or RF.

With regard to the first prediction, we found that being in the zone was associated with increased RF in LO and, to a lesser extent, PPA, supporting the notion that sustained and transient attention enhance stimulus processing in a similar manner—at least in visual processing regions. Specifically, sustaining attention to a task critically depends on constructing and maintaining attentional templates (Rothlein et al., 2018). Such templates would identify task-relevant features and act to facilitate the identification of such features—perhaps even distorting or warping representations along select feature dimensions to boost detection sensitivity or mitigate interference from distracting visual information (Geng, DiQuattro, & Helm, 2017; Nastase et al., 2017; Peelen & Kastner, 2014).

Maintaining such templates boosts the fidelity of stimulus representations as we and others have observed. The content of such templates likely depends on the nature of the task—selecting the set of features, regardless of the level of abstraction, that maximizes the efficiency of the visual search or classification (Rothlein et al., 2018; Hout & Goldinger, 2015; Schmidt & Zelinsky, 2009; Vickery, King, & Jiang, 2005). Previous research using the same task and stimuli found that target–nontarget similarity of exemplar-specific pixel intensity features best explained stimulus-level differences in RT and accuracy, outperforming more abstract category attribute-based features (Rothlein et al., 2018). This could explain why

LO, believed to represent complex visual features and shapes that compose the content of a scene, had an overall greater fidelity and larger fidelity difference across attentional state than the PPA—which is believed to represent the more global and structural attributes of visual scenes that enable the categorization of novel scene exemplars (Park, Brady, Greene, & Oliva, 2011). Further investigation into the representational content contained within these templates is certainly merited. However, because the RSMs only included city exemplars, we can establish that the granularity of the information contained within these regions was specific enough to establish an exemplar-specific representational geometry that was reliable across participants. As such, Boolean representations indicating stimulus category (city vs. mountain) or response decision (press vs. withhold) would be insufficient to explain either the RF or RC results as they would be identical across the city exemplars.⁴

We also found, consistent with our prediction, that being in the zone was associated with an increase in the RC between perceptual processing regions (LO and PPA) and from both of these regions to the DAN.⁵ In other words, when participants were in the zone, the similarity structure of the set of city stimuli were relatively synchronized across these three regions—suggesting stimulus information was being communicated between the LO, PPA, and DAN. Although expected, this finding was novel as, to our knowledge, no one has examined how attention—transient or sustained—has influenced the communication of stimulus-specific information. The DAN has conventionally been implicated in the top-down modulation of stimulus processing; however, the DAN overlaps heavily with many neurotopographical descriptions of a global workspace or multiple demand networks (Fedorenko, Duncan, & Kanwisher, 2013; Dehaene & Changeux, 2011), suggesting the DAN also plays a critical role in actively processing and broadcasting incoming stimulus information. Indeed, multivoxel pattern analyses have shown that the IPS—the parietal nodes within the DAN network—represents stimulus information maintained in working memory (Xu, 2017). Our results support the view that the DAN, particularly the right IPS, actively processes and communicates stimulus information and that this processing is associated with optimal attention. However, ambiguities regarding the directionality of the RC results limit the ability to differentiate top-down versus bottom-up processing contributions of the DAN. Characterizing the role of the DAN with respect to attention and working memory remains an active area of research (e.g., Scimeca, Kiyonaga, & D’Esposito, 2018; Gayet, Paffen, & Van der Stigchel, 2017; Xu, 2017; Gazzaley & Nobre, 2012).

When considering the influence of reward on RF and RC, we predicted that reward would benefit performance through top-down control over the maintenance of attention (Thomson, Besner, & Smilek, 2015; Kurzban, Duckworth, Kable, & Myers, 2013) or enhanced repre-

sentations of the task more generally (Etzel et al., 2016). According to this prediction, increased effort would not change where or how stimulus information was processed other than showing information processing patterns akin to those characterizing the in-the-zone attentional state (greater LO/PPA RF and great LO/PPA-DAN RC). In other words, fluctuations in attention would directly influence how information is processed and effort would directly modulate the mechanisms underlying attention. To our surprise, reward influenced RC and, to a lesser extent, RF in a manner that was distinct and nearly complementary to the influence of intrinsic fluctuations (i.e., zone). We characterized this shift in two ways: first, reward shifted the locus of information processing from the perceptual processing regions (LO and PPA) to the domain-general, flexible processing networks in the DAN and interestingly the DMN. To compare with the zone contrasts, although being in the zone increased the RF and the RC within and between the LO and PPA, reward—failing to influence the LO or PPA—had an analogous effect on the DAN and DMN. Second, reward shifted RC with the perceptual ROIs from the DAN to the DMN. Specifically, although intrinsic fluctuations modulated the degree to which stimulus information was communicated between the perceptual processing regions and DAN, reward shifted this communication toward the DMN. As can be seen in Figure 6, these effects were largely driven by the PCC. The PCC, which is thought of as the core of the DMN, is often implicated in spontaneous thought (Kucyi et al., 2016); however, it overlaps with the retrosplenial cortex, a component of the scene processing network implicated in orienting and locating oneself in a scene (Marchette, Vass, Ryan, & Epstein, 2014). Furthermore, the retrosplenial cortex/PCC has been implicated in the retrieval of autobiographical memories (Vann, Aggleton, & Maguire, 2009) as well as the comprehension of complex narrative structures (Simony et al., 2016). Such studies provide ample evidence that the DMN is capable of representing and flexibly processing a wealth of information.

We speculate that reward boosts performance by shifting task-related processes from a mode that is more automatic and unconscious to a mode that is more deliberate and self-aware. This likely occurs at the expense of task-unrelated thoughts (e.g., mind wandering) that are thought to be generated in the DMN. It may be that stimulus information in the DMN is beneficial because it “replaces” the processing that generates task-unrelated thoughts and therefore inhibits the intrusion of such thoughts into the global, task-positive networks. Alternatively, the DMN could share the load of task-directed stimulus processing with the DAN—effectively boosting the task-directed processing capacity. Comparing the FC results with the RC results further supports this characterization. When participants were in suboptimal attentional states (either unrewarded or out of the zone), the DMN acted as an FC hub with FC between the DMN and the other ROIs (LO, PPA, and particularly DAN)

increasing. Conversely, RC between these regions decreased when participants were not rewarded. Importantly, although reward was associated with decreased FC between the DAN and DMN, it was associated with increased RC between these regions. Taken in tandem, this suggests that reward decreased the overall level of communication between the DAN and DMN. However, the proportion of that the communication between the DAN and DMN that specifically pertains to stimulus information actually increased during rewarded blocks. The fact that reward had the opposite effect on FC and RC highlights the challenges involved in interpreting FC as well as the benefit of including information-based approaches with interparticipant validation to more directly observe cognitive processes from neural activity (e.g., Simony et al., 2016).

Examining the RF and RC across attentional states provides novel evidence toward understanding how fluctuations in sustained attention (as well as reward) reflect changes in the constant flow of information coming through the senses and how this information interacts with internally generated thoughts. Although other measures like FC and synchronized oscillations have been shown to be associated with acts of transient and sustained attention, the underlying information contained within these signals is ambiguous. Representational connectivity analysis or RCA (Henriksson, Khaligh-Razavi, Kay, & Kriegeskorte, 2015; Kriegeskorte et al., 2008) is one of a growing number of measures of connectivity that uses multivoxel patterns to measure information-based correlations across brain regions (Ito et al., 2017; Li, Richardson, & Ghuman, 2017; Anzellotti, Caramazza, & Saxe, 2016; Coutanche & Thompson-Schill, 2013; Walther, 2013; Chai, Walther, & Beck, 2009), and each of these approaches provides complementary information regarding the complex nature of neural communication (Anzellotti & Coutanche, 2018). The biggest advantage of this particular implementation of RCA was that it isolated a single facet of the neural signals across two regions and provided a relatively unambiguous measure of connectivity (see Walther, 2013, for a similar approach). Specifically, RC and RF were computed using the conservative approach of interparticipant reliability, which ensured that whatever was shared in the RSM must have been specific to the set of city exemplars used in this experiment. Indeed, cross-region RSM correlations derived from the same participant (intraparticipant RC) were much greater, but the factors driving these correlations could have included task-unrelated thoughts that coincided with stimulus presentations, signal carryover from adjacent trials, and global signal fluctuations from motion or other noise-related causes (Hebart & Baker, 2017; Cai, Schuck, Pillow, & Niv, 2016). Although some components of within-participant correlations could be interesting, like idiosyncratic components of the city representations (Lee & Geng, 2017; Charest, Kievit, Schmitz, Deca, & Kriegeskorte, 2014), for the purposes of this study, we

felt the inclusion of such components did not merit the reduction of interpretability that would come with intra-participant RC.

Consistent with our emphasis on maximizing the interpretability of the results, we opted to narrow the scope of the RCA analyses to a limited set of ROIs that had a strong basis in the attention and perceptual processing literature. However, recent research by Rosenberg et al. (Rosenberg, Hsu, Scheinost, Constable, & Chun, 2018; Rosenberg, Finn, Scheinost, Constable, & Chun, 2017; Rosenberg et al., 2016) has demonstrated that FC across a highly distributed set of nodes (i.e., individual connectomes or neural fingerprints)—many of which do not correspond to conventional attention areas—can predict individual differences in attentional ability with surprising success. This supports the contention that attention should be thought of as a diverse set of interacting cognitive processes that emerge from brain-wide neural dynamics as opposed to the sole function a few canonical networks. Although this characterization is likely accurate, our study focused on how attention relates to the representation and transmission of stimulus-specific information, and this greatly constrained the scope of the attention-related signals we examined. Therefore, many attention-related neural signals that might be captured by FC-based approaches—although critical for a full characterization of attention—would have registered as noise by this RC-based approach. In other words, because FC-based approaches are potentially sensitive to a diverse array of attention related phenomenon (including stimulus processing), it makes it challenging to map specific cognitive processes to specific connections. Because this RC-based approach measured signal that was specific to stimulus representations, we could more confidently attribute our results to the cognitive processes that related to the fidelity and communication of stimulus information (e.g., the construction and maintenance of attentional templates or the broadcasting of stimulus representations into a global workspace). Furthermore, because this RCA methodology is relatively novel, we limited our analyses to brain regions where we had a reasonable expectation that stimulus information would be both represented and modulated by attention. In addition to providing a solid connection to the vast literature relating attention and reward to these networks, the fact that our results largely conformed to the literature provides important evidence in favor of the methodological soundness of the RCA approach. A promising direction for future research would be to examine more extensive parcellations by computing both FC- and RC-based connectomes to compare and contrast how these measures predict intra- and interindividual differences in attention. Such an approach could provide a richer characterization of the functional roles of the highly distributed network of attention-related regions and connections (see Li et al., 2017, for a similar suggestion).

It is important to note that, despite the increase in interpretability that these RCA analyses afford, the results from this study are still susceptible to a number of ambiguities. First, the way RC is computed does not specify anything about the direction of communication or even the causal relationship of the connectivity between two regions. For example, increased RC between the LO and DAN could reflect information transfer that is bottom-up (LO to DAN), top-down (DAN to LO), or some combination of the two. Furthermore, there may be no direct communication between these regions, but perhaps the observed connectivity was mediated by communication to both of these regions from the PPA. Alternatively, RC could result from two regions converging on a similar representational geometry even if there was never any direct or indirect communication between them. Resolving these issues is not insurmountable and future studies could integrate RCA analysis with statistical methods like effective connectivity or mediation analyses to carefully parse out questions of directionality and causality. An additional concern is that the differences in RC and RF that we observed across zone and reward may have been due to eye blinks or saccades instead of directly corresponding to changes in cognitive processing. Indeed, previous research has shown that eye blinks increase in rate and duration as vigilance decreases across the duration of a sustained attention task (McIntire, McKinley, Goodyear, & McIntire, 2014). On this account, eye blinks or eye movements could result in degraded early visual representations, reducing the fidelity and downstream connectivity. Although consistent with the differences in RF and RC due to attentional state (zone differences), this account fails to explain the differences in RC due to reward because both the reward and no-reward conditions showed significant RF in visual processing regions with little difference between the two conditions. Therefore, the observed reward-related difference in RC between DMN and DAN could not be readily explained by degraded visual input due to blinking or saccades. Nonetheless, future studies would benefit from collecting simultaneous eye-tracking data to explicitly assess this possibility.

In this study, we investigated fluctuations in sustained attention—whether due to intrinsic (uncontrolled) or extrinsic (reward-based) causes—from an information processing framework. By using a novel implementation of representational similarity and connectivity analyses (Kriegeskorte et al., 2008), we showed that optimal attentional states (being in the zone) were associated with an increase in the RF within cortical regions associated with visual processing (LO and PPA) as well as an increase in the communication of stimulus-specific information between these regions and the DAN. Reward, on the other hand, reflected increased RC between the higher-level DAN and DMN networks, suggesting reward boosted performance in a manner that was differentiable from natural fluctuations in attention. This study also highlighted the benefit of explicitly examining the underlying representa-

tions that could be driving the FC between brain regions. Such investigations will continue to provide a richer understanding of the intricacies of information processing in the human brain.

Acknowledgments

This work was supported by the U.S. Department of Veteran Affairs through a Clinical Science Research and Development Career Development Award (Grant Number 1IK2CX000706-01A2) to M. E.

Reprint requests should be sent to David Rothlein, VA Boston Health Care System Jamaica Plain Campus, 150 S. Huntington Ave., Boston, MA 02130, or via e-mail: david.rothlein@gmail.com.

Notes

1. All NeuroSynth maps have been corrected for family-wise error by applying a false discovery rate correction with $q < 0.01$. All maps underwent an additional voxel-wise thresholds $Z > 1.96$ and cluster size threshold of at least 100 voxels.
2. Averaging together RSMs helps distill any shared representational structure. As the number of RSMs averaged together increases, so too does the accuracy with which the group RSM approximates the real shared similarity structure if one exists (Nili et al., 2014). This assumes, however, that the distribution of the similarity values in each cell of the RSM stems from a single population value (unimodal) and a symmetric noise distribution (e.g., Gaussian). In this context, the split-half interparticipant similarity procedure provides the best estimate of the true RC effect sizes.
3. It is important to note that, when doing such a label-scramble permutation on an RSM, the full symmetric RSM must be used and the same permuted label reassignment must be applied to the columns and rows of the RSM.
4. Scene category or response decision representations could have been continuous instead of Boolean—namely, a given stimulus could have had a continuous city-ness and mountain-ness value where the largest value established the category assignment and response decision. In such a case, each city stimulus could fall within a unique position along these feature dimensions, and therefore, a reliable representational geometry could be established that does not specifically identify a city exemplar but places it relative to the two category assignments. Indeed, in a previous study, we have shown that exemplar-level differences in RT indicated how mountain-like each city stimulus is (Rothlein et al., 2018). However, our inclusion of the RT-based nRSM—which predicted that exemplars that had similar RTs would also have similar activation patterns—would account for much of this variance, making it an unlikely explanation for the observed RF or RC results.
5. It is worth noting that, when comparing the RF and RC results across in/out of the zone periods, some incongruity emerges wherein the DAN does not appear to have reliable RSMs as made evident by its low overall RF and minimal influence from zone. However, it does appear to contain reliable stimulus information in its correlation to the RSMs from the LO or PPA as shown from RC between the DAN and these regions, particularly when the participants were in the zone (greater RC between LO/PPA and DAN). One possible explanation is that—due to the large size and likely inclusion of irrelevant voxels—RSMs constructed from the DAN were quite noisy. In support of this, the cluster-based analysis revealed overall fidelity in the right IPS but none of the other DAN clusters. In addition, the LO and PPA were quite reliable across participants.

Therefore, instead of correlating two noisy DAN-derived RSMs as was the case when computing RF, RC entailed correlating the noisy DAN RSM with a cleaner LO or PPA RSM. This noise reduction could explain why RC revealed stimulus processing in the DAN while RF did not.

REFERENCES

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, *35*, 1480–1494.
- Anzellotti, S., Caramazza, A., & Saxe, R. (2016). Multivariate pattern connectivity. *BioRxiv Preprint*. <https://doi.org/10.1101/046151>.
- Anzellotti, S., & Coutanche, M. N. (2018). Beyond functional connectivity: Investigating networks of multivariate representations. *Trends in Cognitive Sciences*, *22*, 258–269.
- Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends in Cognitive Sciences*, *12*, 193–200.
- Bracci, S., Daniels, N., & Op de Beeck, H. (2017). Task context overrules object- and category-related representational content in the human parietal cortex. *Cerebral Cortex*, *27*, 310–321.
- Buschman, T. J., & Kastner, S. (2015). From behavior to neural dynamics: An integrated theory of attention. *Neuron*, *88*, 127–144.
- Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*, 1860–1862.
- Cai, M. B., Schuck, N. W., Pillow, J. W., & Niv, Y. (2016). A Bayesian method for reducing bias in neural representational similarity analysis. *Neural Information Processing Systems*, *29*, 1–10.
- Chai, B., Walther, D. B., & Beck, D. M. (2009). Exploring functional connectivity of the human brain using multivariate information analysis. *Psychology*, *22*, 1–9.
- Charest, I., Kievit, R. A., Schmitz, T. W., Deca, D., & Kriegeskorte, N. (2014). Unique semantic space in the brain of each beholder predicts perceived similarity. *Proceedings of the National Academy of Sciences, U.S.A.*, *111*, 14565–14570.
- Chiu, Y.-C., Esterman, M., Han, Y., Rosen, H., & Yantis, S. (2011). Decoding task-based attentional modulation during face categorization. *Journal of Cognitive Neuroscience*, *23*, 1198–1204.
- Cohen, E. H., & Tong, F. (2015). Neural mechanisms of object-based attention. *Cerebral Cortex*, *25*, 1080–1092.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*, 201–215.
- Coutanche, M. N., & Thompson-Schill, S. L. (2013). Informational connectivity: Identifying synchronized discriminability of multi-voxel patterns across the brain. *Frontiers in Human Neuroscience*, *7*, 15.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*, 162–173.
- Davis, T., Larocque, K. F., Mumford, J., Norman, K. A., Wagner, A. D., & Poldrack, R. A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage*, *97*, 271–283.
- Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron*, *70*, 200–227.
- Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*, 1–37.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Duncan, J. (2013). The structure of cognition: Attentional episodes in mind and brain. *Neuron*, *80*, 35–50.
- Eimer, M. (2014). The neural basis of attentional control in visual search. *Trends in Cognitive Sciences*, *18*, 526–535.
- Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*, 598–601.
- Esterman, M., Noonan, S. K., Rosenberg, M., & Degutis, J. (2013). In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. *Cerebral Cortex*, *23*, 2712–2723.
- Esterman, M., Poole, V., Liu, G., & DeGutis, J. (2017). Modulating reward induces differential neurocognitive approaches to sustained attention. *Cerebral Cortex*, *27*, 4022–4032.
- Esterman, M., Reagan, A., Liu, G., Turner, C., & Degutis, J. (2014). Reward reveals dissociable aspects of sustained attention. *Journal of Experimental Psychology: General*, *143*, 2287–2295.
- Esterman, M., Rosenberg, M. D., & Noonan, S. K. (2014). Intrinsic fluctuations in sustained attention and distractor processing. *Journal of Neuroscience*, *34*, 1724–1730.
- Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, *20*, 1245–1253.
- Etzel, J. A., Cole, M. W., Zacks, J. M., Kay, K. N., & Braver, T. S. (2016). Reward motivation enhances task coding in frontoparietal cortex. *Cerebral Cortex*, *26*, 1647–1659.
- Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *110*, 16616–16621.
- Fortenbaugh, F. C., DeGutis, J., & Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Annals of the New York Academy of Sciences*, *1396*, 70–91.
- Fortenbaugh, F. C., Rothlein, D., McGlinchey, R., DeGutis, J., & Esterman, M. (2018). Tracking behavioral and neural fluctuations during sustained attention: A robust replication and extension. *Neuroimage*, *171*, 148–164.
- Gayet, S., Paffen, C. L. E., & Van der Stigchel, S. (2017). Visual working memory storage recruits sensory processing areas. *Trends in Cognitive Sciences*, *22*, 189–190.
- Gazzaley, A., & Nobre, A. C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences*, *16*, 129–135.
- Geng, J. J., DiQuattro, N. E., & Helm, J. (2017). Distractor probability changes the shape of the attentional template. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1993–2007.
- Gregoriou, G. G., Gotts, S. J., Zhou, H., & Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, *324*, 1207–1210.
- Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience*, *27*, 649–677.
- Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *111*, E962–E971.
- Hebart, M. N., & Baker, C. I. (2017). Deconstructing multivariate decoding for the study of brain function. *Neuroimage*. <https://doi.org/10.1016/j.neuroimage.2017.08.005>.
- Henriksson, L., Khaligh-Razavi, S. M., Kay, K., & Kriegeskorte, N. (2015). Visual representations are dominated by intrinsic fluctuations correlated between areas. *Neuroimage*, *114*, 275–286.

- Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, *77*, 128–149.
- Ito, T., Kulkarni, K. R., Schultz, D. H., Mill, R. D., Chen, R. H., Solomyak, L. I., et al. (2017). Cognitive task information is transferred between brain regions via resting-state network topology. *Nature Communications*, *8*, 1027.
- Jackson, J., Rich, A. N., Williams, M. A., & Woolgar, A. (2017). Feature-selective attention in frontoparietal cortex: Multivoxel codes adjust to prioritize task-relevant information. *Journal of Cognitive Neuroscience*, *29*, 310–321.
- Kitzbichler, M. G., Henson, R. N. A., Smith, M. L., Nathan, P. J., & Bullmore, E. T. (2011). Cognitive effort drives workspace configuration of human brain functional networks. *Journal of Neuroscience*, *31*, 8259–8270.
- Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, *17*, 401–412.
- Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 1–28.
- Kucyi, A., Esterman, M., Riley, C. S., & Valera, E. M. (2016). Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proceedings of the National Academy of Sciences, U.S.A.*, *113*, 13899–13904.
- Kucyi, A., Hove, M. J., Esterman, M., Hutchison, R. M., & Valera, E. M. (2017). Dynamic brain network correlates of spontaneous fluctuations in attention. *Cerebral Cortex*, *27*, 1831–1840.
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*, 661–679.
- Lee, J., & Geng, J. J. (2017). Idiosyncratic patterns of representational similarity in prefrontal cortex predict attentional performance. *Journal of Neuroscience*, *37*, 1257–1268.
- Li, Y., Richardson, R. M., & Ghuman, A. S. (2017). Multi-connection pattern analysis: Decoding the representational content of neural communication. *Neuroimage*, *162*, 32–44.
- Marchette, S. A., Vass, L. K., Ryan, J., & Epstein, R. A. (2014). Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nature Neuroscience*, *17*, 1598–1606.
- McIntire, L. K., McKinley, R. A., Goodyear, C., & McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Applied Ergonomics*, *45*, 354–362.
- Nastase, S. A., Connolly, A. C., Oosterhof, N. N., Halchenko, Y. O., Guntupalli, J. S., Di Oleggio Castello, M. V., et al. (2017). Attention selectively reshapes the geometry of distributed semantic representation. *Cerebral Cortex*, *27*, 4277–4291.
- Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology*, *10*, e1003553.
- Park, S., Brady, T. F., Greene, M. R., & Oliva, A. (2011). Disentangling scene content from spatial boundary: Complementary roles for the parahippocampal place area and lateral occipital complex in representing real-world scenes. *Journal of Neuroscience*, *31*, 1333–1340.
- Parks, E. L., & Madden, D. J. (2013). Brain connectivity and visual attention. *Brain Connectivity*, *3*, 317–338.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *108*, 12125–12130.
- Peelen, M. V., & Kastner, S. (2014). Attention in the real world: Toward understanding its neural basis. *Trends in Cognitive Sciences*, *18*, 242–250.
- Qiao, L., Zhang, L., Chen, A., & Egnér, T. (2017). Dynamic trial-by-trial re-coding of task-set representations in frontoparietal cortex mediates behavioral flexibility. *Journal of Neuroscience*, *37*, 11037–11050.
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, *38*, 433–447.
- Reeder, R. R., Perini, F., & Peelen, M. V. (2015). Preparatory activity in posterior temporal cortex causally contributes to object detection in scenes. *Journal of Cognitive Neuroscience*, *27*, 2117–2125.
- Rosenberg, M., Noonan, S., Degutis, J., & Esterman, M. (2013). Sustaining visual attention in the face of distraction: A novel gradual-onset continuous performance task. *Attention, Perception & Psychophysics*, *75*, 426–439.
- Rosenberg, M. D., Finn, E. S., Scheinost, D., Constable, R. T., & Chun, M. M. (2017). Characterizing attention with predictive network models. *Trends in Cognitive Sciences*, *21*, 290–302.
- Rosenberg, M. D., Finn, E. S., Scheinost, D., Papademetris, X., Shen, X., Constable, R. T., et al. (2016). A neuromarker of sustained attention from whole-brain functional connectivity. *Nature Neuroscience*, *19*, 165–171.
- Rosenberg, M. D., Hsu, W.-T., Scheinost, D., Constable, T. R., & Chun, M. M. (2018). Connectome-based models predict separable components of attention in novel individuals. *Journal of Cognitive Neuroscience*, *30*, 160–173.
- Rothlein, D., DeGutis, J., Germine, L., Wilmer, J., McGlinchey, R., & Esterman, M. (2018). Sensitivity to stimulus similarity is associated with greater sustained attention ability. *Attention, Perception, & Psychophysics*. <https://doi.org/10.3758/s13414-018-1504-0>.
- Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology*, *62*, 1904–1914.
- Scimeca, J. M., Kiyonaga, A., & D'Esposito, M. (2018). Reaffirming the sensory recruitment account of working memory. *Trends in Cognitive Sciences*, *22*, 190–192.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., et al. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124.
- Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations: Shapes of states. *Cognitive Psychology*, *1*, 1–17.
- Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., et al. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nature Communications*, *7*, 1–13.
- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*, 82–96.
- Todd, M. T., Nystrom, L. E., & Cohen, J. D. (2013). Confounds in multivariate pattern analysis: Theory and rule representation case study. *Neuroimage*, *77*, 157–165.
- Vann, S. D., Aggleton, J. P., & Maguire, E. A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, *10*, 792–802.
- Vickery, T. J., King, L.-W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision*, *5*, 81–92.
- Walther, D. B. (2013). Using confusion matrices to estimate mutual information between two categorical measurements. In *Proceedings—2013 3rd International Workshop on Pattern Recognition in Neuroimaging, PRNI 2013* (pp. 220–224).
- Wen, X., Yao, L., Liu, Y., & Ding, M. (2012). Causal interactions in attention networks predict behavioral performance. *Journal of Neuroscience*, *32*, 1284–1292.

- Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., & Haynes, J.-D. (2015). The role of the parietal cortex in the representation of task-reward associations. *Journal of Neuroscience*, *35*, 12355–12365.
- Woolgar, A., Hampshire, A., Thompson, R., & Duncan, J. (2011). Adaptive coding of task-relevant information in human frontoparietal cortex. *Journal of Neuroscience*, *31*, 14592–14599.
- Xu, Y. (2017). Reevaluating the sensory account of visual working memory storage. *Trends in Cognitive Sciences*, *21*, 794–815.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods* *8*, 665–670.