Effective Connectivity During Feature-Based Attentional Capture: Evidence Against the Attentional Reorienting Hypothesis of TPJ

Nicholas E. DiQuattro1,2, Risa Sawaki3 and Joy J. Geng1,2

1Department of Psychology, 2Center for Mind and Brain, University of California, Davis, CA, USA and 3School of Psychology, University of Birmingham, Birmingham, UK

Address correspondence to Joy Geng, Center for Mind and Brain, University of California, 267 Cousteau Pl., Davis, CA 95618, USA. Email: jgeng@ucdavis.edu

The most prevalent neurobiological theory of attentional control posits 2 distinct brain networks: The dorsal and ventral attention networks. The role of the dorsal attentional network in top-down attentional control is well established, but there is less evidence for the putative role of the ventral attentional network in initiating stimulus-driven reorienting. Here, we used functional magnetic resonance imaging and dynamic causal modeling (DCM) to test the role of the ventral and dorsal networks in attentional reorienting during instances of attentional capture by a target-colored distractor. In the region of interest analyses, we found that frontal eye field (FEF) was selectively activated by conditions where attention was reoriented (i.e. to spatial cues and target-colored distractors). In contrast, temporoparietal junction (TPJ) responded positively to all stimulus conditions. The DCM results indicated that FEF received sensory inputs earlier than TPJ, and that only the connection from FEF to TPJ was modulated by the appearance of the target-colored distractor. The results provide novel empirical evidence against the idea that TPJ generates stimulus-driven reorientations of attention. We conclude that our results are incompatible with existing theories of TPJ involvement in the stimulus-driven reorientation of attention and discuss alternative explanations such as contextual updating.

Keywords: attention, dynamic causal modeling, effective connectivity, fMRI, TPJ

Introduction

The most influential model of attentional control in the human brain posits 2 parallel networks: A bilateral dorsal system responsible for attentional selection and a right-lateralized ventral system that reorients attention toward currently unattended, but task-relevant stimuli (Corbetta and Shulman 2002; Astafiev et al. 2006; Corbetta et al. 2008). The hypothesis that the dorsal attentional network, composed of the frontal eye field (FEF) and intraparietal sulcus (IPS), encodes attentional priority and provides “top-down” modulatory signals to the sensory cortex is well supported in the literature (Thompson et al. 1997; Hopfinger et al. 2000; Kastner and Ungerleider 2000; Moore and Armstrong 2003; Gong et al. 2006; Buschman and Miller 2007, 2009; Gottlieb 2007; Ruff et al. 2008; Gregoriou et al. 2009; Zhou and Desimone 2011). In contrast, there is far less evidence that the ventral network, which includes the right temporoparietal junction (R TPJ) and ventral frontal cortex (R VFC), is involved in reorienting attention (Shulman et al. 2009). The purpose of the current study was to test the well-known hypotheses that the ventral network, particularly R TPJ, is involved in the stimulus-driven reorientation of attention by sending a “circuit-breaking” signal to regions within the dorsal attentional network such as FEF, which then executes the shift of attention.

Current results from human neuroimaging and neuropsychology consistently find that R TPJ encodes the presence of unexpected, but potentially task-relevant stimuli (Corbetta and Shulman 2002; Downar et al. 2002; Kincade et al. 2005; Lavie and De Fockert 2005; Serences et al. 2005, Indovina and Macaluso 2007; Shulman et al. 2007, 2009; Corbetta et al. 2008; Vossel et al. 2009; Geng and Mangun 2011; Ristic and Giesbrecht 2011; Chang et al. 2013). For example, TPJ activation increases proportionally with the unexpectedness of an invalid target in a typical Posner cueing paradigm (Vossel et al. 2009). R TPJ also responds to the presence of target-similar distractors that capture attention because of their potential (but not actual) task-relevance (Serences et al. 2005). These and similar findings clearly indicate that R TPJ plays an important role in encoding expectations related to the current environmental and task context. However, they are largely agnostic to the computational significance of the response. One possibility is that R TPJ initiates a fast-latency “interrupt” signal to regions of the dorsal network that “circuit-breaks” the current focus of attention and causes a shift to an unattended, but potentially task-relevant, stimulus (Corbetta and Shulman 2002; Corbetta et al. 2008). Another is that TPJ encodes perceptual functions (Downar et al. 2002; Doricchi et al. 2010; Geng and Mangun 2011; Manginelli et al. 2012). A third possibility is that TPJ encodes both (Chambers et al. 2004; Shulman et al. 2009).

Identifying the informational content of the TPJ response is critical not only for understanding the brain mechanisms of attention, but also because the idea that R TPJ is part of a stimulus-driven network that reorients attention is widespread beyond the attentional literature (Mitchell 2008; Cabeza et al. 2012; Shomstein 2012). Despite this, there has been little direct evidence for the causal relationship between R TPJ and the dorsal attentional network during tasks in which attention is reoriented to potentially task-relevant information such as target-similar distractors. This paucity of evidence is partly due to the fact that the blood oxygen level-dependent (BOLD) response in functional magnetic resonance imaging (fMRI) is slow, precluding conclusions about temporal causality using conventional methods. Additionally, the existence of a homolog for R TPJ in the monkey cortex is uncertain, and receptive field models of R TPJ neurons are therefore lacking (Constantinidis and Steinmetz 2001; Greikves and Fink 2005; Caspers et al. 2006; Orban et al. 2006; Rushworth et al. 2006; Husain and Nachev 2007; Uddin et al. 2010; Thiebaut de Schotten et al. 2011).

Direct tests of the hypothesis that R TPJ has a causal role in reorienting attentional focus are important because evidence from methods with higher temporal resolution such as event-related potentials (ERPs), magnetoencephalography, and transcranial magnetic stimulation (TMS) suggest that R TPJ
responses might be temporally too late for the ventral attentional network to send a fast circuit-breaking reorientation signal to the dorsal attentional network (point raised in Corbetta et al. 2008, accompanied by a call to test the plausibility of TPJ as a ‘circuit-breaker’). For example, TPJ is regarded as one of the sources of the P300/P3b ERP that peaks between 300 and 700 ms after stimulus presentation (Knight et al. 1989; Yamaguchi and Knight 1991; Menon et al. 1997; Knight and Scabini 1998; Ferrara et al. 2000; Bledowski et al. 2004; Luck 2005; Polich 2007; Ristic and Giesbrecht 2011). Similar to TPJ, the P300/P3b occurs in response to low-frequency targets, and unexpected stimuli that reorient attention such as target-similar distracters (Watson et al. 2005; Azizian et al. 2006; Sawaki and Katayama 2006). However, in contrast to the literature on TPJ in attention, the P300 is hypothesized to reflect postperceptual functions that occur later in time such as updating environmental schemas, working memory, or mapping responses onto motivationally important stimuli (Donchin 1981; Nieuwenhuys et al. 2005; Polich 2007).

In comparison, the timing of responses in FEF appears to occur earlier. For example, TMS studies of the human FEF report interference with target discrimination at 40 and 80 ms poststimulus presentation (O’Shea et al. 2004), but those of TPJ have later effects starting at 150, 250 ms, or later (Ellison et al. 2004; Meister et al. 2006; Schindler et al. 2008; Chang et al. 2013; although see Chambers et al. 2004). The temporally late effects from ERPs and TMS run contrary to the idea that TPJ initiates a fast-latency stimulus-driven reorientation signal to dorsal network regions as hypothesized by Corbetta et al. (2002, 2008).

In the present experiment, we used a paradigm that produces feature-based attentional capture in which attention is reoriented automatically toward objects that possess a target feature, such as color (Yantis and Jonides 1990; Folk et al. 1992, 2002; Serences et al. 2005; Sawaki and Katayama 2006). A previous study reported the coactivation of dorsal and ventral network regions by the target-similar distracters, suggesting that these regions were involved in the detection and reorientation of attention to the target-similar distracter (Serences et al. 2005). Although consistent with the idea that TPJ sends a reorientation signal, the data were uninformative regarding the pattern of connectivity between regions.

Thus, our goal was to test the effective connectivity between dorsal and ventral regions in response to a target-similar distracter that produces a shift in spatial attention. We used dynamic causal modeling (DCM), which allows for inferences regarding causality since the model is based at the neuronal level (Friston et al. 2003; Penny et al. 2004). The models included TPJ and FEF as representative nodes of the 2 networks due to the clear hypotheses that TPJ is involved in reorienting attention (Shulman et al. 2009) and the considerable evidence for FEF in shifting attention (see above). Although we limit our models to these 2 regions for parsimony, the models are neutral to the presence of “third-party” regions (e.g. the middle frontal gyrus) that may mediate interactions between networks. If TPJ and the ventral attentional network circuit-break FEF and the dorsal attentional control network through an fast interrupt signal, then we should see that TPJ receives sensory input first, and that connectivity from TPJ to FEF is strengthened by the presence of a target-colored distractor; however, if the signal in TPJ is too late to initiate a reorientation signal, then we might expect excitatory sensory inputs to perturb the model through FEF and modulation of the connection from FEF to TPJ. Such a result would suggest that signals related to the reorientation of attention occur first in dorsal attentional network regions, and that the response in TPJ might be better characterized by postperceptual processes.

Materials and Methods

Participants

Twenty-five adults participated for payment. Five were excluded from analysis due to excessive head motion during scanning or performing at chance level of accuracy, leaving data from 20 participants (9 males, 1 left handed, mean age = 23.4 years, SD = 3.08). Each provided written informed consent in accordance with the local ethics clearance as approved by the National Institutes of Health. All had normal or corrected-to-normal vision. Handedness was determined by a shortened version of the Edinburgh handedness inventory (Oldfield 1971).

Experimental Procedure

Each trial began with a centrally presented cue (0.46° of visual angle) lasting 500 ms that directed the participant’s attention to 1 of 2 bilateral dots (0.07° of visual angle). The dots served as placeholders for spatial attention and remained visible until the bilateral circle stimuli appeared. After a 1500 to 6500 ms delay, 2 colored circles (each subtending at 1.4° of visual angle) appeared bilaterally for 200 ms (3.3° of visual angle from the screen center to the nearest edge of circle) after which the participant indicated the presence or absence of a target with a manual key press (Fig. 1A). Following the offset of the stimuli, a fixation cross was presented for 2000–7000 ms,

Figure 1. (A) Trial procedure: each trial began with the presentation of a 500-ms cue indicating which of the bilateral dots the participant should attend. Following the cue, a variable fixation period occurred for 1500–6500 ms. Finally, 2 colored circles appeared for 200 ms followed by a 2000 to 6000 ms inter-trial interval before the start of the next trial. (B) Examples of our 3 stimulus conditions. Neutral: target color is not present, target-present: target color appears in the cued location, and target-colored distracter: target color appears in the uncued location. O = Orange, G = Green, P = Purple.
which led into the start of the next trial. The target was defined as a conjunction of color and spatial location (i.e. the target was only considered present when the predefined color was presented in the cued side of the display). Each colored circle could be 1 of 3 colors (orange (30.03 cd/m^2); international commission on illumination (CIE) color space: 38.95, 29.72), green (24.5 cd/m^2; CIE color space: 17.42, 26.86), or purple (24.11 cd/m^2; CIE color space: 64.53, 24.08)). Prior to each scanning run, participants were shown a screen presenting the target color for that run. At the start of the experiment, target-color order was randomly determined with the constraints that the same color was never the target color consecutively, and all colors were the target once every 3 runs.

The possible combinations of colors presented resulted in 3 display configurations (Fig. 1B): neutral (i.e. target-color absent), target-present (i.e. target color in the cued location), and target-colored distracter (i.e. target color in the uncued location). The 3 display conditions were crossed with the cue direction (left and right), resulting in 6 experimental conditions. Critically, the target-colored distracter condition required a “target-absent” response despite the target color appearing in the display. Each stimulus pair had an equal probability of occurrence.

**Imaging Data**

MRI data were acquired from a 3-T Siemens Trio scanner equipped with a 32-channel phased-array head coil. A \( T_2^* \)-weighted echo-planar imaging (EPI) sequence was used to acquire volumes of 31 axial slices of 3 mm thickness (3.4 × 3.4 mm in-plane resolution), with a distance factor of 10%, every 1750 ms. Slices were oriented to achieve whole-brain coverage. Two hundred and four volumes were collected in each of 9 runs. Image data were analyzed using SPM8 (Welcome Department of Imaging Neuroscience, London, UK; Friston et al. 1995). Images were realigned and unwarped to correct for interactions between movement and field inhomogeneities (Andersson et al. 2001), normalized to the Montreal Neurological Institute (MNI) EPI template available in SPM8, and resampled to a resolution of 2 × 2 × 2 mm. The data were additionally smoothed with a 3-dimensional 6-mm full-width half-maximum Gaussian kernel. High-resolution \( T_1 \)-weighted structural images were acquired using a magnetization-prepared rapid gradient-echo sequence, coregistered with each subject’s EPI images, and normalized to the MNI template brain. Results are displayed on an average structural image created from normalized \( T_1 \)-weighted images from the current participants.

The data were modeled for each voxel using a general linear model (GLM) that included regressors obtained by convolving each event-related unit impulse (“stick function”) with a canonical hemodynamic response function. The main GLM included 6 experimental conditions given by crossing the cue direction (left and right) and stimuli configuration (neutral, target-present, and target-colored distracter). In addition to the 6 conditions of interest cue onsets, errors, scan session, and 6 movement parameters were modeled separately as variables of noninterest. Condition-specific effects estimated by the GLM were entered into a group-level analysis as contrast images, which were then entered into independent 1-sample \( t \)-tests. A second GLM was created identical to the first but with the addition of reaction time (RT) entered as a trial-specific parametric regressor for each condition. RTs for each condition were scaled by Euclidean normalization and then mean-centered. The 6 RT regressors were created by scaling the expected conditional BOLD response by trial RT; the resultant parameter estimates represent the degree to which BOLD activation in response to a particular stimulus condition is scaled by trial RT.

**Region of Interest Definition**

Data were extracted for regions of interest (ROIs) in both right FEF and TPJ for each individual as representative key structures of the dorsal and ventral attentional networks, respectively. Although there are many regions in the dorsal attentional network that tend to coactivate, there is substantial evidence that right FEF has a particularly strong role in top-down attentional selection (Ruff et al. 2008; Hung et al. 2011).

Each ROI was a spherical volume with a radius of 3 mm (i.e. 19 voxels). Mean-adjusted data (i.e. first eigenvariate of the time series) from each participant were extracted from all voxels within each ROI. The selection of ROIs within each individual was based on a combination of anatomical definitions and results from group random-effects analyses of target-colored distracter trials minus neutral trials. An ROI center within each individual subject was determined by the local maximum (\( P < 0.05 \), uncorrected) closest to peak coordinates from the corresponding group random-effects analysis, within the appropriate anatomical landmark: FEF was defined by the intersection between the superior frontal sulcus and the precentral sulcus (Paus 1996); and TPJ by the centers of the lateral fissure and superior temporal sulcus (Mort et al. 2003) within a small-volume mask created from the findings of Serences et al. (2005). The combined criterion for TPJ was used to constrain our results given the variability in the anatomical location of regions referred to as TPJ (Gillebert et al. 2012). The group coordinates were used for individuals without identifiable clusters (3 for TPJ and 2 for FEF). The mean difference of coordinates between individual ROIs and the group coordinates was <4 mm in all dimensions. The mean ± SD xyz MNI coordinates for each ROI were as follows: FEF: 28.4 ± 3.5, −9.3 ± 6.39, 50.4 ± 3.4; and TPJ: 49.7 ± 4.55, −40.2 ± 4.25, 20 ± 4.86. Using group statistics and anatomical criteria to define ROIs provide individual specificity while still permitting generalization to the population (Ikai and Curtis 2008; Geng and Mangun 2009).

**Dynamic Causal Modeling**

Effective connectivity analyses were conducted using DCM10 as implemented in SPM8. DCM is a method for assessing effective connectivity between brain areas in fMRI. In this study, we used DCM to test models of the functional connectivity between FEF and TPJ. DCM uses a bilinear equation that models neural activity on a millisecond time scale based on stimulus inputs from the experimental paradigm (Friston et al. 2003; Stephan et al. 2010). The model uses a Bayesian framework with biophysically plausible priors that accounts for the complexity of neuronal activity while maintaining the ability to distinguish between competing models (Friston et al. 2003; Daunizeau et al. 2011). The neuronal model time series is then fed into a feed-forward balloon model of the hemodynamic response function, which creates estimated BOLD data based on the neuronal model. Model parameters are then adjusted to
minimize differences between estimated and measured data. The result is a model of neuronal activity that allows for inference to be made about the directional connectivity between model regions.

DCM models are described by 3 sets of parameters: (1) Intrinsic parameters represent the average level of connectivity between regions across the entire experiment. Importantly, these do not represent anatomical connections, but rather the effective functional connectivity between regions. (2) Modulatory parameters describe condition-dependent modulations of intrinsic connectivity in response to an experimental event. (3) Driving inputs characterize the response of each region to stimulus presentation.

The approach we used to determine the best fitting model was to create model families that represent a specific hypothesis. By defining and comparing alternative models, the respective likelihoods of the theories they represent can be observed (Penny et al. 2004, 2010; Stephan et al. 2010). We defined 3 model families based on the location of driving inputs (into TPJ, only, FEF only, or both). Model families were compared using a random-effects Bayesian model selection (BMS) framework (Penny et al. 2004). BMS chooses the winning family based on the likelihood of the models within a family to have produced our observed data (exceedance probability, xp). Once a winning family is chosen, a Bayesian model average (BMA) of models within that family is computed to produce average parameter values weighted by the xp value of each model (Penny et al. 2010; Stephan et al. 2010). The last step of a DCM analysis stream involves entering the intrinsic, modulatory, and driving input parameters from the BMA into classical statistical tests of significance.

DCM is growing in popularity and has been used successfully to test models of human brain connectivity in a number of cognitive domains (e.g. Mechelli et al. 2004; Stephan et al. 2006; Noppeney et al. 2007), including in visual attention (DiQuattro and Geng 2011; Vossel et al. 2012). Additionally, there have now been several studies that provide evidence for the biological plausibility of DCM models (e.g. Stephan et al. 2004; David et al. 2008; David 2011; Murta et al. 2012), although some aspects are debated (Daunizeau et al. 2011; Roebroeck et al. 2011; Valdes-Sosa et al. 2011). However, the application of the method is strongly hypothesis-driven, and interpretation of results is therefore limited by the strength of the initial hypotheses. In particular, the BMS procedure is only able to minimize differences between estimated and measured data. Therefore, evidence for a priori functional models is strongly hypothesis-driven, and interpretation of results is therefore limited by the strength of the initial hypotheses. Thus, there is no way to identify the "true" model if it was not included in the comparison set—almost an inevitability given the complexity of human brain function. For example, DCM models cannot identify "third-party" contributions from regions that are not modeled (i.e. a change in connectivity between regions A and B might be mediated by region C). Thus, comparison of DCM models is most useful in providing a metric of evidence for a priori functional models that would be expected to exist from the literature.

The purpose of the DCM models in this experiment was to test a primary prediction based on the existing Corbetta et al. (2002, 2008) model in which RT to TPJ is hypothesized to initiate a stimulus-driven reorientation signal that circuit-breaks the current attentional focus maintained in dorsal attentional network regions such as FEF (see also, Corbetta and Shulman 2002; Astafiev et al. 2006; Shulman et al. 2009). If TPJ initiates a reorientation signal, then the model that best fits the data should include excitatory stimulus input into TPJ and an increase in effective connectivity from TPJ to FEF when a target-colored distracter appears. However, if TPJ is responding relatively late to the presence of task-relevant stimuli in response to contextual updating (or another postperceptual process), then we would expect to find inputs to FEF and the direction of stimulus-driven modulation, if any, to be on the connection from FEF to TPJ. To be clear, the lack of evidence for an early stimulus-driven reorientation signal from TPJ to FEF, which initiates the first shift of attention, does not preclude the possibility that TPJ is involved in a later stage of attentional control.

**Results**

**Eye-Tracking**

Eye-position data were collected with an Applied Science Laboratories (ASL) EyeTrac 6 operating at 60 Hz. The data were used to ensure that subjects did not make condition-specific eye movements in response to the spatial cue and in preparation of the appearance of the bilateral stimuli. Samples found to be greater than ±5.5° of visual angle from fixation (an area that would be more eccentric than the stimuli) were considered to be blinks or tracker error and were removed from further analysis. The remaining horizontal eye-position data from each trial were baseline-corrected by subtracting out the mean eye position during the 200-ms precue period (which occurred 2–8 s prior to the onset of the bilateral stimuli). The data were then mean-centered and divided into the 6 experimental conditions.

The mean proportion of samples across conditions that fell within ±2° of fixation was 81%, indicating participants’ ability to maintain fixation during the critical stage of the task. To be sure that even the few samples that fell outside of the 2° fixation area were not differentially skewed based on the condition, samples from each condition were further analyzed based on measured eye-position distance from fixation (4° to 4° in steps of 2°) and entered into separate 1-way repeated-measures analysis of variance (ANOVARs). There were no significant differences found due to experimental condition at any eye position (all main effects of side: F1,15 < 2, P > 0.05; all main effects of condition: F1,30 < 1.43, P > 0.05; all interactions: F2,30 < 1.36, P > 0.05). The eye-position data demonstrate that participants were able to maintain fixation during the stimulus presentation, and that there were no differences in the eye position between experimental conditions.

**Behavior**

RT and accuracy data were each entered into separate 1-way repeated-measures ANOVAs with the factor stimulus condition (neutral, target-present, and target-colored distracter). RT analyses were limited to trials with correct responses within ±2 SDs of the individual conditional mean (Fig. 2). The main effect of the stimulus condition was significant, F2,30 = 18.85, P < 0.001. Post hoc paired t-tests (all P-values given with Bonferroni correction unless otherwise stated) showed that RT was significantly longer in the target-colored distracter condition than in the neutral, t(19) = 5.59, P < 0.05, and target-present conditions, t(19) = 4.69, P < 0.05, trials. The neutral and target-present conditions were not significantly different,
$t_{(19)} = -0.54, P = 0.6$. Longer RT in the target-colored distracter condition compared with neutral and target-present replicated findings of feature-based attentional capture (Folk et al. 1992; Serences et al. 2005); this confirmed that the target-colored distracter produced a stimulus-driven shift of visual attention that was absent in the other 2 conditions.

Overall accuracy was high (mean = 92.2%; Fig. 2), demonstrating that participants understood and were able to perform the task well. There was a significant main effect of stimulus condition in accuracy, $F_{2,38} = 7.55, P < 0.05$. Post hoc paired $t$-tests revealed significantly higher accuracy in the neutral (93.8%) condition than in the target-present (91.5%), $t_{(19)} = 3.39, P < 0.05$ and target-colored distracter (91.2%) conditions, $t_{(19)} = -3.33, P < 0.05$, but not between target-present and target-colored distracter, $t_{(19)} = -0.49, P = 0.63$. Lower accuracy in the target-present trials than in the neutral may have been influenced by the relative infrequency of the target-present response (33% compared with 66% target-absent responses). However, only correct trials were included by in the condition-specific fMRI analyses, and therefore the RT data, which replicated typical findings of feature-based attentional capture, were the relevant behavioral correlate of the brain results.

**Brain Imaging Results**

We first identified the critical brain structures in the dorsal and ventral attentional networks expected to be involved in reorienting attention toward a target-colored distracter and assessing its task-relevance. The critical contrasts were therefore between the target-colored distracter condition, which elicited a stimulus-driven attentional shift to the uncued location, and 2 other conditions, which did not systematically produce shifts in attention (see Behavioral RT Results above). To do this, we first conducted a whole-brain contrast of target-colored distracter trials minus neutral trials using a statistical threshold of $P < 0.001$ (cluster corrected at false discovery rate (FDR) $P < 0.05$). This resulted in a number of significant regions, including right FEF, bilateral IPS, and bilateral precuneus (Fig. 3A and Table 1). The activation of FEF and IPS were expected given their known role in encoding shifts of attention in similar paradigms (Serences et al. 2005; Natale et al. 2010). Similarly, the contrast of target-colored distracters minus target-present trials resulted in significant activations in

---

**Figure 2.** RT in each stimulus condition. An overall effect of stimulus condition was present in both RT and accuracy (see text for details). Asterisk indicates significant difference. Error bars are SEM.

**Figure 3.** (A) Dorsal frontoparietal activations from the target-colored distracter minus target-present (white), target-colored distracter minus neutral (black), and the overlap between the 2 contrasts (gray). All clusters shown are significant at $P < 0.001$, with false discovery rate (FDR) cluster correction at $P < 0.05$. (B) $\beta$-values from R FEF in each stimulus condition. (C) R TPJ activation found in the target-colored distracter minus neutral at $P < 0.001$ uncorrected and in a small-volume correction at $P < 0.05$ familywise error corrected (FWE). (D) $\beta$-values from R TPJ in each stimulus condition. Error bars are SEM.
Table 1

<table>
<thead>
<tr>
<th>Name</th>
<th>Coordinates (MNI xyz)</th>
<th>Cluster size (number of voxels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target-colored distractor &gt; neutral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R middle temporal gyrus</td>
<td>36, −68, 22</td>
<td>131</td>
</tr>
<tr>
<td>R intraparietal sulcus</td>
<td>36, −52, 54</td>
<td>158</td>
</tr>
<tr>
<td>R frontal eye fields</td>
<td>28, −12, 50</td>
<td>110</td>
</tr>
<tr>
<td>L intraparietal sulcus</td>
<td>−26, −60, 52</td>
<td>280</td>
</tr>
<tr>
<td>L intraparietal sulcus 2</td>
<td>−40, −38, 38</td>
<td>157</td>
</tr>
<tr>
<td>Precuneus</td>
<td>2, −54, 50</td>
<td>661</td>
</tr>
<tr>
<td>Target-colored distractor &gt; target-present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R frontal eye fields</td>
<td>30, −6, 50</td>
<td>1283</td>
</tr>
<tr>
<td>R intraparietal sulcus</td>
<td>22, −58, 32</td>
<td>986</td>
</tr>
<tr>
<td>L frontal eye fields</td>
<td>−24, −4, 52</td>
<td>134</td>
</tr>
</tbody>
</table>

All activations significant at $P < 0.001$ corrected, with FDR cluster correction at $P < 0.05$.

bilateral FEF and R IPS (Table 1) at a threshold of $P < 0.001$ (cluster corrected at FDR $P < 0.05$). Taken together, these results demonstrated that the core regions of the dorsal attentional network, particularly FEF and IPS and in the right hemisphere, were more activated when a target-colored distracter was present. This suggests that these regions were involved in generating a stimulus-driven shift of attention to the uncued location.

To further quantify the relationship between these brain areas and attentional shifting, we next conducted a separate GLM with RT added as a parametric regressor. RT was significantly longer in the target-colored distracter condition, but some trial-to-trial variability in RT was expected in all conditions (although stimulus-driven shifts should occur most reliably in the target-colored distracter condition). Therefore, regions that correlate with RT should include those involved in the control of attention orienting across conditions. A conjunction analysis of RT covariation in all conditions produced significant activations in R FEF (peak at MNI $xyz$: 30, −2, 52, $k = 9422$), L IPS (MNI: −22, −76, 54, $k = 4196$), and L middle temporal gyrus (MNI: −50, −52, −12, $k = 194$) ($P < 0.001$, cluster corrected at FDR $P < 0.05$). This result confirms that FEF activation was directly related to performance across stimulus conditions, as well as showing overall greatest activation in the target-colored distracter, which elicited the most reliable shifts of attention.

Notably absent from our whole-brain results were any clusters from the ventral attention network at the FDR cluster-corrected threshold of $P < 0.05$. However, given the strongly purported expectation for TPJ involvement in reorienting attention when a target-colored distractor or invalid target is present (Corbetta and Shulman 2002; Macaluso et al. 2002; Serences et al. 2005; Vossel et al. 2009; Doricchi et al. 2010; Natale et al. 2010), we conducted a more targeted small-volume correction procedure using a spherical volume with a 10-mm radius centered on the R TPJ coordinates reported in Serences et al. (2005) (MNI: 56, −46, 24, converted from Talairach space using the “tal2mni” function by M. Brett, http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach). We chose the coordinates from Serences et al. (2005) due to the similarities in the 2 task designs.

A single TPJ cluster was significant (MNI: 50, −40, 20, $k = 23$; $P < 0.05$ familywise error corrected (FWE); this same region was also significant at the more liberal whole-brain level at $P < 0.001$, uncorrected; Fig. 3C). In addition to being located within the a priori functionally defined ROI, the TPJ cluster was also within the anatomical definitions of right TPJ as delineated by Mort et al. (2003) (i.e. between the middle of the lateral fissure and the middle of the superior temporal sulcus), thus giving us confidence that the activated region was typical of reports of TPJ in the literature (Corbetta and Shulman 2002; Serences et al. 2005).

To understand the specific response profiles within the canonical dorsal and ventral network regions, beta ($\beta$) values from R FEF and R TPJ (Fig. 3B,D) were entered into 2 analyses: The first was an 1-way repeated-measures ANOVA of data from the stimulus conditions. The second was a 2-tailed t-test of the cue response. The results from R TPJ showed a significant main effect of stimulus condition, $F_{2,38} = 20.54$, $P < 0.001$. Post hoc t-tests showed significant differences between all conditions, all $t_{(19)} > 3.51$, $P < 0.05$, with Bonferroni correction; the results followed a staircase pattern such that responses were greatest in the target-colored distracter, target-present, and neutral condition. Interestingly, activation was significantly above zero in all 3 conditions, all $t_{(19)} > 4.77$, $P < 0.05$, Bonferroni corrected. This suggests that TPJ contributed to stimulus processing even when a reorientation of attention did not occur. This result runs contrary to the idea that R TPJ responds only to reorientations of attention to unattended, but relevant stimuli (see Discussion). Importantly, R TPJ did not respond indiscriminately to all visual stimuli: Activation in response to cue onsets was not significantly different from zero, $t_{(19)} = 0.54$, $P = 0.59$ (Fig. 3D). This result was consistent with the idea that the dorsal attentional network suppresses R TPJ during periods of top-down attentional control (Shulman et al. 2007; Anticevic et al. 2010).

Similar to TPJ, there was a significant main effect of stimulus condition in R FEF, $F_{2,38} = 32.44$, $P < 0.001$. However, post hoc t-tests revealed that the only significant differences were between the target-colored distracter and the other 2 conditions, both $t_{(19)} > 6.05$, $P < 0.05$ Bonferroni corrected. There was no difference between the neutral and target-present conditions, $t_{(19)} = 0.92$, $P = 0.37$. Additionally, activation was only significantly above zero in the target-colored distracter condition, $t_{(19)} = 4.29$, $P < 0.05$ and in response to cue presentation, $t_{(19)} = 3.05$, $P < 0.05$. These results demonstrated greatest FEF activation in the 2 situations where attention would be expected to be reoriented, whether in response to an endogenous spatial cue or an exogenous target-colored object. Further, FEF’s pattern of activation exemplified a region that is engaged in reorienting specifically, in contrast to activations found in TPJ that are responding qualitatively similarly to all stimulus conditions. However, these results provide no information about the temporal dynamics between regions and therefore, we next used DCM (see Materials and Methods) to further investigate the temporal relationship between stimulus presentation, activity in R TPJ, and R FEF.

**DCM Results**

The primary goal for this study was to understand the connectivity dynamics between structures of the dorsal and ventral attentional networks in reorienting attention and assessing the contextual relevance of stimuli. To do so, we performed a DCM analysis on time series extracted from representative ROIs identified in the whole-brain analyses described above. To limit the number of model parameters, we chose a single representative region from each network for these analyses. Right FEF was
chosen for the dorsal attentional network because of its well-established role in controlling shifts of spatial attention (Moore and Fallah 2004; O’Shea et al. 2004; Thompson et al. 2005; Buschman and Miller 2007; Zhou and Thompson 2009; Ronconi et al. 2014; Vossel et al. 2012; Wardak et al. 2012), and our analogous results (see above). Right TPJ was the obvious choice for the ventral network region given that the goal of this study was to test the specific hypothesis for its role in stimulus-driven attentional reorienting (Corbetta and Shulman 2002; Corbetta et al. 2008; Shulman et al. 2009). Although DCM models are agnostic to the anatomical connections between these regions, there is evidence that the right posterior parietal cortex, including our TPJ ROI, is connected to the frontal component of the dorsal network by the superior longitudinal fasciculus II (Thiebaut de Schotten et al. 2011), thus supporting the plausibility of our DCM models.

Twelve models were constructed to test the specific pattern of connectivity between R FEF and R TPJ during target-colored distractor trials. All models had reciprocal intrinsic connections, but different driving inputs and modulatory parameters (see Materials and Methods). Driving inputs from both the target-colored distracter and target-present conditions, which were matched in terms of visual stimuli, entered the model through FEF, TPJ, or both and were used to define 3 families of models. Within each family, the models differed in the pattern of how the target-colored distracter modulated the connectivity between regions. The model space included a factorial crossing of regions with driving inputs and connections with modulatory parameters.

The hypothesis that TPJ circuit-breaks spatial attention by sending a reorientation signal to FEF predicts that the final model should have 2 characteristics: Driving inputs into TPJ, signifying an excitatory stimulus-driven effect on TPJ activity; and a positive modulatory parameter on the connection from TPJ to FEF, signifying an upward modulation of the connection strength. These model features would be expected if the target-colored distracter produced a condition-specific reorientation signal from TPJ to FEF.

In contrast, characteristics consistent with alternative hypotheses in which TPJ is involved in late-stage attentional or decisional processes (Downar et al. 2000; Polich 2007; Dorich et al. 2010; Geng and Mangun 2011) are found in models with excitatory driving inputs to FEF and modulation of the connection from FEF to TPJ. Note that in addition to the model structure, the parameter values (i.e. positive or negative) are critical for interpretation of the causal effects between stimuli and ROIs. The random-effects BMS family comparison procedure resulted in selection of Family 3 (driving inputs into both TPJ and FEF) as the best fit to our data (xp value = 0.97; Family 1 xp = 0.004, Family 2 xp = 0.021). The next step was to perform a BMA of the models in the winning family in order to determine the specific parameter values for the driving inputs, intrinsic connections, and modulatory parameters (Fig. 4A). All parameter statistics were based on 2-tailed one-sample t-tests.

Although the winning family of models had driving inputs to both FEF and TPJ, the parameter values were very different. The driving input into FEF was significant and positive, \(t_{19} = 5.08, P < 0.05\) (mean: 0.72 Hz), but the inputs to TPJ were negative and only marginally significant, \(t_{19} = -2.04, P = 0.057\) (mean: -0.22 Hz). This suggests that the stimulus-driven effect on FEF was excitatory, but the immediate effect on TPJ was weakly suppressive. This may be consistent with the idea of TPJ filtering task-irrelevant stimuli during top-down attentional control (Shulman et al. 2003, 2007).

Intrinsic connections were both significant, but the connection from FEF to TPJ was positive, \(t_{19} = 2.17, P < 0.05\) (mean: 0.60 Hz), and the connection from TPJ to FEF was negative, \(t_{19} = -4.12, P < 0.05\) (mean: -0.91 Hz). The only significant modulatory parameter was on the connection from FEF to TPJ, \(t_{19} = -2.4, P < 0.05\) (mean: -0.19 Hz); this modulation was negative, suggesting a partial reduction in the strength of connectivity with the appearance of a target-colored distracter. The modulatory parameter on the connection from TPJ to FEF was not significant, \(t_{19} = 1.45, P = 0.16\). These results together suggest that stimulus activity had an excitatory effect on FEF, and that FEF activity had a causal effect on TPJ activity in our task (Shulman et al. 2007) that was reduced in the target-colored distracter condition. Interestingly, none of the model parameters were consistent with the idea that TPJ has an early excitatory response that is used to send a reorientation signal to FEF. The results are more consistent with a TPJ role in a later process such as contextual updating of the task context based on current stimuli (Donchin 1981; Polich 2007; Dorich et al. 2010; Geng and Mangun 2011).

To test the condition specificity of the DCM parameters, we conducted an analogous set of family comparisons using the target-present trials for modulatory parameters. Similar to our previous corpus, Family 3 (driving inputs into FEF and TPJ) also had the highest exceedance probability (xp value = 0.98; Family 1 xp = 0.0002, Family 2 xp = 0.024). The pattern for the driving and intrinsic connections were qualitatively identical to the models of the target-colored distracter (driving inputs to: FEF: \(t_{19} = 5.22, P < 0.05\) (mean: 0.72 Hz; TPJ: \(t_{19} = -2.49, 0.05\).
P<0.05, mean = −0.28 Hz; intrinsic connections FEF to TPJ: $t_{(19)} = 2.15$, $P<0.05$, mean = 0.58 Hz; TPJ to FEF: $t_{(19)} = -4$, $P<0.05$, mean = −0.94 Hz; Fig. 4B). However, none of the modulatory parameters were significant (FEF to TPJ: $t_{(19)} = 1.11$, $P=0.28$; TPJ to FEF: $t_{(19)} = -1.13$, $P = 0.27$). Results from this control analysis were consistent with expectations that the condition-nonspecific parameters would be similar, but that the conditional ones would change. In particular, targets did not produce any condition-specific changes in connectivity between TPJ and FEF. This was in contrast to the target-colored distracter, which reduced connectivity from FEF to TPJ.

### Discussion

In this experiment, we used fMRI and DCM to investigate the response profile and connectivity dynamics between the dorsal and ventral attentional control networks in response to a target-colored distracter. The paradigm is well known in the psychological literature for eliciting feature-based attentional capture and the associated performance costs. Our RT results replicated these findings and showed longer RT in the target-colored distracter condition than either the neutral or target-present conditions. The longer RT was presumably due to stimulus-driven attention shifts away from the cued location toward the target-colored distracter.

The appearance of the target-colored distracter produced reliable activations in regions of the dorsal attentional network, including bilateral FEF, IPS, and the precuneus. These activations were similar when the target-colored distracter was contrasted with neutral and target-present trials. This suggests that the increased activation was due to the additional shift in attention to the target-similar distracter in the uncued location. Activation was also found in R TPJ, as expected. However, in contrast to expectations that R TPJ selectively responds to unexpected stimuli, the pattern of response in TPJ was positive in all 3 conditions (Fig. 3D). The response in R TPJ was not qualitatively different to target-colored distracters, but only quantitatively so (although the pattern was qualitatively different from the response to a cue stimulus). It may be that R TPJ was activated by all 3 stimulus conditions in our task, because the probability of each event was equal. Thus, while the cue was informative about the task (i.e. partially defined the target), subjects could not build a true statistical expectation of whether the trial would be valid (i.e. the target color in the cued location), or invalid (i.e. a target in the uncued location). We return to this point below in discussion of an alternative theory for TPJ activation.

The main purpose of the experiment was to use DCM to test the effective connectivity between the ventral and dorsal regions. To do so, we selected TPJ and FEF as representative ROIs from each network. We used a 2-tiered approach to select the final model. First, we divided the models based on driving inputs and also used Bayesian model averaging to determine the parameter values for the final model. Using this procedure, the DCM analysis produced a model with positive driving inputs into FEF, but negative driving input parameters into TPJ. This suggested that stimuli had an immediate stimulus-driven excitatory effect on FEF, as expected based on neurophysiology. However, stimulus inputs had a delayed effect on TPJ. This result runs counter to the idea that TPJ could provide the dorsal attentional network (here represented by FEF) with a temporally early stimulus-driven reorientation signal. Specifically, it does not seem plausible that TPJ is sending the dorsal network an interrupt signal that precedes the reorientation (i.e. initial shift) of attention to the target-colored distracter. However, this does not preclude the possibility that TPJ provides a later stimulus-driven signal that contributes to updating attentional control.

A second feature of the model was a significant condition-specific negative modulation of the connection from FEF to TPJ when the target-colored distracter appeared. There was no such significant modulation when the target appeared (Fig. 4A). This decrease in the strength of the connection specifically when the target-colored distracter appeared suggests that FEF may reduce control over TPJ when attention is shifted toward unattended, but potentially task-relevant information (i.e. nontarget item sharing target features). This interpretation is consistent with the idea that FEF exerts control over TPJ under conditions of strong top-down attentional set (Shulman et al. 2007), but reduces its influence when there is a stimulus-driven change in attention (DiQuattro and Geng 2011). This interpretation is also consistent with our results at cue onset when top-down mechanisms would be expected to dominate attentional control: TPJ activation was not significantly different from zero, but FEF activation increased.

In summary, our data are inconsistent with the idea that R TPJ encodes stimulus-driven shifts of attention and sends a re-orientation signal to FEF. One possible alternative explanation that is consistent with our current data is from the P300 ERP literature and suggests that TPJ engages in “contextual updating” of internal models (Geng and Vossel under review). Although there is no consensus regarding the computational specifics of the internal model, there is evidence that R TPJ may encode expectations regarding the relationship between a sensory stimulus and the context-appropriate action (Downar et al. 2002; Doricchi et al. 2010; Geng and Mangun 2011). This shares similarity with the idea that the ventral system may be involved in slower adjustments to task responses under conditions of changing task sets or expectations (Corbetta et al. 2008). Interestingly, all of these hypotheses that TPJ and the ventral system are engaged in postperceptual updating are consistent with the finding that TPJ activity is suppressed during periods of top-down attentional control, such as in visual search (Shulman et al. 2003, 2007) or in response to a top-down cue (as in the data reported here). Although the idea of TPJ being involved in contextual updating is uncommon in the neuroimaging literature, a similar hypotheses for TPJ have been made in the domain of embodied perception (Silani et al. in review; Decety and Lamm 2007; Tsakiris et al. 2008). Future experiments are necessary to further flesh out and test new hypotheses for a TPJ role in postperceptual updating.

In conclusion, we used fMRI and DCM to investigate the connectivity dynamics between key nodes of the dorsal and ventral attentional networks. We found concurrent activation of both networks in response to target-colored distracters confirming their role in instances of feature-based attentional capture, but interestingly, R TPJ responded to all stimulus conditions, suggesting a nonspecificity for target-colored distracters, per se. In contrast, FEF responded selectively to stimuli that produced a reorientation of attention, either endogenously (i.e. the cue) or exogenously (i.e. by a target-colored distracter). The DCM results indicated that FEF had an excitatory stimulus-driven response, but that the activation seen in TPJ was temporally delayed. Furthermore, the only significant modulation of effective connectivity in the target-colored distracter condition was a reduction in the strength of
coupling between signals from FEF to TPJ. We conclude that R TPJ does not engage in stimulus-driven reorientation of attention by circuit-breaking activity in the dorsal network (e.g. FEF), but rather suggest that TPJ engages in informing internal contextual models of sensory-action relationships in the current task.

Funding
This work was supported by the Hellman Foundation and the National Science Foundation (BCS-1230777–0).

Notes
We would like to thank Katerina Monlux for her help in scanning participants. Conflict of Interest: None declared.

References


