Behavioral/Cognitive

Bottom-Up and Top-Down Factors Differentially Influence Stimulus Representations Across Large-Scale Attentional Networks

 Nicole M. Long and Brice A. Kuhl

Department of Psychology, University of Oregon, Eugene, Oregon 97402

Visual attention is thought to be supported by three large-scale frontoparietal networks: the frontoparietal control network (FPCN), the dorsal attention network (DAN), and the ventral attention network (VAN). The traditional view is that these networks support visual attention by biasing and evaluating sensory representations in visual cortical regions. However, recent evidence suggests that frontoparietal regions actively represent perceptual stimuli. Here, we assessed how perceptual stimuli are represented across large-scale frontoparietal and visual networks. Specifically, we tested whether representations of stimulus features across these networks are differentially sensitive to bottom-up and top-down factors. In a pair of pattern-based fMRI studies, male and female human subjects made perceptual decisions about face images that varied along two independent dimensions: gender and affect. Across studies, we interrupted bottom-up visual input using backward masks. Within studies, we manipulated which stimulus features were goal relevant (i.e., whether gender or affect was relevant) and task switching (i.e., whether the goal on the current trial matched the goal on the prior trial). We found that stimulus features could be reliably decoded from all four networks and, importantly, that subregions within each attentional network maintained coherent representations. Critically, the different attentional manipulations (interruption, goal relevance, and task switching) differentially influenced feature representations across networks. Whereas visual interruption had a relatively greater influence on representations in visual regions, goal relevance and task switching had a relatively greater influence on representations in frontoparietal networks. Therefore, large-scale brain networks can be dissociated according to how attention influences the feature representations that they maintain.

Key words: attention; cognitive control; decoding; parietal cortex; prefrontal cortex; resting-state networks

Significance Statement

Visual attention is supported by multiple frontoparietal attentional networks. However, it remains unclear how stimulus features are represented within these networks and how they are influenced by attention. Here, we assessed feature representations in four large-scale networks using a perceptual decision-making paradigm in which we manipulated top-down and bottom-up factors. We found that top-down manipulations such as goal relevance and task switching modulated feature representations in attentional networks, whereas bottom-up manipulations such as interruption of visual processing had a relatively stronger influence on feature representations in visual regions. Together, these findings indicate that attentional networks actively represent stimulus features and that representations within different large-scale networks are influenced by different forms of attention.

Introduction

Visual attention is thought to be supported by several frontoparietal networks (Posner and Petersen, 1990; Corbetta and Shul-

man, 2002; Dosenbach et al., 2008). The idea that the brain is comprised of multiple functional networks has been inspired and elaborated by resting-state analyses of human fMRI data (Yeo et al., 2011), which reveal three networks of particular importance to attentional control: the frontoparietal control network (FPCN), the dorsal attention network (DAN), and the ventral

Correspondence should be addressed to either Nicole Long or Brice Kuhl, Department of Psychology, 1227 University of Oregon Eugene, OR 97403. E-mail: niclong@uoregon.edu or bkuhl@uoregon.edu.

attention network (VAN). Traditionally, these networks have been thought to support visual attention by biasing and evaluating sensory representations within visual cortical areas (Desimone and Duncan, 1995; Egner and Hirsch, 2005; Serences and Yantis, 2006; Gazzaley and Nobre, 2012). However, recent evidence from pattern-based fMRI studies has blurred the distinction between sensory representations in visual cortical areas and control processes in frontoparietal regions. Namely, there is accumulating evidence that frontoparietal regions actively represent stimulus features during visual attention and working memory (Ester et al., 2015; Lee and Kuhl, 2016; Xu, 2017). These findings suggest a potentially transformative approach for understanding the functional role of frontoparietal networks in visual attention: that frontoparietal networks can be characterized – and dissociated from visual cortical regions – in terms of how they represent stimuli in relation to bottom-up and top-down factors.

Attention manipulations that may dissociate stimulus representations in frontal and parietal cortices from those in visual cortex include robustness to interruption of visual processing, goal relevance, and task switching. For example, working memory representations in the intraparietal sulcus (Bettencourt and Xu, 2016) and prefrontal cortex (Miller et al., 1996) are more robust to distraction than are representations in visual cortical areas. Likewise, multiple frontal and parietal regions preferentially represent goal-relevant stimulus information, as shown via electrophysiological recordings in monkeys (Rainer et al., 1998; Swaminathan and Freedman, 2012; Roy et al., 2014; Sarma et al., 2016) and pattern-based fMRI in humans (Kuhl et al., 2013; Sreenivasan et al., 2014; Ester et al., 2015; Bracci et al., 2017). Finally, dorsal frontal and parietal regions show increased univariate activation on trials when goals change (task switching) (Braver et al., 2003; Monsell, 2003; Yeung et al., 2006; Ravizza and Carter, 2008; Bode and Haynes, 2009; Esterman et al., 2009). Collectively, these findings suggest that different forms of attention potentially differentiate stimulus representations across frontoparietal and visual regions. However, there has been limited application of "representation-based" analyses to large-scale networks. Do frontoparietal networks actively and coherently represent stimulus information? Are representations in different networks influenced by different forms of attention?

Here, we conducted a pair of pattern-based fMRI studies to determine how various attention-related manipulations (interruption of visual processing, goal relevance, and task switching) influence feature representations in frontoparietal attentional networks (FPCN, DAN, and VAN) and, as a comparison, within a network of visual regions (VisN). In both studies, subjects viewed faces that varied along two independent dimensions: gender (male vs female) and affect (happy vs grumpy). On each trial, subjects made a perceptual decision related to a cued face feature (e.g., "Male?"). Using pattern classification analyses, we first tested for representation of stimulus features within each network and, importantly, tested whether frontoparietal regions within a common network maintained "coherent" representations. Next, we determined how each attention manipulation influenced feature representations across networks. Visual interruption was manipulated across studies: in Study 1, stimuli were followed by a visual mask; in Study 2, there was a longer stimulus duration and no mask. Goal relevance was manipulated by varying the dimension (gender/affect) that was currently relevant. Finally, task switching was manipulated by alternating between goals on a trial-by-trial manner, resulting in trials in which goals repeated (stay trials) and trials in which goals changed (switch

trials). We predicted that frontoparietal networks would actively and coherently represent stimulus features and that frontoparietal representations would be relatively more sensitive to top-down manipulations (goals, task switching), whereas representations in VisN would be relatively more sensitive to bottom-up manipulations (interruption).

Materials and Methods

Subjects
Thirty-two (19 female; mean age = 22 years) right-handed, native English speakers from the University of Oregon community participated in the fMRI studies. Sixteen subjects participated in Experiment 1 and 16 participated in Experiment 2. Four total subjects were excluded. In Experiment 1, one subject was excluded for poor task performance (65% accuracy, which was >4 SDs away from the average performance of 95%) and one was excluded for excessive head motion. In Experiment 2, two subjects were excluded for exiting the scanner before task completion (one subject complained of nausea and the other began coughing repeatedly; each missing two of six runs). Therefore, for each experiment, there was a final set of 14 subjects included for analyses. All subjects had normal or corrected-to-normal vision. Informed consent was obtained in accordance with the University of Oregon Institutional Review Board. The raw, de-identified data and the associated experimental and analysis codes used in this study can be accessed via the Kuhl laboratory website (http://kuhllab.com/publications/).

An additional 14 (5 female; mean age = 19) right-handed, native English speakers from the University of Oregon community participated in a follow-up behavioral study. All subjects had normal or corrected-to-normal vision. Informed consent was obtained in accordance with the University of Oregon Institutional Review Board and subjects received course credit for participating.

Materials
Stimuli consisted of 88 face images drawn from various internet sources. From this pool, 16 faces were designated as "target" faces, 24 as "filler" faces, and 48 as "localizer" faces. Within each set (target, filler, localizer) gender and affect were balanced: 1/4 of the faces were "happy, males"; 1/4 were "happy, females"; 1/4 were "grumpy, males"; and 1/4 were "grumpy, females." Gender and affect were determined, with unanimous agreement, by two independent raters.

Experimental design and statistical analysis
Procedure and design. Each trial began with the presentation of one of four "goal cues" or questions: "Male?," "Female?," "Happy?," or "Grumpy?" The goal was presented for 1400 ms (Fig. 1A) and was immediately followed by a face stimulus. In Experiment 1, the face was presented for 100 ms and immediately followed by a visual mask composed of scrambled face parts for 500 ms. In Experiment 2, the face was presented for 600 ms with no mask. After the mask (Experiment 1) or face (Experiment 2), there was a 6 s interstimulus interval (ISI) during which a fixation cross was shown. On each trial, the subject's task was to respond "Yes" or "No" via button box as to whether the face matched the goal/question. Subjects could respond at any point after the onset of the face, including during the ISI.

Trials were organized into blocks. Each block contained 17 trials, with the first trial in each block representing a "filler" face. The designation of the first trial as a filler face was particularly important for analyses of "stay" versus "switch" trials because the first trial cannot be designated as either a stay or switch trial (see below). The stimuli for the remaining 16 trials in each block were target faces. Each of the 16 target faces appeared exactly once per block. Within each block, only two of the four goals were presented (either "Male?" or "Female?" or either "Happy?" or "Grumpy?"). The goals alternated in an A-A-B-B-A-A-B-B manner (Fig. 1C). Therefore, there was a constant alternation between switch trials (goal change) and stay trials (goal repeat). A 3 s get ready screen at the start of each block informed subjects which two goals would occur within that block.

Each of the six scan runs contained four blocks. The four blocks in each run comprised each of the four possible combinations of goals (Male/
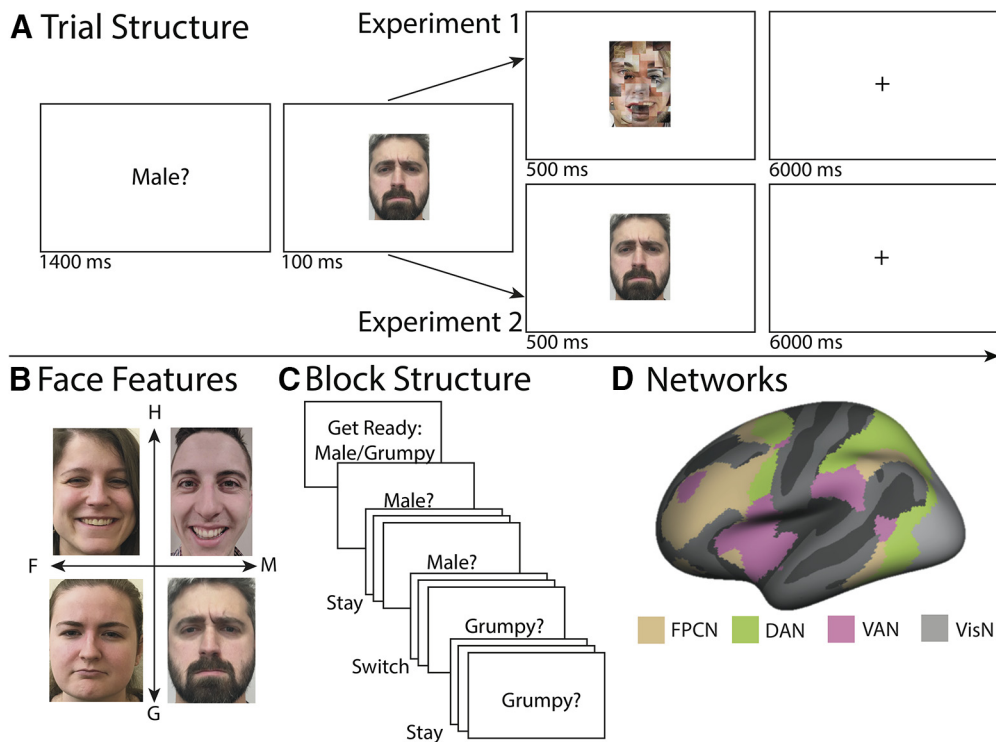
Figure 1. Experimental design. *A*, In a given trial, a goal cue was presented for 1400 ms. There were one of four possible goal cues ("Male?," "Female?," "Happy?," or "Grumpy?"). After the goal, a face was presented for 100 ms. In Experiment 1, the face was followed by a mask image for 500 ms. In Experiment 2, the face was presented for an additional 500 ms. A fixation cross was then presented for 6000 ms. Subjects responded "Yes" or "No" as to whether the face matched the goal. *B*, Faces varied along two dimensions: gender and affect. Note that faces shown here were not those used in the experiment. Permission was given for these faces to be published. *C*, Each block followed the same general structure. At the start of each block, subjects were cued as to which goals would be relevant for that block. Only two goals were relevant in each block (one from each dimension) and these goals alternated in a fixed AABBAABB schedule. This enforced a constant alternation between switch and stay trials. *D*, Four independently defined resting-state networks of a priori interest (Yeo et al., 2011).

Grumpy, Female/Grumpy, Male/Happy, and Female/Happy). Block order was randomized across runs and across subjects. The order of the 16 target faces was also randomized in each block. Across the six scan runs, there was a grand total of 408 trials (including filler trials). To familiarize subjects with the task, before entering the scanner, subjects completed a 24-trial practice round. Subjects practiced each goal combination (e.g., "Male?/Happy?") and the stimuli used were the 24 filler faces.

After the main experiment, subjects completed two localizer runs. However, because data from the localizer scan are not reported here, details of the task are not included.

The design of the behavioral follow-up study was identical to Experiment 1, with the exception of the ISI (3 s instead of 6 s) and the inclusion of a post-task questionnaire, in which subjects were asked increasingly specific questions as to whether they noticed the task switching structure. We first asked subjects if they noticed any patterns throughout the task. We then noted that there was a pattern to the goal cues and asked if they were aware of such a pattern. Finally, we explicitly asked subjects if they noticed the A-A-B-B-A-A-B-B structure of the goal cues.

fMRI data acquisition
Imaging data were collected on a Siemens 3 T Skyra scanner at the Robert and Beverly Lewis Center for NeuroImaging at the University of Oregon. Before the functional imaging, a whole-brain high-resolution anatomical image was collected for each subject using a T1-weighted protocol (grid size 256 × 256; 176 sagittal slices; voxel size 1 × 1 × 1 mm). Whole-brain functional images were collected using a T2*-weighted multiband accelerated EPI sequence (TR = 2 s; TE = 25 ms; flip angle = 90°; 72 horizontal slices; grid size 104 × 104; voxel size 2 × 2 × 2 mm). For the main experiment, six functional scans were collected, each consisting of 280 volumes. For the localizer task, two functional scans were collected each consisting of 225 volumes.

fMRI data preprocessing
Preprocessing of the functional data was conducted using FSL 5.0 (FMRIB Software Library, http://www.fmrib.ox.ac.uk/fsl; Smith et al., 2004) and custom scripts. Images were first corrected for head motion using MCFLIRT (Jenkinson et al., 2002). Motion-corrected images were smoothed with a Gaussian kernel with 1.7 mm SD (∼4 mm FWHM).

Network selection
We assessed feature representations in four resting-state networks defined from a large, independent sample of subjects (Yeo et al., 2011): the FPCN, DAN, VAN, and VisN (Fig. 1D). The resting-state networks were generated for each subject using their high-resolution anatomical image and the FreeSurfer cortical parcellation scheme (http://surfer.nmr.mgh. harvard.edu). The networks were then coregistered to the functional data.

Univariate analyses
Univariate data analyses were conducted under the assumptions of the general linear model (GLM) using SPM12 (http://www.fil.ion.ucl.ac.uk/ spm). To test for univariate effects of switch versus stay trials, we defined a model with separate regressors for switch and stay trials. The model also included regressors for scan run and six motion parameters for each run. Switch versus stay trials were contrasted using paired-samples t tests resulting in subject-specific statistical parametric maps. These t values were then averaged within network, resulting in a single mean t statistic per network and subject. For each network, one-sample t tests were then applied across subjects to test for effects of switch versus stay trials at the group level.

Pattern classification analyses
For pattern classification analyses, functional data were detrended, high-pass filtered (0.01 Hz), and z-scored within scan (mean response of each

voxel within each scan = 0). Next, data were temporally compressed via a weighted averaged of TRs 3, 4, and 5 (40%, 40%, 20%, respectively) relative to trial onset (representing 4–10 s after the goal was presented). TRs 3 and 4 were given greater weight because the hemodynamic response tends to peak at these TRs. The temporally compressed data resulted in a single spatial pattern of activity for each trial. Before pattern classification analyses were performed, an additional round of z-scoring was applied across voxels to the trial-specific spatial patterns. This final round of z-scoring resulted in each trial-specific spatial volume having a mean activation equal to 0. Therefore, mean univariate activity was matched precisely across all conditions and trial types (Kuhl and Chun, 2014; Long et al., 2016). Pattern classification analyses were performed using penalized (L2) logistic regression (penalty parameter = 1) implemented via the Liblinear toolbox (Fan et al., 2008) and custom MATLAB (RRID:SCR 001622) code. Classifier performance was assessed in two ways. "Classification accuracy" represented a binary coding of whether the classifier successfully guessed the queried feature of the face. We used classification accuracy for general assessment of classifier performance (i.e., whether features could be decoded). "Classifier evidence" was a continuous value reflecting the logit-transformed probability that the classifier assigned the correct feature each trial. Classifier evidence was used as a trial-specific, continuous measure of feature information, which was used to assess trial-level correlations in feature representations within and between networks (see below).

For each subject, four separate classifiers were trained to decode stimulus features. A given classifier was trained to discriminate either male from female faces (gender classifier) or happy from grumpy faces (affect classifier). Additionally, separate gender classifiers were applied for trials in which the goal was either "Male?" or "Female?" (i.e., trials in which gender was relevant) versus trials in which the goal was either "Happy?" or "Grumpy?" (i.e., trials in which affect was relevant). Likewise, separate affect classifiers were applied for trials in which affect was the relevant dimension versus trials in which gender was the relevant dimension. Goal-relevant feature representations were indexed by performance of the gender classifier on gender-relevant trials and performance of the affect classifier on affect-relevant trials. Goal-irrelevant feature representations were indexed by performance of the gender classifier on affect-relevant trials and performance of the affect classifier on gender-relevant trials. All classification analyses were performed using leave-one-run-out cross-validation. A critical element of our design and implementation of classification analyses is that we deliberately orthogonalized feature information from behavioral responses. In other words, there was no consistent mapping between feature information and motor response. As an example, during trials in which gender was relevant, 50% of the time, a female face would require a "Yes" response (i.e., on "Female?" trials) and, 50% of the time, a female face would require a "No" response (i.e., on "Male?" trials). Therefore, successful decoding of goal-relevant feature information cannot be attributed to decoding of the planned or executed motor responses. A second critical element of our design is that, across trials, goals and features were not always matched. In fact, they were independent. That is, when presented with a goal cue of "Male?," the subsequently presented stimulus was equally likely to be a male or female face. This design feature critically allowed us to deconfound goal and feature information. In other words, decoding of stimulus features cannot be driven exclusively by goal information.

To decode top-down goals, an additional four classifiers were applied. The four classifiers corresponded to the four different possible combinations of goals in a given block: "Male?" versus "Grumpy?," "Male?" versus "Happy?," "Female?" versus "Grumpy?," and "Female?" versus "Happy?" Classification was performed using leave-one-run-out cross-validation. Classification accuracy was averaged across all four classifiers to provide a single measure of goal decoding accuracy. Because goal decoding could potentially be driven by low-level information such as visual word form or subvocal articulation, we also tested whether goal classifiers generalized to different goals corresponding to the same dimensions. For example, a classifier trained to discriminate between "Male?" versus "Happy?" faces can also be thought of as a classifier that is discriminating between the gender versus affect dimensions. If so, then a classifier trained on "Male?" versus "Happy?" may successfully transfer to the discrimination

of "Female?" versus "Grumpy?" Successful transfer would suggest that the goal representation is, at least in part, a representation of the relevant dimension as opposed to the specific word cue per se. Therefore, for each of the four goal classifiers described above, we tested for transfer to the "complementary" pair of goals; i.e., goal pairs that corresponded to discrimination between the same dimensions. This dimension decoding analysis was again performed using leave-one-run-out cross-validation so as to match the goal decoding analyses in terms of statistical power. Classification accuracy was averaged across all four transfer tests to provide a single aggregate value of dimension decoding accuracy.

Representational coherence analysis
To assess the coherence of representations within and between networks, we decomposed each of the three attentional networks into separate frontal and parietal ROIs: frontal-FPCN, parietal-FPCN, frontal-DAN, parietal-DAN, frontal-VAN, and parietal-VAN. None of the voxels from the frontal ROIs were contiguous with voxels from the parietal ROIs. These ROIs were generated using the average subject brain in Free Surfer. The ROIs were then projected from this volume space to surface space and then converted from average surface space to subject space. Because the number of voxels in these regions varied both within and across subjects and differences in ROI size are likely to influence classifier performance, classification analyses were performed by randomly subsampling 500 voxels from each of the six frontal and parietal ROIs. This process was repeated for 100 iterations for each ROI and subject, with each iteration involving a different random sample of 500 voxels. Classification of goal-relevant and goal-irrelevant features was performed using the same approach described above. Here, however, classifier evidence (a continuous value reflecting the strength of classifier information) was the critical dependent measure. For each subject, trial-by-trial measures of classifier evidence were correlated within network (e.g., a correlation between classifier evidence from frontal-FPCN and parietal-FPCN) and between network (e.g., frontal-FPCN and parietal-DAN). The correlation analyses were separately performed for relevant and irrelevant feature evidence. In total, we obtained 50,400 ρ values: 100 iterations × 9 correlations × 28 subjects × 2 relevance conditions. All ρ values were Fisher-z transformed before averaging across iterations and goal relevance. Finally, z-ρ values were averaged across the three "within" network correlations (e.g., correlations between frontal-FPCN and parietal-FPCN) and the six "between" network correlations (e.g., correlations between frontal-FPCN and parietal-DAN), resulting in two z-ρ values for each subject, mean within network correlation, and mean between network correlation.

Statistical analyses
One-sample t tests were used to compare representational coherence measures (z-ρ values) to zero. Paired-sample t tests were used to compare classification accuracy across subjects to chance decoding accuracy, as determined by permutation procedures. Namely, for each subject and network, we shuffled the condition labels of interest (e.g., "male" and "female" for the gender feature classifier) and then calculated classification accuracy. We repeated this procedure 1000 times for each network and subject and then averaged the 1000 shuffled accuracy values for each network and subject. These mean values were used as network- and subject-specific empirically derived measures of chance accuracy. Paired samples t tests compared the true (unshuffled) accuracy values to the shuffled accuracy values. For these paired-sample t tests, we report uncorrected p-values; however, all of the p-values exceeded the threshold for significance after Bonferroni correction; that is, after adjusting for the fact that we tested effects across four networks (i.e., a threshold of p = 0.0125). Mixed-effects ANOVAs were used to compare conditions and/or networks; experiment was always included as a factor when data from both experiments were included.

Results

Large-scale networks represent stimulus features
Motivated by recent evidence that activity patterns in frontoparietal regions represent stimulus features (Swaminathan and Freedman, 2012; Ester et al., 2015; Lee and Kuhl, 2016; Bracci et
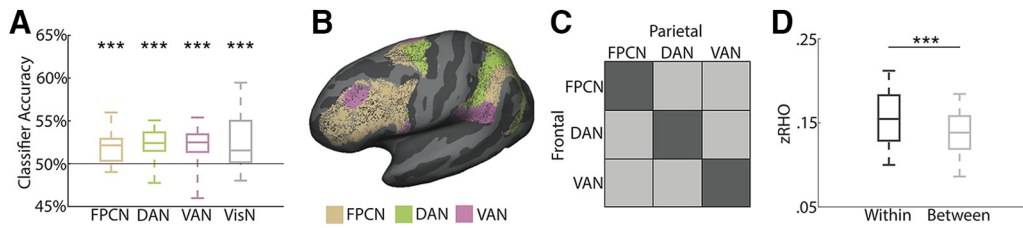
Figure 2. Feature decoding across networks. **A**, Feature decoding averaged across experiments and goal relevance. Feature decoding was reliably above chance in all networks. **B–D**, Representational coherence within and between networks. **B**, Each attentional network was divided into discontiguous frontal and parietal subregions (here shown for a single subject). **C**, Correlations were computed for trial-level measures of goal-relevant feature evidence for each pair of frontal and parietal subregions. Within-network correlations (dark gray) corresponded to correlations between pairs of subregions from the same network (e.g., frontal-FPCN and parietal-FPCN). Between-network correlations (light gray) corresponded to correlations between pairs of subregions from different networks (e.g., frontal-FPCN and parietal-DAN). This analysis was then repeated for goal-irrelevant feature evidence and Fisher Z-transformed ρ values were then averaged across the goal-relevant and goal-irrelevant correlation analyses. **D**, Within-network correlations were reliably stronger than between-network correlations. ***$p < 0.001$.

al., 2017), we first tested for representations of face features (gender, affect) within each of the three attentional networks (FPCN, DAN, VAN) and, for comparison, within VisN. Importantly, and as described in the Materials and Methods, our procedure deliberately deconfounded feature information from behavioral responses, so decoding accuracy cannot be explained by decoding of response preparation or execution. Averaging across the gender and affect classifiers (producing a single value per subject reflecting feature decoding), accuracy was above chance in each of the four networks (FPCN, $t_{(27)} = 5.82$, $p < 0.001$; DAN, $t_{(27)} = 5.37$, $p < 0.001$; VAN, $t_{(27)} = 5.42$, $p < 0.001$; VisN, $t_{(27)} = 4.34$, $p < 0.001$; Figure 2A), confirming the sensitivity of these large-scale networks to feature information. There were no significant differences in decoding accuracy for the gender versus affect classifiers for any of the four networks ($t < 1$, $p > 0.35$). A mixed-effects ANOVA with factors of network (FPCN, DAN, VAN, VisN) and experiment (1, 2) did not reveal a main effect of network ($F_{(3,78)} = 0.49$, $p = 0.69$), indicating that overall feature decoding was comparable across networks. The effects of experiment are considered below.

Representational coherence within networks
The preceding results demonstrate that activity patterns in resting-state networks represent stimulus features. However, the use of large-scale networks as ROIs is predicated on the notion that individual components (brain regions) within these networks are acting together in a functionally relevant way to support behavior. To test for coherence of representations within the attentional networks, we divided each of the three attentional networks into frontal and parietal subregions (see Materials and Methods and ROIs from a sample subject in Fig. 2B) and correlated trial-by-trial feature evidence within and between each network (Fig. 2C). Correlations were significantly >0 both within networks (M = 0.15, SD = 0.03; $t_{(27)} = 24.7$, $p < 0.001$) and between networks (M = 0.14, SD = 0.03; $t_{(27)} = 27.4$, $p < 0.001$; Fig. 2D). A 2 × 2 mixed-effects ANOVA with factors of network pairing (within, between) and experiment (1, 2) revealed a main effect of network pairing ($F_{(1,26)} = 23.9$, $p < 0.001$), with stronger correlations within networks than between networks. This main effect of network pairing did not interact with experiment ($F_{(1,26)} = 0.03$, $p = 0.86$). Therefore, although feature representations were present across all of the attentional networks, there was greater representational coherence within networks than between networks, validating the use of these networks as large-scale ROIs for pattern-based analyses.

Robustness of feature representations to visual interruption
We next tested the relative sensitivity of feature representations in each network to the interruption of visual processing by compar-
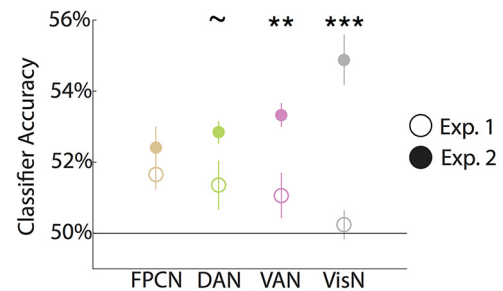

Figure 3. Feature decoding as a function of interruption. Interruption of visual processing was manipulated across experiments. In Experiment 1 (clear circles), stimuli were followed by a visual mask. In Experiment 2 (filled circles), there was a longer stimulus duration and no mask. There was a significant network-by-experiment interaction ($p < 0.001$). Significant differences between experiments were observed in VAN and VisN. Error bars indicate SEM. $\sim p < 0.10$, **$p < 0.01$, ***$p < 0.001$.

ing decoding accuracy across experiments. Sensitivity to visual interruption would be reflected by relatively lower decoding accuracy in Experiment 1 (100 ms stimulus + 500 ms visual mask) than in Experiment 2 (600 ms stimulus + no visual mask; Fig. 3) mixed-effects ANOVA with factors of experiment (1, 2) and network (FPCN, DAN, VAN, VisN) revealed a main effect of experiment ($F_{(1,26)} = 26.5$, $p < 0.001$), with lower decoding accuracy in Experiment 1 than Experiment 2, consistent with a disruptive influence of interruption. Relative sensitivity to visual interruption markedly varied across networks, as reflected by a significant experiment by network interaction ($F_{(3,78)} = 6.2$, $p < 0.001$).

Post hoc independent sample t tests revealed no reliable differences between experiments in FPCN ($t_{(26)} = 1.1$, $p = 0.29$), a trend in DAN ($t_{(26)} = 2.0$, $p = 0.06$), and reliably greater feature decoding in Experiment 2 than Experiment 1 in VAN ($t_{(26)} = 3.2$, $p = 0.003$) and VisN ($t_{(26)} = 5.8$, $p < 0.001$). Therefore, the experimental manipulation of including a brief visual mask (Experiment 1) versus extended stimulus presentation (Experiment 2), the only difference between the two experiments, robustly diminished feature representations in VisN and VAN, but had relatively less influence on feature representations in DAN and FPCN.

Goal relevance influences feature representations
We next tested the relative sensitivity of feature representations in each network to trial-specific behavioral goals; that is, whether goal-relevant feature information was stronger than goal-irrelevant feature information (Kuhl et al., 2013; Roy et al., 2014; Sarma et al., 2016; Bracci et al., 2017). For example, on trials in which gender was the relevant dimension ("Male?" or "Female?"
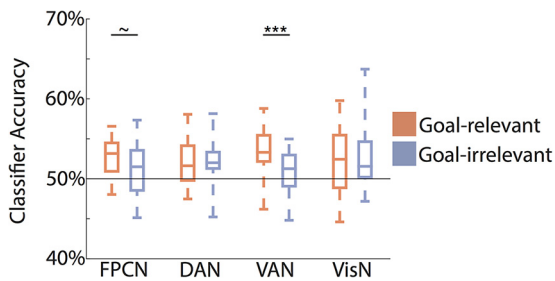
Figure 4. Feature decoding as a function of goal relevance. Orange indicates goal-relevant feature decoding (e.g., decoding whether a face was male vs female when the goal was "Male?" or "Female?"). Blue indicates goal-irrelevant feature decoding (e.g., decoding whether a face was male vs female when the goal was "Happy?" or "Grumpy?"). There was a significant interaction between network and goal relevance ($p = 0.03$). Goal-relevant feature decoding was reliably greater than goal-irrelevant feature decoding in VAN, with a similar trend in FPCN. $\sim p < 0.10$, ***$p < 0.001$.

trials), goal-relevant feature information was indexed by accuracy of the gender classifier and goal-irrelevant feature information was indexed by accuracy of the affect classifier. To be clear, for this analysis, decoding accuracy refers only to the feature information in the actual face image. Therefore, if gender was the relevant dimension and the face image was a female face, then the classifier was accurate if it guessed "Female" regardless of whether the current goal was "Male?" or "Female?" This approach ensured that feature information was not confounded with goal information. Separate classifiers were trained/tested for relevant and irrelevant feature representations. That is, one set of classifiers was specifically trained and tested on goal-relevant feature information and a separate set was trained and tested on goal-irrelevant feature information. Separate classifiers were used for goal-relevant and goal-irrelevant feature decoding so that there was no assumption that goal-relevant and goal-irrelevant feature representations are encoded in a common (generalizable) way.

A mixed-effects ANOVA with factors of experiment, goal relevance (relevant, irrelevant), and network (FPCN, DAN, VAN, VisN) revealed a marginally significant main effect of relevance ($F_{(1,26)} = 3.93$, $p = 0.06$), reflecting a trend for stronger decoding of goal-relevant than goal-irrelevant feature information (Fig. 4). The effect of relevance did not vary by experiment ($F_{(1,26)} = 0.07$, $p = 0.80$). More critically, there was a significant interaction between network and goal relevance ($F_{(3,78)} = 3.12$, $p = 0.03$), indicating that the sensitivity of feature representations to top-down goals varied across the networks. This interaction between network and relevance did not vary by experiment ($F_{(3,78)} = 1.24$, $p = 0.30$).

For each network, we ran follow-up mixed-effects ANOVAs with factors of goal relevance (relevant, irrelevant) and experiment to test for effects of goal relevance within each network. A significant main effect of goal relevance (relevant > irrelevant) was observed in VAN ($F_{(1,26)} = 13.9$, $p < 0.001$), with a similar trend in FPCN ($F_{(1,26)} = 3.7$, $p = 0.07$). There was no effect of relevance in DAN ($F_{(1,26)} = 0.03$, $p = 0.87$) or VisN ($F_{(1,26)} = 0.12$, $p = 0.73$). None of the main effects of relevance interacted with experiment ($F < 2.5$, $p > .1$).

Task switches influence feature representations
In the preceding section, we considered whether goal-relevant features were more strongly represented than goal-irrelevant features. However, goal cues (and, therefore, the relevant feature dimension) changed every two trials, creating both switch (goal change) and stay (goal repeat) trials. To measure the influence of

goal cue switches, we first probed the behavioral data for effects of task switching. Reaction times (RTs) were compared as a function of trial type (switch, stay) and experiment in a mixed-effects ANOVA. As expected, RTs were significantly greater for switch trials (M = 1423 ms, SD = 259 ms) than stay trials (M = 1398 ms, SD = 269 ms; $F_{(1,26)} = 12.7$, $p = 0.001$; Fig. 5A). This effect did not interact with experiment ($F_{(1,26)} = 3.2$, $p = 0.08$). Accuracy did not significantly differ across switch and stay trials (switch, M = 96.92%, SD = 3.17%; stay, M = 97.52%, SD = 2.17%; $t_{(27)} = -1.5$, $p = 0.13$). There was a trend toward an interaction between switch/stay and experiment ($F_{(1,26)} = 4.03$, $p = 0.06$), reflecting a relatively greater effect of accuracy (stay > switch) in Experiment 2.

To determine whether the predictable nature of goal cues (i.e., the A-A-B-B-A-A-B-B alternation) was detected by subjects and/or influenced switch costs (Rogers and Monsell, 1995; Monsell, 2003), we ran an independent behavioral study in which 14 subjects completed an experiment virtually identical to Experiment 1 (see Materials and Methods). Critically, however, this behavioral experiment also included a postexperiment questionnaire to assess whether subjects became aware of the alternating pattern of goal cues. Across all subjects, RTs were significantly greater for switch trials (M = 1298 ms, SD = 227 ms) than stay trials (M = 1257 ms, SD = 204 ms; $t_{(13)} = 2.58$, $p = 0.02$), consistent with the data from Experiments 1 and 2. Of the 14 subjects, four were able to explicitly describe the task structure (i.e., they were aware of the pattern of alternation). Among these four subjects, switch costs were numerically greater (M = 59 ms) relative to subjects who were unaware of the task structure (M = 33 ms). These data indicate that, despite the highly structured task alternation, most subjects did not become aware of this structure. In addition, at least in this small sample of subjects, there was no evidence that awareness of the structure reduced switch costs. Therefore, it is unlikely that explicit awareness of the task structure had a major influence on fMRI-based effects of task switching.

Prior univariate fMRI studies of task switching have consistently reported increased activation during switch versus stay trials in dorsal frontoparietal regions (overlapping with FPCN and DAN), whereas more ventral regions (overlapping with VAN) are generally insensitive to switch effects (Kimberg et al., 2000; Rushworth et al., 2002; Braver et al., 2003; Brass and Von Cramon, 2004; Corbetta et al., 2008). In light of these prior studies, and as a comparison point for our decoding analyses, we tested for univariate effects of task switching in each of the four resting-state networks. Consistent with prior findings, we found reliably greater activation for switch than stay trials in FPCN ($t_{(27)} = 4.0$, $p < 0.001$) and DAN ($t_{(27)} = 2.5$, $p = 0.02$) and no effect in VAN ($t_{(27)} = 1.2$, $p = 0.23$) or VisN ($t_{(27)} = 0.23$, $p = 0.82$; Fig. 5B). *Post hoc* independent-samples t tests within each network did not reveal any between-experiment differences in switch effects in FPCN, DAN, or VAN ($t < 1.5$, $p > .1$). In VisN, there was a reliable difference between experiment ($t_{(26)} = -2.3$, $p = 0.03$), although neither experiment alone showed a reliable difference between switch and stay trials. In Experiment 1, activation was relatively greater for stay than switch trials ($t_{(13)} = -1.7$, $p = 0.12$), whereas in Experiment 2, activation was relatively greater for switch than stay trials ($t_{(13)} = 1.6$, $p = 0.14$).

Next, we considered the novel question of whether task switches influenced the strength of feature representations. For this analysis, we used the same classifiers described above and then back-sorted trial-specific accuracy values as a function of trial type (switch vs stay). Decoding accuracy was entered into a
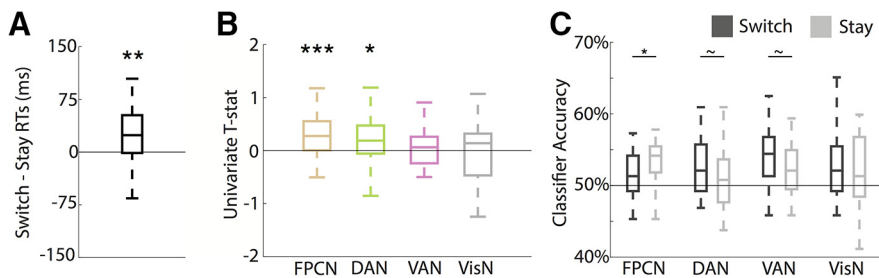
Figure 5. Effects of task switching. ***A****,* Difference in RTs (ms) between switch and stay trials. RTs were significantly slower for switch trials than stay trials. ***B****,* Univariate contrast of switch versus stay trials. Activation was significantly greater for switch than stay trials in both FPCN and DAN. ***C****,* Goal-relevant feature decoding separated by switch and stay trials. There was a significant interaction between switch/stay and network, reflecting greater goal-relevant feature decoding on stay than switch trials in FPCN; the opposite trends were true for DAN and VAN. $\sim p < 0.10$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

mixed-effects ANOVA with factors of task switch (switch, stay), relevance (relevant, irrelevant), network, and experiment. The main effect of task switch was not significant ($F_{(1,26)} = 0.09$, $p = 0.77$). However, there was a robust 3-way interaction between the effects of task switch, relevance, and network ($F_{(3,78)} = 4.50$, $p = 0.006$). Considering goal-relevant feature information alone, a mixed-effects ANOVA with factors of task switch, network, and experiment revealed a significant interaction between task switch and network ($F_{(3,78)} = 4.87$, $p = 0.004$). For goal-irrelevant feature information, a similar mixed-effects ANOVA did not reveal an interaction between task switch and network ($F_{(3,78)} = 0.60$, $p = 0.62$). Therefore, the influence of task switches on feature representations varied across the networks, but only when considering goal-relevant feature representations.

The interaction between task switch and network for goal-relevant feature representations was driven by relatively stronger representations on switch versus stay trials in DAN and VAN and an opposite pattern in FPCN. This interaction did not significantly vary by experiment ($F_{(3,78)} = 2.25$, $p = 0.09$). Follow-up mixed-effects ANOVAs with factors of task switch (switch, stay) and experiment revealed that, within FPCN, there was a significant effect of task switch ($F_{(1,26)} = 6.3$, $p = 0.02$), reflecting relatively lower feature information on switch than stay trials. In contrast, there were opposite trends (greater feature information on switch than stay trials) in DAN ($F_{(1,26)} = 4.2$, $p = 0.05$) and VAN ($F_{(1,26)} = 4.2$, $p = 0.05$). In VisN, there was no effect of task switch ($F_{(1,26)} = 1.1$, $p = 0.31$). The effect of task switch interacted with experiment in FPCN ($F_{(1,26)} = 6.6$, $p = 0.02$), but not in DAN, VAN, or VisN ($F < 2.2$, $p > .15$). For FPCN, the interaction between experiment and task switch reflected a greater decoding advantage on stay versus switch trials in Experiment 2 compared with Experiment 1. Collectively, these data indicate that task switches had an effect on goal-relevant feature information, but this effect varied across networks and was most evident in the attentional networks.

Network representations of task goals
Although our central aim was to assess feature representations across networks, a complementary question is how/whether these networks represent top-down behavioral goals. Previous research has revealed top-down goal effects within multiple frontoparietal regions (Esterman et al., 2009; Greenberg et al., 2010; Liu and Hou, 2013; Waskom et al., 2014; Waskom and Wagner, 2017). Here, to test for goal representations, we trained and tested four separate pairwise classifiers to distinguish each combination of gender and affect goals. For example, one classifier was trained to dissociate "Male?" versus "Grumpy?" goal cues. Averaging

across the pairwise classifiers, goal decoding was significantly above chance in each of the four networks (FPCN, $t_{(27)} = 8.8$, $p < 0.001$; DAN, $t_{(27)} = 10.6$, $p < 0.001$; VAN, $t_{(27)} = 5.9$, $p < 0.001$; VisN, $t_{(27)} = 10.8$, $p < 0.001$; Fig. 6). A mixed-effects ANOVA with factors of experiment and network revealed no main effect of experiment ($F_{(1,26)} = 0.76$, $p = 0.39$), a main effect of network ($F_{(3,78)} = 8.53$, $p < 0.001$), and no interaction between network and experiment ($F_{(3,78)} = 0.31$, $p = 0.81$). Adding task switch (switch, stay) as a factor (Waskom et al., 2014) revealed no main effect of switch on goal decoding ($F_{(1,26)} = 0.01$, $p = 0.92$), nor was there an interaction between the factors of task switch and network ($F_{(3,78)} = 0.70$, $p = 0.56$).

Because goals were communicated to subjects via word cues, successful goal decoding potentially reflects multiple influences: (1) "true" abstract goal representations, (2) representations of the visual word form of each goal cue, and/or (3) subvocal rehearsal of the word cues. Therefore, as a complementary measure, we also assessed "dimension" decoding; that is, decoding of the dimension (gender or affect) that was relevant on each trial. Dimension decoding was assessed by testing for transfer of classifiers across goal cues but within dimensions. For example, a classifier trained to discriminate "Male?" versus "Grumpy?" goals would be tested on "Female?" versus "Happy?" goals. Successful transfer would occur if the representation of the "Male?" goal is similar to the representation of the "Female?" goal and, likewise, if "Grumpy?" is similar to "Happy?" Although this analysis does not entirely rule out effects of visual word form or subvocal rehearsal, it at least reduces these concerns. One-sample t tests revealed reliable dimension decoding in all four networks (FPCN, $t_{(27)} = 7.9$, $p < 0.001$; DAN, $t_{(27)} = 7.5$, $p < 0.001$; VAN, $t_{(27)} = 5.1$, $p < 0.001$; VisN, $t_{(27)} = 10.0$, $p < 0.001$). A mixed-effects ANOVA with factors of experiment and network revealed no main effect of experiment ($F_{(1,26)} = 1.09$, $p = 0.31$), a main effect of network ($F_{(3,78)} = 5.53$, $p = 0.002$), and no reliable interaction between network and experiment ($F_{(3,78)} = 0.17$, $p = 0.92$).

Finally, to compare goal decoding and dimension decoding directly, we ran a mixed-effects ANOVA with factors of goal level (goal decoding vs dimension decoding), network, and experiment. There was a significant main effect of goal level ($F_{(1,26)} = 25.86$, $p < 0.001$), reflecting lower dimension decoding accuracy than goal decoding accuracy, which is not surprising given that dimension decoding requires transfer of the classifier to nonidentical goal cues. There was a significant main effect of network ($F_{(3,78)} = 8.32$, $p < 0.001$), reflecting (as with the separate ANOVAs for goal decoding and dimension decoding) relatively lower decoding accuracy in VAN than in the other networks (Fig. 6). The interaction between goal level and network was also significant ($F_{(3,78)} = 2.95$, $p = 0.04$), indicating that the "cost" of transferring across goal cues (but within dimensions) varied across networks. Namely, the difference between goal decoding and dimension decoding was most pronounced in VisN (Fig. 6).

Discussion
Here, using pattern-based fMRI methods and a perceptual decision-making task, we compared representations of stimulus features across multiple resting-state networks. We specifically
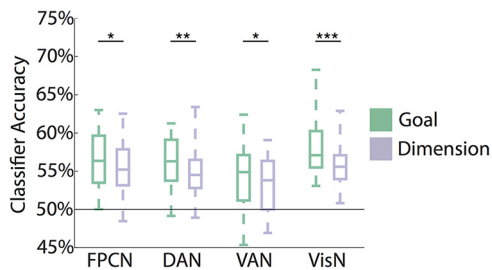
Figure 6. Goal and dimension decoding. For goal decoding (teal), four different classifiers were trained to dissociate each of the possible pairs of goals that appeared in a given block. The goal pairs always consisted of one goal from each dimension (e.g., "Male?" vs "Grumpy?" goals). For dimension decoding (lavender), the classifier training was identical to goal, but the classifiers were tested on different goal pairs corresponding to the same dimensions. For example, a classifier trained to dissociate Male/Grumpy goals would be tested on Female/Happy goals. All networks showed reliable goal and dimension decoding. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

targeted networks that are thought to contribute to attention and cognitive control (FPCN, DAN, and VAN; Dosenbach et al., 2008; Cole et al., 2013; Sestieri et al., 2017) and, as an important comparison, visual cortical networks. Although stimulus features were reliably decoded from each network, of critical interest was how feature representations in each network were influenced by three attentional manipulations: (1) interruption of visual processing, (2) goal relevance, and (3) task switching. Whereas bottom-up manipulations (interruption) had a relatively greater influence on feature representations within VisN, top-down manipulations (goals and task switches) had a relatively greater influence on representations in attentional networks. These findings reveal an important dissociation between feature representations in higher-level attentional networks and feature representations in visual cortex.

Whereas most decoding studies use spatially contiguous ROIs or searchlight analyses, we decoded from large, spatially discontiguous networks identified from independent studies of resting-state connectivity (Vincent et al., 2008; Yeo et al., 2011). We used these resting-state networks because they have been linked specifically to attentional processes (Fox et al., 2006; Corbetta et al., 2008; Gratton et al., 2017). That said, these networks have been defined based on correlations in univariate responses during rest and it was an open question whether regions within these networks would exhibit correlated feature representations, particularly after mean univariate responses were removed (see Materials and Methods). When specifically considering correlations between frontal and parietal regions from each attentional network, we found that decoded feature representations were more coherent within networks than between networks. Beyond validating our approach, this finding provides novel evidence that brain regions within attentional networks represent common stimulus information.

The traditional view of frontoparietal regions is that they bias and/or evaluate stimulus representations held in perceptual regions (Kastner et al., 1999; Zanto et al., 2010; Liu et al., 2011; Gazzaley and Nobre, 2012). However, our findings, along with other recent evidence (Kuhl et al., 2013; Bettencourt and Xu, 2016; Ester et al., 2016; Bracci et al., 2017), challenge this view by demonstrating active stimulus representations within frontoparietal regions. This raises the critical question of how frontoparietal representations differ from those in perceptual regions. We show that frontoparietal representations were sensitive to different forms of attention than representations in visual cortical ar-

eas. The fact that frontoparietal representations were relatively more sensitive to top-down factors helps to reconcile evidence that frontoparietal regions represent stimulus information with the putative role of these regions in top-down processing. Specifically, our results suggest a transformation of information from perceptual regions to frontoparietal regions that selectively represent and/or evaluate stimulus features. We next briefly consider the pattern of results for each attentional manipulation.

Previous work has shown that visual distraction disrupts working memory representations in visual cortex to a greater degree than representations in frontoparietal regions (Miller et al., 1996; Suzuki and Gottlieb, 2013; Woolgar et al., 2015; Bettencourt and Xu, 2016). Based on these findings, we predicted that feature representations in the attentional networks would be relatively less influenced by the across-experiment manipulation of visual interruption. Indeed, the effect of interruption (Experiment 1 vs Experiment 2) significantly differed across networks, with feature representations in VisN most strongly "suffering" from the interruption in bottom-up visual input (Fig. 3). Among the attentional networks, VAN showed a significant effect of interruption and there was a trend for DAN, whereas feature representations in FPCN were not influenced by interruption. The fact that VAN feature representations were influenced by interruption is consistent with the idea that VAN is involved in bottom-up attentional capture (Corbetta and Shulman, 2002).

One reason why frontoparietal regions may actively represent stimulus features is so that behaviorally relevant decisions can be made. Indeed, several recent studies have found that stimulus representations in frontoparietal regions are biased by task demands (Swaminathan and Freedman, 2012; Kuhl et al., 2013; Ester et al., 2016; Sarma et al., 2016; Bracci et al., 2017). We specifically designed our stimuli to be multidimensional so that we could test for flexible representation of individual features. As with visual interruption, we found that the influence of behavioral goals varied across networks (Fig. 4). In VAN, there was significantly greater representation of goal-relevant than goal-irrelevant features, with a similar trend in FPCN. At first pass, the effect of goal relevance in VAN seems at odds with the putative role of VAN in bottom-up attentional orienting (Corbetta and Shulman, 2002). For example, univariate responses in VAN increase for oddball stimuli or targets that appear at invalid locations (Bledowski et al., 2004; Kincade et al., 2005). However, more recent evidence suggests that VAN plays a role in comparing bottom-up input to top-down goals (Vossel et al., 2014; Gratton et al., 2017). In fact, a recent meta-analysis found greater VAN responses to task-relevant than task-irrelevant oddballs (Kim, 2014), suggesting that VAN's response to exogenous stimuli is moderated by top-down goals. Therefore, the present finding of preferential decoding of goal-relevant features in VAN is consistent with the view that VAN plays a role in orienting attention to task-relevant perceptual input.

VisN feature representations were unaffected by goals. Although other studies have clearly found that top-down factors can influence stimulus representations in early visual areas (Jehee et al., 2011; Sprague and Serences, 2013; Ester et al., 2016), the present findings suggest that frontoparietal regions can impose their own attentional biases to favor goal-relevant features as opposed to simply inheriting biases imposed in visual cortical regions. Potentially, the lack of attentional bias in VisN in the present study reflects the fact that we used stimuli (faces) that are processed holistically. With different stimulus types and/or attentional manipulations, it is likely easier to gate processing at earlier stages.

A large body of previous research indicates that switching between tasks (goals) is associated with increased univariate activity in dorsal frontoparietal regions (Corbetta et al., 2008). Our finding of greater univariate activation for switch versus stay trials in FPCN and DAN, but not VAN or VisN, is consistent with this literature. However, we are not aware of prior pattern-based fMRI studies that have compared frontoparietal feature representations across switch versus stay trials, so our analysis of task switching effects on feature representations was necessarily more exploratory. We found that the influence of task switching varied across networks (Fig. 5) and, in particular, that switching effects were more apparent in the attentional networks than VisN. Given that task switching requires reconfiguration of top-down attention, this finding is consistent with the argument that feature representations in attentional networks are relatively more sensitive to top-down attention. Among the attentional networks, FPCN showed greater feature decoding on stay than switch trials, whereas DAN and VAN showed opposite trends. Although we did not predict this pattern a priori, the tendency for feature representations to be stronger on switch than stay trials (in DAN and VAN) is reminiscent of recent evidence for greater decoding of top-down goals on switch than stay trials (Waskom et al., 2014). For FPCN, it is notable that switch trials were associated with relatively greater univariate activity but relatively lower feature decoding, raising the possibility that these measures reflect opposing processes. Given the exploratory nature of this analysis, we believe this question requires additional investigation.

Although our primary focus was on frontoparietal representations of stimulus features, several prior studies have reported frontoparietal representations of top-down goals (Waskom et al., 2014; Hanson and Chrysikou, 2017; Loose et al., 2017; Waskom and Wagner, 2017; Qiao et al., 2017) Consistent with these findings, we observed significant goal decoding in all four networks. One open question is whether the strength of goal representations is influenced by task switches. Although there is some evidence that goal representations are relatively stronger on switch trials than stay trials (Waskom et al., 2014), others have failed to observe switch-related effects in task representations (Loose et al., 2017). Like Loose et al. (2017), we did not observe a significant difference in goal decoding as a function of task switching; however, we did find task-switching effects in the decoding of goal-relevant features. Therefore, additional research will be needed to better understand how and when task switching influences the strength of top-down goals and/or goal-relevant feature representations.

Given that classifiers trained on one pair of goals (e.g., "Male" vs "Grumpy" goals) reliably transferred to nonidentical goals that shared the same dimensions (e.g., "Female" vs "Happy" goals), this suggests that goal representations reflected, at least in part, information about or attention to the goal-relevant stimulus dimension (gender or affect). The fact that the "transfer cost" (goal vs dimension decoding) was relatively greatest in VisN is consistent with the idea that goal decoding in VisN was more closely related to low-level properties of the goals (e.g., the visual word form of the cue). Considering goal-decoding performance across networks also provides a useful comparison point for the feature decoding results. For example, comparing overall goal decoding versus decoding of goal-relevant features across the three attentional networks (a 2 × 3 ANOVA) revealed a highly significant interaction ($F_{(2,52)} = 7.40$, $p = 0.001$), reflecting a dissociation between the networks that best represented goal-relevant features (VAN) versus the goals themselves (FPCN/DAN). Therefore, theoretical accounts of how these networks contribute to attention

will benefit from considering, not only how feature representations vary across networks, but also the hierarchical organization of feature and goal representations (Koechlin and Summerfield, 2007; Badre, 2008).

In summary, we show that resting-state networks implicated in attentional control actively and coherently represent stimulus features and that network-based feature representations can be dissociated in terms of their sensitivity to various forms of attention (interruption of visual processing, goal relevance, and task switching). Whereas feature representations in visual cortical areas are sensitive to low-level manipulations (visual interruption), feature representations in attentional networks are sensitive to higher-level manipulations (goal relevance and task switching). At a broad level, these findings indicate that multiple networks actively represent stimulus features, with the nature of these feature representations providing insight into each network's functional role.

References

Badre D (2008) Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. Trends Cogn Sci 12:193–200. CrossRef Medline

Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nat Neurosci 19:150–157. CrossRef Medline

Bledowski C, Prvulovic D, Goebel R, Zanella FE, Linden DE (2004) Attentional systems in target and distractor processing: a combined ERP and fMRI study. Neuroimage 22:530–540. CrossRef Medline

Bode S, Haynes JD (2009) Decoding sequential stages of task preparation in the human brain. Neuroimage 45:606–613. CrossRef Medline

Bracci S, Daniels N, Op de Beeck H (2017) Task context overrules object- and category-related representational content in the human parietal cortex. Cereb Cortex 27:310–321. CrossRef Medline

Brass M, von Cramon DY (2004) Decomposing components of task preparation with functional magnetic resonance imaging. J Cogn Neurosci 16:609–620. CrossRef Medline

Braver TS, Reynolds JR, Donaldson DI (2003) Neural mechanisms of transient and sustained cognitive control during task switching. Neuron 39:713–726. CrossRef Medline

Cole MW, Reynolds JR, Power JD, Repovs G, Anticevic A, Braver TS (2013) Multi-task connectivity reveals exible hubs for adaptive task control. Nat Neurosci 16:1348–1355. CrossRef Medline

Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3:201–215. CrossRef Medline

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306–324. CrossRef Medline

Desimone R, Duncan J (1995) Neural mechanisms of selective visual attention. Annu Rev Neurosci 18:193–222. CrossRef Medline

Dosenbach NU, Fair DA, Cohen AL, Schlaggar BL, Petersen SE (2008) A dual-networks architecture of top-down control. Trends Cogn Sci 12:99–105. CrossRef Medline

Egner T, Hirsch J (2005) Cognitive control mechanisms resolve conict through cortical amplification of task-relevant information. Nat Neurosci 8:1784–1790. CrossRef Medline

Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87:893–905. CrossRef Medline

Ester EF, Sutterer DW, Serences JT, Awh E (2016) Feature-selective attentional modulations in human frontoparietal cortex. J Neurosci 36:8188–8199. CrossRef Medline

Esterman M, Chiu YC, Tamber-Rosenau BJ, Yantis S (2009) Decoding cognitive control in human parietal cortex. Proc Natl Acad Sci U S A 106:17974–17979. CrossRef Medline

Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ (2008) Liblinear: a library for large linear classification. The Journal of Machine Learning Research 9:1871–1874.

Fox MD, Corbetta M, Snyder AZ, Vincent JL, Raichle ME (2006) Spontane-

ous neuronal activity distinguishes human dorsal and ventral attention systems. Proc Natl Acad Sci U S A 103:10046–10051. CrossRef Medline

Gazzaley A, Nobre AC (2012) Top-down modulation: bridging selective attention and working memory. Trends Cogn Sci 16:129–135. CrossRef Medline

Gratton C, Neta M, Sun H, Ploran EJ, Schlaggar BL, Wheeler ME, Petersen SE, **Nelson SM (2017) Distinct stages of moment-to-moment processing in the cinguloopercular and frontoparietal networks. Cereb Cortex 27:2403–2417. CrossRef Medline

Greenberg AS, Esterman M, Wilson D, Serences JT, Yantis S (2010) Control of spatial and feature-based attention in frontoparietal cortex. J Neurosci 30:14330–14339. CrossRef Medline

Hanson GK, Chrysikou EG (2017) Attention to distinct goal-relevant features differentially guides semantic knowledge retrieval. J Cogn Neurosci 29:1178–1193. CrossRef Medline

Jehee JF, Brady DK, Tong F (2011) Attention improves encoding of task-relevant features in the human visual cortex. J Neurosci 31:8210–8219. CrossRef Medline

Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841. CrossRef Medline

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron 22:751–761. CrossRef Medline

Kim H (2014) Involvement of the dorsal and ventral attention networks in oddball stimulus processing: a meta-analysis. Hum Brain Mapp 35:2265–2284. CrossRef Medline

Kimberg DY, Aguirre GK, D'Esposito M (2000) Modulation of task-related neural activity in task-switching: an fMRI study. Brain Res Cogn Brain Res 10:189–196. CrossRef Medline

Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M (2005) An event-related functional magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. J Neurosci 25:4593–4604. CrossRef Medline

Koechlin E, Summerfield C (2007) An information theoretical approach to prefrontal executive function. Trends Cogn Sci 11:229–235. CrossRef Medline

Kuhl BA, Chun MM (2014) Successful remembering elicits event-specific activity patterns in lateral parietal cortex. J Neurosci 34:8051–8060. CrossRef Medline

Kuhl BA, Johnson MK, Chun MM (2013) Dissociable neural mechanisms for goal-directed versus incidental memory reactivation. J Neurosci 33:16099–16109. CrossRef Medline

Lee H, Kuhl BA (2016) Reconstructing perceived and retrieved faces from activity patterns in lateral parietal cortex. J Neurosci 36:6069–6082. CrossRef Medline

Liu T, Hou Y (2013) A hierarchy of attentional priority signals in human frontoparietal cortex. J Neurosci 33:16606–16616. CrossRef Medline

Liu T, Hospadaruk L, Zhu DC, Gardner JL (2011) Feature-specific attentional priority signals in human cortex. J Neurosci 31:4484–4495. CrossRef Medline

Long NM, Lee H, Kuhl BA (2016) Hippocampal mismatch signals are modulated by the strength of neural predictions and their similarity to outcomes. J Neurosci 36:12677–12687. CrossRef Medline

Loose LS, Wisniewski D, Rusconi M, Goschke T, Haynes JD (2017) Switch independent task representations in frontal and parietal cortex. J Neurosci 37:8033–8042. CrossRef Medline

Miller EK, Erickson CA, Desimone R (1996) Neural mechanisms of visual working memory in prefontal cortex of the macaque. J Neurosci 16:5154–5167. Medline

Monsell S (2003) Task switching. Trends Cogn Sci 7:134–140. CrossRef Medline

Posner MI, Petersen SE (1990) The attention system of the human brain. Annu Rev Neurosci 13:25–42. CrossRef Medline

Qiao L, Zhang L, Chen A, Egner T (2017) Dynamic trial-by-trial recoding of task-set representations in the frontoparietal cortex mediates behavioral exibility. J Neurosci 37:11037–11050. Medline

Rainer G, Asaad WF, Miller EK (1998) Selective representation of relevant information by neurons in the primate prefrontal cortex. Nature 393:577–579. CrossRef Medline

Ravizza SM, Carter CS (2008) Shifting set about task switching: behavioral and neural evidence for distinct forms of cognitive exibility. Neuropsychologia 46:2924–2935. CrossRef Medline

Rogers RD, Monsell S (1995) Costs of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology General 124:207. CrossRef

Roy JE, Buschman TJ, Miller EK (2014) PFC neurons reect categorical decisions about ambiguous stimuli. J Cogn Neurosci 26:1283–1291. CrossRef Medline

Rushworth MF, Hadland KA, Paus T, Sipila PK (2002) Role of the human medial frontal cortex in task switching: a combined fMRI and TMS study. J Neurophysiol 87:2577–2592. CrossRef Medline

Sarma A, Masse NY, Wang XJ, Freedman DJ (2016) Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices. Nat Neurosci 19:143–149. CrossRef Medline

Serences JT, Yantis S (2006) Selective visual attention and perceptual coherence. Trends Cogn Sci 10:38–45. CrossRef Medline

Sestieri C, Shulman GL, Corbetta M (2017) The contribution of the human posterior parietal cortex to episodic memory. Nat Rev Neurosci 18:183–192. CrossRef Medline

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 23:S208–S219. CrossRef Medline

Sprague TC, Serences JT (2013) Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. Nat Neurosci 16:1879–1887. CrossRef Medline

Sreenivasan KK, Vytlacil J, D'Esposito M (2014) Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. J Cogn Neurosci 26:1141–1153. CrossRef Medline

Suzuki M, Gottlieb J (2013) Distinct neural mechanisms of distractor suppression in the frontal and parietal lobe. Nat Neurosci 16:98–104. CrossRef Medline

Swaminathan SK, Freedman DJ (2012) Preferential encoding of visual categories in parietal cortex compared with prefrontal cortex. Nat Neurosci 15:315–320. CrossRef Medline

Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. J Neurophysiol 100:3328–3342. CrossRef Medline

Vossel S, Geng JJ, Fink GR (2014) Dorsal and ventral attention systems distinct neural circuits but collaborative roles. Neuroscientist 20:150–159. CrossRef Medline

Waskom ML, Kumaran D, Gordon AM, Rissman J, Wagner AD (2014) Frontoparietal representations of task context support the exible control of goal-directed cognition. J Neurosci 34:10743–10755. CrossRef Medline

Waskom ML, Wagner AD (2017) Distributed representation of context by intrinsic subnetworks in prefrontal cortex. Proc Natl Acad Sci U S A 114:2030–2035. CrossRef Medline

Woolgar A, Williams MA, Rich AN (2015) Attention enhances multi-voxel representation of novel objects in frontal, parietal and visual cortices. Neuroimage 109:429–437. CrossRef Medline

Xu Y (2017) Reevaluating the sensory account of visual working memory storage. Trends Cogn Sci 21:794–815. CrossRef Medline

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165. CrossRef Medline

Yeung N, Nystrom LE, Aronson JA, Cohen JD (2006) Between-task competition and cognitive control in task switching. J Neurosci 26:1429–1438. CrossRef Medline

Zanto TP, Rubens MT, Bollinger J, Gazzaley A (2010) Top-down modulation of visual feature processing: the role of the inferior frontal junction. Neuroimage 53:736–745. CrossRef Medline