HHS Public Access

Representational Similarity of Social and Valence Information in the Medial Prefrontal Cortex

Robert S. Chavez and **Todd F. Heatherton**

Dartmouth College

Abstract

The human brain is remarkably adept at integrating complex information to form unified psychological representations of agents, objects, and events in the environment. Two domains in which this ability is particularly salient are the processing of social and valence information and are supported by common cortical areas in the medial prefrontal cortex (MPFC). Because social information is often embedded within valenced emotional contexts, it is possible that activation patterns within the MPFC may represent both of these types of cognitive processes when presented simultaneously. The current study tested this possibility by employing a large-scale automated meta-analysis tool, together with multi-voxel pattern analysis to investigate the representational similarity of social and valence information in the MPFC during functional magnetic resonance imaging. Using a representational similarity analysis, we found a high degree of representational similarity both within social dimensions and within valence dimensions, but not across them (e.g., positive social information was highly dissimilar to negative non-social information), in a ventral portion of the MPFC. These results were significantly correlated with a behaviorally measured similarity structure of the same stimuli, suggesting that a psychologically meaningful representation of social and valence information is reflected by multi-voxel activation patterns in the ventral MPFC.

Keywords

One of the most remarkable features of the brain is the ability to synthesize and integrate complex information into meaningful psychological representations. With just a quick glance, we can immediately differentiate a happy person from a sad person or a dangerous situation from a safe one. Doing so requires the brain not only to encode the lower-order sensory features of a stimulus but also to integrate those perceptions into higher-order information such as the content and emotional features of a scene. As a highly social species, humans are particularly adept at the ability to perceive information about other individuals, rapidly and effortlessly (Ambady & Rosenthal, 1992). Similarly, people are also able to automatically process information about the attractiveness or aversiveness of a stimulus to form valenced evaluations of objects, events, or agents in the environment.

Address correspondence to: Robert S. Chavez (robert.s.chavez.GR@dartmouth.edu), Dartmouth College, Department of Psychological & Brain Sciences, Moore Hall, 6207, Hanover NH, 03755..

Though social and valence processing reflect ostensibly dissociable cognitive functions, there is a surprising degree of correspondence in the cortical regions serving each of these processes converging in the medial prefrontal cortex (MPFC).

Despite the inherent complexity of social cognition, researchers have demonstrated that people are able to form lasting impressions of other individuals within a fraction of a second. In less than 40ms, people make accurate personality trait judgments of presented faces with high between-subject reliability (Bar, Neta, & Linz, 2006). Furthermore, brief interactions produce trait inferences spontaneously and are independent of allotted attentional resources (Todorov & Uleman 2003). Using neuroimaging techniques such as functional magnetic resonance imaging (fMRI), researchers have shown that these types of rapid social inferences are supported by a network of brain regions including the MPFC, temporal parietal junction (TPJ), posterior cingulate cortex (PCC), and superior temporal sulcus (STS; see Lieberman 2010). However, although the TPJ, PCC, and STS code for important aspects of social cognition such as cued attention and beliefs (Mitchell, 2007; Saxe & Kanwisher, 2003), autobiographical episodic memory retrieval (Cavanna & Trimble, 2006), and biological motion perception (Puce & Perrett, 2003), these regions are also recruited in processing information devoid of any social content (Martin, 2007). Indeed, although the TPJ, PCC, STS, and associated brain areas play an essential role in the ability to processes social information in the environment, recent theories suggest that some of these regions may have been re-purposed over human evolutionary history to serve social functions (Parkinson & Wheatley, 2013), rather than being areas dedicated to social cognition per se. By contrast, a growing body of evidence suggests that abstracting person knowledge is supported primarily by the MPFC (Wagner, Haxby, & Heatherton, 2012). A study by Mitchell (2007) found that the MPFC shows robust activation when people are engaged in mental state attribution (i.e., theory-of-mind), whereas similar TPJ activations were confounded by attentional reorienting demands of the task. Moreover, the MPFC is particularly responsive when making social inferences based on minimal or ambiguous information (Jenkins & Mitchell, 2011) and when referencing one's own beliefs in order to make sense of the minds of others (Heatherton, 2011). This has led to the notion that the MPFC may function to gather and simulate information from multiple sources to make probabilistic estimates of other people's experience (Mitchell, 2009). From this perspective, the MPFC may serve as an important region for combining multiple processes together to form higher-order perceptions of other people in specific contexts. In summary, although a network of regions support social cognitive processes in the brain, the MPFC serves as the critical area of importance when the perception of others needs to be integrated with additional information from the environment and the self.

Though our brains are able to seamlessly perceive and respond to other people within our environments, this ability does not exist in a vacuum. Another important feature of human social life is that our interactions are often grounded in the valenced emotional contexts in which we experience them—people fall in love, behave aggressively toward enemies, or provide social support to friends in need. In contrast to the cortical networks implicated in social cognition, researchers investigating the neural basis of emotional valence processing have largely focused on subcortical areas such as the amygdala and ventral striatum. Though early work conceptualized the amygdala as a center for negative emotions such as fear

processing (LeDoux, 1996), this view has been revised to reflect findings that amygdala activity can indicate both positive and negative emotional processing (Garavan, Pendergrass, Ross, Stein, & Risinger, 2001), as well as promote vigilance when stimuli are ambiguously valenced (Whalen, 1998). Conversely, a wealth of research in animals and humans has demonstrated that activity in the ventral striatum is most consistently associated with processing positively valenced stimuli (Burgdorf & Panksepp, 2006), particularly with respect to reward anticipation (Knutson, Adams, Fong, & Hommer, 2001).

Other research has focused on areas beyond subcortical areas to describe valence processing, and it is widely recognized that areas within the MPFC processes information related to both positive and negative valence. On the one hand, a review by Etkin et al. (2011) highlighted the role of the MPFC in negative emotional contexts, arguing for differential functions of the dorsal and ventral MPFC in processing appraisal and regulation, respectively. On the other hand, a meta-analysis of 142 fMRI studies found that areas of the MPFC showed preferential activation to positively valenced rewards, especially during the outcome or receipt phases of reward delivery (Liu, Hairston, Schrier, & Fan, 2011). Moreover, valence processing is beginning to be described from a network perspective. Additional evidence linking MPFC to valence processing comes from studies demonstrating that the disruption of amygdalo-frontal and striato-frontal connectivity impairs emotional processing in both healthy and clinical populations. For example, Kim and Whalen (2009) found that amygdala reactivity to fearful faces correlated with the structural integrity of a white matter pathway connecting amygdala to the ventral MPFC. In turn, the structural integrity of this pathway then predicted individual differences in trait-anxiety within the same subjects. Another study by Furman et al. (2011) found that resting-state functional connectivity of the ventral striatum to the MPFC was significantly attenuated in depressed patients when compared to healthy controls, suggesting that this disconnection may contribute to anhedonia. Furthermore, disruptions of these connections impact the ability of MPFC to communicate with the amygdala and ventral striatum and are thought to directly contribute to the symptomology of anxiety disorders (Kim et al., 2011) and depression (Heller et al., 2009). Overall, there are multiple lines of evidence that the amygdala and ventral striatum do not act alone in the processing of valence information, but instead rely on input from the MPFC to evaluate both positive and negative information about a given stimulus. This suggests that neural activity within the MPFC may encode valence-related information, and activity within this region can potentially be used to distinguish different types of affective processing.

Taken together, there is a broad literature pointing to the MPFC as an intersection in the neural basis of both social cognition and valence processing. Because social perception is often embedded within emotionally laden contexts, having a common cortical mechanism by which to integrate social and valence information could help support a unified psychological representation of these factors. The MPFC receives input from all sensory modalities, and it has been speculated that this region may serve as a processing hub that binds internally and externally generated information (Moran, Kelley, & Heatherton, 2013). To this end, it is possible that distributed activation patterns within the MPFC could be used to differentiate the presence of socially relevant information under different emotionally valenced contexts. By utilizing multivariate information in the fMRI signal, multi-voxel

pattern analysis (MVPA) methods allow for the investigation of this possibility (Haxby, 2012; Haxby et al., 2001).

Rather than simply describing the conditions under which a brain region shows the greatest activation, cognitive neuroscientists have become increasingly interested in the notion of *representation*: how distributed patterns of neural activity contain informational content reflecting dissociable mental processes (Norman, Polyn, Detre, & Haxby, 2006). An MPVA approach called representational similarity analysis (RSA; Kriegeskorte et al. 2008) uses voxel activity patterns to calculate quantitative distance measures between stimulus types. These distances are then used to compute dissimilarity matrices (DSMs) characterizing the stimulus representation in the brain. Not only does RSA provide a method by which to measure the similarity patterns among different categories of stimuli, it also provides the ability to compare similarity structures across modalities (Kriegeskorte et al., 2008). This allows researchers to relate underlying brain models of stimulus representation to explicit psychological representations measured behaviorally. For example, Connolly et al. (2012) used RSA to show that activity patterns in the ventral temporal cortex produced a similarity structure of biological classes (e.g., insects, birds, monkeys) that was highly correlated to behavioral similarity judgments of the same stimuli. The similarity structure of these classes in early visual areas of the brain did not correlate with these behavioral judgments, suggesting that psychologically meaningful information about biological classes of these animals is primarily represented in the ventral temporal cortex. If the MPFC serves as hub for the integration of social and valence information, RSA may be able to dissociate the representational content of both these features when they are presented concurrently. We explore this possibility here.

In the current study, we tested the hypothesis that social and valence information are simultaneously represented in the MPFC and that these neural representations have corresponding behaviorally measured psychological representations. During an fMRI scanning session, participants were shown a series of photos matched for arousal and basic visual properties from the International Affective Picture System (IAPS). Each picture's content varied on both sociality and valence dimensions. Subjects were kept naive to the purpose of the study to observe implicit responses. Capitalizing on the broad literature in both domains described above, we employed a large-scale automated meta-analysis tool to isolate areas within the MPFC where social and valence information share overlapping spatial activations in the neuroimaging literature. These areas were then used as common cortical regions of interest (ROIs) for RSA. Neural representations were then compared to behavioral similarity ratings of the same stimuli in an independent sample of subjects. We hypothesized that behavioral similarity ratings would be correlated with representational similarity of social and valence information in the MPFC, supporting the notion that these brain activity patterns reflect psychologically meaningful representations.

Methods

Participants

For the fMRI portion of the study, 50 subjects between the ages of 18 and 19 were recruited from the local Dartmouth community. These subjects were screened to be right-handed and

reported no abnormal neurological history. Two subjects were excluded for having head motion exceeding 3 mm, leaving a total of 48 subjects (28 female) in the final analysis. For the behavioral portion of the study, similarity ratings were assessed separately in an independent sample of 32 subjects (23 female) age 18-20. All subjects gave informed consent in accordance with the guidelines set by the Committee for the Protection of Human Subjects at Dartmouth College and received course credit or were paid for their participation.

Procedure

For the fMRI procedure, subjects were scanned while viewing images selected from the IAPS photo set (Lang, Bradley, & Cuthbert, 2008). In keeping with previous research on implicit social inference (Powers, Wagner, Norris, & Heatherton, 2011; Wagner, Kelley, & Heatherton, 2011), participants were kept naive to the purpose of the task and were simply asked to use a button response to categorize each image as either an indoor or outdoor scene. Pictures varied on both sociality (social and non-social) and valence (negative, neutral and positive) dimensions, leaving a total of six stimulus categories matched for arousal and low-level visual properties (i.e., luminance, hue, saturation and RGB values). Example stimuli include: social negative (depictions of pain; people crying), social neutral (person at a computer; person on the phone), social positive (children playing; friends laughing), non-social negative (dirty toilets; polluted water), non-social neutral (a stack of books; a spoon) and non-social positive (picturesque landscape; appetizing dessert). Critically, across valence categories, the non-social pictures were defined by being devoid of any depiction of people. A total of 150 pictures (25 per category) were presented for 2500 ms each along with intermittent passive fixation trials of variable durations (0–7500 ms) to introduce jitter in to the fMRI time series. The order of the pictures was pseudo-randomized and counterbalanced across participants.

The behavioral portion of the experiment measured pairwise similarity judgments of each of the six categories using the same set of IAPS pictures. For the task, subjects were presented with two of the 150 images side-by-side, along with a ratings bar at the bottom of the screen. Before beginning the experiment, subjects were instructed to note both the content and mood of each of the photos and asked to rate the similarity of each pair of pictures on a 1 to 7 scale (1 = very dissimilar; 7 = very similar). A complete set of all pairwise combinations of the 150 pictures yields 11,175 comparisons per subject, making this unfeasible for the time constraints of the experiment. Instead, each of the 25 pictures from one of the six social/valence conditions was randomly paired with one of the 25 pictures from another. This produced a final total of 375 similarity ratings across the 15 social/valence condition comparisons. Within social/valence condition pairings, pictures were sampled without replacement to ensure that the same pictures were not used twice within the same comparison category.

To calculate the behavioral DSM, we averaged similarity ratings of the picture pairs within each of the 15 social/valence condition comparisons for each subject. Since subjects were asked to rate the similarity of each pair of pictures (as opposed to dissimilarity), we reversed scored the similarity ratings between to calculate dissimilarity values between each of the

social/valence condition pairs. Thus, the values in the DSM reflect the average dissimilarity score between pairs of pictures in each social/valence comparison category per subject. Finally, these dissimilarity values were then averaged across the 32 subjects to create the final 6×6 behavioral DSM used for comparison with the RSA in the fMRI analysis.

Imaging Acquisition and Analysis

Magnetic resonance imaging was conducted with a Philips Achieva 3.0 Tesla scanner using a thirty-two channel phased array head coil. Structural images were acquired using a T1-weighted MP-RAGE protocol (220 sagittal slices; TR: 8.176 ms; TE: 3.72 ms; flip angle: 8°; 1 mm isotropic voxels). Functional images were acquired using a T2*-weighted echoplanar sequence (TR: 2500 ms; TE: 35 ms; flip angle: 90°), consisting of 270 whole-brain volumes (36 axial slices per volume, 3 mm thick, 0.5 mm gap, 33 mm in-plane resolution). Stimuli were presented via an Epson ELP-7000 LCD projector onto a screen at the end of the MRI bore. Subjects viewed stimuli via an angled mirror mounted on the head coil.

The fMRI data were analyzed using fMRI Expert Analysis Tool in FSL (Smith et al., 2004). First, all slices were interpolated to a common time point (i.e., slice-time correction) to correct for differences in slice acquisition. Next, we applied mean-based intensity normalization of all volumes by the same factor and high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 80.0 s). Time-series statistical analyses were carried out using local autocorrelation correction. In accordance with standard practices for MVPA studies (Haxby et al., 2001), fMRI images were left unsmoothed and statistical parametric maps were analyzed in each subject's unregistered native space. Activation patterns were then calculated by convolving the stimulus timing files with a double-gamma hemodynamic response function. A multiple regression analysis was performed to estimate the hemodynamic response parameters for each of the six experimental conditions. Brain activation patterns were measured using contrast of parameter estimate (COPE) values for each of the six experimental conditions compared to fixation.

For the RSA procedure, we were interested in investigating cortical areas that had been previously implicated in social and valence processing from the neuroimaging literature described above. To do this, we employed a large-scale automated meta-analysis tool (Neurosynth; Yarkoni et al. 2011) to isolate overlapping spatial extents for both domains. The terms "social" and "valence" yielded 472 and 250 studies, respectively, used in the meta-analysis. These studies produced composite spatial extents for each term using the quantitative reverse inference procedure thresholded at $z > 3$. These maps were then binarized, and added to one another. Finally, overlapping voxels from both social and valence maps were used as ROIs in the RSA. Because brain activity patterns were assessed in each subject's native-space fMRI image, standard-space ROIs needed to be transformed into each subject's unregistered fMRI image. To achieve this, we aligned functional data to each subject's anatomical scan before registering it to the MNI template using FNIRT non-linear registration in FSL. Transformation matrices and non-linear warps from this procedure were then inverted and ROI masks were transformed into each subject's native-space brain image. Multi-voxel COPE activation measurements from these ROIs were then

imported and analyzed using R statistical software (Ihaka & Gentleman, 1996). Neural DSMs were calculated within each ROI for each subject by calculating the correlation distance between all pairs of the six social/valence categories. This approach differs from previous RSA studies (Connolly et al., 2012; Kriegeskorte et al., 2008) in that we averaged the signal across each of the categories rather than having an item-wise similarity structure. The motivation for this approach was to allow us to compare across the behavioral and imaging similarity structures in the absence of having all pairwise dissimilarity values between stimuli in the behavioral task due to the combinatorial explosion issue. This resulted in a 6×6 DSM per ROI for each of the 48 subjects. These were then averaged across all subjects for each ROI to create the final neural DSMs used in the analysis.

Though our hypotheses were specific to the multivariate representations, we also conducted standard univariate tests in order to contrast the two methods. The preprocessing steps for the univariate tests were identical for that of the RSA with the exception of using a 6mm FWHM Gaussian smoothing kernel. Average parameter estimates were estimated for each of the six social/valence conditions. Finally, within each of the four ROIs from the automated meta-analysis, univariate statistical analyses were conducted using a 2 (social, non-social) $\times$ 3 (positive, neutral, negative) within-subjects ANOVA.

Results

From the behavioral portion of the study, the similarity structure of the behavioral ratings showed that subjects were able to dissociate information related to both social and valence dimensions when comparing pairs of images. Cross-subject results yielded a behavioral DSM with a high degree of similarity within both the social dimension and the valence dimension but not across them. Specifically, social images were rated as more similar to each other than non-social images, regardless of the valence. Likewise, each valence category was rated as more similar to itself than to other valence categories, irrespective of the social content of the image. On the other hand, there was a low degree of similarity across social and valence dimensions. For example, negative non-social images were rated as highly dissimilar to positive social images, and neutral social images were highly dissimilar to positive non-social images. This intuitive similarity structure reflects the psychological representation of social and valence information that is present in each image as rated by the subjects. The results of this analysis are plotted in the behavioral DSM heat map in Figure 1.

Consistent with previous observations, results from the automated meta-analysis revealed overlapping activation extents of social and valence processing in the MPFC from the neuroimaging literature. Specifically, there were two large, discrete clusters of cortical overlap situated in the dorsal MPFC and the ventral MPFC, each with cluster-extents exceeding ten contiguous voxels (see Figure 2). The dorsal MPFC cluster extended from the frontal pole to an area adjacent to the superior frontal gyrus. The ventral MPFC cluster extended from the frontal pole to an area just superior to the medial orbitofrontal cortex, ending just anterior to the subgenual cingulate cortex. Neither MPFC cluster extended more than 14 mm laterally from the medial longitudinal fissure. In addition to these MPFC clusters, other small points of overlapping voxels were found in the ventral anterior

cingulate cortex, right inferior frontal gyrus, right fusiform, and left medial temporal lobe. Because of their small cluster-extents (i.e., less than ten contiguous voxels) and our desire to focus on the hypothesized ROIs, these additional voxels were excluded from the RSA procedure. Finally, the meta-analysis also revealed two large subcortical clusters centered in the bilateral amygdala. In total, the results of the meta-analysis produced four final ROIs used to isolate brain regions in which to calculate the neural DSMs: dorsal MPFC, ventral MPFC, right amygdala, and left amygdala. Similar to the behavioral results, each of these regions produced a DSM corresponding to each of the experimental conditions.

To compare the behavioral results with fMRI results, the non-diagonal half of the symmetrical behavioral DSM was then correlated with the matching DSM from the RSA procedure in each ROI. Because little is known about the underlying null distribution for dissimilarity values, we used both standard correlation and non-parametric permutation tests to assess the relationship between the behavioral DSM and the neural DSMs. Permutation tests randomly shuffle data labels to create an empirically derived distribution for the test statistic underlying the null hypothesis. Using this distribution, permutation p-values are calculated for the original data. From these analyses, we found that representational similarity in the ventral MPFC was correlated to the behavioral similarity ratings using both parametric and non-parametric tests (Pearson $r = .539$, $p = .038$; permutation $p = .037$). No other ROIs from the RSA procedure produced a significant relationship to the behavioral similarity ratings using either parametric or non-parametric tests. Results from the dorsal MPFC yielded a non-significant relationship to the behavioral DSM (Pearson $r = .44$, $p = .101$; permutation $p = .091$), but may reflect a trend towards significance at more liberal statistical thresholds. Neither of the neural DSMs in the amygdala ROIs showed a significant relationship to the behavioral DSM: right amygdala (Pearson $r = .182$, $p = .516$; permutation $p = .556$), left amygdala (Pearson $r = .05$, $p = .858$; permutation $p = .824$). Moreover, with high degrees of dissimilarity across all social and valence categories the representational content of the amygdala ROIs was noticeably more homogenous than either MPFC area. DSM heat maps for all four ROIs from the RSA are displayed in Figure 3. In sum, only the neural similarity structure of social and valence categories in the ventral MPFC showed a statistically significant relationship to the similarity structure from the behavioral ratings.

Results from the univariate analyses for each of the four ROIs are displayed in Figure 4. For the ventral MPFC, there was no main effect of social category [$F(1, 47) = 0.63$, $p= .43$] or valence [$F(2, 94) = 0.94$, $p= .4$], but there was a significant interaction [$F(2, 94) = 5.13$, $p= .008$]. For the dorsal MPFC, there was no main effect of social category [$F(1, 47) = 2.4$, $p= .13$], but there was a significant main effect of valence driven by greater responses in positive social condition [$F(2, 94) = 5.95$, $p= .004$] and producing a significant interaction [$F(2, 94) = 5.02$, $p= .008$]. For the left amygdala, there was no main effect of valence [$F(2, 94) = 2.19$, $p= .12$], but there was a significant main effect of social category driven by greater responses in positive social condition [$F(1, 47) = 9.59$, $p= .003$] and a near-significant interaction [$F(2, 94) = 2.82$, $p= .064$]. Similarly, for the right amygdala, there was no significant main effect of valence [$F(2, 94) = 1.11$, $p= .33$], but there was a near-significant

main effect of social category [$F(1, 47) = 3.95$, $p= .052$] but showed a significant interaction [$F(2, 94) = 3.44$, $p= .036$].

Discussion

Because it is so effortless, we often take for granted the ability to combine complex information together to form coherent representations of the social and emotional content of our environment. Using behaviorally measured similarity ratings as a metric for the psychological representation of social and valence information, the current study shows that the neural representations of these processes are reflected in multi-voxel activation patterns in the ventral MPFC. Moreover, the similarity structure of these activation patterns suggest that social and valence information are represented concurrently within the ventral MPFC. This extends previous research showing that areas of the MPFC are involved in processing both social and valence information in isolation of one another. Instead, our results suggest that the ventral MPFC may act as a hub region for the integration of social and valence information when they are presented simultaneously. This function of the ventral MPFC may be critical for embedding social information within emotional contexts, allowing for a unified perception of this kind of high-level information.

Although we hypothesized that social and valence information would be supported by the MPFC more generally, the results of our analysis implicate a specific area of the ventral MPFC underlying these representations. This may be surprising given that social cognitive processes such as mentalizing are often associated with greater activation of the dorsal MPFC (Wagner et al., 2012). However, compared to the dorsal MPFC, the ventral MPFC has been shown to have a much higher degree of connectivity to the amygdala and ventral striatum (Di Martino et al., 2008; Stein et al., 2007). Because we know that these structures are involved in positive and negative valence processing, greater connectivity of these areas to the ventral MPFC may facilitate the relaying of social and valence information between these cortical and subcortical networks. Previous MVPA studies have also indicated that value coding in the ventral MPFC and orbitofrontal cortex (but not the dorsal MPFC) has a distributed neural representation that includes increased and decreased activations to the anticipation and receipt of rewarding stimuli (Kahnt, Heinzle, Park, & Haynes, 2010). Moreover, activation in these regions has been linked to the valuation of both gains and losses of incommensurable outcomes, suggesting that these regions may use common currency of valuation irrespective of what is being evaluated (FitzGerald, Seymour, & Dolan, 2009). Valence, by definition, is an evaluation of both the attractiveness and aversiveness of stimuli in the environment. Thus, the results of our RSA analysis suggest that the representation of valence information in the ventral MPFC is maintained across social and non-social situations and is consistent with previous evidence that this region codes for both positive and negative valuation across contexts.

The results from our analysis may have implications for related areas of social cognition as well. Compared to the dorsal MPFC, the ventral MPFC has been more consistently implicated in studies of self-referential cognition (see: Wagner et al., 2012). However, the ventral MPFC has also been shown to be more responsive to information about others who are closer or more similar to the self. For example, Mitchell, Macrae, & Banaji (2006) found

that the ventral MPFC was recruited when subjects were mentalizing about individuals with a similar political ideology, but the dorsal MPFC was recruited when subjects were mentalizing about individuals with opposing political ideologies. They suggested that this finding may be evidence that people use information about the self when simulating the thoughts of similar-others (but not dissimilar-others) during mentalizing. In another set of studies, Krienen, Tu, & Buckner (2010) found that, when thinking about others, ventral MPFC activity was driven not by similarity but rather by interpersonal closeness (i.e., friends vs. strangers). The authors suggested that this may be due to the increased salience or value one attaches to people with whom one has personal involvement. Given that the results of the current study suggest the representation of value (in this case, emotional valence) is maintained across social and non-social contexts, it may be that closeness and similarity increase the amplitude of the ventral MPFC response to others but do not fundamentally shift the underlying representation. Put differently, the ventral MPFC may be recruited to dissociate the positive and negative characteristics of both a friend and stranger, even if its overall response is greater to a friend. However, the current study did not directly test the effect of closeness or similarity on the representation of social and valence categories, and this possibility remains speculative.

Outside of the MPFC, we also found that activation similarity patterns in bilateral amygdala areas did not support the simultaneous representation of social and valence information. There are at least two possible explanations for this finding. First, at the voxel resolution used here, subcortical regions may not be as amenable to RSA methods due to their small size. Similarly, given the differences between cortical and subcortical anatomical structure, it is also possible that the informational content in the amygdala may not be as rich as it is in the cortical areas typically studied using RSA. This could explain the relatively homogeneous dissimilarity content seen in the DSMs of both amygdala ROIs. Second, a disproportionate number of studies on negatively valenced emotions (e.g., threat or fear) are studied exclusively using social stimuli (e.g., faces). This skew could potentially over-represent the degree to which social information is processed by the amygdala more generally, biasing the meta-analysis to include of the amygdala ROIs. Additionally, studies using backwards masking paradigms have shown that that the amygdala is privy to valenced emotional expressions that are not consciously accessible (Whalen et al., 1998). Thus, it is also possible that the behavioral DSM does not reflect concurrent social and valence information in the amygdala due to differences in conscious and unconscious representations of these stimuli.

More broadly, the results of our RSA analyses add to a body of work attempting to build a parsimonious description of the general function of the MPFC. One compelling account for the function of this region is the "gateway hypothesis," which proposes that the MPFC may act as a critical waypoint for binding internally and externally generated information by altering the flow of this information among other systems that process them (Burgess, Dumontheil, & Gilbert, 2007). Consistent with this idea, social and valence processing is supported by multiple brain regions and areas within the MPFC, which has known anatomical and functional connections to all sensory modalities. To the extent that valence reflects internally generated emotional processing and is coupled with externally cued social perception, our results provide tentative support for the gateway hypothesis. However,

additional work will be needed to support other aspects of this hypothesis, which are beyond the scope of the current study. Nonetheless, our results suggest that specific combinations of social and valence information can be decoded from activation patterns in the ventral MPFC, suggesting that this region may be critical for the integration of complex cognitive processing from an extended network of brain areas.

In addition to the primary analyses, we also conducted a series of univariate tests to supplement the results of the RSA procedure. As illustrated in Figure 4, the pattern of univariate responses is similar among each of the brain regions, with the exception of the ventral MPFC. This is notable because the ventral MPFC was the only region from the RSA analysis that was significantly correlated with the behavioral DSM. It is tempting to conclude that the RSA analyses are simply more sensitive than the univariate contrasts to the differences between conditions. However, though classification-based MVPA methods can be more sensitive to the mean activation differences between conditions, it is unclear if the same principle applies to RSA-based MVPA methods as well. It is also important to keep in mind that RSA uses the topography of voxel responses to calculate dissimilarity values as its metric of interest and is blind to the regional-average response. Moreover, a recent study by Davis et al., (2014) highlights the difficultly with drawing conclusions about the differences between MVPA and univariate methods without careful studies specifically designed to target these questions. That being said, we cannot rule out the possibility that RSA patterns in the ventral MPFC were influenced by mean differences between conditions (as reflected in the univariate responses) making these patterns more dissociable from one another when compared to the other ROIs.

Another notable finding from the univariate contrasts was the activation patterns in the bilateral amygdala ROIs. Specifically, although non-social scenes showed a linear increase in activity as the images became more negative, both positive and negative images of social scenes increased bilateral amygdala activation. These results are consistent with another recent study by Cunningham & Kirkland (2014), who found that the bilateral amygdala showed greater response to both positively and negatively valenced IAPS pictures when compared to neutral pictures. However, their study did not take the social content of the pictures into account. Thus, the results of the current study may suggest that the activation to positively valenced pictures in the amygdala seen in Cunningham and Kirkland (2014) may have been driven by IAPS images depicting social scenes. This is a potentially important effect that warrants further investigation. All told, the findings from the univariate analyses reflect a different kind of information than the RSA procedure, but considered together, they offer a broader interpretation of the results.

It is important to point out that one limitation of the current study was that the behavioral and neural DSMs were calculated from independent samples of individuals. Although we believe this may enhance the generalizability of our findings in the ventral MPFC, it may reduce the sensitivity to detect relationships in the dorsal MPFC and amygdala if these regions process social or valence information more idiosyncratically. Our results cannot speak to this possibility. However, we feel it is worth noting that the sample size in both our fMRI analysis (N=48) and behavioral analysis (N=32) is several times larger than many of the most highly cited MVPA studies, including Kriegesorte et al. (2006; N=11) and Haxby

et al. (2001; N=6), and other RSA studies such as Connolly et al. (2012; N=12). So although our ability to capture idiosyncratic effects was limited, the current study does benefit from additional between-subject statistical power not seen in similar analyses.

Finally, though sometimes characterized as a well-defined, homogeneous region, the MPFC is a large, liberally defined cortical area that serves many diverse cognitive functions. Our results benefitted greatly from the use of the automated meta-analysis to efficiently isolate areas of the MPFC that had previously been implicated in social and valence processing. There is growing criticism of fMRI studies that continue to employ whole-brain analysis techniques for the sole purpose of mapping cognitive processes to brain regions, especially when statistical power tends to be low (see: Yarkoni et al. 2010; Poldrack 2012; Moran and Zaki 2013). This criticism also extends to MVPA studies employing whole-brain searchlight procedures (Etzel, Zacks, & Braver, 2013). As such, our methodology served not only to leverage the extant literature to target hypothesized areas of interest, but also to exclude potentially spurious findings resulting from whole-brain explorations. Though limiting our ability to speak to the role of other regions in social and valence representation in the brain, our confidence in these results is strengthened by this hypothesis-driven approach.

In conclusion, we provide evidence that the similarity structure of brain activation patterns in the ventral MPFC reflects the representation of social and valence information in the brain. Moreover, these neural representations were correlated to a behaviorally measured similarity structure of the same stimuli, suggesting that they reflect meaningful psychological representations. This work adds to a body of literature in the neural basis of social and valence processing and suggests that the ventral MPFC may integrate information from both of these domains to support a unified experience of person perception within emotional contexts.

Acknowledgements

References

Ambady N, Rosenthal R. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. Psychological Bulletin. 1992; 111(2):256–274.

Bar M, Neta M, Linz H. Very first impressions. Emotion. 2006; 6(2):269–78.. [PubMed: 16768559]

Burgdorf J, Panksepp J. The neurobiology of positive emotions. Neuroscience & Biobehavioral Reviews. 2006; 30(2):173–87.. [PubMed: 16099508]

Burgess PW, Dumontheil I, Gilbert SJ. The gateway hypothesis of rostral prefrontal cortex (area 10) function. Trends in Cognitive Sciences. 2007; 11(7):290–8. [PubMed: 17548231]

Cavanna AE, Trimble MR. The precuneus: a review of its functional anatomy and behavioural correlates. Brain. 2006; 129(3):564–83.. [PubMed: 16399806]

Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu Y-C, Haxby JV. The representation of biological classes in the human brain. The Journal of Neuroscience. 2012; 32(8): 2608–18. [PubMed: 22357845]

Cunningham WA, Kirkland T. The joyful, yet balanced, amygdala: moderated responses to positive but not negative stimuli in trait happiness. Social Cognitive and Affective Neuroscience. 2014; 9(6): 760–6. [PubMed: 23563851]

Davis T, Larocque KF, Mumford J, Norman KA, Wagner AD, Poldrack RA. What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. NeuroImage. 2014; 97:271–283. [PubMed: 24768930]

Di Martino, a; Scheres, A.; Margulies, DS.; Kelly, a M. C.; Uddin, LQ.; Shehzad, Z.; Milham, MP. Functional connectivity of human striatum: a resting state FMRI study. Cerebral Cortex. 2008; 18(12):2735–47. [PubMed: 18400794]

Etkin A, Egner T, Kalisch R. Emotional processing in anterior cingulate and medial prefrontal cortex. Trends in Cognitive Sciences. 2011; 15(2):85–93. [PubMed: 21167765]

Etzel JA, Zacks JM, Braver TS. Searchlight analysis: promise, pitfalls, and potential. NeuroImage. 2013; 78:261–9. [PubMed: 23558106]

FitzGerald THB, Seymour B, Dolan RJ. The role of human orbitofrontal cortex in value comparison for incommensurable objects. The Journal of Neuroscience. 2009; 29(26):8388–95. [PubMed: 19571129]

Furman DJ, Hamilton JP, Gotlib IH. Frontostriatal functional connectivity in major depressive disorder. Biology of Mood & Anxiety Disorders. 2011; 1(1):11. [PubMed: 22737995]

Garavan H, Pendergrass JC, Ross TJ, Stein EA, Risinger RC. Amygdala response to both positively and negatively valenced stimuli. Neuroreport. 2001; 12(12):2779–83. [PubMed: 11522965]

Haxby JV. Multivariate pattern analysis of fMRI: the early beginnings. NeuroImage. 2012; 62(2):852–5. [PubMed: 22425670]

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 2001; 293(5539):2425–30. [PubMed: 11577229]

Heatherton TF. Neuroscience of self and self-regulation. Annual Review of Psychology. 2011; 62:363–90..

Heller AS, Johnstone T, Shackman AJ, Light SN, Peterson MJ, Kolden GG, Davidson RJ. Reduced capacity to sustain positive emotion in major depression reflects diminished maintenance of fronto-striatal brain activation. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106(52):22445–50. [PubMed: 20080793]

Ihaka R, Gentleman R. R: A Language for Data Analysis and Graphics. Journal of Computational and Graphical Statistics. 1996; 5(3):299–314.

Jenkins AC, Mitchell JP. Medial prefrontal cortex subserves diverse forms of self-reflection. Social Neuroscience. 2011; 6(3):211–8. [PubMed: 20711940]

Kahnt T, Heinzle J, Park SQ, Haynes J-D. The neural code of reward anticipation in human orbitofrontal cortex. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107(13):6010–5.. [PubMed: 20231475]

Kim MJ, Loucks RA, Palmer AL, Brown AC, Solomon KM, Marchante AN, Whalen PJ. The structural and functional connectivity of the amygdala: from normal emotion to pathological anxiety. Behavioural Brain Research. 2011; 223(2):403–10.. [PubMed: 21536077]

Kim MJ, Whalen PJ. The structural integrity of an amygdala-prefrontal pathway predicts trait anxiety. The Journal of Neuroscience. 2009; 29(37):11614–8. [PubMed: 19759308]

Knutson B, Adams CM, Fong GW, Hommer D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. The Journal of Neuroscience. 2001; 21(16):RC159. [PubMed: 11459880]

Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience. 2008; 2(November):4. [PubMed: 19104670]

Krienen FM, Tu P-C, Buckner RL. Clan mentality: evidence that the medial prefrontal cortex responds to close others. The Journal of Neuroscience. 2010; 30(41):13906–15. [PubMed: 20943931]

Lang, PJ.; Bradley, MM.; Cuthbert, BN. Technical Report A-8. University of Florida; Gainesville, FL: 2008. International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual..

LeDoux, JE. The Emotional Brain. Simon & Schuster; New York: 1996.

Lieberman, MD. Social cognitive neuroscience.. In: Fiske, ST.; Gilbert, DT.; Lindzey, G., editors. Handbook of Social Psychology. 5th ed.. McGraw-Hill; New York, NY: 2010. p. 143-193.

Liu X, Hairston J, Schrier M, Fan J. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neuroscience & Biobehavioral Reviews. 2011; 35(5):1219–36. [PubMed: 21185861]

Martin A. The representation of object concepts in the brain. Annual Review of Psychology. 2007; 58:25–45.

Mitchell JP. Activity in right temporo-parietal junction is not selective for theory-of-mind. Cerebral Cortex. 2007; 18(2):262–71. [PubMed: 17551089]

Mitchell JP. Social psychology as a natural kind. Trends in Cognitive Sciences. 2009; 13(6):246–51. [PubMed: 19427258]

Mitchell JP, Macrae CN, Banaji MR. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. Neuron. 2006; 50(4):655–63.. [PubMed: 16701214]

Moran JJM, Zaki J. Functional neuroimaging and psychology: What have you done for me lately? Journal of Cognitive Neuroscience. 2013; 25(6):834–842. [PubMed: 23469884]

Moran JM, Kelley WM, Heatherton TF. What can the organization of the brain's default mode network tell us about self-knowledge? Frontiers in Human Neuroscience. 2013; 7:391. [PubMed: 23882210]

Norman, K. a; Polyn, SM.; Detre, GJ.; Haxby, JV. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends in Cognitive Sciences. 2006; 10(9):424–30.. [PubMed: 16899397]

Parkinson C, Wheatley T. Old cortex, new contexts: re-purposing spatial perception for social cognition. Frontiers in Human Neuroscience. 2013; 7:645. [PubMed: 24115928]

Poldrack RA. The future of fMRI in cognitive neuroscience. NeuroImage. 2012; 62(2):1216–20. [PubMed: 21856431]

Powers KE, Wagner DD, Norris CJ, Heatherton TF. Socially excluded individuals fail to recruit medial prefrontal cortex for negative social scenes. Social Cognitive and Affective Neuroscience. 2011:1–7. [PubMed: 21247932]

Puce A, Perrett D. Electrophysiology and brain imaging of biological motion. Philosophical Transactions of the Royal Society of London, Series B. 2003; 358(1431):435–45.. [PubMed: 12689371]

Saxe R, Kanwisher N. People thinking about thinking peopleThe role of the temporo-parietal junction in "theory of mind.". NeuroImage. 2003; 19(4):1835–1842. [PubMed: 12948738]

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage. 2004; 23(Suppl 1):S208–19. [PubMed: 15501092]

Stein JL, Wiedholz LM, Bassett DS, Weinberger DR, Zink CF, Mattay VS, Meyer-Lindenberg A. A validated network of effective amygdala connectivity. NeuroImage. 2007; 36(3):736–45. [PubMed: 17475514]

Todorov A, Uleman JS. The efficiency of binding spontaneous trait inferences to actors' faces. Journal of Experimental Social Psychology. 2003; 39(6):549–562.

Wagner DD, Haxby JV, Heatherton TF. The Representation of Self and Person Knowledge in the Medial Prefrontal Cortex. Wiley Interdisciplinary Reviews: Cognitive Science. 2012; 3(4):451–470. [PubMed: 22712038]

Wagner DD, Kelley WM, Heatherton TF. Individual differences in the spontaneous recruitment of brain regions supporting mental state understanding when viewing natural social scenes. Cerebral Cortex. 2011; 21(12):2788–96. [PubMed: 21527789]

Whalen PJ. Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current Directions in Psychological Science. 1998; 7(6):177–188.

Whalen PJ, Rauch SL, Etcoff NL, McInerney SC, Lee MB, Jenike M. a. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. The Journal of Neuroscience. 1998; 18(1):411–8. [PubMed: 9412517]

Yarkoni T, Poldrack RA, Nichols TE, Essen DC, Van, Wager TD, Van Essen DC. Large-scale automated synthesis of human functional neuroimaging data. Nature Methods. 2011; 8(8):665–70.. [PubMed: 21706013]

Yarkoni T, Poldrack RA, Van Essen DC, Wager TD. Cognitive neuroscience 2.0: Building a cumulative science of human brain function. Trends in Cognitive Sciences. 2010; 14(11):489–96. [PubMed: 20884276]

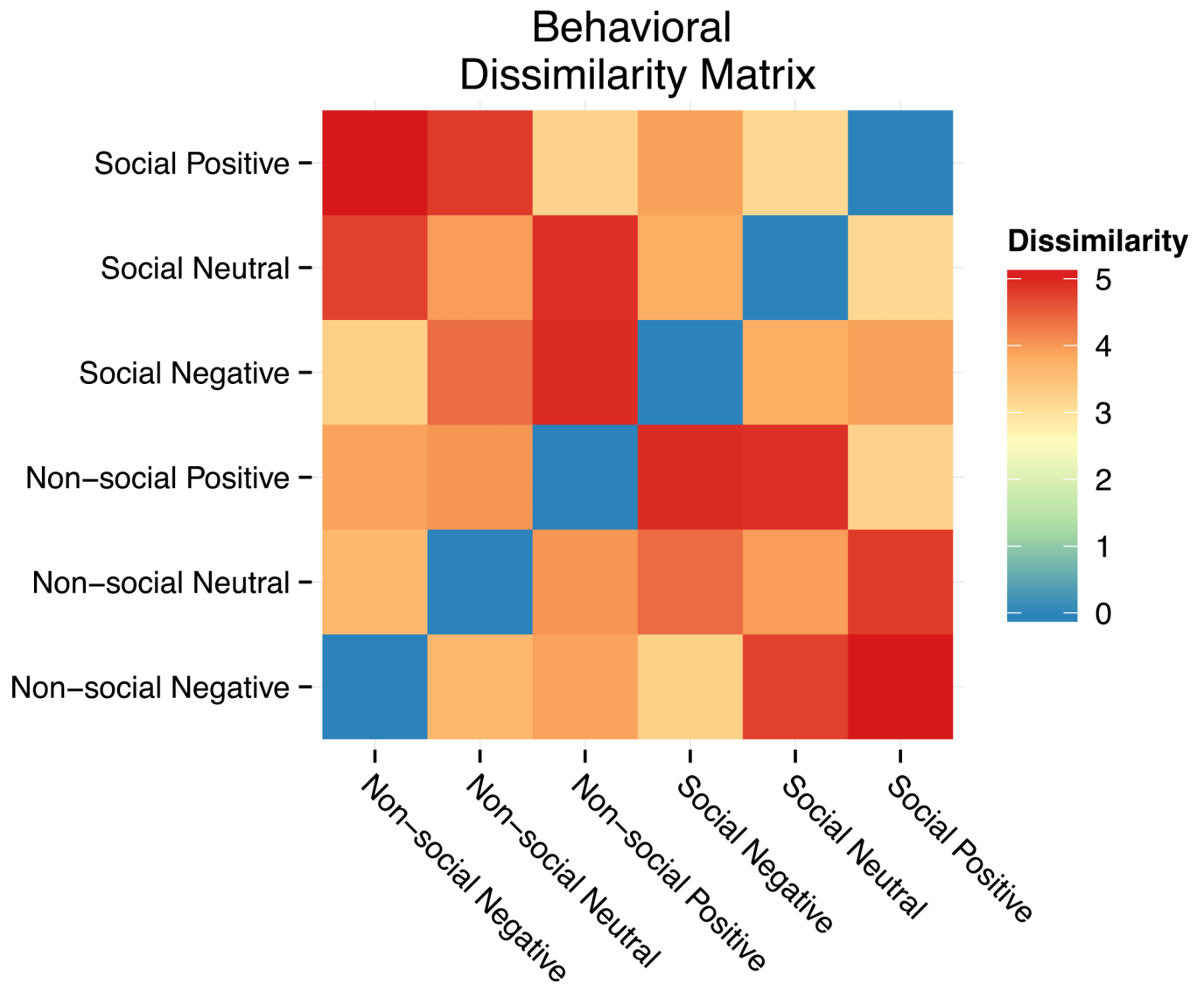## Behavioral Dissimilarity Matrix

Figure 1.
Dissimilarity matrix for the behavioral ratings portion of the study showing a high degree of similarity within both social and valence dimensions. Subjects rated pictures with social content as more similar to each other than pictures devoid of social content, irrespective of the valence of the photos. Likewise, subjects rated each valence category as more similar to itself, irrespective of the social content of the pictures. Cross-dimensional social and valence categories were rated as highly dissimilar to one another.
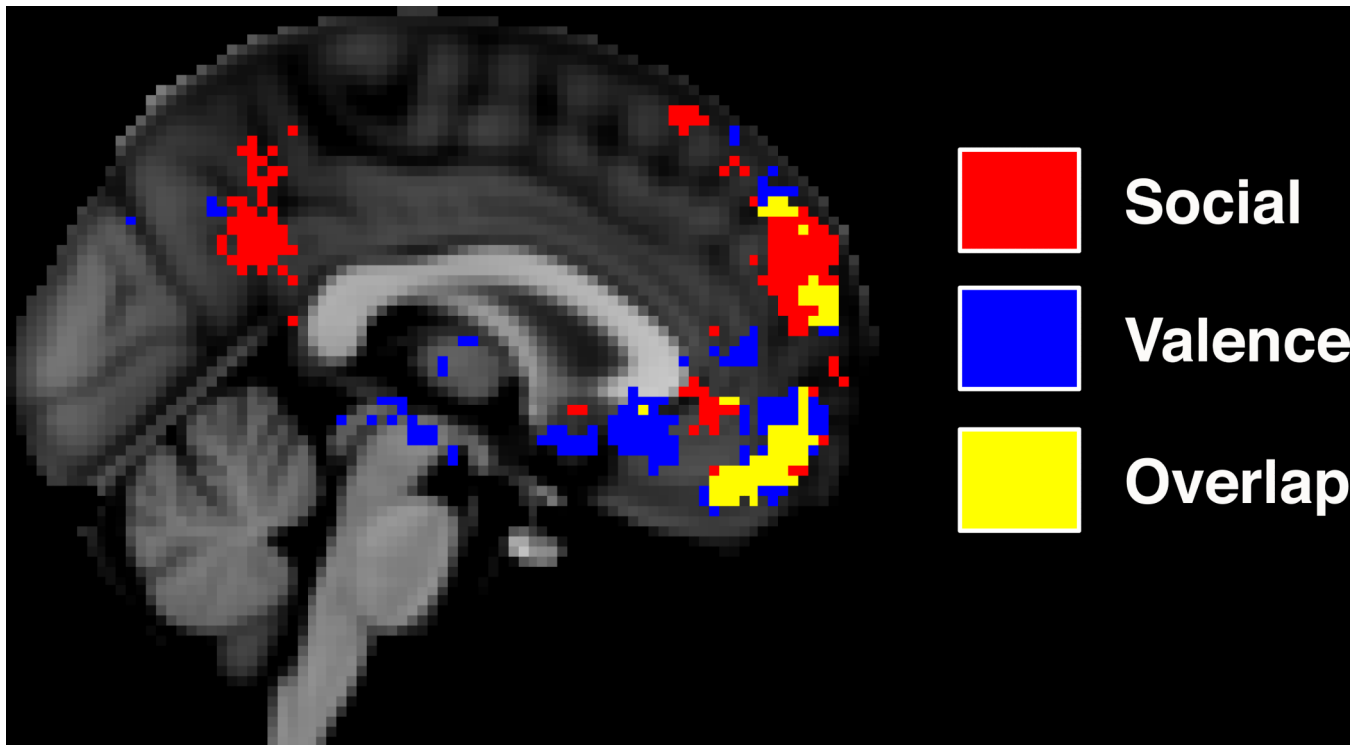
Figure 2.
Sagittal slice from the results of the automated meta-analysis showing overlapping spatial extents of brain areas involved in social and valence processing. Two large clusters of overlap can be seen in both the dorsal and ventral medial prefrontal cortex. Two additional clusters were found in the bilateral amygdala. These four areas were used as regions of interest for the representational similarity analysis.
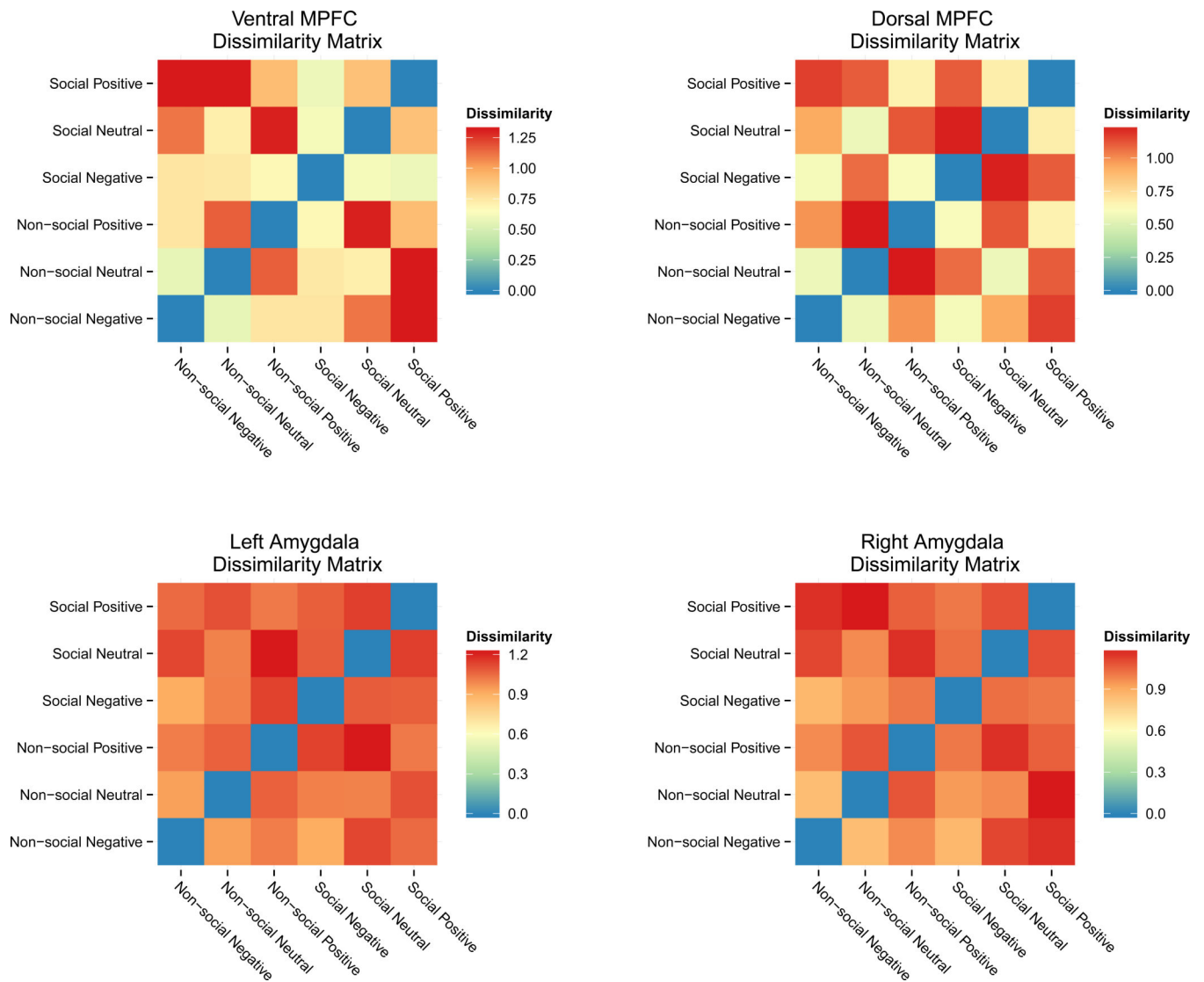
Figure 3.
Neural dissimilarity matrices for the four regions of interest from the representational similarity analysis. The ventral and dorsal medial prefrontal cortex showed a richer degree of representational content than the amygdala areas, which were markedly more homogeneous. The ventral medial prefrontal cortex was the only neural dissimilarity matrix to be significantly correlated with the behaviorally measured similarity structure.
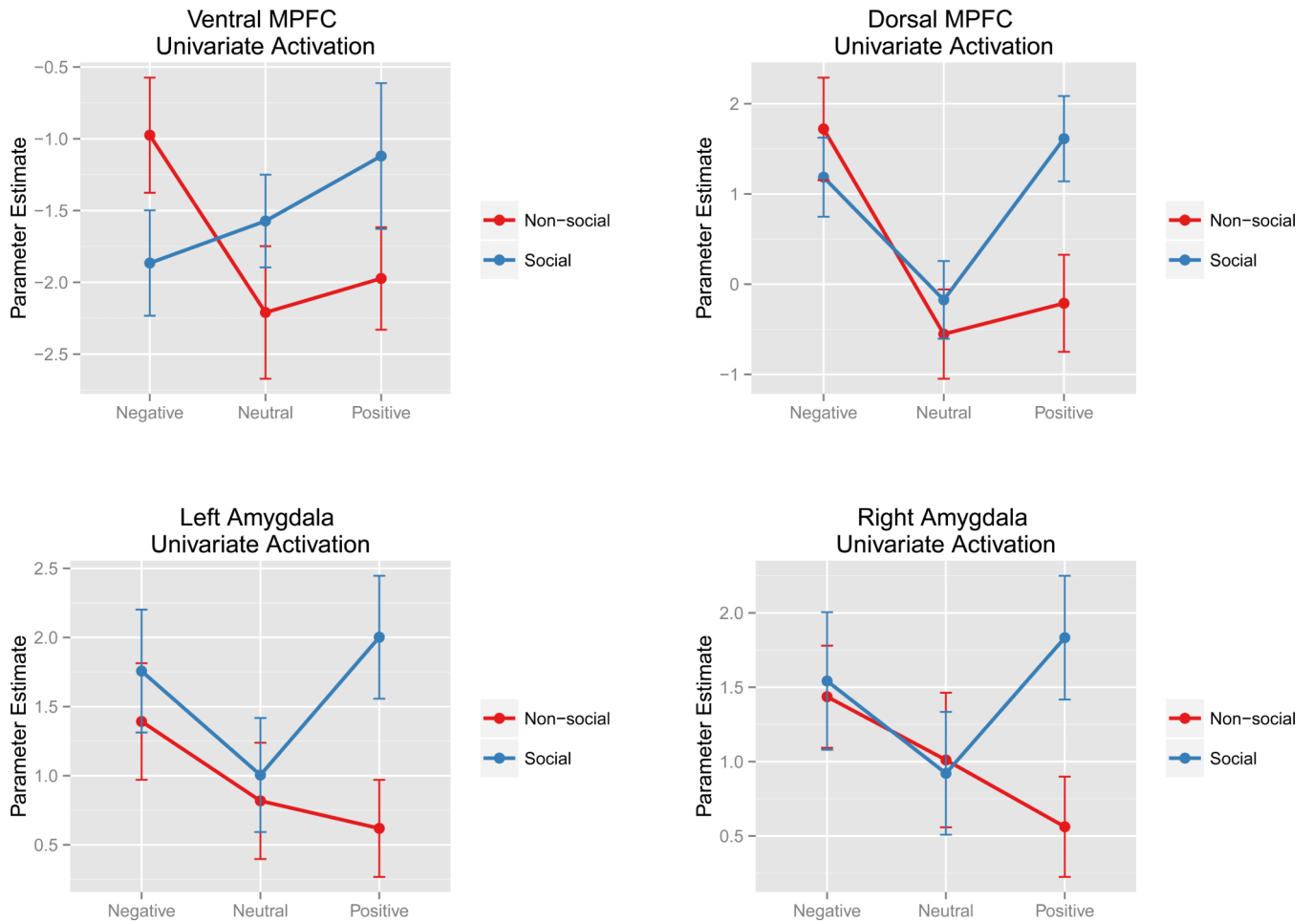
Figure 4.
Univariate contrasts of parameter estimates for each of the six social/valence conditions within the four ROIs from the RSA analysis. Error bars represent the standard error of the mean.