

The neuroscience of group membership

Samantha Morrison^a, Jean Decety^{b,c}, Pascal Molenberghs^{a,*}

^a The University of Queensland, School of Psychology, Qld 4072, Australia

^b Department of Psychology and Centre for Cognitive and Social Neuroscience, The University of Chicago, IL, United States

^c Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, IL, United States

ARTICLE INFO

Article history:

Received 9 January 2012

Received in revised form

11 May 2012

Accepted 11 May 2012

Available online 17 May 2012

Keywords:

fMRI

Group membership

Social identity

Social neuroscience

ABSTRACT

The present study aimed to uncover the neural activity associated with specific in-group and out-group word related stimuli, to examine the neuroanatomical basis of group membership concept representation, and investigate to what extent neural processes represent 'in-group' differently from 'out-group'. Participants' brain activity was measured with functional MRI while they had to categorize social, in-group and out-group words and non-social, living and non-living words. The results showed that a network of brain regions previously identified as the 'social brain', including the cortical midline structures, tempo-parietal junction and the anterior temporal gyrus showed enhanced activation for social words versus non-social words. Crucially, the processing of in-group words compared to the out-group words activated a specific network including the ventral medial prefrontal and anterior and dorsal cingulate cortex. These regions correspond to a neural network previously identified as the 'personal self'. Our results suggest that the 'social' and 'personal self' are closely related and that we derive our self image from the groups we belong to.

Crown Copyright © 2012 Published by Elsevier Ltd. All rights reserved.

1. Introduction

The need to belong to a group is an intrinsic and defining quality of human nature and this is reflected in the human tendency to perceive socially relevant categories, think in terms of stereotypes, and join groups (Ridley, 1996). Unusually large brains of primates and humans have been associated with living in complex social groups (Dunbar, 2011), and although much is known about the conditions under which people form groups (Hogg & Abrams, 1988; Tajfel & Turner, 1985), the specific neural processes that represent group membership remain largely unknown. On the other hand, social neuroscience studies have found evidence for distributed neural networks involved in social cognition in general (Amodio & Frith, 2006; Blakemore, 2008; Cacioppo & Decety, 2011), and recent evidence suggests the existence of two large-scale interacting neural networks that specifically represent the self and others (Uddin et al., 2007). The first network includes frontoparietal areas involved in the action observation network and is responsible for embodied cognition used to decode actions performed by others through simulation (Grafton, 2009; Molenberghs et al., 2012; Rizzolatti & Sinigaglia, 2010). Another brain network, underpinned by the tempo-parietal junction (TPJ) and cortical midline structures

(CMS) such as the medial prefrontal cortex, the cingulate cortex, and the precuneus is responsible for abstract perspective taking. The CMS regions are critical to the representation, monitoring, evaluation, and integration of self-referential stimuli (Northoff & Bermpohl, 2004) and the TPJ for Theory of Mind (Saxe, 2006) and complex social cognitive reasoning (Decety & Lamm, 2007; Decety et al., 2012).

Brain regions in the CMS have been consistently linked with cognitive operations that serve a 'self' function (Frith, 2007; Kelley et al., 2002). Specifically, the medial prefrontal cortex has been shown to be crucial for social judgements and self-referential processing (D'Argembeau et al., 2007, 2008, 2010; Gusnard et al., 2001; Jenkins & Mitchell, 2011; Macrae et al., 2004). For example, in an fMRI experiment, Kelley and colleagues (2002) monitored the neural activity of participants as they categorised a series of trait adjectives based on their relevance to the self (e.g., 'Does this trait describe you?'), other (e.g., 'Does this trait describe George Bush?') or case (e.g., 'Is this trait adjective presented in lowercase?'). During self-referential processing significant neural activity was found in the medial prefrontal cortex, the precuneus, and posterior cingulate cortex. These three regions are often activated together (Uddin et al., 2007, 2009), with the medial prefrontal cortex associated with mental state attribution and the posterior cingulate and precuneus linked with episodic memory retrieval (Cavanna & Trimble, 2006; Maddock, 1999; Mitchell et al., 2002; Ochsner et al., 2005), which suggest that these areas belong to a network through which personal identity

* Corresponding author.

E-mail address: p.molenberghs@uq.edu.au (P. Molenberghs).

and personal experiences are interlinked (Vogeley & Fink, 2003). In addition to the CMS regions, the TPJ has been associated with trait inferences about others, and is theorised to support the ability to reason about the contents of mental states (Saxe, 2006; Van Overwalle, 2009; Vogeley et al., 2001). It is worth noting that lower-level computations in the TPJ may underlie this complex social cognitive reasoning (Decety & Lamm, 2007). Recent fMRI research has also shown that social concepts such as “honor” or “brave” as opposed to animal function concepts such as “nutritious” or “useful” activate a selective anterior temporal lobe region, in addition to brain regions within the CMS (Zahn et al., 2007).

Recently, brain activity in the medial prefrontal cortex has been observed specifically when individuals make evaluative decisions about their in-group (Volz et al., 2009). Social identity theory acknowledges that both individual characteristics and socially shared characteristics (e.g., identification with a particular group) define a person's place in society (Tajfel & Turner, 1985). Both parts of the self concept are to some extent derived from favourable evaluative comparisons. Depending on what aspect of the self is salient, people will compare themselves to other individuals or compare their in-group to the relative out-group. Based on this theory, Volz and colleagues (2009) predicted that the social self (i.e., identity based on group membership) is derived by the same cognitive mechanisms as the personal self and that therefore significant medial prefrontal cortex activation was expected for situations in which the social self was addressed. Their prediction was supported by data from an fMRI experiment using a minimal group paradigm in which participants who showed the most in-group bias also had increased activity in the medial prefrontal cortex (Volz et al., 2009).

Together, these neuroimaging studies provide a clear illustration as how in-group processing can become part of self processing. Previous fMRI studies have already shown that we perceive faces (Cunningham et al., 2004; Harris & Fiske, 2007; Hart et al., 2000; Van Bavel et al., 2008; 2011) and actions (Molenberghs et al., *in press*) of in- and out-group members differently. What is missing in the literature, however, are investigations that specifically examine how neural mechanisms represent group-related concept words. The goal of the current study was to identify the neuroanatomical location of the neurodynamic response associated with stored group social categories, and to specifically investigate if in-group categories are stored differently from out-group categories. In order to

test this hypothesis, participants were presented with social in-group and out-group words, and non-social, living and non-living words while being scanned. We predicted that social words compared to non-social words would activate a specific network of brain regions previously associated with the ‘social brain’ including the medial prefrontal cortex, cingulate cortex, precuneus, temporo-parietal junction and superior anterior temporal lobe. In addition we predicted that in-group words compared to out-group words would additionally activate regions involved in the ‘personal self’, especially the medial prefrontal cortex.

2. Methods

2.1. Participants

Twenty healthy volunteers (six males, mean age=22.9 years, range: 18–33 years) completed the experiment. Participants were deemed healthy after passing the MRI medical checklist (i.e., no pacemaker, no brain clips, no major surgery, no mental history, etc.). Of the 20 participants, 17 had received tertiary qualifications, and the remaining 3 received secondary education. All but two participants were right-handed. Participants received a reimbursement of \$30 for their time. All participants gave written informed consent. No participants were excluded from the experiment. The study was approved by the Behavioural & Social Sciences Ethical Review Committee of the University of Queensland.

2.2. Experiment

Participants were required to choose a list of seven groups that they felt they belonged to and a list of seven groups they felt they did not belong to. It was explained that group membership could include any group that the participant was affiliated with and could be based on broad social concepts including gender, nationality, religious affiliation, or occupation (e.g., male, Australian, Muslim). The participants were then taken into the MRI scanner. All of the experimental stimuli appeared in a similar format: white coloured text on a black background. E-prime software (Psychology Software Tools, Inc.) was used to run the task on a PC. The task consisted of five different conditions: My team (MT), other team (OT), living (LI), non-living (NL), and a baseline condition. For each participant the 14 groups (seven that they felt they personally belonged to and the seven that they felt they personally did not belong to) listed during the interview constituted the MT and OT conditions. During the task the participants had to press either the left or the right button, respectively with their left or right hand, to indicate that they did or did not belong to this group. Participants also categorised a series of 14 different (7 living and 7 non-living) non-social control words (e.g., table, computer, elephant) Fig. 1.

The words were presented randomly for three seconds at the centre of the screen. A fixation cross followed each display and remained on the screen for one second.

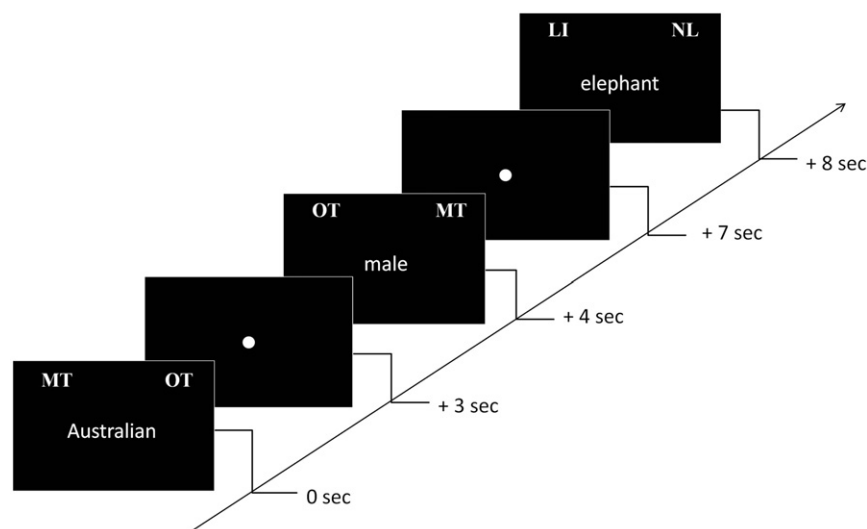


Fig. 1. Schematic representation of a section of the experimental task. Participants had to categorize own team (OT), other team (OT), living (LI) and non-living (NL) words by pressing a left or right button to indicate the side of the matching stimulus. Each slide was presented for 3 s followed by a 1 s fixation point. During the baseline the fixation point remained on the screen for 4 s.

When the group words were presented, the category labels: 'MT' and 'OT' appeared at the top of the screen. Half the time MT was on the left and OT was on the right and vice versa. For the non-group words the category labels: 'LI' and 'NL' followed a similar pattern. A null event was also presented to participants in which the fixation cross remained on the screen for an additional four seconds. The null event was used as a low level baseline to contrast the conditions against in the fMRI analysis. The entire task was conducted in 5 repeated fMRI runs, each lasting approximately 6 min in duration. Each run consisted of 14 trials per condition. A high resolution structural MRI scan was conducted after the third run.

2.3. fMRI image acquisition

A 3-Tesla Siemens MRI scanner with 32-channel head volume coil was used to obtain the data. Functional images were acquired using gradient-echo planar imaging (EPI) with the following parameters: repetition time (RT) 3 s, echo time (TE) 30 ms, flip angle (FA) 90°, 64 × 64 voxels at 3 × 3 in-plane resolution. Whole brain images were acquired every 3 s. The first two T_R periods from each functional run were removed to allow for steady-state tissue magnetization. T1-weighted image covering the entire brain was also acquired after the third run and used for anatomical reference (TR=1900, TE=2.32 ms, FA=9°, 192 cubic matrix, voxel size=0.9 cubic mm, slice thickness=0.9 mm).

2.4. fMRI analysis

SPM8 software (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London) run through Matlab (Mathworks Inc., USA) was used to analyse the data. To counter any head-movements all the EPI images were realigned to the first scan of each run. The anatomical image was then coregistered to the mean functional image. To correct for variation in brain size and anatomy between participants, each structural scan was normalised to the MNI T1 standard template (Montreal Neuropsychological Institute, Montreal, Canada) using segmentation. Spatial normalisation of all the EPI images was then conducted, using a standard stereotaxic space with a voxel size of 3 × 3 × 3 mm. This process mathematically transformed each participant's brain image to match the template so that any chosen brain region should refer to the same region across all participants. Before further analysis, all images were smoothed with an isotropic Gaussian kernel of 6 mm. As part of the first level of analysis a general linear model was created for each participant. For each participant, in each of the four conditions (e.g., MT, OT, LI, and NL), an event related design identified the regions with significant BOLD changes in each voxel. The BOLD changes in each condition were compared to the baseline. In the second level of analysis contrast images for each condition minus baseline across all participants were included in a factorial design. Follow up tests were created for each research hypothesis to determine if the differences in brain activation between conditions were significant. A cluster-level threshold with a familywise error rate (FWE) of $p < 0.05$, was used to define significant activation for all analyses, and a voxel-level probability threshold of $p < .001$ was used to define each cluster.

3. Results

3.1. Behavioural results

Mauchley's test indicated that the assumption of sphericity had been violated in the analysis of reaction time, $\chi^2=12.00$, $p=.035$. Therefore degrees of freedom were corrected using Greenhouse-Geiser estimates of sphericity ($\epsilon=.67$). The assumption of sphericity was not violated in the analysis of accuracy, $\chi^2=5.85$, $p=.32$. To correct for multiple comparisons, a Bonferroni corrected threshold was applied to all post hoc tests.

3.2. Reaction time

Reaction times were recorded for the four conditions (MT, OT, LI, and NL). The Shapiro–Wilk Test indicated that reaction times for all the conditions were normally distributed: MT ($p=.67$), OT ($p=.23$), LI ($p=.50$) and NL ($p=.50$). A one-way repeated measures ANOVA revealed a significant difference between the conditions, $F(2.02, 38.43)=37.62$, $p<.001$, $\eta^2=.66$. Post hoc pairwise comparisons revealed that the mean reaction time for the OT condition ($M=1204$, $SD=177$) was significantly slower than all the other conditions: MT condition ($M=1119$, $SD=163$, $p=.001$), LI condition ($M=965$, $SD=131$, $p<.001$) and NL condition

($M=1011$, $SD=121$, $p<.001$). Participants were significantly slower in the MT condition, compared to the LI condition ($p<.001$), and NL condition ($p=.004$). All other comparisons were non-significant.

3.3. Accuracy

The percentage of correct responses was recorded for each of the four conditions.

The Shapiro–Wilk Test indicated that the accuracy was not normally distributed for all the conditions: MT ($p<.001$), OT ($p=.002$), LI ($p=.22$) and NL ($p=.28$). Therefore we used non-parametric testing for the accuracy data. The Related-Samples Friedman's Two-Way Analysis of Variance by Ranks indicated a significant difference between the four conditions, $p=.002$. Post hoc pairwise Wilcoxon Signed Rank comparisons revealed that the mean percentage of correct responses in the MT condition ($M=97.7\%$, $SD=2.64$) was significantly higher than in the OT ($M=96.0\%$, $SD=4.03$, $p=.03$), LI ($M=93.6\%$, $SD=4.41$, $p=.002$) and NL conditions ($M=95.3\%$, $SD=3.29$, $p=.004$). All other comparisons were non-significant.

Because accuracy alone does not take into account response biases we also calculated the unbiased hit rate (H_u). This measure is the joint probability that a stimulus category is correctly identified given that it is presented at all and that the response is correctly used given that it is used at all (Wagner, 1993). The Shapiro–Wilk Test indicated that the H_u was not normally distributed for all the conditions: MT ($p<.001$), OT ($p=.003$), LI ($p=.47$) and NL ($p=.54$). Therefore we used non-parametric testing for the H_u accuracy data. The Related-Samples Friedman's Two-Way Analysis of Variance by Ranks indicated a significant difference between the four conditions, $p<.001$. Post hoc pairwise Wilcoxon Signed Rank comparisons revealed that the H_u index in the MT condition ($M=0.944$, $SD=0.053$) was significantly higher than in the, LI ($M=0.896$, $SD=0.059$, $p=.003$) and NL conditions ($M=0.897$, $SD=0.055$, $p=.003$). The H_u index in the OT condition ($M=0.941$, $SD=0.053$) was also significantly higher than in the, LI ($p=.002$) and NL conditions ($p=.003$). All other comparisons were non-significant.

3.4. Functional MRI results

The implicit mask image produced by SPM for the random effects analysis which shows the coverage of the EPI images is shown in Fig. 2. The image shows that the whole brain was covered, apart from some small parts in the ventral frontal and anterior and middle temporal regions which are known to be affected by air pockets close to the brain.

3.5. Social versus non-social words

To identify the network specifically involved in social word representation we contrasted the social (MT and OT) minus non-social (LI and NL) words. Increased hemodynamic activity was found in the left and right temporo-parietal junction (TPJ), the precuneus and adjacent posterior cingulate cortex, the ventromedial and dorsomedial prefrontal cortex, the left anterior and posterior temporal gyrus and the left orbitofrontal cortex (Table 1, Fig. 3). The reverse contrast did not elicit any significant brain activation.

3.6. My team versus other team words

To identify the network specifically involved in own-group membership we compared the "My Team" condition minus the "Other Team" condition. Increased activity was found in the anterior cingulate cortex, the ventromedial prefrontal cortex

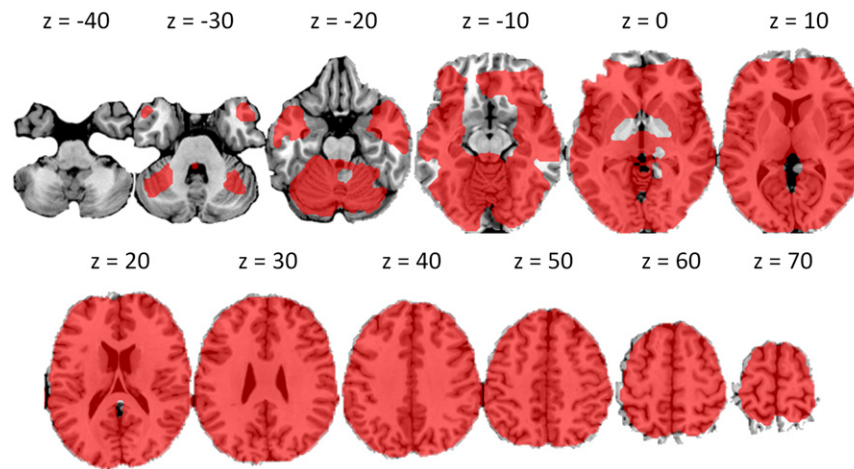


Fig. 2. The implicit mask image produced by SPM for the random effects analysis showing the coverage of the EPI images displayed on transversal slices (labelled with MNI coordinates) of the ch2better.nii.gz template using MRIcron.

Table 1

Cluster size and associated peak values for the significant contrasts: social versus non-social words and my team versus other team words.

Contrast and brain region	FWE corrected <i>p</i> value (cluster-level)	Cluster size	Peak <i>Z</i>	MNI coordinates		
				<i>x</i>	<i>y</i>	<i>z</i>
Social versus non-social words						
Left temporo-parietal junction	< 0.001	249	6.47	−45	−64	31
Precuneus and posterior cingulate cortex	< 0.001	357	5.36	−3	−46	28
Ventromedial prefrontal cortex	0.022	64	5.09	−3	44	−17
Left anterior middle temporal gyrus	0.008	81	4.70	−63	−13	−14
Left orbitofrontal cortex	0.005	88	4.38	−45	35	−5
Left posterior middle temporal gyrus	0.002	108	4.35	−66	−34	1
Dorsomedial prefrontal cortex	0.015	70	3.91	−6	56	16
Right temporo-parietal junction	0.016	69	3.84	36	−64	43
My team versus other team words						
Anterior cingulate cortex	0.012	74	4.91	−3	32	7
Ventromedial prefrontal cortex	0.018	67	4.33	12	50	−5
Dorsal midcingulate cortex	0.016	69	4.31	−18	−1	52

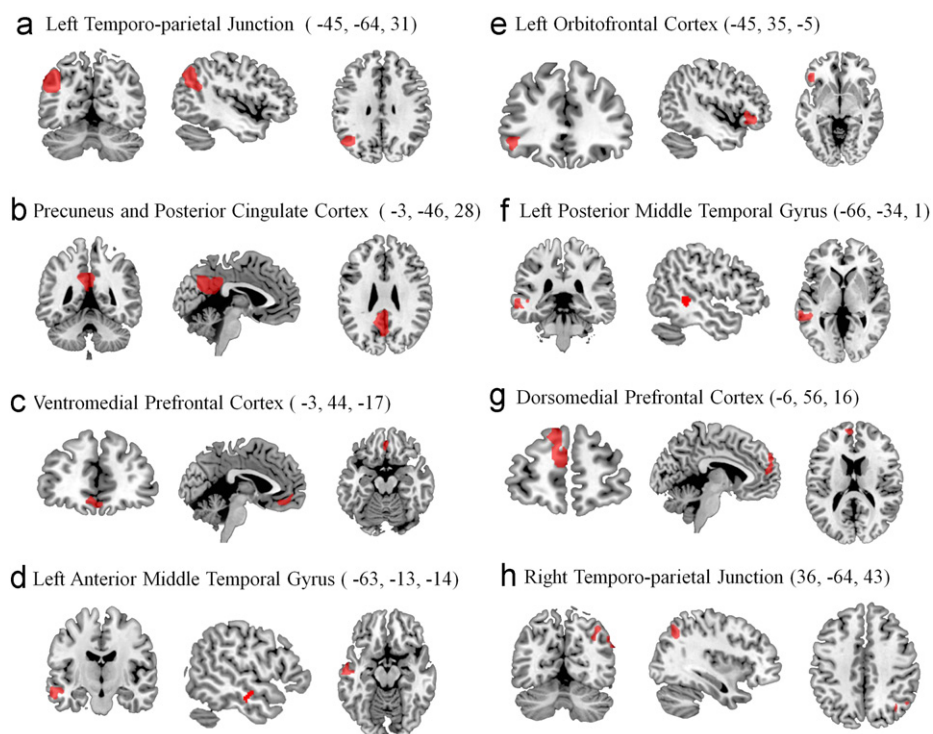


Fig. 3. Significant brain activation in the social words (MT and OT) versus non-social words (LI and NL) contrast. Individual figures are labelled with MNI coordinates. Activations (thresholded at $p < 0.001$ uncorrected) are displayed on a ch2better.nii.gz template using MRIcron.

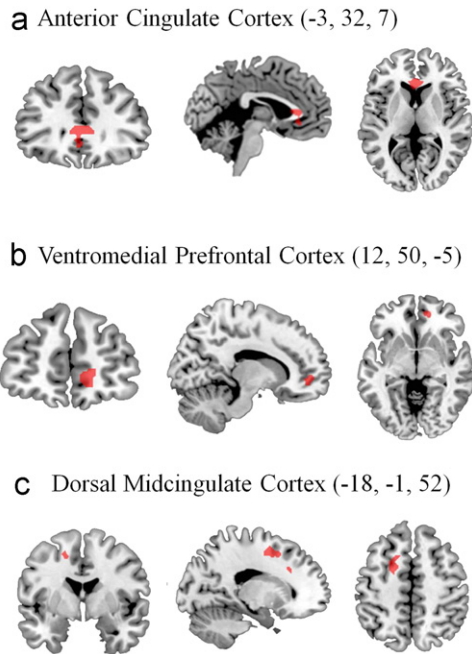


Fig. 4. Significant brain activation in the 'My Team' (MT) minus 'Other Team' (OT) contrast. Individual figures are labelled with MNI coordinates. Activations (thresholded at $p < 0.001$ uncorrected) are displayed on a *ch2better.nii.gz* template using MRIcron.

and the dorsal midcingulate cortex (Table 1, Fig. 4). The reverse contrast did not reveal any significant brain activation. No significant differences were found between the two non-social (LI and NL) conditions.

To further examine the specificity of our findings and to make sure the fMRI results were not related to differences in reaction time and accuracy between conditions, we performed our secondary data analyses in which we partialled out the effects of reaction time and accuracy by modelling them as parametric modulations in our fMRI design. We then tested if in our new design the results described above were still significant by performing for the two contrasts ("social minus non-social words" and "My Team minus Other Team words") region of interest analyses for each of the significant clusters listed in Table 1. All the results remained highly significant (voxel-level threshold of $p_{FWE} \leq 0.005$) which shows that differences in reaction time and accuracy between conditions did not cause the effects.

4. Discussion

To date, this study is the first to examine the neural response associated with the presentation of specific in-group and out-group related word stimuli. The study aimed to identify the brain regions that were representative of group membership concept representation, and to compare the neural activity elicited for in-group words to that of the out-group words. The differences that were found offer an anatomically based explanation for social group distinctions. Social words versus non-social words elicited a unique pattern of brain response within the CMS, orbitofrontal cortex, TPJ and the anterior temporal gyrus. Brain activity centred around the CMS, including the medial prefrontal cortex, the cingulate cortex, and the precuneus, has consistently been associated with social cognition (Northoff & Bermpohl, 2004). For example, Mitchell and colleagues (2002) previously identified unique neural activity in the medial prefrontal lobe, a region involved in self-monitoring and mental state attribution (Frith, 2007), which was associated with person judgments as opposed to object judgments. Additionally, Zahn and colleagues (2007)

detected brain activity in the superior anterior temporal lobe region specifically for tasks that involved abstract social semantic knowledge. They also found brain regions such as the orbitofrontal cortex, the medial prefrontal cortex, and the temporo-parietal junction to be activated with the presentation of words that were associated with social concepts, as opposed to animal function concepts. Interestingly the anterior temporal region, contrary to the medial prefrontal region, responded independently of emotional valence, which suggest that this regions just provides abstract conceptual knowledge to the frontolimbic regions where these concepts are associated with either positive or negative connotations (Zahn et al., 2007, 2009). Van Overwalle (2009) also argued for the importance of the medial prefrontal cortex and the TPJ in social cognition with the latter being involved in attributing mental states to oneself and others (Saxe, 2006), as well as increased allocation of attention and the sense of agency (Decety & Lamm, 2007). The current results align with the literature, verifying that social knowledge is associated with distinct neural substrates. The number of neural regions that were found to be associated with social cognition is also supportive of the notion of the 'social brain', and highlights the fact that a large part of the human brain is 'social'. Given that virtually all human activity is shaped by social context and has some sort of social implication (Iacoboni et al., 2004), it is not surprising that the human ability to attend to, and process social relations, is facilitated by a number of cortical areas.

Several regions were found to be uniquely associated with the in-group while no specific neural activity was associated with the out-group. It could be argued that participants were more familiar and had more positive associations with the in-group and this could partially explain the difference in activation. However we think this is unlikely given the fact that both the in-group and out-group words presented to each participant were elected by each participant and as a result both had a certain level of relevance and familiarity for each individual. In addition, it was not necessarily the case that the out-groups were perceived as less positive, for example participants picked out-groups such as 'tennis players', 'art students' and 'full-time workers' that are not stigmatized groups or likely to elicit innate prejudices. In addition, previous fMRI studies have shown that faces (Van Bavel et al., 2008, 2011) and actions (Molenberghs et al., in press) in newly created in-group members are processed in more depth than out-group members although participants had the same amount of exposure and affinity to both groups. Future research should clarify how the comparison between in-group and out-group word representation differs in newly created and existing groups.

The areas of activation specifically associated for processing in-group words were located within the CMS, including the anterior cingulate cortex, the ventromedial prefrontal cortex, and the dorsal midcingulate cortex. Previous research has linked these regions to cognitions that serve a self function. For example, Johnson and colleagues (2002) specifically identified neural activity in the medial prefrontal cortex and anterior cingulate cortex that was associated with self-reflective thought and Kelley and colleagues (2002) recorded distinct neural activity in the medial prefrontal lobe and cingulate cortex when participants categorised self-referential stimuli. Given that the in-group words activated areas of the brain that are associated with self-referential processing, the overlapping neural activity in brain areas that are associated with personal and social identity, demonstrates that the in-group can be represented neuroanatomically as part of the self (Volz et al., 2009). The overlapping cortical regions are also a possible explanation for why close others are very often perceived to be similar to the self (Heatherton et al., 2006) and why the medial prefrontal cortex is not activated for extreme low status out-groups (Harris & Fiske, 2006).

In terms of the activation detected in the cingulate cortex, it is worth noting that neuroimaging studies have documented the functional heterogeneity of this region. Activity in the dorsal cingulate cortex has been associated with tasks that require heightened, discriminative attention, whereas the anterior ventral cingulate mediates affective responses (Drevets & Raichle, 1998; Zubieta et al., 2003). The fact that social group words activate the dorsal medial prefrontal cortex and in-group words specifically a ventral part of the medial prefrontal cortex is interesting because brain activity accompanying the comprehension of triadic relationships (i.e., understanding the relationship between two minds and an object; Me, You, and This) has been previously associated with the dorsal part of the medial prefrontal cortex (Saxe, 2006), while the ventral part of the medial prefrontal cortex has been implicated in 'emotional empathy', which is the ability to experience a congruent observed emotion (Saxe, 2006) which happens more easily in own group members (Xu et al., 2009). Similarly Volz and colleagues (2009), have suggested that the dorsal part of the medial prefrontal cortex is involved in abstract social judgments and mentalizing, whereas the ventral part of the MPFC is involved in more general affective functioning. These views correspond well with the observation of specific ventral medial prefrontal cortex and anterior cingulate cortex activation for emotional significant own-group words in our study. Ventral parts of the medial prefrontal cortex have also been shown to become engaged when people make trait inferences about familiar others, such as close family members or friends as well as with judgements made about the self (D'Argembeau et al., 2007; Van Overwalle, 2009).

Overall, the current results extend and add to the existing body of research by directly comparing the pattern of neural response elicited when individuals make distinctions between their relative in-groups and out-groups. The findings show that only in-groups words, and not out-groups words, are associated with unique neural activity. The neural activity identified for in-group words, allude to a higher level of sophistication in the areas of the social brain which enables people to differentially represent the in-group from the out-group. In consideration of the perceptual bias that is the accentuation effect, perhaps it is not surprising that we perceive in-group members as more like the self, and out-groups members as less individuated. Not only does the human brain seem to have more cortical facilities that are used to process in-group related stimuli, but these cortical areas are within the CMS regions, which as previously stated, have been established to serve a self function.

Acknowledgements

This work was supported by a UQ Postdoctoral Fellowship and a UQ Early Career Research Grant awarded to PM.

References

- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *7*, 268–277. *Nature Reviews Neuroscience*, *7*, 268–277.
- Blakemore, S.-J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*, 267–277.
- Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*, 564–583.
- Cacioppo, J. T., & Decety, J. (2011). Social neuroscience: Challenges and opportunities in the study of complex behavior. *Annals of the New York Academy of Sciences*, *1224*, 162–173.
- Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science*, *15*, 806–813.
- D'Argembeau, A., Feyers, D., Majerus, S., Collette, F., Van der Linden, M., Maquet, P., et al. (2008). Self-reflection across time: cortical midline structures differentiate between present and past selves. *Social Cognitive and Affective Neuroscience*, *3*, 244–252.
- D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Baetens, E., Luxen, A., et al. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *Journal of Cognitive Neuroscience*, *19*, 935–944.
- D'Argembeau, A., Stawarczyk, D., Majerus, S., Collette, F., Van der Linden, M., & Salmon, E. (2010). Modulation of medial prefrontal and inferior parietal cortices when thinking about past, present, and future selves. *Social Neuroscience*, *5*, 187–200.
- Decety, & Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist*, *13*, 580–593.
- Decety, K. J., Michalska, & K. D., Kinzler (2012). The contribution of emotion and cognition to moral sensitivity: a neurodevelopmental study. *Cerebral Cortex*, *22*, 209–220.
- Drevets, W. C., & Raichle, M. E. (1998). Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: implications for interactions between emotion and cognition. *Cognition & Emotion*, *12*, 353–385.
- Dunbar, R. (2011). Evolutionary basis of the social brain. In: J. Decety, & J. T. Cacioppo (Eds.), *The Oxford Handbook of Social Neuroscience* (pp. 28–38). New York: Oxford University Press.
- Frith, C. D. (2007). The social brain?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 671–678.
- Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. In: M. B. K. A. Miller (Ed.), *Year in Cognitive Neuroscience 2009* (pp. 97–117).
- Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences USA*, *98*, 4259–4264.
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: neuroimaging responses to extreme out-groups. *Psychological Science*, *17*, 847–853.
- Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social Cognitive and Affective Neuroscience*, *2*, 45–51.
- Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport*, *11*, 2351–2355.
- Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., & Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience*, *1*, 18–25.
- Hogg, M. A., & Abrams, D. (1988). *Social Identifications: A Social Psychology Of Intergroup Relations And Group Processes*. London: Routledge.
- Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., et al. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage*, *21*, 1167–1173.
- Jenkins, A. C., & Mitchell, J. P. (2011). Medial prefrontal cortex subserves diverse forms of self-reflection. *Social Neuroscience*, *6*, 211–218.
- Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain*, *125*, 1808–1814.
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785–794.
- Macrae, C. N., Heatherton, T. F., & Kelley, W. M. (2004). A self less ordinary: the medial prefrontal cortex and you. In: M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 1067–1077). London: MIT Press.
- Maddock, R. J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. *Trends in Neurosciences*, *22*, 310–316.
- Mitchell, J. P., Heatherton, T. F., & Macrae, C. N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences USA*, *99*, 15238–15243.
- Molenberghs, P., Cunningham, R., & Mattingley, J. B. (2012). Brain regions with mirror properties: a meta-analysis of 25 human fMRI studies. *Neuroscience and Biobehavioral Reviews*, *36*, 341–349.
- Molenberghs, P., Halász, V., Mattingley, J. B., Vanman, E., & Cunningham, R. Seeing is believing: neural mechanisms of action perception are biased by team membership. *Human Brain Mapping*. doi: <http://dx.doi.org/10.1002/hbm.22044>, in press.
- Northoff, G., & Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, *8*, 102–107.
- Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D. E., Kihlstrom, J. F., et al. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage*, *28*, 797–814.
- Ridley, M. (1996). *The origins of virtue*. London: Penguin Books.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, *16*, 235–239.
- Tajfel, H., & Turner, J. C. (1985). Social identity theory and intergroup behaviour. In: S. Worchel, & W. G. Austin (Eds.), *Psychology of Intergroup Relations* (pp. 7–24). Chicago: Nelson-Hall.

- Uddin, L. Q., Iacoboni, M., Lange, C., & Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences*, **11**, 153–157.
- Uddin, L. Q., Kelly, A. M. C., Biswal, B. B., Castellanos, F. X., & Milham, M. P. (2009). Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Human Brain Mapping*, **30**, 625–637.
- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: a functional magnetic resonance imaging investigation. *Psychological Science*, **19**, 1131–1139.
- Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience*, **23**, 3343–3354.
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, **30**, 829–858.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., et al. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, **14**, 170–181.
- Vogeley, K., & Fink, G. R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, **7**, 38–42.
- Volz, K. G., Kessler, T., & von Cramon, D. Y. (2009). In-group as part of the self: in-group favoritism is mediated by medial prefrontal cortex activation. *Social Neuroscience*, **4**, 244–260.
- Wagner, H. L. (1993). On measuring performance in category judgment studies of nonverbal behavior. *Journal of Nonverbal Behavior*, **17**, 3–28.
- Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? racial group membership modulates empathic neural responses. *Journal of Neuroscience*, **29**, 8525–8529.
- Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences USA*, **104**, 6430–6435.
- Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., et al. (2009). The neural basis of human social values: evidence from functional MRI. *Cerebral Cortex*, **19**, 276–283.
- Zubieta, J. K., Ketter, T. A., Bueller, J. A., Xu, Y. J., Kilbourn, M. R., Young, E. A., et al. (2003). Regulation of human affective responses by anterior cingulate and limbic mu-opioid neurotransmission. *Archives of General Psychiatry*, **60**, 1145–1153.