Behavioral/Systems/Cognitive

Encoding-Specific Effects of Social Cognition on the Neural Correlates of Subsequent Memory

Jason P. Mitchell, C. Neil Macrae, and Mahzarin R. Banaji

To examine whether social cognition recruits distinct mental operations, we measured brain activity during social (“form an impression of this person”) and relatively nonsocial (“remember the order in which person information is presented”) orienting tasks. Extending previous research on the neural basis of social cognition, the impression formation task differentially engaged an extensive region of the dorsomedial prefrontal cortex (PFC). In contrast, the nonsocial sequencing task differentially engaged the superior frontal and parietal gyri, precentral gyrus, and the caudate. In addition, we compared encoding activations for subsequently remembered (i.e., hits) to subsequently forgotten (i.e., misses) items. The brain regions in which the blood oxygenation level-dependent signal distinguished subsequent hits from subsequent misses depended on which orienting task was performed at encoding; subsequent memory was correlated with encoding activity only in the medial PFC for impression formation trials but in the right hippocampus for sequencing trials. These data inform two interrelated cognitive issues. First, results underscore the neuroanatomical distinctiveness of social cognition and suggest that previous psychological theories may have neglected important functional differences in how the human brain instantiates social and nonsocial cognitive processes. Second, by demonstrating that activity in different brain regions correlates with subsequent memory as a function of the orienting task performed at encoding, these data provide evidence of the neural basis for encoding specificity, the principle that memory is critically determined by the cognitive process engaged by the initial study episode.

Key words: hippocampus; memory; functional magnetic resonance imaging; medial prefrontal cortex; social cognition; theory of mind

Introduction

Recent neuroimaging and neuropsychological research has suggested that social cognition, processes that support an understanding of the psychological characteristics of other people, such as their beliefs, feelings, or personality, may be subserved by a system of brain regions that is neuroanatomically distinct from that giving rise to our understanding of inanimate objects (Frith and Frith, 2001; Gallagher and Frith, 2003). For example, Mitchell et al. (2002) demonstrated that whereas judgments about inanimate objects engage brain regions associated with semantic memory tasks (left ventrolateral prefrontal and inferotemporal cortex), comparable judgments about social agents modulate a distinct set of brain regions that include the dorsal and ventral medial prefrontal cortex (PFC). These results parallel previous neuroimaging demonstrations that the medial PFC is selectively engaged during social–cognitive tasks that require consideration of the mental states of either another person (for review, see Gallagher and Frith, 2003) or of oneself (Gusnard et al., 2001; Kelley et al., 2002).

In the current study, we capitalized on these previous observations to address two interrelated questions. First, can the putative segregation of social and nonsocial processes help resolve a long-standing theoretical issue within behavioral studies of social cognition? Tasks that direct attention to the socially relevant aspects of items typically produce memory performance that differs from nonsocial tasks both quantitatively (more accurate memory) and qualitatively (e.g., recall that clusters around spontaneously inferred personality traits of the targets) (Hastie and Kumar, 1979; Hamilton et al., 1980, 1989; Wyer et al., 1984; Srull and Wyer, 1989). Extant cognitive theories have posited that these differences emerge because social–cognitive processing prompts deeper, more elaborative, encoding of the sort that generally supports episodic memory (e.g., generation of schemas). Here, we suggest that, although parsimonious, these models may have failed to consider the possibility that social cognition relies instead on specialized processes that are distinct from those supporting memory in other domains.

Second, we capitalized on the predicted dissociation between social and nonsocial tasks to examine the neural basis of encoding specificity (Tulving and Thomson, 1973; Tulving, 1979). The principle of encoding specificity states that the manner in which information is initially encoded dictates the kinds of memory traces subsequently available. Although previous neuroimaging studies have demonstrated both (1) that processing different materials (e.g., faces, words) may engage different brain regions (Brewer et al., 1998; Kelley et al., 1998; Wagner et al., 1998) and
that different orienting tasks (e.g., semantic vs phonological encoding) are associated with activations in distinct brain regions (Fiez, 1997; Poldrack et al., 1999), strong evidence for the neural basis of the encoding specificity principle also requires that activation in distinct brain regions should correlate with subsequent memory success as a function of orienting task (Otten and Rugg, 2001a,b; Otten et al., 2002; Fletcher et al., 2003). Although a full test of the encoding specificity principle requires manipulation of both encoding and retrieval tasks, in the current study, we exploit differences between social and nonsocial orienting tasks to demonstrate encoding-specific differences in the neural correlates of successful remembering.

Materials and Methods

Participants. Participants were 17 (10 males) right-handed, native English speakers with no history of neurological problems (mean age, 21 years; range, 18–23). Informed consent was obtained in a manner approved by the Human Studies Committee of the Massachusetts General Hospital.

Stimuli and behavioral procedure. Stimuli consisted of 180 descriptive statements that suggested 18 different personality traits (10 statements per trait); for example, friendly (“he went out of his way to meet someone from a different background”), fun-loving (“at the party, he was the first to start dancing on the table”), inconsiderate (“he refused to loan his extra blanket to the other campers”), and lazy (“he slept through the lecture because he knew that he could watch it online”). During encoding, statements were paired with 18 faces (Caucasian males photographed against a blue background). Each face was paired with 10 descriptive statements. To replicate previous behavioral procedures as closely as possible, five of these statements converged on a single descriptive trait (e.g., friendly), whereas the remaining statements were chosen randomly from among the other traits. No behavioral or functional magnetic resonance imaging (fMRI) differences were observed between these types of statements; accordingly, all analyses are reported collapsed across this factor.

Each encoding trial consisted of a face–statement pair presented for 5500 msec. Each pair was accompanied by one of two cues (form impression, remember order) that indicated, respectively, whether the impression formation or sequencing task was to be performed on that trial. In line with behavioral studies that have investigated social–cognitive effects on memory, for impression formation trials, participants were instructed to use the statement to infer the personality traits of the target person; participants were told that, for these trials, their opinion about each person would later be measured. For sequencing trials, participants were instructed to encode the order in which statements were paired with each face (e.g., they were encouraged to consider the statements as describing the activities performed by each person at various times of the previous month and to encode the order in which they occurred). Participants were told that, for these trials, their memory for the sequences would later be tested. In fact, no such tests were administered.

In each of three functional runs, six faces were each presented 10 times (30 impression formation and 30 sequencing trials). A different descriptive statement accompanied each presentation of a face; however, across presentations, a given face was consistently associated with the same orienting task. To optimize estimation of the event-related fMRI response, trials were intermixed in a pseudorandom order and separated by a variable interstimulus interval (500–7500 msec) (Dale, 1999). During interstimulus intervals, participants passively viewed a fixation crosshair.

Approximately 10 min after completing the last functional run, participants performed an associative retrieval task. During the retrieval task, each of the 180 previously encountered statements was presented, and participants were asked to indicate which of six faces had earlier been paired with that statement. On each trial, one face was the correct target, whereas the five distractor faces were those that had appeared within the same functional run as the target. Encoding data were conditionalized on the basis of both subsequent memory performance and orienting task, resulting in four trial types: impression-hits, impression-misses, sequencing-hits, sequencing-misses.

Imaging procedure. Imaging was conducted using a 1.5 Tesla Sonata scanner (Siemens, Erlangen, Germany). We first collected a high-resolution T1-weighted structural scan (MP-RAGE), followed by three functional runs of 240 volume acquisitions (25 axial slices; 5 mm thick; 1 mm skip). Functional scanning used a gradient-echo planar pulse sequence (repetition time, 2 sec; echo time, 40 msec; 3.75 × 3.75 in-plane resolution). Stimuli were projected onto a screen at the end of the magnet bore that participants viewed by way of a mirror mounted on the head coil. A pillow and foam cushions were placed inside the head coil to minimize head movements.

fMRI data were preprocessed and analyzed using SPM99 (Wellcome Department of Cognitive Neurology, London, UK). First, functional data were time-corrected for differences in acquisition time between slices for each whole-brain volume and realigned to correct for head movement. Functional data were then transformed into a standard anatomical space (3 mm isotropic voxels) based on the ICBM 152 brain template (Montreal Neurological Institute, Montreal, Quebec, Canada), which approximates Talairach and Tournoux atlas space. Normalized data were then spatially smoothed (8 mm full-width at half-maximum) using a Gaussian kernel.

Statistical analyses were performed using the general linear model in which the event-related design was modeled using a canonical hemodynamic response function, its temporal derivative, and covariates of no interest (a session mean and a linear trend). Comparisons of interest were implemented as linear contrasts using a random effects model. A voxel-based statistical threshold of p < 0.001 was used for all comparisons; regions of interest (ROIs) were required to exceed 25 contiguous voxels in extent (providing an α level of p < 0.05; corrected) for contrasts between tasks (e.g., impression formation > sequencing) and five contiguous voxels (i.e., p < 0.001; uncorrected) for subsequent memory comparisons (e.g., impression-hits > impression-misses). Peristimulus hemodynamic time courses for each of these ROIs were extracted on a subject–by-subject basis using a selective averaging procedure (R. A. Poldrack, University of California Los Angeles, Los Angeles, CA). Statistical comparisons between conditions were conducted using ANOVA procedures on the parameter estimates associated with each trial type.

Results

Behavioral data

At test, participants correctly matched 58.8% of statements to the face with which it was originally paired (chance performance, 16.7%). Replicating previous findings that better memory performance follows social than nonsocial orienting tasks, associative memory was significantly better after impression formation (62.9% hits) than sequencing (54.6% hits) (t(16) = 2.57; p < 0.025; Cohen’s d = 0.64).

fMRI data: impression formation > sequencing

We adopted several complementary analytic strategies to examine differences in neural activation among conditions. First, we directly contrasted impression formation > sequencing, regardless of subsequent memory (Fig. 1). Impression formation trials were associated with reliably greater activation (compared with sequencing trials) in a single location: dorsomedial PFC. This region was distributed as an extensive arc (comprising 298 voxels) along the medial banks of the superior frontal gyrus bilaterally that extended dorsally in more posterior areas (Fig. 1). No other brain regions were observed in this comparison.

An automated search algorithm was used to identify discrete ROIs clustered around local maxima within this large dorsomedial PFC region. Using a radius of 8 mm, 13 such regions were identified (Table 1). A selective averaging procedure was then used to extract peristimulus hemodynamic time courses for each of these ROIs on a subject–by-subject basis, and each ROI was interrogated for differences between experimental conditions. In seven of these regions, the blood oxygenation level-dependent
BOLD signal correlated with subsequent memory success, but only for impression formation trials. That is, in these regions, we observed greater activation for impression-hits than impression-misses but no difference between sequencing-hits and sequencing-misses. In the remaining six regions, the BOLD signal was not correlated with subsequent memory for either orienting task. Intriguingly, these two sets of regions were neuroanatomically distinct. The seven ROIs demonstrating a difference in memory \( (Dm) \) effect for impression formation items were all anterior to \( y = 45 \) and ventral to \( z = 45 \), whereas six of the seven ROIs in which we observed no difference between impression-hits and impression-misses were posterior to and all were dorsal to these planes (Fig. 2A, B).

To confirm the selectivity of this subsequent memory effect at the whole-brain level (i.e., when including all voxels in our brain volume), we also directly contrasted impression-hits > impression-misses. This contrast yielded a single cluster of voxels, located in the dorsomedial PFC \((-9, 60, 33)\). As shown in Figure 2C, whereas the BOLD signal in this region differentiated between impression-hits and impression-misses \( (F_{(1,16)} = 18.27; p < 0.001) \), no such difference was observed for sequencing items \( (F_{(1,16)} = 1.24; p > 0.25) \). That activation in this region correlated more strongly with subsequent memory for impression formation than for sequencing was confirmed by the presence of a significant two (orienting task: impression formation, sequencing)–by-two (subsequent memory: hit, miss) interaction \( (F_{(1,16)} = 5.16; p < 0.05) \).

**fMRI data: sequencing > impression formation**

As for impression formation trials, we first directly contrasted sequencing > impression formation, regardless of subsequent memory (Table 2). Sequencing trials were associated with reliably greater activation with impression formation trials in a distributed set of regions that included the bilateral superior frontal gyrus, left superior parietal gyrus, right precentral gyrus, and the head of the caudate. However, BOLD signal in none of these regions correlated with subsequent memory [i.e., the hemodynamic time courses in each of these ROIs did not differentiate between hits and misses for either the sequencing or the impression formation trials (all \( p \) values > 0.20)].

Nevertheless, a direct contrast of sequencing-hits > sequencing-misses yielded a single brain region that demonstrated a \( Dm \) effect for sequencing trials, located in the right hippocampus \( (21, -12, -21) \). As shown in Figure 3, whereas the BOLD signal in this region of the right hippocampus differentiated between sequencing-hits and sequencing-misses \( (F_{(1,16)} = 21.27; p < 0.001) \), no such difference was observed for impression formation items \( (F_{(1,16)} = 0.18; \text{NS}) \). That activation in this region correlated more strongly with subsequent memory for sequencing than impression formation was confirmed by the presence of a significant two (orienting task)-by-two (subsequent memory) interaction \( (F_{(1,16)} = 8.25; p < 0.02) \). Finally, following the practice of previous researchers who frequently used more relaxed statistical thresholds when examining activations in the medial temporal lobe because of the relatively low signal/noise ratio observed in this region (Ojemann et al., 1997; Davauchi and Wagner, 2002; Strange et al., 2002), we also contrasted sequencing-hits > sequencing-misses at a slightly more lenient statistical threshold \( (p < 0.0025; \text{uncorrected}) \). This contrast yielded additional activations along the length of the right hippocampus as well as in several other medial temporal lobe regions bilaterally (Table 3; Fig. 3). None of these regions demonstrated a \( Dm \) effect for impression formation trials (all \( p \) values > 0.55).

Finally, even at a very relaxed statistical threshold \( (p < 0.10; \text{uncorrected}) \), we did not observe subsequent memory effects in the ventrolateral PFC for either orienting task, despite the ubiquity of \( Dm \) effects previously observed in this region (for review, see Paller and Wagner, 2002). The current study used two orienting tasks that differ substantially from the two-alternative semantic or phonological decisions or visual imagery tasks used in most extant \( Dm \) research, and we speculate that the differences between the processing demands of the current encoding tasks and those used in previous studies may have contributed to the observation that only the dorsomedial PFC and hippocampus, and not the ventrolateral PFC, correlated with subsequent memory.

**Discussion**

**Implications for social cognition**

Psychological research on social cognition has generally sought an account of how general purpose mechanisms of categorization, memory, and attention are recruited for thinking about and perceiving other people. In interpreting the memory differences produced by social and nonsocial tasks, researchers have generally theorized that social processing simply provokes differential use of memory-related constructs introduced previously by cognitive psychologists, such as inter-item associations, schemas, and elaborative encoding (Hastie and Kumar, 1979; Hamilton et
Figure 2. Regions of the medial PFC that were relatively anterior and ventral generally demonstrated a significant Dm effect for impression formation trials, whereas regions of the medial PFC that were relatively posterior and dorsal (B) failed to differentiate between hits and misses. A, Mean hemodynamic time course for all voxels obtained from the contrast of impression formation > sequencing that were anterior to y = 45 and ventral to z = 45. B, Same data for voxels posterior and dorsal to these coordinates. In addition, a single brain region, located in the dorsomedial PFC, was obtained from the direct contrast of impression-hits > impression-misses (C). No Dm effects were observed for sequencing items in any medial PFC region. Each panel displays hemodynamic time courses for each trial type as a function of memory: impression-hits (solid closed circles), impression-misses (dashed open circles), sequencing-hits (solid closed triangles), sequencing-misses (dashed open triangles). Within each panel, the voxels comprising each ROI are overlaid on the anterior portion of a sagittal (x = 6) slice of the subjects’ mean normalized brain.

Table 2. Peak voxel and number of voxels for ROIs obtained from the contrast of sequencing > impression formation (p < 0.05; corrected)

<table>
<thead>
<tr>
<th>Anatomical label</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxel(s)</th>
<th>Maximum t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior frontal gyrus</td>
<td>24</td>
<td>3</td>
<td>54</td>
<td>39</td>
<td>7.72</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>18</td>
<td>54</td>
<td>14</td>
<td>4.03</td>
</tr>
<tr>
<td></td>
<td>-21</td>
<td>6</td>
<td>57</td>
<td>61</td>
<td>6.51</td>
</tr>
<tr>
<td>Superior parietal gyrus</td>
<td>-6</td>
<td>-69</td>
<td>54</td>
<td>34</td>
<td>5.71</td>
</tr>
<tr>
<td>Head of caudate</td>
<td>12</td>
<td>9</td>
<td>-6</td>
<td>13</td>
<td>5.10</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>18</td>
<td>21</td>
<td>-6</td>
<td>13</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>-39</td>
<td>54</td>
<td>12</td>
<td>4.65</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>-48</td>
<td>51</td>
<td>12</td>
<td>4.17</td>
</tr>
</tbody>
</table>

al., 1980, 1989; Wyer et al., 1984; Srull and Wyer, 1989; but see Macrae et al., 2004).

However, the current data suggest the need to reconsider this approach. Here, we demonstrated that a task typically used to invoke social processing, impression formation, engaged brain regions that were neuroanatomically distinct from a nonsocial task that shared many of the same formal processing demands (e.g., integrating faces and statements into an associative memory representation). This neural dissociation suggests corresponding differences in the cognitive processes underlying the social and nonsocial tasks and extends previous suggestions that social cognition relies on a consistent, yet distinct, set of mental processes subserved in part by the medial PFC (Mitchell et al., 2002; Mason et al., 2004).

Moreover, this neural dissociation did not passively reflect differences between the two tasks but also appeared to be functionally significant for successful memory formation. When participants encoded information in a socially relevant manner (impression formation), activation in one brain region, dorsomedial PFC, correlated with later memory performance, but when participants encoded information less socially (sequencing), memory was correlated with activation in a different region, right hippocampus. This subsequent memory dissociation directly challenges previous theoretical assumptions that patterns of memory performance after social processing result simply from “deeper” or selective use of the same encoding operations engaged by nonsocial processing. Rather, the act of processing information in a socially relevant manner may engage specialized social–cognitive processes, which in turn may produce the unique patterns of memory performance (e.g., better memory, clustered recall, etc.) that have been reported after impression formation (Hastie and Kumar, 1979; Hamilton et al., 1980, 1989; Wyer et al., 1984; Srull and Wyer, 1989).

Activity in the medial PFC plays a critical role in understanding the minds of other people, the precise contribution of this region to social cognition remains relatively unspecified (Frith and Frith, 1999; Gallagher and Frith, 2003). One intriguing possibility is suggested by recent neuroimaging research demonstrating that regions of the medial PFC are also engaged during self-referential processing (Gusnard et al., 2001; Kelley et al., 2002; Macrae et al., 2004). For instance, Gusnard et al. (2001) reported a region of the dorsomedial PFC, nearly identical to the one observed in the present study, that was differentially activated during a self-referential task. That the processes subserving knowledge of oneself may share the same neural basis as those subserving understanding of others suggests the plausibility of so-called “simulation” accounts of social cognition (Humphrey, 1984, 1986), which posit that much of our knowledge about the mental states of other people may derive from first-hand consideration of our own internal experience (for review, see Nickerson, 1999; Meltzoff and Brooks, 2001).

Resulting in part from the relative dearth of animal models and appropriate neuropsychological patients, few organizing principles have emerged regarding the functional neuroanatomy of the medial PFC. Additional research is needed to resolve at least two issues regarding medial PFC organization. First, it is currently unclear to what extent functionally discrete subregions exist within the medial PFC and for what cognitive processes these subregions might be specialized. One possibility is suggested by the functional organization of neighboring anterior
cingulate regions, which have been segregated into a dorsal region that is engaged during demanding cognitive tasks (e.g., the Stroop color-naming task) and a more ventral region that is engaged during affective tasks (e.g., the emotional Stroop task) (Bush et al., 2000). Although other researchers (Gusnard et al., 2001) have suggested that a similar cognitive–affective dimension might be respected along the dorsoventral axis of the medial PFC and various bits of empirical evidence support such a distinction (Gusnard and Raichle, 2001), this hypothesis has yet to be examined directly. Second, developing theories must also account for the frequent observation in some medial PFC regions of “deactivations” relative to resting baseline, including the observation that some areas of the medial PFC, notably the dorsal extent, have alternately yielded both activations above and deactivations below baseline across studies (Gusnard et al., 2001; Gusnard and Raichle, 2001). Indeed, although the dorsomedial PFC region observed in the current study was associated with activations above baseline, previous work has observed deactivations in a very similar region (Mitchell et al., 2002). It is as yet unclear what cognitive processes characterize such functional differences between signal increases and decreases in the medial PFC.

**Implications for memory**

By demonstrating that different brain regions give rise to subsequent memory as a function of the manner in which material is originally processed, the present study contributes to a small number of empirical demonstrations suggesting a neural basis for the phenomenon of encoding specificity (Otten and Rugg, 2001a; Otten et al., 2002; Fletcher et al., 2003). For example, Otten et al., (2002) demonstrated that activity in distinct brain regions correlated with subsequent memory as a function of whether words were encoded as part of an animacy (does the word refer to the property of a living entity) or syllable-counting (odd or even number of syllables) task. In line with this previous work, the results of the present study provide evidence of a subsequent memory dissociation when two different tasks are used to encode the same stimulus materials. When participants were oriented toward the socially relevant aspects of a stimulus, strength of activation in the medial PFC correlated with later memory performance. Previous researchers have also reported a region of the dorsomedial PFC that predicted subsequent recognition after a semantic (but not nonsemantic) encoding task (Otten et al., 2001; Otten and Rugg, 2001a). Interestingly, the particular semantic task used in these studies required participants to judge whether or not a word referred to a property of living entities, a task that may also prompt the engagement of social–cognitive processing.

In contrast, when participants attempted to encode the temporal order in which statements appeared, a task that requires minimal social–cognitive processing, the strength of activation in the right hippocampus correlated with subsequent memory. This hippocampal Dm effect was obtained despite significantly lower overall memory for items in the sequencing condition and is consistent with observations that suggest a key role for the hippocampus in memory for sequences (Agster et al., 2002; For- tin et al., 2002). It is also possible that, because the sequencing task prompted participants to encode statements in relation to one another, this hippocampal effect may reflect the episodic retrieval of earlier statements during sequencing trials. Consistent with this possibility, previous research has demonstrated that medial temporal lobe activity during a retrieval task may correlate with subsequent memory as indexed by a second retrieval test (Stark and Okado, 2003).

When first introduced to cognitive psychology (Tulving and Thomson, 1973; Tulving, 1979), the encoding specificity principle was used to demonstrate that episodic retrieval should not be considered the simple product of the “strength” of a stored memory trace. Rather, Tulving and Thomson (1973) pointed out that varying encoding contexts could establish fundamentally different mnemonic representations, for which a correspondingly different constellation of test cues would be needed for successful retrieval. Recent neuroimaging investigations of encoding specificity (Otten and Rugg, 2001a; Otten et al., 2002; Fletcher et al., 2003) have sought to make a parallel point regarding the neural basis of memory: successful encoding should not depend on the strength of processing in a fixed set of brain regions, but rather the processing demands of the encoding context should critically determine which brain regions will be functionally significant for successful subsequent memory. The present study provides evidence for such a neural dissociation. Tests of other aspects of the encoding specificity principle, most notably the prediction that patterns of Dm effects should also differ as a function of the nature of the retrieval task, await future research.

**References**


Bozeat S, Gregory CA, Ralph MA, Hodges JR (2000) Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of


