Person-specific Theory of Mind in Medial pFC

B. Locke Welborn and Matthew D. Lieberman

Abstract

■ Although research on theory of mind has strongly implicated the dorsomedial pFC (incuding medial BA 8 and BA 9), the unique contributions of medial pFC (MPFC; corresponding to medial BA 10) to mentalizing remain uncertain. The extant literature has considered the possibility that these regions may be specialized for self-related cognition or for reasoning about close others, but evidence for both accounts has been inconclusive. We propose a novel theoretical framework: MPFC selectively implements "person-specific theories of mind" (ToM_p) representing the unique, idiosyncratic traits or attributes of well-known individuals. To test this hypothesis, we used fMRI to assess MPFC responses in Democratic and Republican participants as they evaluated more or less subjectively well-known political figures. Consistent with the ToM_p account, MPFC showed greater activity to subjectively well-known targets, irrespective of participants' reported feelings of closeness or similarity. MPFC also demonstrated greater activity on trials in which targets (whether politicians or oneself) were judged to be relatively idiosyncratic, making a generic theory of mind inapplicable. These results suggest that MPFC may supplement the generic theory of mind process, with which dorsomedial pFC has been associated, by contributing mentalizing capacities tuned to individuated representations of specific well-known others. ■

INTRODUCTION

The capacity to make sense of the mental states and traits of others is an extraordinary ability that allows us to plot a course through the complexities of social life. Countless studies have identified regions of the medial pFC (MPFC) involved in thinking about the mental states of others and of ourselves (Amodio & Frith, 2006). Within this region, an asymmetry has been observed (Lieberman, 2010; Van Overwalle, 2009), such that dorsomedial pFC (DMPFC; medial aspects of BA 8/BA 9) is more often found in studies of mentalizing (i.e., thinking about the mental states and traits of others), whereas a more ventral area of MPFC (medial aspect of BA 10) is more often found in studies of self-reflection (i.e., thinking about one's own states and traits; Denny, Kober, Wager, & Ochsner, 2012). A number of recent investigations have attempted to clarify the ways in which MPFC might also contribute to mentalizing.

Early neuroimaging studies of trait self-knowledge, a form of self-reflection, typically included famous targets as controls (e.g., George Bush), and more often than not, the comparison of self to famous targets yielded activity in MPFC (Kelley et al., 2002). A few studies also included a well-known close other as a control, and unlike the famous targets, the close others often produced MPFC activity similar to that of self-reference (Vanderwal, Hunyadi, Grupe, Connors, & Schultz, 2008; Ochsner et al., 2005; cf. Heatherton et al., 2006). Research from Mitchell, Banaji, and Macrae (2005) provides a potential account of these results in terms of the perceived similarity of close others. If close others are perceived to be similar to the self, then reflecting on one's own reactions to a query and projecting this on to the other person would be an efficient strategy for estimating the other's reactions.

Consistent with this hypothesis, activity in MPFC during mentalizing judgments of a social target has been shown to vary parametrically with perceived similarity to the self (Mitchell et al., 2005). Social targets with a political orientation similar to the self (liberal/conservative) also elicit greater MPFC response during judgments of preferences than do politically dissimilar others (Mitchell, Macrae, & Banaji, 2006). By this account, the essential function of MPFC is self-knowledge, but this self-knowledge can be used strategically to make inferences about similar others.

A second account suggests that the closeness of others to oneself is the key factor driving MPFC activity. More specifically, Krienen, Tu, and Buckner (2010) have suggested that MPFC is primarily sensitive to the social relevance to or social distance from oneself, signaling friendship and kinship affiliation rather than abstract similarity. In their research, MPFC was consistently more responsive to real-world friends than to unknown strangers, even when those strangers were judged to be more similar to the self than comparable friends. Thus, similarity and kinship accounts of MPFC contributions to mentalizing have both garnered a fair amount of empirical support.

In the current research, we propose a novel characterization of MPFC's contribution to mentalizing that would

University of California, Los Angeles

Journal of Cognitive Neuroscience 27:1, pp. 1–12
doi:10.1162/jocn_a_00700

simultaneously account for existing data on self-reflection, similarity, and kinship. DMPFC has often been characterized as supporting our theory of mind (Wimmer & Perner, 1983), a generic model of how minds react to various situations and experiences. This generic theory of mind (ToM_g) can be applied to anyone in a real or imagined context—allowing us, for example, to confidently predict how a typical male adult with a gun to his head would respond to a request to express his undying love of Justin Bieber's music (compared with when the gun is absent from the same scene). Nearly all neuroimaging studies of theory of mind have focused on strangers or imaginary characters for which a ToM_g is sensible to apply.

However, in our daily life, we interact with friends, family, and coworkers repeatedly and often learn that their distinctive personalities mean that our ToM_g does not always apply. Instead, we may generate a person-specific theory of mind (ToM_p) that is tailored to a particular individual. Whereas there is a considerable body of evidence showing that DMPFC supports generic mentalizing (ToM_g), we hypothesize that MPFC supports ToM_p and aimed to test this notion empirically. More specifically, we predicted that MPFC would be more active for social targets about whom we have extensive knowledge, particularly when this knowledge is both idiosyncratic to that target and relevant to a judgment to be made.

The ToM_p hypothesis can account for results associated with the existing theoretical perspectives discussed above. Close relationships with well-known others naturally furnish us with a diverse array of interpersonal experiences, from which we can generate a unique ToM_p. In the case of similar others, we may draw upon an especially rich ToM_p, that of the self. Thus, a ToM_p account of MPFC function is congruent with the findings associated with the similarity and closeness approaches. This approach also suggests that representations of the self are not qualitatively distinct from other person-specific representations, but rather are the most well-developed exemplars of the kinds of social representations handled by MPFC more broadly.

The ToM_p hypothesis can also be distinguished from these other approaches on empirical grounds. Only the ToM_p hypothesis predicts that MPFC will be recruited when thinking about individuals about whom we have a great deal of idiosyncratic knowledge but who are also disliked and perceived as dissimilar from us. To test this hypothesis, we asked individuals with a strong political affiliation (Democrat/Republican) to make trait judgments about four political figures, two in the participant's own political party (Own Party) and two in the opposing party (Opposition Party). We assumed that the political targets in the participant's own party would be seen as both more similar and closer to oneself than the political targets in the opposing party. Critically, we also manipulated the political targets used such that the amount of prior knowledge about them varied (see Figure 1). Within each political party, participants nominated one political figure about whom they knew a considerable amount (high-knowledge targets) and one political figure about whom they knew relatively little (low-knowledge targets).

We hypothesized that regardless of political affiliation, participants would generate greater MPFC activity to well-known political targets, because of the recruitment of a ToM_p, than to the less well-known political targets, for which only the ToM_g would be available. We also hypothesized that, for each trait judgment, the extent to which the political target is judged to be distinctive from both the self and the typical person reflects the likelihood that a ToM_p is being applied to a particular trait, because neither projection from the self or a ToM_g would be

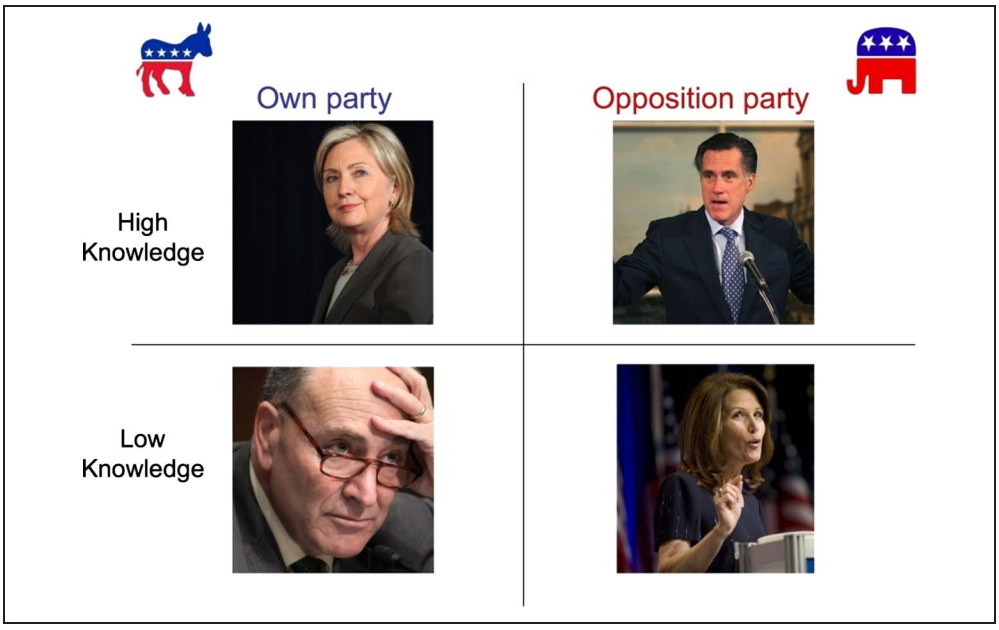

Figure 1. Example targets for a hypothetical Democratic participant, reflecting an experimental crossing of the Knowledge and Party Affiliation factors. Targets were selected independently for each participant, based on self-reported knowledge of Own Party and Opposition Party political figures. See Table 1 for descriptive data related to each target.

applicable. Thus, MPFC should be more active during judgments of more idiosyncratic traits for a given target. Finally, we also hypothesized that MPFC during the retrieval of self-knowledge reflects the use of one ToM_p among several, rather than a "self" mechanism, per se. Consequently, we predicted that MPFC would be more active during self-judgments on traits for which the self is judged to be idiosyncratic (i.e., distinct from the typical person).

METHODS

Participants

Sixteen participants (8 women) were recruited by a combination of e-mail solicitations and in-person presentations at undergraduate Democratic and Republican clubs at University of California, Los Angeles (UCLA). All participants

(a) indicated strong affiliation with either the Democratic or Republican party by selecting either 0–2 or 8–10, respectively, on an 11-point Likert scale anchored on either end at either "Strongly Republican" and "Strongly Democratic";
(b) indicated that they considered themselves to be at least moderately knowledgeable regarding current American politics (>5 on a 9-point Likert scale anchored at 1 = *not at all knowledgeable*, 5 = *moderately knowledgeable*, and 9 = *extremely knowledgeable*); and
(c) met target selection criteria detailed below.

Participants were judged ineligible for participation if they did not meet the above criteria. In addition, participants were ineligible if they were left-handed, using psychoactive medications or drugs, had been diagnosed with a neurological or psychiatric disorder, were pregnant, had a history of claustrophobia, or presented any other condition that would render participation in fMRI research hazardous.

Participants were all young adults between 18 and 29 years (M = 22.1, SD = 3.2). All participants were compensated $40 for their contribution to this research. Participants provided written informed consent approved by the UCLA Institutional Review Board. One participant's data are not included in these analyses because of partial data acquisition failure.

Target Selection

Four political figures (one high-knowledge target and one low-knowledge target from each party) were selected ideographically for each participant based on a screening questionnaire that queried participants about each of 50 contemporary politicians. All potential targets included in the screening questionnaire were active political figures who presently or formerly (<5 years prior) served in one more of the following offices: President of the United States, Senator, Congressional Representative, or State Governor. No targets were included that might be well known for reasons unrelated to politics and governance (e.g., Arnold Schwarzenegger, Jesse Ventura, or Al Franken) or who had been associated with highly publicized scandals or controversies (e.g., Bill Clinton or Larry Craig).

For each screening target, participants indicated their degree of knowledge and liking on 9-point Likert scales and reported the target's party affiliation. Responses on both scales were used to select four targets for the scanning session (one high-knowledge target and one low-knowledge target from each party) subject to the following conditions:

(1) Each participant's high-knowledge targets were rated between *Very knowledgeable* and *Extremely knowledgeable* (knowledge scale values of 7–9).
(2) Each participant's low-knowledge targets were rated between *Slightly knowledgeable* and *Moderately knowledgeable* (knowledge scale values of 2–5) and at least 4 points lower than both high-knowledge targets.
(3) Each participant's own party candidates were liked (liking scale values of >5). Opposition party candidates were disliked (liking scale values of <5).
(4) All targets' party affiliations were correctly identified.

These criteria accomplished the following objectives: (a) ensure appropriate like/dislike attitudes toward own party and opposite party targets, (b) provide targets with desired variation in knowledge for the political targets, and (c) accommodate individual differences in the use of screening scales (see Figure 1 for example targets for a hypothetical Democratic participant). Participants were excluded if appropriate targets could not be identified.

This individualized target selection procedure subjected each participant to a 2 (target politician knowledge, High vs. Low) × 2 (target politician party affiliation, Own Party vs. Opposition Party) within-subject factorial design, yielding four cells (see Figure 1). As detailed below, participants judged the applicability of trait words to each political figure while undergoing fMRI. Planned comparisons between hemodynamic activity associated with assessments of target politicians in different experimental conditions (e.g., High Knowledge > Low Knowledge, Own Party > Opposition Party) allowed for a direct test of our primary hypothesis regarding the function of MPFC.

Behavioral Measures

Given that self-reports of knowledge assessments may be biased and/or self-serving, participants were asked (after scanning) to write essays demonstrating their political knowledge of each target. Ratings of these essays by independent evaluators blind to experimental hypotheses were used as an alternative (unbiased) measure of target knowledge.

In their essays, participants were instructed to describe the target's activities in politics and government, stances on contemporary political issues, and important

accomplishments. These essays were rated by trained evaluators (research assistants) who were blind to the participant's party affiliation and the experimental condition of the target political figure. Evaluators were instructed to judge how knowledgeable the participant was about the target in the domain of politics and government, using a 7-point Likert scale anchored at *not at all knowledgeable* and *extremely knowledgeable*. Essay-rated knowledge scores reflecting raters' independent judgments of participants' target-specific knowledge were thus available for all four political figure targets. Evaluators did not possess any special political expertise and were not affiliated with either political party. Parametric modulation of activity in MPFC by essay-rated target knowledge is employed as a complementary test of the ToM_P hypothesis. Lastly, participants completed self-report measures indicating their degree of overall similarity, closeness, and connectedness to each of the four political targets selected for the scanner task (range 0–10, inclusive). Participants also completed a measure of perceived personal overlap with each of the political targets (range 1–7, inclusive).

fMRI Paradigms

While undergoing fMRI, participants completed a trait judgment task (Politician Judgment Task) in which they rated the applicability of 30 personality traits to each of the four target political figures over three functional runs. In two separate functional runs, participants rated the applicability of these 30 traits to the self as well as to the ordinary American (Self/Ordinary American Judgment Task). A case judgment control task (uppercase/lowercase) using the same trait words and phrases was included in all runs as an experimental control.

Trait words and phrases were selected for the above tasks from a larger list on the basis of pilot testing within the UCLA undergraduate population, ensuring that items were comprehensible and meaningful to participants. All trait items were relevant to the domain of politics and governance (e.g., "patriotic," "able to take command," "opportunistic"). A mixed block/event-related design was employed to best explore hypotheses concerning both target level and trial level effects, with trial events grouped into superordinate blocks based on target identity. During each trait judgment trial, participants used an on-screen 9-point Likert-type scale to indicate the applicability of the specified trait to the current target.

Trait judgments of political figures were spread over three functional runs, with each run containing eight experimental blocks and two control blocks. Block order within and between runs was randomized for each participant, with the following constraints: (1) the same target was never selected for two consecutive blocks, (2) each political figure appeared at least once in each run, (3) no political figure appeared more than twice in any single run, and (4) no more than two blocks featuring targets from the same party were presented consecutively.

Trait judgments regarding the self and the ordinary American were spread over two functional runs, with each run containing six experimental blocks (3 Self and 3 Ordinary American) and two control blocks. Block order for Self/Ordinary American judgments was similarly randomized, with no two blocks of the same type occurring consecutively. In total, participants completed six blocks of trait judgment for each political figure, for the self and for the ordinary American, and 10 control blocks.

Each block consisted of a 2-sec introduction specifying the target, followed by five trait judgment trials, and concluding with a 5-sec rest period between blocks. During each trait judgment trial, participants used an on-screen 9-point Likert-type scale to indicate the applicability of the specified trait to the current target. Trait words and phrases appeared on-screen for 5 sec, during which participants moved the scale to the appropriate response value and confirmed their selection. Scale movement and stimulus duration were determined on the basis of pilot testing, such that participants could comfortably make their judgments, move the on-screen scale, and confirm their responses in the allotted time. Trait words and phrases, as well as the scale selection indicator, were removed from the screen following response selection. Each trial was followed by a jittered ISI drawn from an exponential random distribution with a mean of 2 sec. Stimulus presentation was identical for Self/Ordinary American runs, except that the targets of judgment differed. During the control task, the same 30 trait words were presented in either all uppercase or all lowercase letters, and participants were required to make a binary case judgment. This task was designed to control for basic perceptual and motor processing associated with the use of the on-screen scale, as well as spontaneous lexicosemantic processing unrelated to the trait judgments themselves.

fMRI Data Acquisition

All imaging data were acquired using a 3.0-T Siemens Trio scanner at the Ahmanson-Lovelace Brain Mapping Center at UCLA. Across five functional runs, 761 T2*-weighted echo-planar images were acquired during completion of experimental tasks described above (slice thickness = 3 mm, gap = 1 mm, 36 slices, repetition time [TR] = 2000 msec, echo time [TE] = 25 msec, flip angle = 90°, matrix = 64 × 64, field of view = 200 mm). An oblique slice angle was used to minimize signal dropout in ventral medial portions of the brain. In addition, a T2-weighted, matched-bandwidth anatomical scan was acquired for each participant (TR = 5000 msec, TE = 34 msec, flip angle = 90°, matrix = 128 × 128; otherwise identical to EPIs). Lastly, we acquired a T1-weighted magnetically prepared rapid acquisition gradient-echo anatomical image (slice thickness = 1 mm, 176 slices, TR = 2530 msec, TE = 3.31 msec, flip angle = 7°, matrix = 256 × 256, field of view = 256 mm).

fMRI Data Preprocessing and Analysis

Preprocessing and ROI Definition

Functional data were analyzed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK). Within each functional run, image volumes were corrected for slice acquisition timing, realigned to correct for head motion, segmented by tissue type, and normalized into standard MNI stereotactic space (resampled at 3 × 3 × 3 mm) using a diffeomorphic registration algorithm (Ashburner, 2007). Finally, images were smoothed with an 8-mm Gaussian kernel, FWHM. Given our specific hypotheses regarding the role of MPFC in deploying ToM_p, all principal analyses were conducted on a predefined MPFC ROI. This ROI was constructed using the Automated Anatomical Labeling toolbox (Tzourio-Mazoyer et al., 2002) of the Wakeforest University Pickatlas (Maldjian, Laurienti, Burdette, & Kraft, 2003), encompassing BA 10. This base region was dilated and constrained to the medial aspect ($-15 < x < 15$). All ROI analyses reported below interrogate only voxels within this region.

Two analytical strategies were employed to assess the role of MPFC in ToM_p. First, we used selective averaging for hypothesis testing on the ROI considered as a whole functional unit. Given the relatively large volume of the MPFC, we also use an alternative exploratory approach, searching within the same ROI for significant clusters and correcting for multiple comparisons within that search space. Monte Carlo simulations implemented in 3dClustSim (from AFNI; Cox, 1996) were used to determine appropriate cluster size thresholds given the smoothness of the images (16 contiguous voxels) to ensure overall false discovery rate (FDR) of less than 0.05, when combined with a voxel-wise significance threshold of $p < .005$. All results reported exceed these joint voxel-wise and cluster-extent thresholds.

Target Level Analyses

For both the Politician Judgment Task and the Self/Ordinary American Judgment Task, we defined a general linear model (GLM) for each participant. Blocks were modeled as variable epochs spanning the duration between the onset of the first trial and the offset of the final trial, convolved with the canonical (double-gamma) HR. For the Politician Judgment Task, five regressors of interest were modeled (High Knowledge, Own Party; High Knowledge, Opposition Party; Low Knowledge, Own Party; Low Knowledge, Opposition Party; Case Judgment Control). For the Self/Ordinary American Judgment Task, three regressors of interest were modeled (Self; Ordinary American; Case Judgment Control). All analyses controlled for 18 motion parameters (3 translations and rotations, as well as their squares and first-order derivatives). The time series was high-pass filtered using a cutoff period of 128 sec, and serial autocorrelations were modeled as an AR(1) process. Contrast images were averaged across runs for each participant and entered into a mixed effects analysis at the group level. Parameter estimates were extracted from the MPFC ROI using MarsBaR (Brett, Anton, Valabregue, & Poline, 2002) and subjected to a repeated-measures ANOVA. In the results presented below and the accompanying figures, parameter estimates reflect contrasts between the appropriate experimental condition and the case judgment control task, unless otherwise noted.

Target level factors (overall similarity, closeness, connectedness, personal overlap, scored knowledge, mean similarity, mean idiosyncrasy, and mean positivity) not subject to experimental manipulation were analyzed as parametric modulators of activity in the MPFC ROI. The effects of each factor (when appropriate, controlling for other factors, see below) were assessed based on the appropriate parameter estimates from a GLM, identical to the above except for the different regressors of interest (that is, regressors reflected continuous scores on target level variables, rather than target identity).

Trial-by-trial Trait Level Analyses

The effects of trial-to-trial trait level perceived similarity, idiosyncrasy, and positivity were assessed using analyses of parametric modulation of the hemodynamic response to each of three trial level predictors (similarity, idiosyncrasy, and positivity) implemented using an event-related GLM. Similarity was operationalized as the absolute value of the difference between self-judgments and political target judgments on a given trait, with small values indicating high similarity. Idiosyncrasy was computed as the absolute value of the difference between judgments of the political target and judgments of the ordinary American, with large values indicating high idiosyncrasy. Judgment positivity simply reflected the participant's rating of a given political target on a specific trait (reverse-coded for trait words or phrases judged by a pilot sample to be negative). For self-judgments, we also computed an index analogous to target idiosyncrasy, reflecting the distinctiveness of the self on a given trait, relative to the ordinary American. This self-idiosyncrasy was operationalized as the absolute value of the difference between judgments of the self and judgments of the ordinary American on a particular trait, with large values indicating high self-idiosyncrasy. The GLMs used to assess the effects of trait level factors differed from those used to assess target level factors in that specific trials (rather than the superordinate block) were modeled as discrete events, using a variable epoch spanning the duration from trial onset to response.

RESULTS

Behavioral Results

Eight participants identified as strong members of the Democratic Party, and seven identified as strong

members of the Republican Party. See Table 1 for a summary of descriptive data.

Target Differences in Self-reported Knowledge, Closeness, Connectedness, Personal Overlap, and Overall Similarity

Participants indicated that they possessed greater knowledge of High-knowledge targets than Low-knowledge targets (M_{high} = 7.90, M_{low} = 3.43, t = 19.31, p < .001). This result is direct consequence of the target selection procedures employed (see Methods above for details). Self-reported knowledge did not differ significantly between Own Party and Opposition Party targets (M_{Own} = 5.77, M_{Opp} = 5.57, t = 1.71, *ns*).

Participants judged themselves to be more similar (M_{Own} = 5.44, M_{Opp} = .66, t = 16.51, p < .001), closer (M_{Own} = 4.61, M_{Opp} = .14, t = 10.67, p < .001), and more connected (M_{Own} = 5.28, M_{Opp} = .25, t = 17.09, p < .001) to Own Party than to Opposition Party targets. In addition, participants perceived greater personal overlap with Own Party relative to Opposition Party targets (M_{Own} = 5.00, M_{Opp} = 1.50, t = 14.34, p < .001). Participants did not see themselves as more similar to High-knowledge targets than to Low-knowledge targets and did not perceive greater overlap (all *p*s *ns*). However, participants did feel closer (M_{high} = 3.00, M_{low} = 1.75, t = 5.49, p < .001) and more connected (M_{high} = 3.32, M_{low} = 2.21, t = 3.60, p = .003) to high-knowledge targets than to low-knowledge targets. For this reason, relevant target level analyses reported below control for closeness and connectedness to target.

Target Differences in Knowledge, Assessed by Individual Essays

Participants' essays regarding each political figure were coded by three research assistants who were blind to the target's experimental condition. The Spearman–Brown reliability of these ratings across raters is 0.75, suggesting that they represent a reasonably reliable measure of target knowledge. Consistent with self-reports, participants demonstrated greater essay-rated knowledge of High-knowledge relative to Low-knowledge targets (M_{high} = 4.08, M_{low} = 3.31, t = 6.55, p < .001). In contrast, Own Party and Opposition Party targets did not differ in essay-rated knowledge (M_{Own} = 4.13, M_{Opp} = 3.92, t = 1.00, *ns*).

Target Differences in Trait Level Idiosyncrasy, Similarity, and Positivity

On the basis of participants' trait ratings of political figure targets during the scanner session, aggregate indices of target idiosyncrasy, positivity, and similarity were computed (see Methods). On average, participants judged High-knowledge targets to be more idiosyncratic than Low-knowledge targets (M_{high} = 2.33, M_{low} = 1.77, t = 5.15, p < .001), but not more positive (M_{high} = 0.69,

Table 1. Descriptive Behavioral Data Summarized for Each of Four Political Figure Targets

	High Knowledge		*Low Knowledge*	
	Own Party	*Opposition Party*	*Own Party*	*Opposition Party*
Self-Report				
Knowledge (0–10)	8.00	7.80	3.53	3.33
Closeness (0–10)	5.93	0.07	3.29	0.21
Connectedness (0–10)	6.50	0.14	4.07	0.36
Overall similarity (0–10)	5.86	0.39	5.02	0.93
Personal overlap (1–7)	5.29	1.43	4.71	1.57
Rated				
Knowledge (1–7)	4.86	4.62	3.40	3.21
Trial-based				
Idiosyncrasy (0–8)	2.41	2.26	1.83	1.71
Similarity (0–8)	6.57	4.95	6.77	5.44
Positivity (−4 to 4)	2.17	−0.79	1.48	−0.31

Self-report items reflect participant responses to questionnaire items. Scored knowledge data are derived from coder ratings of individual essays on the accomplishments, positions, values, and career of each political figure target. Trial-based items represent aggregate means of idiosyncrasy, similarity, and positivity across trait judgments completed during the scanning session (see Methods for details regarding these indices).

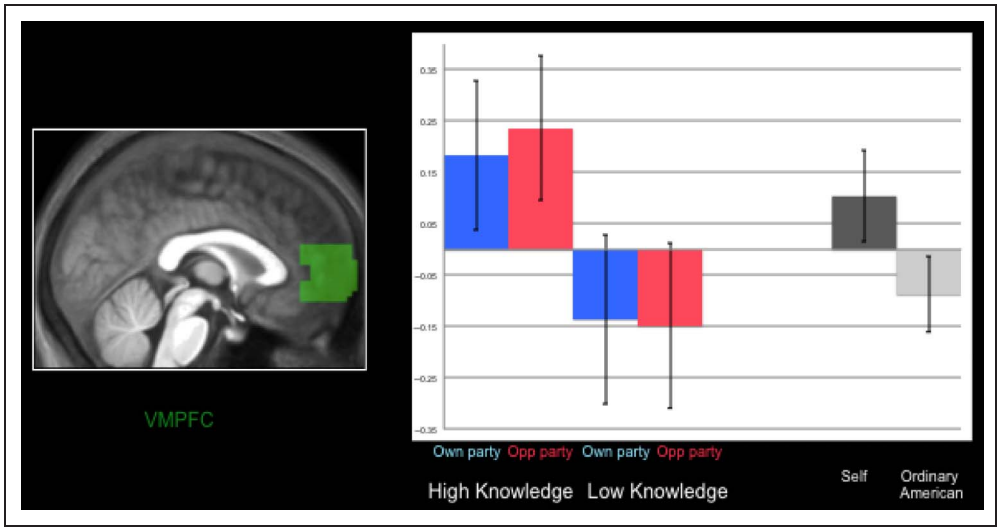

Figure 2. MPFC demonstrated greater activity to High-knowledge than to Low-knowledge targets. There was no significant effect of Party Affiliation on activity in the MPFC and no interaction between Party Affiliation and target Knowledge. MPFC also demonstrated greater activity to Self trials than to Ordinary American trials. Parameter estimates (relative to control) are plotted separately for each target. Error bars reflect the *SEM*. See also Table 2 for associated search-within analysis.

$M_{low} = 0.59$, $t = 0.64$, *ns*). Low-knowledge targets were judged to be more similar to the Self than High-knowledge targets ($M_{high} = 5.76$, $M_{low} = 6.11$, $t = 2.61$, $p = .02$). Own Party targets were not judged to be more idiosyncratic than Opposition Party targets ($M_{Own} = 2.12$, $M_{Opp} = 1.98$, $t = 0.62$, *ns*). However, as expected, on a trial-by-trial basis, participants judged Own Party targets to be more similar to the Self than Opposition Party targets ($M_{Own} = 6.67$, $M_{Opp} = 5.20$, $t = 5.97$, $p < .001$) and more positive overall ($M_{Own} = 1.83$, $M_{Opp} = -0.55$, $t = 8.58$, $p < .001$).

fMRI Results

Target Level Experimental Factors: Knowledge and Party Affiliation

Our primary hypothesis was that MPFC activity would be greater for High-knowledge targets than Low-knowledge targets, as the former would rely on ToM_p more than the latter. Additionally, we hypothesized that this effect would not be moderated by Own/Other Party status of the targets. Repeated-measures ANOVA indicated that MPFC was more responsive to High-knowledge than to Low-knowledge targets ($F(1, 13) = 16.799$, $p = .001$), but not more responsive to Own Party than to Opposition Party targets ($F(1, 13) = 0.025$, *ns*), as shown in Figure 2. (One participant was dropped from target level analyses. One ordered pair [essay-rated target knowledge, MPFC ROI parameter estimate] from this participant was identified as an extreme multivariate outlier. Cook's *D* for this case was 1.00, exceeding our threshold of 0.07.)

No interaction between target Knowledge and Party Affiliation was observed ($F(1, 13) = 0.130$, *ns*). Follow-up comparisons between specific cells showed greater activity to High-knowledge Own Party targets than to Low-knowledge Own Party targets ($t(13) = 2.705$, $p = .018$) and to High-knowledge Opposition Party targets than to Low-knowledge Opposition Party targets ($t(13) = 3.004$, $p = .010$). Each of these effects was also identified when we searched within the MPFC mask for significant clusters (see Table 2). In a whole-brain conjunction analysis (of High Knowledge Own Party > Low Knowledge Own Party ∩ High Knowledge Opposition Party > Low Knowledge Opposition party), only the MPFC demonstrated activity sensitive to target knowledge across both parties (peak MNI: −6, 53, 13; $t = 6.77$, $k = 76$).

Participants also showed greater activity in the MPFC to Self trials than to Ordinary American trials ($t(13) = 4.06$, $p = .001$), replicating self-reference effects observed in

Table 2. Summary of Search-within Analyses for the MPFC ROI

Test Effect	*x*	*y*	*z*	*t*	*k*
Target Level Analyses					
High > low knowledge (both parties)	−9	56	19	6.01	320
	−12	50	−11	5.08	
	−3	53	−2	4.73	
High > low knowledge (Own party)	6	59	13	4.17	59
	−6	53	−5	3.98	28
High > low knowledge (Opp party)	−3	56	13	6.7	165
Essay-rated target knowledge	−6	50	13	7.27	173
Trial-by-trial Trait Level Analyses					
Trait level political figure idiosyncrasy	−6	42	3	5.29	73
Trait level self-idiosyncrasy	−15	48	−12	4.45	16

All results are FDR-corrected $p < .05$ with combined voxel-wise p and cluster size thresholds (see Methods). Coordinates reported are from local maxima separated by at least 20 mm. *x*, *y*, and *z* = MNI coordinates in left–right, anterior–posterior, and inferior–superior dimensions; peak $t = t$ statistic value at the each local maxima; k = cluster voxel extent; cluster p(FWE) = cluster level FWE probability.

previous research (Denny et al., 2012; Heatherton et al., 2006; Kelley et al., 2002). The MPFC did not differentiate between the Self and High-knowledge Own Party targets ($t(13) = 0.718$, *ns*) or High-knowledge Opposition Party targets ($t(13) = 0.434$, *ns*). However, this region was significantly more active to Self trials than to Low-knowledge Own Party targets ($t(13) = 2.22$, $p = .0451$) and Low-knowledge Opposition Party targets ($t(13) = 2.490$, $p = .027$). This pattern of results is sensible if MPFC is responding to available knowledge concerning the target (including the Self) but is more difficult to explain if MPFC is encoding information concerning similarity or party affiliation.

Target Level Nonexperimental Factors

Consistent with our primary hypothesis, essay-rated target knowledge (i.e., participant knowledge of each target, defined by ratings of the essays the participant wrote about each target) demonstrated a significant linear relationship with activity within the MPFC ROI when controlling for closeness to the target ($t(13) = 3.339, p = .005$), but closeness did not significantly predict activity in this region ($t(13) = 1.584$, *ns*). Separate analyses show a consistent positive relationship between essay-rated target knowledge and activity in the MPFC ROI, controlling separately for connectedness to target ($t(13) = 3.157, p = .008$), perceived personal overlap ($t(13) = 4.244$, $p < .001$), and overall similarity ($t(13) = 3.472$, $p = .004$). In contrast, there was no relationship between MPFC activity and target level closeness ($t(13) = 1.584$, *ns*), connectedness ($t(13) = 1.192$, *ns*), overlap ($t(13) = 0.238$, *ns*), or overall similarity ($t(13) = 0.721$, *ns*). As can be seen in the plots of essay-rated target knowledge against parameter estimates from the MPFC ROI (target > control; Figure 3), MPFC activity is positively associated with essay-rated target knowledge for political figures from Own and Opposition Parties. Each of these effects was also identified when we searched within the MPFC ROI for significant clusters (see Table 2).

These results are not easily explained by similarity and closeness accounts of MPFC function. The MPFC showed sensitivity to within-subject variation in target knowledge (both as assessed using experimental contrasts and using independent, essay-rated target knowledge) that was not attenuated by controlling for self-reported closeness, connectedness, personal overlap, or overall similarity. In contrast, none of the other target level factors significantly predicted activity in this region.

Trial-by-trial Trait Level Analyses of Idiosyncrasy

Trial-by-trial analyses allow characterization of MPFC response to factors such as idiosyncrasy and similarity that vary from trait to trait both within and across targets (i.e., a participant may view the same candidate as idiosyncratic on some traits but not others). High idiosyncrasy trait judgments cannot depend on the deployment of ToM_g, as the target is perceived to differ from the ordinary American on the trait in question and therefore ought instead to depend on a ToM_p. To rule out other explanations of hemodynamic response associated with our trial-by-trial index of target idiosyncrasy, analyses reported below employ statistical controls for perceived similarity to self and RT. These analyses were run across all trials, ignoring the identity of the target and target level factors (e.g., closeness, party affiliation). The MPFC ROI demonstrated a significant linear relationship with trait level idiosyncrasy ($t(14) = 2.257, p = .040$), controlling for trait similarity, positivity, and RT, such that greater activity was observed for trials on which targets were judged to be more idiosyncratic. Similarity and positivity were each marginally significant as unique predictors of MPFC activity ($t(14) = 1.919, p = .076$ and $t(14) = 1.798, p = .094$, respectively), but were not significant controlling for idiosyncrasy.

Conjunction Analysis: Target Level Essay-rated Knowledge and Trait Level Idiosyncrasy

To examine the relationship between essay-rated target knowledge and trait level idiosyncrasy, we performed a conjunction analysis using the minimum statistic (Nichols,

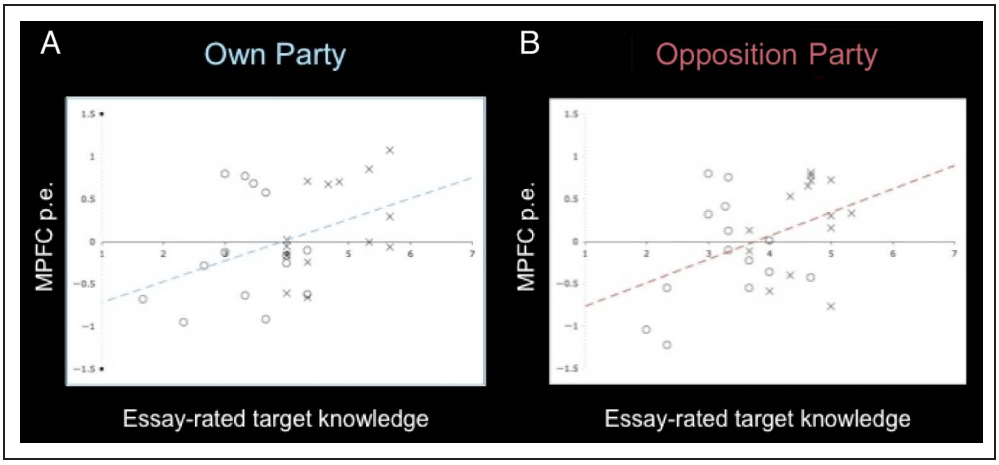

Figure 3. Activity within the MPFC ROI demonstrated a significant linear relationship with essay-rated target knowledge for both Own Party (A) and Opposition Party (B) political figures. To present this relationship graphically, we plot MPFC parameter estimates (p.e.) from the block target > control comparison against essay-rated target knowledge. "X" marks represent self-reported High-knowledge targets; "O" marks represent self-reported Low-knowledge targets. See also Table 2 for associated search-within analysis.

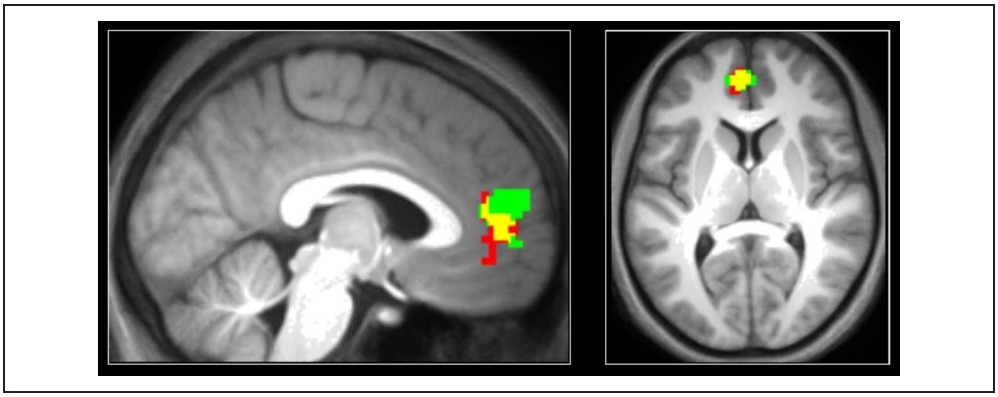

Figure 4. Essay-rated target knowledge and trait level idiosyncrasy demonstrate overlapping neural correlates. Results of a conjunction analysis constrained to the MPFC are displayed, showing activity associated with essay-rated target knowledge (green), trait-by-trait idiosyncrasy (red), and their conjunction (yellow). The conjunction cluster consisted of 41 voxels, with peak at MNI −6, 47, 7. See also Table 2 and Figure 5.

Brett, Andersson, Wager, & Poline, 2005) from these analyses (described above), constrained to the MPFC ROI. A significant cluster was detected within this region (peak MNI: −6, 47, 7; $t = 4.53$, $k = 41$; see Figure 4), suggesting that MPFC supports both person-specific mentalizing (ToM_p) as well as judgments of particularly idiosyncratic traits, regardless of overall target knowledge.

Self-idiosyncrasy

The ToM_p hypothesis of MPFC suggests that this region is not functionally devoted to self-processes per se. Rather, according to the hypothesis, the representation of oneself is typically the most idiosyncratic person representation one has, and thus, accessing self-knowledge recruits this region quite reliably. If this is the case, MPFC should be more active during self-judgments to the extent that a person views himself idiosyncratically on a particular trait. Contrary to expectations, a parametric modulation of self-judgments by trait level self-idiosyncrasy (controlling for RT) failed to produce a significant response in MPFC ROI as a whole ($t = 1.397$, *ns*). However, when we searched within the ROI, we did observe a cluster whose activity was significantly associated with self-idiosyncrasy (MNI: −15, 48, −12; $t = 4.45$, $k = 16$; see Figure 5).

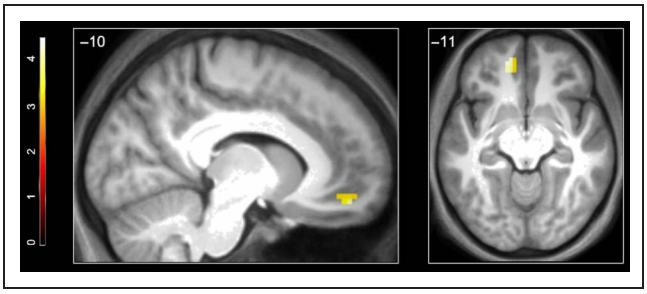

Figure 5. Trial-by-trial self-idiosyncrasy (difference from ordinary American) predicted activity in the VMPFC ROI (while controlling for judgment positivity) when assessing the applicability of trait words to the self. Peak MNI: −15, 48, −12, FDR-corrected $p < .05$.

DISCUSSION

A variety of hypotheses have been advanced to explain the unique contribution of MPFC processes to human mentalizing (Lieberman, 2012). Although each of the approaches adopted thus far are plausible, the results of the present research argue in favor of the ToM_p account of MPFC. This account is consistent with a wide range of prior results and integrates them under a more coherent umbrella account of MPFC's social cognitive functions. Crucially, the MPFC was more responsive to High-knowledge than to Low-knowledge targets, regardless of whether the target was seen as similar or close to oneself (i.e., in one's own political party or the opposing party). In more fine-grained analysis, MPFC was more active when judging a target on a particular trait to the extent that the target was seen as idiosyncratic on that dimension, differing from both the typical person and the participant making the judgment. Finally, MPFC was also more active when judging the self on a particular trait to the extent that the self was seen as more idiosyncratic on that trait, consistent with the view that MPFC is not a self-knowledge region per se. The self, on this view, is merely the paradigmatic instance of an intimately well-known social target, and self-relevant cognition is therefore extremely likely to be associated with MPFC response.

Prior similarity findings are therefore consistent with the ToM_p account of MPFC. Personal similarity may lead the social thinker to project her own attributes and preferences on to others and engage in mental simulation where appropriate when gauging similar others' mental states (Mitchell et al., 2005, 2006). According to the ToM_p account, this social reasoning strategy relies on the most person-specific representation one has: the self. However, it is just one of many person-specific mental models represented by MPFC, and in other cases, we may recruit this region when relying on another other non-self ToM_p. Unlike the similarity account, the ToM_p account also predicts that MPFC can be involved in projecting from individuals other than the self. If a new target is deemed similar to a friend or family member for whom we have ToM_p ("you remind me of my mother"), then MPFC

should be recruited as that ToM_p is projected on to the novel target.

Prior closeness findings (Krienen et al., 2010) are also consistent with the ToM_p account of MPFC, insofar as we typically represent the minds of close others with detailed, idiosyncratic models. However, only the ToM_p account can explain why MPFC is more active when thinking about targets who are well known but neither similar nor close to oneself, such as a well-known politician in an opposing political party. In addition, of the three accounts, only the ToM_p account also predicts that thinking about more idiosyncratic aspects of the self would differentially recruit MPFC.

Prior research has suggested that information about specific dispositional traits, which may serve as a scaffold for ToM_p, is encoded in the MPFC in the form of a trait code (Ma et al., 2014; Ma, Baetens, Vandekerckhove, Van der Cruyssen, & Van Overwalle, 2013; Ma, Vandekerckhove, Van Overwalle, Seurinck, & Fias, 2011). Rather than simply encoding the valence of a given judgment, this research implies that traits are represented in a discrete fashion within MPFC. The availability of trait-specific information is a necessary precursor for the formation of a ToM_p, which in addition requires the association of specific traits with specific targets. This study suggests that MPFC does in fact pair trait information with specific targets to produce consistent representations of the minds of specific individuals. Evidence concerning the existence of a trait code in the MPFC enhances the plausibility of the ToM_p account.

Against the interpretation of our present results as evidence for person-specific mentalizing processes in MPFC, one might argue that this region is simply responding to the greater affective salience or importance of the High-knowledge targets relative to the Low-knowledge targets. However, we do not believe that this account is consistent with the most recent work on the neural representation of subjective value. Bartra, McGuire, and Kable (2013) have recently considered these issues in a quantitative meta-analysis and find that DMPFC and anterior insula do tend to respond in a U-shaped fashion to subjective value, whether positive or negative. In contrast, the MPFC seems to respond preferentially to positive stimuli, rather than to general arousal or saliency. In light of this work, selective MPFC activity to High-knowledge targets (even when disliked) is unlikely to reflect processing of their affective salience.

ToM_p and Group Size

The ToM_p model of MPFC function is also consistent with findings inspired by Dunbar's (1998) social brain hypothesis, which links brain size across primate species with group size. Although the group effects are often conceptually linked to the mentalizing network, including DMPFC, most of the group size findings observed in human and primate neuroimaging instead implicate MPFC. For example, a number of recent studies show a linear relationship between MPFC volume and social network size (Powell, Lewis, Roberts, Garcia-Finana, & Dunbar, 2012; Lewis, Rezaie, Brown, Roberts, & Dunbar, 2011; Powell, Lewis, Dunbar, Garcia-Finana, & Roberts, 2010). Such results imply that MPFC may be essential for encoding information about multiple individuals in complex social arrangements. Successfully navigating a relatively large social network may depend on the ability to deploy a diverse repertoire of ToM_p, which may be facilitated by greater MPFC volume.

Most strikingly, Sallet et al. (2011) recently observed increased gray matter in MPFC when macaques were moved from smaller to larger living groups. If successful group living depended solely on generic forms of social cognition, it would scale easily to any group size because the same generic knowledge applies to all individuals. Instead, successful group living is partially about keeping track of idiosyncratic social information about each relevant individual and the distinctive relationships between individuals in the group. Thus, the canonical mentalizing system may support a basic capacity for group living, whereas MPFC may drive the size of the manageable group. Groups with complex social dynamics are characteristic of the human species, and indeed BA 10 within MPFC is one of the only frontal regions known to be disproportionately larger in humans than other primates (Semendeferi, Armstrong, Schleicher, Ziles, & Van Hoessen, 2001).

Why a Second System?

Evolutionary pressures may have expanded the volume of MPFC in primates and humans, supplying them with more robust social cognitive resources for living in increasingly complex groups. However, such considerations do not explain why person-specific and generic forms of mentalizing ought to be functionally and structurally differentiated in distinct subregions of the MPFC to begin with. Insight into this issue may be gained by consideration of analogous distinctions in the cognitive neuroscience of declarative memory.

The memory literature has long distinguished the roles of neocortical versus hippocampal systems in the formation, consolidation, and retrieval of memories over time (Squire, Stark, & Clark, 2004; Alvarez & Squire, 1994). Long-term, relatively permanent representations of a semantic nature are thought to depend primarily on the neocortex and to change only slowly over time. In contrast, the hippocampus serves a crucial but time-limited function in declarative memory, conjoining the distributed neocortical representations that together constitute the memory as a whole (Squire, 1992).

Although the existence of multiple memory systems and their interaction during learning has thus been well characterized, the cognitive utility of encoding memory in distinct, overlapping forms has not always been evident.

Using computational models, McClelland, McNaughton, and O'Reilly (1995) presented an intriguing account of the complementary contributions of hippocampal and neocortical memory to the formation of categorical knowledge. They found that the attempt to encode information about the category membership of novel exemplars in a single, unitary connectionist system produced "catastrophic interference." In brief, if modeled knowledge structures were too malleable, they were ineffective at forming long-term categorical representations and integrating information from different observations over time. Instead, they were disproportionately biased by the unique features of each newly presented exemplar. In contrast, more rigid knowledge structures were unable to evolve and incorporate novel information, because inflexible representations were insufficiently sensitive to new or atypical instances (e.g., learning that a penguin is a bird, despite phenotypic dissimilarity). McClelland et al. theorized that the hippocampal formation provides an intermediary system, capable of storing new information in a form that does not interfere destructively with established knowledge. Over time, the aggregation of relevant information then shapes enduring category representations in the more slowly adapting neocortex.

We propose that a mechanism of this type might help to explain the utility of different systems for generic and person-specific mentalizing processes. The social thinker cannot help but take for granted her extensive and sophisticated understanding of how the minds of others form beliefs, experience emotions, and make decisions. This wealth of information constitutes our generic theory of the human mind and, as such, must be applicable to a diversity of situations. Precisely because its content has to apply broadly to countless situations, its representational features ought to be more fixed. By abstracting from the idiosyncrasies of any particular individual or experience, our ToM_g retains an essential generalizability that facilitates the understanding of novel mentalizing episodes. ToM_g ought therefore to develop slowly and change only when many experiences considerably reshape our fundamental assumptions and expectations regarding human thought, feeling, and action. Consistent with this notion, prior research has suggested DMPFC in abstract mentalizing contexts in which the application of a ToM_g is appropriate (Baetens, Ma, Steen, & Van Overwalle, 2014; Spunt, Satpute, & Lieberman, 2011; Spunt, Falk, & Lieberman, 2010).

In contrast, a ToM_p would allow for the efficient and rapid encoding of unique, idiosyncratic information regarding particular, important individuals. Whether concerning friends or enemies, similar or dissimilar others, the use of ToM_p may allow more precise inferences concerning their mental states and more accurate predictions of their future behavior. Separate systems for person-specific and generic mentalizing would allow us to take advantage of idiosyncratic knowledge in thinking about those we know well, without resulting in destructive interference with our more general theory of mind.

Conclusions

The results reported above substantiate a novel account of MPFC function, according to which this region implements ToM_p and enables nuanced, adaptive responding to the idiosyncratic characteristics of particular, well-known individuals. ToM_p can account for the empirical findings associated with other approaches to mentalizing (e.g., similarity, closeness) but also provides new insights and testable predictions that go beyond these perspectives. Future research might elucidate the circumstances under which ToM_p and ToM_g are deployed and better characterize the psychological and neural correlates of the formation of person-specific theories. In addition, further work may help to uncover the relationship between ToM_p and other sources of mental state inference.

Acknowledgments

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship (grant DGE-1144087).

Reprint requests should be sent to Matthew D. Lieberman, Department of Psychology, 4611 Franz Hall, University of California, Los Angeles, Los Angeles, CA 90095-1563, or via e-mail: lieber@ucla.edu.

REFERENCES

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences, U.S.A., 91,* 7041–7045.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7,* 268–277.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage, 38,* 95–113.

Baetens, K., Ma, N., Steen, J., & Van Overwalle, F. (2014). Involvement of the mentalizing network in social and non-social high construal. *Social Cognitive and Affective Neuroscience, 9,* 817–824.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage, 76,* 412–427.

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox [abstract]*. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan. Available on CD-ROM in *Neuroimage 16*(2).

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers in Biomedical Research, 29,* 162–173.

Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in the medial prefrontal cortex. *Journal of Cognitive Neuroscience, 24,* 1742–1752.

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology, 6,* 178–190.

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.
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.
Krienen, F. M., Tu, P.-C., & Buckner, R. L. (2010). Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *Journal of Neuroscience, 30,* 13906–13915.
Lewis, P. A., Rezaie, R., Brown, R., Roberts, N., & Dunbar, R. I. M. (2011). Ventromedial prefrontal volume predicts understanding of others and social network size. *Neuroimage, 57,* 1624–1629.
Lieberman, M. D. (2010). Social cognitive neuroscience. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 143–193). New York: McGraw-Hill.
Lieberman, M. D. (2012). Self-knowledge: From philosophy to neuroscience to psychology. In S. Vazire & T. D. Wilson (Eds.), *Handbook of self-knowledge* (pp. 63–76). New York: Guilford.
Ma, N., Baetens, K., Vandekerckhove, M., Kestemont, J., Fias, W., & Van Overwalle, F. (2014). Traits are represented in the medial prefrontal cortex: An fMRI adaptation study. *Social Cognitive and Affective Neuroscience, 9,* 1185–1192.
Ma, N., Baetens, K., Vandekerckhove, M., Van der Cruyssen, L., & Van Overwalle, F. (2013). Dissociation of a trait and a valence representation in the mPFC. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nst143.
Ma, N., Vandekerckhove, M., Van Overwalle, F., Seurinck, R., & Fias, W. (2011). Spontaneous and intentional trait inferences recruit a common mentalizing network to a different degree: Spontaneous inferences activate only its core areas. *Social Neuroscience, 6,* 123–138.
Maldjian, J. A., Laurienti, P. J., Burdette, J. B., & Kraft, R. A. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19,* 1233–1239.
McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102,* 419–457.
Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience, 17,* 1306–1315.
Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron, 50,* 655–663.
Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage, 25,* 653–660.
Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D., Kihlstrom, J. F., et al. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage, 28,* 797–814.
Powell, J. L., Lewis, P. A., Dunbar, R. I. M., Garcia-Finana, M., & Roberts, N. (2010). Orbital prefrontal cortex volume correlates with social cognitive competence. *Neuropsychologia, 48,* 3554–3562.
Powell, J. L., Lewis, P. A., Roberts, N., Garcia-Finana, M., & Dunbar, R. I. M. (2012). Orbital prefrontal cortex volume predicts social network size: An imaging study of individual differences in humans. *Proceedings of the Royal Society B: Biological Sciences, 279,* 2157–2162.
Sallet, J., Mars, R. B., Noonan, M. P., Andersson, J. L., O'Reilly, J. X., Jbabdi, S., et al. (2011). Social network size affects neural circuits in macaques. *Science, 334,* 697–700.
Semendeferi, K., Armstrong, E., Schleicher, A., Ziles, K., & Van Hoessen, G. W. (2001). Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology, 114,* 224–241.
Spunt, R. P., Falk, E. B., & Lieberman, M. D. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Science, 21,* 1593–1598.
Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2011). Identifying the what, why, and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience, 23,* 63–74.
Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review, 99,* 195–231.
Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annual Review of Neuroscience, 27,* 279–306.
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15,* 273–289.
Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30,* 829–858.
Vanderwal, T., Hunyadi, E., Grupe, D. W., Connors, C. M., & Schultz, R. T. (2008). Self, mother and abstract other: An fMRI study of reflective social processing. *Neuroimage, 41,* 1437–1446.
Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of false beliefs in young children's understanding of deception. *Cognition, 13,* 103–128.