

Research Articles | Behavioral/Cognitive

## Decomposing Cognitive Processes in the mPFC During Self-Thinking

<https://doi.org/10.1523/JNEUROSCI.2378-24.2025>

Received: 17 December 2024

Revised: 4 April 2025

Accepted: 8 April 2025

Copyright © 2025 Levorsen et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

---

*This Early Release article has been peer reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.*

**Alerts:** Sign up at [www.jneurosci.org/alerts](http://www.jneurosci.org/alerts) to receive customized email alerts when the fully formatted version of this article is published.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18

**Decomposing Cognitive Processes in the mPFC During Self-Thinking**

**Abbreviated Title:** self-reference in the mPFC

**Marie Levorsen<sup>1</sup>, Ryuta Aoki<sup>2,3</sup>, Constantine Sedikides<sup>1</sup> and Keise  
Izuma<sup>1,4,5,\*</sup>**

<sup>1</sup> School of Psychology, University of Southampton, Southampton, SO17 1BJ, UK

<sup>2</sup> Graduate School of Humanities, Tokyo Metropolitan University, Tokyo, 192-0397, Japan

<sup>3</sup> Human Brain Research Center, Graduate School of Medicine, Kyoto University, Kyoto, 606-8501, Japan

<sup>4</sup> School of Economics & Management, Kochi University of Technology, Kochi, 780-8515, Japan

<sup>5</sup> Research Center for Mind, Brain, and Behavior, Kochi University of Technology, Kochi, 780-8515, Japan

\*Keise Izuma, School of Economics & Management, Kochi University of Technology, Kochi 780-8515, Japan. Email: [izuma.keise@kochi-tech.ac.jp](mailto:izuma.keise@kochi-tech.ac.jp) ORCID: 0000-0003-0256-3571

19

## Abstract

20 Past cognitive neuroscience research has demonstrated that thinking about both the self and  
21 other activate the medial prefrontal cortex (mPFC), a central hub of the default mode network.  
22 The mPFC is also implicated in other cognitive processes, such as introspection and  
23 autobiographical memory, rendering elusive its exact role during thinking about the self.  
24 Specifically, it is unclear whether the same cognitive process explains the common mPFC  
25 involvement or distinct processes are responsible for the mPFC activation overlap. In this  
26 preregistered functional magnetic resonance imaging (fMRI) study with 35 male and female  
27 human participants, we investigated whether and to what extent mPFC activation patterns  
28 during self-reference judgment could be explained by activation patterns during the tasks of  
29 other-reference judgment, introspection, and autobiographical memory. Multi-voxel pattern  
30 analysis (MVPA) showed that only in the mPFC were neural responses both concurrently  
31 different and similar across tasks. Furthermore, multiple regression and variance partitioning  
32 analyses indicated that each task (i.e., other-reference, introspection, memory) uniquely and  
33 jointly explained significant variances in mPFC activation during self-reference. These findings  
34 suggest that the self-reference task engages multiple cognitive processes shared with other  
35 tasks, with the mPFC serving as a crucial hub where essential information is gathered and  
36 integrated to support judgments based on internally constructed representations.

37

38

### Significance statement

39 This study advances our understanding of the medial prefrontal cortex (mPFC), a central hub of  
40 the default mode network, in self-referential thinking. By using fMRI, multi-voxel pattern  
41 analysis, and variance partitioning, we demonstrate that mPFC activation during self-reference  
42 judgment is explained by shared and unique contributions from other cognitive processes,  
43 including other-reference, introspection, and autobiographical memory. Importantly, the mPFC  
44 is the only region where neural responses were concurrently similar and different across these  
45 tasks, suggesting its role in integrating diverse cognitive processes. These findings highlight the  
46 mPFC's critical function in gathering and integrating information for judgments based on internal  
47 representations, shedding light on its multifaceted role in self-related cognition.

48

JNeurosci Accepted Manuscript

50 Thinking about the self and expressing who one is to others are fundamental aspects of  
51 human experience. The self has fascinated researchers for more than a century (James, 1890;  
52 Cooley, 1902). Reflecting this enduring interest, the intricate neural architecture of the self has  
53 been a persistent focus of inquiry (Wagner et al., 2019; Frewen et al., 2020). Using  
54 neuroimaging methods such as fMRI, studies have established that the midline structures, the  
55 mPFC and posterior cingulate cortex (PCC), are active during the self-reference task in which  
56 individuals judge if a presented personality trait or attitudinal statement describes them (Denny  
57 et al., 2012; Murray et al., 2012).

58 Although the robust association between the mPFC and self-reference processing has  
59 raised the possibility that the mPFC's primary function is processing self-relevant information  
60 (Kelley et al., 2002; Northoff, 2016), the mPFC is also involved in thinking about other people  
61 (Denny et al., 2012; Murray et al., 2012). Based on these observations, some researchers  
62 (Gillihan and Farah, 2005; Legrand and Ruby, 2009) criticized the self-specific view of the  
63 mPFC, arguing that some general cognitive processes are common to self-reference and other-  
64 reference processing. For example, inferential processing and memory recall are common to  
65 both (Legrand and Ruby, 2009). In other words, mPFC activation during the self-reference task  
66 might not be related to the self specifically, but rather it is a result of general cognitive processes  
67 which take place during the self-reference task, as well as during other tasks. Indeed, the mPFC  
68 and PCC are also known to be activated by autobiographical memory (Kim, 2012; Martinelli et  
69 al., 2013) and by decision-making based on internal or subjective criteria such as moral  
70 reasoning (Nakao et al., 2012).

71 From a broader perspective, the mPFC and PCC are considered the core hubs of the  
72 default mode network – a network of brain regions that show heightened activation at rest  
73 (Andrews-Hanna et al., 2010). These regions are activated by a variety of tasks that depend on

74 internally constructed representations, including not only self-reference and other-reference  
75 processing or autobiographical memory, but also introspection (i.e., thinking about one's own  
76 emotional states), episodic future thinking, creativity, affective decision-making, and spatial  
77 navigation (Buckner and DiNicola, 2019; Menon, 2023). For the past two decades, researchers  
78 have attempted to identify a key common cognitive process that explains mPFC's involvement  
79 in these distinct tasks. However, these attempts are often based on univariate activation overlap  
80 (or meta-analyses), and univariate activation overlap does not constitute strong evidence for a  
81 common cognitive process across tasks (Levorsen et al., 2023; Woo et al., 2014). Thus,  
82 experimental evidence on the extent to which different tasks share a common cognitive  
83 process(es) is lacking.

84         Recently, fMRI studies using MVPA and representational similarity analysis (RSA)  
85 approaches have compared patterns of activation for self-reference processing to a few other  
86 tasks. For example, Chavez et al. (2017) demonstrated that self-reference processing evoked  
87 similar activation patterns in ventral mPFC as positive affect (see also Yankouskaya et al.,  
88 2017). When comparing self-reference to other-reference, studies have demonstrated distinct  
89 patterns of activation in mPFC (Feng et al., 2018; Courtney and Meyer, 2020; Koski et al., 2020;  
90 Parelman et al., 2022). Although these results begin to clarify scholarly understanding of  
91 affective and cognitive processes during self-thinking, the degree to which other internally  
92 focused processes (namely, introspection and memory) explain self-reference, remains  
93 unknown.

94         In the present study, using RSA and MVPA, we aimed to test similarities and differences  
95 in neural responses between the self-reference task and three other tasks that also rely on  
96 internal representation and are known to robustly activate the mPFC. These are the other-  
97 reference, autobiographical memory (Addis et al., 2007; Summerfield et al., 2009), and  
98 introspection (Goldberg et al., 2006; Gusnard et al., 2001) tasks. Furthermore, through variance

99 partitioning analysis, we sought to quantify the extent to which explainable variance in mPFC  
100 activation patterns during self-thinking can be attributed to activation patterns from the other  
101 three tasks.

102

103

## Materials and Methods

### 104 Preregistration

105 We preregistered the sample size, hypotheses, participant exclusion criteria, and data analysis  
106 plan at the Open Science Framework (<https://osf.io/mn9fz>). Unless otherwise noted, we  
107 analyzed the data in accord with the preregistration.

108

### 109 Participants

110 The experiment was approved by the Kochi University of Technology ethics committee. Before  
111 the online autobiographical memory session, participants checked a box to indicate their  
112 consent. We obtained written consent prior to the fMRI experiment.

113 The final sample comprised 35 Kochi University of Technology students (8 women, 27  
114 men), ranging in age from 18 to 22 years ( $M = 19.47$ ,  $SD = 1.08$ ). The sample size was based on  
115 similar previous studies (Chavez et al., 2017; Wen et al., 2020; Yankouskaya et al., 2017). We  
116 remunerated them with 2,500 Japanese yen. Participants were right-handed, had no history of  
117 psychiatric disorders, and had normal or corrected-to-normal vision. We excluded data from one  
118 additional participant due to excessive head movement (preregistered exclusion criteria of  $>3$   
119 mm).

120

121 **Experimental procedure**

122 The experiment consisted of two parts: (a) online autobiographical memory survey, and (b) fMRI  
123 experiment. The two sessions took place on separate days, 6.97 days apart on average ( $SD =$   
124 2.54).

125

126 **Online autobiographical memory session.** We adapted the autobiographical memory task  
127 from Wen et al. (2020). Prior to the fMRI scan, we instructed participants to write down 15  
128 autobiographical memories, corresponding to one of 15 events each. These memories should  
129 pertain to an event bound to a specific time and context that occurred more than one year ago,  
130 but after their age of 10 years. The memories ought to be clear so that participants be in a  
131 position to remember the relevant people, objects, and location in detail.

132

133 **Stimuli preparation.** We selected for each participant 10 of the 15 listed event memories. We  
134 used the selected memories as stimuli in the autobiographical memory task during the fMRI  
135 experiment. We based memory selection on the amount of detail and number of characters  
136 included in each description. We matched the number of characters with stimuli in the general  
137 knowledge task (see below). For each memory, we removed critical words and replaced them  
138 with blank underscores prior to the fMRI experiment.

139

140 **The fMRI experiment.** The fMRI experiment consisted of the following three tasks (Figure 1):  
141 (a) self/other trait judgement, (b) introspection, and (c) autobiographical memory. The self/other  
142 trait judgment task had three conditions (Figure 1a-c), whereas the introspection (Figure 1d &  
143 1e) and autobiographical memory (Figure 1f & 1g) tasks had two conditions each. Thus, there



144 was a total of seven conditions. Participants completed five fMRI runs, with each run lasting  
145 approximately 6.5 minutes. Each run included two blocks of seven conditions for a total of 14  
146 blocks. We pseudorandomized the block order within each run, so that the same task block was  
147 not presented twice in a row. At the beginning of each block, participants viewed a cue for 1  
148 second indicating that the task that was about to commence. All text stimuli were in Japanese.  
149 We programmed all tasks in Psychtoolbox (<http://psychtoolbox.org/>) with Matlab software  
150 (version 2018a; <http://www.mathworks.co.uk>).

151

152 *Self/other trait judgement task:* The stimuli comprised 40 trait adjectives from a pool of  
153 normalized trait adjectives (Anderson, 1968), which we translated into Japanese. The stimuli  
154 consisted of an equal number of positive (e.g., “honest,” “trustworthy”) and negative (e.g.,  
155 “mean,” “greedy”) traits. For each trial, we presented a trait in the middle of the screen. In the  
156 self-reference block (Figure 1a), we asked participants to judge whether each trait describes  
157 them. In the other-reference block (Figure 1b), before fMRI scanning, we asked participants to  
158 write down the name of one of their close friends on a piece of paper. During scanning, we  
159 instructed them to judge whether each trait describes this specific friend. In the semantic  
160 judgment block (Figure 1c), we instructed them to judge whether each trait is positive or  
161 negative. The same 40 adjectives were used across the three tasks. We presented each trial for  
162 2 sec, followed by a 1 second fixation cross, and we presented four traits in each block (12 sec  
163 per block). We randomly determined for each participant the order of traits in each of the self-  
164 reference, other-reference, and semantic conditions, but each block always included two  
165 positive and two negative words. We presented a fixation cross for 12 sec before the next block.

166

167 *Introspection task:* We adapted the introspection task from (Gusnard et al., 2001). It consisted of  
168 two conditions: introspection and categorization. We downloaded 40 picture stimuli (i.e., images  
169 of objects, animals, or sceneries) from the Open Affective Standardized Image Set (Kurdi et al.,  
170 2017). Half of the stimuli were negative and half positive. For each trial in the introspection block  
171 (Figure 1d), we presented participants with an image and asked them how the image made  
172 them feel. They could respond “positive” or “negative.” In the categorization block (Figure 1e),  
173 we asked participants to judge whether each picture depicted a scene that was “indoors” or  
174 “outdoors.” The same 40 images were used across the two tasks. For each participant, the  
175 order of images was randomly determined in each of the introspection and categorization tasks,  
176 but each block always included two positive and two negative images. We presented each  
177 image for 2 sec and displayed a fixation cross for 1 sec before the next image appeared (each  
178 block lasted 12 sec). After an introspection/categorization block, we displayed a fixation cross  
179 for 12 sec before the next block.

180  
181 *Autobiographical memory task:* The autobiographical memory task (Wen et al., 2020) comprised  
182 two conditions: memory and knowledge. For each trial in the memory condition (Figure 1f),  
183 participants encountered one of the memories they had previously listed in the online  
184 autobiographical memory session. Each memory consisted of, on average, 67.6 Japanese  
185 characters ( $SD = 7.25$ ), which we matched with the length of the stimuli used in the knowledge  
186 condition. Within each memory, we replaced three critical words with blank underscores. We  
187 asked participants to recall the memory and fill in the blanks for the missing words, but do so in  
188 their mind rather than by pressing a button (i.e., we recorded no responses during this task).

189 In the knowledge condition (Figure 1g), we presented participants with text related to  
190 general knowledge ( $M = 67.8$  characters,  $SD = 8.11$  characters), such as a description of a

191 common topic (e.g., Mt. Fuji, football, seatbelt), in which we replaced certain words with blank  
192 underscores. We instructed participants to think of appropriate words to fill in the blanks.

193 In both conditions, we presented each text stimulus for 14 sec and followed it by a fixation  
194 cross (4-6 sec). Next, we asked: "Were you recollecting a specific event?" (1 = *not at all*, 5 =  
195 *extremely vividly*). Participants had up to 6 sec to respond. We presented a fixation cross for 10  
196 sec before the next block.

197

### 198 **Behavioral data analysis**

199 A one-way Analysis of Variance was conducted to compare reaction time (RT) and  
200 response rates across the self-reference, other-reference, and semantic judgment tasks. Given  
201 that the RT data were not normally distributed, we log-transformed them beforehand. We  
202 followed up significant effects with a Bonferroni-corrected tests. All reported p values were two-  
203 sided.

204 We also ran a multiple regression analysis, with RT in the other-reference condition as the  
205 dependent variable and response similarity as the primary independent variable (1 = same  
206 responses to the same trait, -1 = different responses). Additional independent variables included  
207 participant response (1 = yes, -1 = no), trait valence (1 = positive, -1 = negative), number of  
208 characters of each word stimulus, and the interaction between participant response and trait  
209 valence. We ran the same regression analysis for RT in the self-reference condition. Due to  
210 high multicollinearity, we removed the interaction term for some participants. Also, we excluded  
211 the valence term for two participants, because it was highly positively correlated with responses,  
212 indicating strong self-enhancing and other-enhancing tendencies.

213

## 214 **fMRI data acquisition**

215 We acquired images using a 3.0 T Prisma Siemens MRI scanner with a 64-channel  
216 phased-array head coil. For functional imaging, we used T2\*-weighted gradient-echo echo-  
217 planar imaging (EPI) sequences. We acquired 42 contiguous transaxial slices (covering almost  
218 the entire cerebrum) with a thickness of 3 mm, in an interleaved order. We acquired the images  
219 with the following parameters: Time repetition (TR) = 2500 ms, echo time (TE) = 25 ms, flip  
220 angle (FA) = 90°, field of view (FOV) = 192 mm<sup>2</sup>, matrix = 64 × 64. Additionally, we acquired a  
221 T1-weighted structural image (with 1mm isotropic resolution) from each participant.

222

## 223 **fMRI data preprocessing**

224 We carried out preprocessing and statistical analysis in SPM 12 (Wellcome Department of  
225 Imaging Neuroscience), implemented in MATLAB (Math Works). To allow for T1 equilibration,  
226 we discarded the first four volumes before preprocessing and data analyses. We used SPM  
227 12's preproc\_fmri.m script to perform preprocessing of the fMRI data. We spatially realigned all  
228 functional images within each run to the mean using 7th-degree B-spline interpolation. We  
229 normalized the volumes to MNI space using a transformation matrix that we obtained from the  
230 EPI normalization of the first participant to the EPI template. We resampled the volumes to a  
231 voxel size of 3 × 3 × 3 mm<sup>3</sup>, that is, we retained the original voxel size. We used the 7th-degree  
232 B-spline interpolation option for normalization. We applied spatial smoothing (of 8 mm FWHM)  
233 to the data for the whole brain univariate analysis. To maintain fine-grained activation patterns,  
234 we did not apply smoothing to the data for representational similarity analysis nor for  
235 multivariate pattern analysis.

236

## 237 **Univariate fMRI analysis**

238 *General Linear Model (GLM):* We first ran a conventional GLM analysis, modeling  
239 separately each of the seven task blocks (i.e., conditions) with duration of 12 sec, except for the  
240 autobiographical memory and general knowledge tasks which had a duration of 14 sec. The  
241 memory and knowledge tasks had a rating phase which we modeled separately as a nuisance  
242 regressor (duration = participant's response time). We also included six head motion  
243 parameters as nuisance regressors. To examine mPFC activation, we created the following six  
244 contrast images for each participant: (a) self > semantic, (b) other > semantic, (c) self > other,  
245 (d) introspection > categorization, (5) memory > knowledge, (6) rest > semantic + knowledge +  
246 categorization. We used the last contrast to identify regions that showed increased activations  
247 during passive rest compared to externally focused tasks (Gusnard et al., 2001; Shulman et al.,  
248 1997; Wen et al., 2020). Furthermore, we created seven additional contrast images (each of the  
249 seven tasks relative to the implicit baseline [i.e., rest]). The spmT images from these contrasts  
250 were used in the subsequent MVPA and RSA analyses (details below).

251 *Group analysis:* We conducted a second-level whole brain group analysis for each of the  
252 contrasts. We set the statistical threshold at  $p < 0.001$  voxel-wise (uncorrected) and cluster  $p <$   
253  $0.05$  (FWE corrected for multiple comparisons).

254

### 255 **Representational similarity analysis**

256 We conducted the RSA to test the similarity in activation patterns between the self and  
257 each of the other-reference, introspection, and memory conditions. For each participant, neural  
258 data were extracted from the spmT image of each contrast, and we computed neural  
259 representational similarity matrix (RSM; Figure 2a) based on Pearson correlation across  
260 activation patterns in each pair of conditions across the five runs. There are three model RSMs  
261 (Figure 2b-d), each of which addresses the similarity between the self and: (a) other, (b)

262 introspection, (c) memory. Given that we are interested in the similarity between the self and  
263 other, independently of similarities across the remaining conditions, we excluded from analyses  
264 the irrelevant conditions. For example, when testing the self = introspection model (Figure 2c),  
265 we excluded the other-reference, memory, and knowledge conditions so that pattern similarities  
266 involving those irrelevant conditions would not affect the results. We evaluated the fit between  
267 the neural RSM and model RSM via Kendall's tau-a for each participant (Nili et al., 2014).  
268 Activations of any two conditions within the same run are likely to be positively correlated largely  
269 due to shared physiological noises (Alink et al., 2015); as such, we excluded correlations  
270 between any pairs of conditions within the same run to the model RSM. We also excluded  
271 correlations between neural responses of the same conditions (Ritchie et al., 2017). We ran  
272 these RSAs using a searchlight approach (explained below).

273

#### 274 **Classifier-based MVPA**

275 The above RSA tests whether activation patterns are similar between two conditions. We  
276 proceeded to conduct classifier-based MVPA to examine whether activations patterns in the two  
277 conditions were distinct. We implemented a linear support vector machine (SVM), carried out  
278 via MATLAB in combination with LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>) (Levorsen  
279 et al., 2021; Wake and Izuma, 2017), with a cost parameter of  $c = 1$  (default).

280 We used MVPA to find out if the activation patterns for the following contrasts were  
281 distinct: (a) self > semantic versus other > semantic, (b) self > semantic versus introspection >  
282 categorization, and (c) self > semantic versus memory > knowledge. For each participant,  
283 neural data were extracted from the spmT image of each of these contrasts. To evaluate  
284 classification performance, we employed a leave-one-run-out cross-validation procedure. Thus,  
285 we first left out one run in each cross-validation, and, using the data from the rest of runs, we

286 trained a classifier that discriminates (e.g., activation patterns between self > semantic versus  
287 introspection > categorization contrasts). Subsequently, we tested the classifier performance  
288 using the data from the left-out run. We repeated this procedure five times leaving out a different  
289 run each time, and we averaged the five classification accuracy values. Like the RSA, we ran  
290 the classifier-based MVPA using a searchlight approach (below).

291

### 292 **Searchlight analysis**

293 We conducted the RSA and MVPA with a searchlight approach (Kriegeskorte et al., 2006).  
294 For the RSA, we extracted local patterns of neural activity from searchlights with a three-voxel  
295 radius, so that each searchlight consisted of a maximum of 123 voxels (and less on the edges  
296 of the brain). We made a neural RSM from each searchlight and computed Kendall's tau-a  
297 between neural versus each of the three model RSMs (Figure 2), which we saved for a center  
298 voxel, resulting in three correlation maps for each participant.

299 Similarly, for the classifier-based MVPA, we carried out MVPA within each searchlight,  
300 and we saved a classification accuracy for a center voxel, resulting in a total of three  
301 classification accuracy maps for each participant. Within each searchlight, we removed mean  
302 activity by subtracting the mean value of a searchlight sphere from values of the individual voxel  
303 so that mean activation difference across conditions could not account for MVPA results.

304 *Group analysis:* We applied smoothing before the group analysis of the RSA and MVPA  
305 outputs (with a Gaussian kernel of 4-mm FWHM). Following the smoothing, we entered the  
306 Kendall's tau-a maps and classification accuracy maps into a second-level permutation-based  
307 analysis (with 5,000 permutations). We used the Statistical Non-Parametric Mapping toolbox for  
308 SPM (Nichols and Holmes, 2002). Within the preregistered mPFC region of interest (ROI), we  
309 set a statistical threshold (i.e., voxel-level) at  $p < 0.005$ , and a cluster-level threshold at  $p < 0.05$

310 (FWE corrected). Outside of the mPFC, we set a statistical threshold at  $p < 0.001$ , and a cluster-  
311 level threshold at  $p < 0.05$  (FWE corrected).

312

### 313 **ROI analysis**

314 We further investigated the role of the mPFC in thinking about the self by running a ROI  
315 analysis. We used Neurosynth (<https://neurosynth.org/>; Yarkoni et al., 2011) to define our mPFC  
316 ROI independently of our data. We downloaded an association map (thresholded at  $q < .01$ ,  
317 False Discovery Rate corrected), which we generated from a term-based meta-analysis with the  
318 label “self-referential” (downloaded on October 10th, 2023). The mPFC ROI included 308  
319 voxels. We ran the following multivariate pattern regression analyses within the ROI.

320

### 321 **Multivariate pattern regression**

322 The above RSA and classifier-based MVPA address neural pattern similarity and  
323 difference separately for each pair of tasks. We conducted a multivariate pattern regression  
324 analysis to compare pattern similarity across multiple tasks within the same framework. We ran  
325 a multiple regression analysis where activation patterns of the self > semantic contrast were a  
326 dependent variable, whereas those of (a) the other > semantic, (b) introspection >  
327 categorization, and (c) memory > knowledge contrasts were independent variables (Figure 3).

328 As stated above, given that activation patterns of any two conditions within the same run  
329 are likely to be positively correlated likely due to shared physiological noise, we ran the  
330 regression analysis 20 times (i.e., all possible pairs of five runs excluding pairs from the same  
331 run) so that independent and dependent variables were always from two different runs. We



332 averaged all outputs (i.e., beta values and adjusted  $R^2$ ) across the 20 regression analyses  
333 within each participant.

334 *Noise ceiling model:* To provide an estimate of how much systematic variation in activation  
335 patterns of the self > semantic contrast could be explained in the data given measurement  
336 noise, we included a noise-ceiling model. This model simply included the data from the self >  
337 semantic contrast as both a dependent and independent variable (although they were from  
338 different runs) in the multivariate pattern regression. Thus, the only difference between the noise  
339 ceiling model and original full model (illustrated in Figure 3) was the inclusion of activation  
340 patterns of the self > semantic contrast as another independent variable in the noise ceiling  
341 model.

342 *Variance Partitioning Analysis:* Following the multivariate pattern regression analysis, we  
343 carried out variance partitioning analysis to infer the amount of unique and shared variance  
344 between three different predictors. We conducted seven multiple regression analyses: one with  
345 all three independent variables as predictors (illustrated in Figure 3), three with different pairs of  
346 two independent variables as predictors, and three with individual independent variables as  
347 predictors. Comparing the explained variance ( $R^2$ ) of a model used alone with the explained  
348 variance when used with other models would allow us to infer the amount of unique and shared  
349 variance between different predictors.

350 *Permutation test:* To assess the significance of the findings from the multivariate pattern  
351 regression analyses and variance partitioning analysis, we ran permutation tests where voxels  
352 were randomly shuffled. The self > semantic contrast and other > semantic contrast have the  
353 semantic condition as a common control condition, and this common control condition is likely to  
354 bias a beta value associated with the other > semantic activation patterns to a positive direction.  
355 Thus, our permutation test randomly shuffled beta activation map of the self-reference condition  
356 (i.e., self > implicit rest contrast). We computed a randomly-shuffled-self > semantic contrast

357 image (and a corresponding t-statistics map) so that the effect of the similarity in neural  
358 responses between the semantic task versus each of the remaining five tasks remained intact in  
359 each permutation. We repeated this step 1,000 times to estimate null distributions. Furthermore,  
360 shuffling voxels may overly destroy spatial autocorrelation in the original data, which might bias  
361 results of the permutation test. Thus, we smoothed shuffled data via a Gaussian kernel with the  
362 standard deviation of 0.86 before conducting a multiple regression analysis (see Burt et al.,  
363 2020 for a similar approach). We selected a standard deviation of 0.86, because it produced the  
364 smallest sum of square error between the smoothness (quantified as Moran's I based on an  
365 inverse Euclidean distance matrix; Moran, 1950) of the original data versus that of shuffled-and-  
366 then-smoothed data (repeated 1,000 times; we tried all standard deviation values ranging from  
367 0 to 2.0 with an increment of 0.02).

368

### 369 **Deviations from preregistration**

370 We deviated from the preregistration as follows. First, we preregistered and conducted  
371 MVPA testing pattern generalizability (i.e., cross-task classification) which, like the RSA, aims to  
372 examine the similarity in activation patterns between two conditions. However, we do not report  
373 relevant results, because they were similar to the results of the RSA described below; also, this  
374 analysis is inappropriate when testing the similarity between the self-references and other-  
375 reference conditions due to their common control condition. Second, we did not preregister the  
376 following: behavioral data analyses, reaction time (RT)-controlled MVPA, multivariate pattern  
377 regression, and variance partitioning analyses.

378

379

## **Results**

### 380 **Behavioral results**

381 During the self/other trait judgment conditions, participants pressed one of the two keys in  
382 almost all trials in the self (99.6%), other (99.9%), and semantic (99.9%) conditions. There was  
383 a significant difference in RT across the three conditions ( $F_{(2,68)} = 19.37, p < 0.001$ ). Pairwise t-  
384 tests revealed that RTs were significantly different from each other across conditions. RTs in the  
385 self-reference condition ( $M = 1.21$  sec,  $SD = 0.18$  sec) were significantly longer than those in  
386 the other-reference condition ( $M = 1.14$  sec,  $SD = 0.25$  sec;  $p_{corrected} = 0.001$ ) and in the  
387 semantic condition ( $M = 1.08$  sec,  $SD = 0.19$  sec;  $p_{corrected} < 0.001$ ). RTs in the other condition  
388 were significantly longer than those in the semantic condition ( $p_{corrected} = 0.046$ ).

389 We next examined if RTs in the other-reference condition were influenced by response  
390 similarity between the self and other, as reported in a previous study (Thornton and Mitchell,  
391 2018). We obtained a significant effect of response similarity ( $t_{(34)} = -3.80, p = 0.003$ ). RTs were  
392 shorter when the self- and other-reference judgments for the same trait were identical (i.e., both  
393 yes or both no). Although this result suggests egocentric anchoring and adjustment in other-  
394 reference judgment, we observed a similar effect in the self-reference condition (see below).  
395 Number of characters was significantly related to RTs, meaning the more characters a word  
396 had, the slower the participant responded ( $t_{(34)} = 3.83, p = 0.003$ ). We also obtained a significant  
397 Participant Response  $\times$  Trait Valence interaction ( $t_{(22)} = -4.14, p = 0.002$ ). Participants were  
398 slower to respond yes than no when judging if a negative trait described their friend, whereas  
399 they did not differ in their responses to positive traits. No other significant effect emerged.

400 We conducted the same regression analysis for the self-reference condition to test if RTs  
401 in the self-reference condition were influenced by response similarity between the self and  
402 other. We found significant effects of response similarity ( $t_{(34)} = -2.81, p = 0.041$ ) and number of  
403 characters ( $t_{(34)} = 2.97, p = 0.027$ ). When we compared beta values for the self-reference versus  
404 other-reference conditions, we observed no significant difference in the effect of response  
405 similarity on RT ( $t_{(34)} = 1.30, uncorrected p = 0.20$ ), suggesting that the significant effect of the

406 response similarity obtained in the other-reference condition might be at least partially explained  
407 by unknown stimulus features.

408 Consistent with prior research (Moran et al., 2006), participants were more likely to  
409 endorse a positive trait as self-descriptive and a negative trait as not self-descriptive ( $t_{(34)} = 4.28$ ,  
410  $p < 0.001$ ). However, we observed this positivity bias in the other-reference condition as well  
411 ( $t_{(34)} = 8.46$ ,  $p < 0.001$ ); indeed, this bias was stronger for the other-reference than the self-  
412 reference condition, indicating that participants were more other-enhancing than self-enhancing  
413 ( $t_{(34)} = 3.48$ ,  $p = 0.001$ ). These results are largely consistent with some findings suggesting that  
414 self-enhancement is weaker for East Asian compared to Western individuals (Heine &  
415 Hamamura, 2007; but see Cai et al., 2016).

416 During the introspection task, participants pressed one of the two keys in almost all trials  
417 in the introspection (99.9%) and categorization (99.7%) conditions. RTs during the introspection  
418 condition ( $M = 1.08$  sec,  $SD = 0.21$ ) were significantly faster than those during the  
419 categorization ( $M = 1.14$  sec,  $SD = 0.17$ ) condition ( $t_{(34)} = 3.79$ ,  $p < 0.001$ ), likely because some  
420 pictures were ambiguous as to whether they were taken indoors or outdoors.

421 During the autobiographical memory task, participants successfully gave their vividness  
422 rating within the time limit of 6 sec for almost all trials in the memory (99.0%) and knowledge  
423 (98.7%) conditions. Vividness ratings were significantly higher in the memory ( $M = 4.37$ ,  $SD =$   
424  $0.40$ ) compared to the knowledge ( $M = 2.45$ ,  $SD = 0.79$ ) condition ( $t_{(34)} = 17.73$ ,  $p < 0.001$ ),  
425 testifying to the effectiveness of our memory manipulation.

426

## 427 **fMRI results**

428 **Univariate analysis results:** Replicating findings from several studies (Denny et al., 2012;  
429 Murray et al., 2012), the self > semantic contrast significantly activated the midline structure

430 including mPFC and PCC (Figure 4a and Table S1). The other > semantic contrast activated  
431 similar regions (Figure 4b and Table S2). Left temporoparietal junction (TPJ) was also  
432 commonly activated by the self and other conditions. Although there were some regions that  
433 were uniquely activated by either the self > semantic or other > semantic contrast when the self  
434 and other conditions were directly compared, the self > other contrast did not lead to any  
435 significant activation. The opposite contrast (other > self) revealed only one significant cluster in  
436 PCC (303 voxels;  $x = 6, y = -64, z = 29$ ).

437 As per prior studies (Goldberg et al., 2006; Gusnard et al., 2001), the introspection >  
438 categorization contrast significantly activated the mPFC (Figure 4c). Other activated areas  
439 included anterior cingulate cortex, temporal pole, lateral temporal cortex, and lateral occipital  
440 cortex (Table S3).

441 The memory versus knowledge contrast significantly activated regions previously  
442 implicated in autographical memory including the mPFC, PCC/precuneus, posterior inferior  
443 parietal lobule (pIPL), and lateral temporal cortex (LTC; Kim, 2012; Martinelli et al., 2013)  
444 (Figure 4d and Table S4).

445 Taken together, the above four contrasts all significantly activated the common region  
446 within the mPFC (1,565 voxels; Figure 4e). Bilateral temporal poles were also commonly  
447 activated by all four contrasts (left  $x = -36, y = 17, z = -22, 109$  voxels; right  $x = 30, y = 14, z = -$   
448  $22, 185$  voxels). No other region was commonly activated. Yet, although the introspection >  
449 categorization contrast activated the PCC (92 voxels), it did not pass our preregistered cluster-  
450 level threshold. When we directly compared the four contrasts to each other within the  
451 commonly activated mPFC areas, no significant difference emerged ( $F_{(2.38, 80.86)} = 0.08, p = 0.94$ ;  
452 Figure 4f).

453 Consistent with prior findings (Gusnard et al., 2001; Shulman et al., 1997; Wen et al.,  
454 2020), the rest > semantic + categorization + knowledge contrast revealed that areas in the  
455 default mode network, including mPFC, PCC, IPL, TPJ/AG, and LTC, were active during rest  
456 compared to the externally focused tasks (Figure 4g).

457

#### 458 **Results of RSA: Are activation patterns evoked by two tasks similar?**

459 The RSA (Figure 2) aims to test whether the self-reference judgment evoked similar  
460 activation patterns with each of the other-reference judgment, introspection, and  
461 autobiographical memory.

462 The Self = Other model was significantly associated with a network of brain regions  
463 involved in self-reference and social cognition including mPFC, PCC/precuneus, bilateral inferior  
464 frontal gyrus (IFG), bilateral superior temporal sulcus, and bilateral temporal pole (Figure 5a).  
465 However, the other two models were associated only with mPFC and left IFG. In particular, the  
466 Self = Introspection model was significantly associated with mPFC ( $x = 0, y = 53, z = 35, 1,930$   
467 voxels; see Figure 5b) and left IFG (extending to temporal pole;  $x = -51, y = 20, z = 2, 379$   
468 voxels). Further, the Self = Memory model was significantly associated with mPFC ( $x = -12, y =$   
469  $44, z = 5, 103$  voxels) (Figure 5c) and left IFG ( $x = -45, y = 26, z = -7, 368$  voxels).

470

#### 471 **Results of MVPA testing pattern discriminability: Are activation patterns evoked by two** 472 **tasks distinguishable?**

473 The classifier-based MVPA tested pattern discriminability with a searchlight approach. It  
474 addressed whether activation patterns evoked by different tasks were distinguishable or linked  
475 to different cognitive processes. Indeed, activation patterns evoked during the self-reference

476 task (relative to the semantic task) were distinguishable from the other-reference task in the  
477 mPFC, PCC, and right superior temporal sulcus (extending to the temporal pole; Figure 5d).  
478 These areas largely overlapped with the areas activated by the self-reference and other-  
479 reference condition relative to the semantic condition (Figure 4a and 4b), indicating that those  
480 areas were commonly activated both by the self and other conditions compared to the semantic  
481 condition, but their activation patterns were systematically different. Given that the self and other  
482 conditions had the semantic condition as common control, the difference between the self and  
483 other conditions is likely to be underestimated in this analysis.

484 In contrast, activation patterns elicited by the self-reference condition were distinguishable  
485 from each of the introspection and memory conditions in a number of regions across the whole  
486 brain including the mPFC, PCC, intraparietal lobule, middle temporal gyrus, and TPJ (Figures  
487 5e and 5f).

488 These results, together with the RSA results reported above, indicate that mPFC  
489 activation patterns during the self-reference judgement were similar to those elicited during each  
490 of the other-reference judgement, introspection, and autobiographical memory (Figure 5a-c).  
491 Nonetheless, they were still distinguishable from activation patterns of each of the three tasks  
492 (Figure 5d-f). In fact, there was one cluster within the mPFC (Figure 5g; a total of 96 voxels)  
493 showing significant association/classification accuracy in all six analyses (Figure 5a-f), and the  
494 mPFC cluster is the only region that showed the 6-way overlap with the cluster size larger than  
495 20 voxels. This 6-way overlap was located in the anterior part of the mPFC (Brodmann Area  
496 [BA] 10) and pregenual anterior cingulate cortex (pgACC; BA 32). Furthermore, this cluster  
497 entirely overlapped with the areas commonly activated by the four contrasts in the univariate  
498 analyses (Figure 5e). It also showed a substantial overlap (53 out of 96 voxels [55.2%]) with the  
499 areas significantly active during the rest (Figure 5g), indicating that most of the 6-way overlap  
500 area (Figure 5e) is located in the mPFC within the default mode network.

501

502 **Results of ROI analyses:** We conducted additional ROI analyses to refine the findings and run  
503 control analyses. We defined the mPFC ROI independently of our own data using Neurosynth  
504 (see Methods). This ROI analysis focused on areas within the mPFC most strongly associated  
505 with self-reference processing (Figure 6a).

506 *Does the difference in activation patterns between the self- and other-reference conditions*  
507 *simply reflect the difference in RTs between them?* According to our behavioral results, RTs  
508 were significantly longer for the self-reference condition compared to the other-reference  
509 condition. Thus, the difference in activation patterns between the two conditions might be  
510 explained by the difference in RTs (e.g., task difficulty). To rule out this possibility, we ran  
511 additional GLM where we categorized self-reference and other-reference task blocks into short  
512 and long RT blocks based on average RTs in each block. We modeled the other five tasks in  
513 the same way as the original GLM. Then, we ran an MVPA analysis testing whether it can  
514 distinguish activation patterns of the mPFC ROI during the short versus long RT blocks.

515 Within the mPFC ROI, the average accuracy for classifying the short and long RT blocks  
516 was 51.71%, which did not differ significantly from the theoretical chance level of 50% (Wilcoxon  
517 signed rank test,  $p = 0.31$ ). Also, it was significantly lower than the accuracy for classifying  
518 actual self- versus other-reference blocks (average = 63.14%; paired-sample Wilcoxon signed  
519 rank test,  $p = 0.002$ ). We additionally ran the same MVPA (short vs. long RT blocks) across the  
520 whole brain with a searchlight approach, but did not find any significant area. Taken together,  
521 the difference in RTs between the self and other conditions is unlikely to explain the difference  
522 in activation patterns between the two conditions.

523 *Which task best explains activation patterns of the self-reference condition?* The results of  
524 the RSA reported above (Figure 5a-c) indicate that mPFC activation patterns during the self-



525 reference condition were similar to those of the other-reference, introspection, and memory  
526 conditions. However, these analyses addressed neural pattern similarity separately for each pair  
527 of tasks. To compare pattern similarity across the tasks within the same framework, we carried  
528 out a multivariate pattern regression analysis where activation patterns of the self > semantic  
529 contrast were a dependent variable, whereas those of the other > semantic, introspection >  
530 categorization, and memory > knowledge contrasts were independent variables (Figure 3).  
531 Activation patterns of each of the three contrasts were significantly associated with mPFC ROI  
532 activation patterns of the self > semantic contrast (Figure 6b; all  $p_{perm} < 0.001$ ), suggesting that  
533 the similarity in mPFC neural responses between the self-reference task and each of the other  
534 three tasks remain significant even after controlling for the effect of neural responses during the  
535 other two tasks.

536 Adjusted  $R^2$  were significantly lower than the that of the noise ceiling model ( $p_{perm} < 0.001$ ;  
537 Figure 6c). Hence, there was still unexplained variance even after considering noise in the fMRI  
538 data, suggesting that there were patterns of activations specific to the self-reference judgment  
539 (not shared by the other three tasks). Activation patterns of the other > semantic, introspection >  
540 categorization, and memory > knowledge contrasts collectively explained, on average, 79.44 %  
541 of explainable variances in the mPFC activation patterns of the self > semantic contrast.

542 *Variance Partitioning Analysis (VPA):* We conducted a variance partitioning analysis to  
543 quantify how much variance in the mPFC ROI responses of the self > semantic contrast is  
544 explained uniquely by activation patterns of each of the other > semantic, introspection >  
545 categorization, and memory > knowledge contrasts, while considered together with the other  
546 two conditions. We present the results in Figure 6d. Each of the seven portions significantly  
547 explained the variance in neural responses of the self > semantic contrast (all  $p_{perm} < 0.001$ ).  
548 The results suggest that the mPFC activation patterns reflect multiple cognitive processes. For  
549 example, a significant amount of variances explained by all three contrasts indicate that there

550 were specific patterns of mPFC neural responses that were shared across self-reference, other-  
551 reference, introspection, and memory tasks, which likely reflects a cognitive process common  
552 for the four tasks. Similarly, a significant amount of variances explained by the other-reference  
553 and introspection conditions indicate that there were specific patterns of mPFC neural  
554 responses that were shared across self-reference, other-reference, and introspection tasks  
555 which likely reflects a cognitive process common for the three tasks, but not the memory task  
556 (see below for more discussion). These patterns of shared and unique variance among the  
557 tasks align with the RSA/MVPA results reported earlier (Figure 5): shared variance accounts for  
558 the similarity between tasks observed in RSA (Figure 5a-c), whereas unique variance explains  
559 the task-specific differences captured by MVPA (Figure 5d-f).

560 We ran the same multivariate pattern regression and variance partitioning analyses in  
561 other regions related to self-reference (based on the same Neurosynth meta-analysis map with  
562 the term “self-referential”) and to the default mode network (based on Andrews-Hanna et al.,  
563 2010). The anterior and dorsal parts of the mPFC (amPFC and dmPFC) of the default mode  
564 network were the only regions that evinced the same pattern of the results as the mPFC  
565 reported above: (a) significantly positive beta values for all three independent variables  
566 (multivariate pattern regression), and (b) significantly positive variance explained for all seven  
567 portions (variance partitioning analysis; Figure 7), indicating a unique and complex role played  
568 by the mPFC during thinking about the self.

569

570

## Discussion

571 We provided a more nuanced and precise picture of the mPFC’s role during thinking about  
572 the self. Replicating prior findings, each of the self-reference, other-reference, introspection, and  
573 autographical memory tasks activated the mPFC compared to their corresponding control

574 condition (Figure 4). Furthermore, we demonstrated that the relationship between activation  
575 patterns during the self-reference task and those of the other three tasks (other-reference,  
576 introspection, and autobiographical memory) was intricate. That is, mPFC neural responses  
577 during the self-reference task were not simply similar to one task and different from the other  
578 two tasks. Instead, the mPFC neural responses during the self-reference task were both similar  
579 and distinct at the same time from each of the other-reference, introspection, and  
580 autobiographical memory tasks (Figure 5). The mPFC was the only region across the whole  
581 brain that showed these patterns of results.

582 Furthermore, the multivariate pattern regression together with the variance partitioning  
583 analyses revealed complex relationships of activation patterns of each of the three other tasks  
584 to mPFC neural responses during the self-reference task (Figure 6). According to the variance  
585 partitioning analyses, not only each of the other-reference, introspection, and memory tasks  
586 uniquely explained significant amounts of variances in mPFC neural responses during the self-  
587 reference task, but also each pair of these tasks and all three tasks jointly explained significant  
588 amounts of variances of the mPFC neural responses during the self-reference task (Figure 6d).  
589 Hence, it suggests that there are cognitive processes common to thinking about the self and (a)  
590 each of the three tasks, (b) each pair of the three tasks, and (c) all three—indicating a total of at  
591 least seven overlapping cognitive components. In addition, adjusted  $R^2$  of the full model were  
592 significantly lower than the that of the noise ceiling model (Figure 6c), suggesting that some  
593 mPFC responses during the self-reference task reflect cognitive processes that are not shared  
594 with the other tasks and may be unique to self-referential cognition. Altogether, our results  
595 indicate that there are at least eight cognitive processes (see Figure 8) at play simultaneously  
596 when performing the self-reference task, some of which are common across tasks. Thus, the  
597 present study goes beyond simply identifying similarities or differences in activation patterns. It  
598 demonstrates that the self-reference task—often treated as a unitary process—involves multiple

599 and dissociable cognitive components. By applying variance partitioning analysis within the  
600 same sample of participants across multiple tasks, we were able to quantify the unique and  
601 shared contributions of each cognitive process to mPFC activity. This approach, especially  
602 when paired with MVPA and RSA, remains underutilized in this domain and represents a  
603 methodological advance for dissecting complex psychological functions.

604 Our variance partitioning analysis indicated that the mPFC is the only region that showed  
605 significantly positive variance explained for all seven portions (Figures 6 & 7). This result  
606 suggests that the mPFC, one of the core hubs of the default mode network (Andrews-Hanna et  
607 al., 2010; Andrews-Hanna, 2012), is a place where necessary information is gathered and  
608 integrated for judgments based on internally constructed representations. Yeshurun et al. (2021)  
609 considered the default mode network as an active and dynamic sense-making network that  
610 integrates incoming extrinsic information with prior intrinsic information to form rich, context-  
611 dependent models of situations as they unfold over time. More recently, Menon (2023) argued  
612 that the default mode network integrates multiple cognitive functions to create a coherent  
613 internal narrative of our experiences (see also Koban et al., 2021). Our findings offer empirical  
614 support for these frameworks by showing that the mPFC, within the default mode network,  
615 serves as a central convergence hub—integrating diverse sources of information to enable task-  
616 relevant judgments.

617 As a metaphor, this integration process is akin to making a soup: ingredients (information)  
618 are gathered from various parts of the kitchen (brain) and brought together in a pot (mPFC),  
619 where they are mixed and transformed into a cohesive dish (judgment). Just as many soups  
620 share core ingredients—e.g., both Italian minestrone and Japanese miso soup use common  
621 vegetables, and all soups require water—cognitive tasks often rely on shared processes, with  
622 some core mechanisms being common across many tasks. Similarly, internally directed tasks  
623 (e.g., self-reference, other-reference, introspection, and autobiographical memory) may share

624 overlapping cognitive components, all integrated within the mPFC to support coherent mental  
625 representations. This integrative role may explain why such a wide variety of social and  
626 cognitive tasks consistently engage the mPFC. In addition to the four tasks examined here, the  
627 mPFC has been implicated in theory of mind, episodic future thinking, and even spatial  
628 navigation (Menon, 2023). These tasks all likely involve constructing internal models by drawing  
629 on multiple sources of information—further underscoring the mPFC’s central role in synthesizing  
630 diverse cognitive inputs into unified, context-sensitive representations.

631 This view of mPFC’s role in the self-reference task invites re-interpretation of prior  
632 findings. Several studies showed that mPFC activation patterns differ depending on the target  
633 person in the self/other-reference tasks (e.g., self vs. close-other vs. distant other) and the  
634 dimension of person knowledge being assessed (e.g., traits, physical attributes, social roles)  
635 (Feng et al., 2018; Courtney and Meyer, 2020; Koski et al., 2020). The present study suggests  
636 that the divergent mPFC neural responses are driven by variations in the extent to which each  
637 task engages distinct types of information—and, by extension, different cognitive processes. For  
638 instance, thinking about close others and acquaintances may rely more on one’s  
639 autobiographical memory, whereas thinking about unfamiliar others (e.g., celebrities) may rely  
640 more on semantic memory (Courtney and Meyer, 2020). The mPFC activation patterns are also  
641 likely to vary depending on whether a context is general or specific (“I am friendly in general vs.  
642 at the university”) (Martial et al., 2018) and on differences in various dimensions of distance  
643 similarity (e.g., temporal, spatial, social, hypothetical) (Tamir and Mitchell, 2011), as these  
644 judgements often rely on divergent informational sources.

645 Our study does not specify what cognitive processes are at play during the self-reference  
646 task (see Figure 8 for ideas on possible candidate processes), leaving this issue open for future  
647 research. Nonetheless, as to the self-specific cognitive process, in our prior work (Levorsen et  
648 al., 2023), we reported that the self-specific activation patterns depend on the importance of the

649 stimuli for the self-concept, and so access to this self-concept information stored in the mPFC  
650 may be responsible for the self-specific mPFC activation patterns we observed here.

651 Relatedly, our results highlight an important conceptual challenge for social/cognitive  
652 neuroscientists; each of many tasks used in the field involves multiple cognitive processes (or  
653 operations), and each of these processes needs to be identified to fully understand the function  
654 of the mPFC (and any other brain regions). For example, our findings indicate that the  
655 difference between the self-reference and semantic tasks is not only the level of self-referential  
656 processing, but also that there are several other additional cognitive processes involved in the  
657 self-reference task, some of which are shared with other-reference, introspection, and memory  
658 tasks (Figure 8). Thus, rather than a traditional brain mapping approach that identifies regions  
659 activated by the self-reference > semantic contrast, a more refined approach is needed—one  
660 that links specific neural activation patterns to underlying cognitive components, rather than to  
661 entire task conditions. Although the utility of a multivariate approach over a univariate approach  
662 has been well recognized (and its methodology has been well developed; Haynes and Rees,  
663 2006; Haxby, 2012), identifying each basic cognitive process involved in a social/cognitive task  
664 remains a challenge (e.g., Schaafsma et al., 2015). Addressing this issue is critical for  
665 advancing our understanding of how complex mental functions are instantiated in the brain.

666 Finally, we note a limitation of our study. The seven tasks that we used vary in terms of  
667 visual stimuli and requirement for button responses. Granted, we made sure to match them  
668 within each task group (i.e., self, other, and semantic tasks employed the same text stimuli), and  
669 we employed contrast values with a corresponding control task (i.e., self > semantic, other >  
670 semantic, introspection > categorization, memory > knowledge) when investigating differences  
671 in neural responses across tasks. However, it is still possible that, due to differences in stimuli  
672 and response requirement, similarities in neural responses across tasks are underestimated  
673 (Figure 3a-c), whereas differences are overestimated (Figure 3d-f). Although these differences

674 are unlikely to explain the main mPFC findings (e.g., Figure 6), future studies should use better-  
675 matched tasks to reveal roles possibly played by other brain regions.

676 In conclusion, the current findings enhance understanding of the mPFC and its  
677 involvement in self-referential thinking by demonstrating its unique role in integrating diverse  
678 cognitive processes. The mPFC is not merely activated by self-reference, but also shows  
679 complex activation patterns that are both similar and distinct from other cognitive tasks such as  
680 other-reference, introspection, and autobiographical memory. Taken together with the role of the  
681 mPFC within the default mode network reported previously, the findings indicate that the mPFC  
682 serves as a hub where information from various brain regions is gathered and integrated,  
683 facilitating tasks that involve constructing internal representations.

684

JNeurosci Accepted Manuscript

685

### **Acknowledgements**

686 This research was supported by a Japan Society for Promotion of Science (JSPS) KAKENHI

687 Grant Number JP19K24680 (to K.I.).

688

JNeurosci Accepted Manuscript



689

### Data availability

690 Unthresholded group-level statistical maps are available on NeuroVault

691 (<https://neurovault.org/collections/19712/>).

692

JNeurosci Accepted Manuscript

693

## References

- 694 Addis DR, Wong AT, Schacter DL (2007) Remembering the past and imagining the future:  
695 common and distinct neural substrates during event construction and elaboration.  
696 *Neuropsychologia* 45:1363-1377.
- 697 Alink A, Walther A, Krugliak A, van den Bosch JJF, Kriegeskorte N (2015) Mind the drift -  
698 improving sensitivity to fMRI pattern information by accounting for temporal pattern drift.  
699 bioRxiv.
- 700 Anderson NH (1968) Likableness ratings of 555 personality-trait words. *J Pers Soc Psychol*  
701 9:272-279.
- 702 Andrews-Hanna JR (2012) The brain's default network and its adaptive role in internal  
703 mentation. *Neuroscientist* 18:251-270.
- 704 Andrews-Hanna JR, Smallwood J, Spreng RN (2014) The default network and self-generated  
705 thought: component processes, dynamic control, and clinical relevance. *Ann N Y Acad*  
706 *Sci* 1316:29-52.
- 707 Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL (2010) Functional-anatomic  
708 fractionation of the brain's default network. *Neuron* 65:550-562.
- 709 Buckner RL, DiNicola LM (2019) The brain's default network: updated anatomy, physiology and  
710 evolving insights. *Nat Rev Neurosci* 20:593-608.
- 711 Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD (2020) Generative modeling of brain maps  
712 with spatial autocorrelation. *Neuroimage* 220:117038.
- 713 Cai H, Wu L, Shi Y, Gu R, Sedikides C (2016) Self-enhancement among Westerners and  
714 Easterners: a cultural neuroscience approach. *Soc Cogn Affect Neurosci* 11:1569-1578.
- 715 Chavez RS, Heatherton TF, Wagner DD (2017) Neural population decoding reveals the intrinsic  
716 positivity of the self. *Cereb Cortex* 27:5222-5229.
- 717 Cooley CH (1902) *Human nature and the social order*. New York: Charles Scribner's Sons.
- 718 Courtney AL, Meyer ML (2020) Self-other representation in the social brain reflects social  
719 connection. *J Neurosci* 40:5616-5627.

720 Denny BT, Kober H, Wager TD, Ochsner KN (2012) A meta-analysis of functional neuroimaging  
721 studies of self- and other judgments reveals a spatial gradient for mentalizing in medial  
722 prefrontal cortex. *J Cogn Neurosci* 24:1742-1752.

723 Feng C, Yan X, Huang W, Han S, Ma Y (2018) Neural representations of the multidimensional  
724 self in the cortical midline structures. *Neuroimage* 183:291-299.

725 Frewen P, Schroeter ML, Riva G, Cipresso P, Fairfield B, Padulo C, Kemp AH, Palaniyappan L,  
726 Owolabi M, Kusi-Mensah K, Polyakova M, Fehertoi N, D'Andrea W, Lowe L, Northoff G  
727 (2020) Neuroimaging the consciousness of self: Review, and conceptual-methodological  
728 framework. *Neurosci Biobehav Rev* 112:164-212.

729 Gillihan SJ, Farah MJ (2005) Is self special? A critical review of evidence from experimental  
730 psychology and cognitive neuroscience. *Psychological bulletin* 131:76-97.

731 Goldberg, II, Harel M, Malach R (2006) When the brain loses its self: prefrontal inactivation  
732 during sensorimotor processing. *Neuron* 50:329-339.

733 Gusnard DA, Akbudak E, Shulman GL, Raichle ME (2001) Medial prefrontal cortex and self-  
734 referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci*  
735 *U S A* 98:4259-4264.

736 Haxby JV (2012) Multivariate pattern analysis of fMRI: the early beginnings. *Neuroimage*  
737 62:852-855.

738 Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. *Nat Rev*  
739 *Neurosci* 7:523-534.

740 Heine SJ, Hamamura T (2007) In search of East Asian self-enhancement. *Pers Soc Psychol*  
741 *Rev* 11:4-27.

742 James W (1890) *The Principles of Psychology*. New York: Henry Holt and Company.

743 Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF (2002) Finding the self?  
744 An event-related fMRI study. *J Cogn Neurosci* 14:785-794.

745 Kim H (2012) A dual-subsystem model of the brain's default network: self-referential processing,  
746 memory retrieval processes, and autobiographical memory retrieval. *Neuroimage*  
747 61:966-977.

748 Koban L, Gianaros PJ, Kober H, Wager TD (2021) The self in context: brain systems linking  
749 mental and physical health. *Nat Rev Neurosci* 22:309-322.

750 Koski JE, McHaney JR, Rigney AE, Beer JS (2020) Reconsidering longstanding assumptions  
751 about the role of medial prefrontal cortex (MPFC) in social evaluation. *Neuroimage*  
752 214:116752.

753 Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping.  
754 *Proc Natl Acad Sci U S A* 103:3863-3868.

755 Kurdi B, Lozano S, Banaji MR (2017) Introducing the Open Affective Standardized Image Set  
756 (OASIS). *Behav Res Methods* 49:457-470.

757 Legrand D, Ruby P (2009) What is self-specific? Theoretical investigation and critical review of  
758 neuroimaging results. *Psychol Rev* 116:252-282.

759 Levorsen M, Ito A, Suzuki S, Izuma K (2021) Testing the reinforcement learning hypothesis of  
760 social conformity. *Hum Brain Mapp* 42:1328-1342.

761 Levorsen M, Aoki R, Matsumoto K, Sedikides C, Izuma K (2023) The self-concept is  
762 represented in the medial prefrontal cortex in terms of self-importance. *J Neurosci*  
763 43:3675-3686.

764 Martial C, Stawarczyk D, D'Argembeau A (2018) Neural correlates of context-independent and  
765 context-dependent self-knowledge. *Brain Cogn* 125:23-31.

766 Martinelli P, Sperduti M, Piolino P (2013) Neural substrates of the self-memory system: new  
767 insights from a meta-analysis. *Hum Brain Mapp* 34:1515-1529.

768 Menon V (2023) 20 years of the default mode network: A review and synthesis. *Neuron*  
769 111:2469-2487.

770 Moran JM, Macrae CN, Heatherton TF, Wyland CL, Kelley WM (2006) Neuroanatomical  
771 evidence for distinct cognitive and affective components of self. *J Cogn Neurosci*  
772 18:1586-1594.

773 Moran PA (1950) Notes on continuous stochastic phenomena. *Biometrika* 37:17-23.

774 Murray RJ, Schaer M, Debbane M (2012) Degrees of separation: A quantitative neuroimaging  
775 meta-analysis investigating self-specificity and shared neural activation between self-  
776 and other-reflection. *Neurosci Biobehav Rev* 36:1043-1059.

777 Nakao T, Ohira H, Northoff G (2012) Distinction between Externally vs. Internally Guided  
778 Decision-Making: Operational Differences, Meta-Analytical Comparisons and Their  
779 Theoretical Implications. *Frontiers in neuroscience* 6:31.

780 Nichols TE, Holmes AP (2002) Nonparametric permutation tests for functional neuroimaging: a  
781 primer with examples. *Hum Brain Mapp* 15:1-25.

782 Nili H, Wingfield C, Walther A, Su L, Marslen-Wilson W, Kriegeskorte N (2014) A toolbox for  
783 representational similarity analysis. *PLoS Comput Biol* 10:e1003553.

784 Northoff G (2016) Is the self a higher-order or fundamental function of the brain? The "basis  
785 model of self-specificity" and its encoding by the brain's spontaneous activity. *Cogn*  
786 *Neurosci* 7:203-222.

787 Parelman JM, Dore BP, Cooper N, O'Donnell MB, Chan HY, Falk EB (2022) Overlapping  
788 functional representations of self- and other-related thought are separable through  
789 multivoxel pattern classification. *Cereb Cortex* 32:1131-1141.

790 Ritchie JB, Bracci S, Op de Beeck H (2017) Avoiding illusory effects in representational  
791 similarity analysis: What (not) to do with the diagonal. *Neuroimage* 148:197-200.

792 Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015) Deconstructing and reconstructing  
793 theory of mind. *Trends Cogn Sci* 19:65-72.

794 Schiller B, Sperl MFJ, Kleinert T, Nash K, Gianotti LRR (2024) EEG Microstates in social and  
795 affective neuroscience. *Brain Topogr* 37:479-495.

796 Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997)  
797 Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J*  
798 *Cogn Neurosci* 9:648-663.

799 Sui J, Gu X (2017) Self as object: Emerging trends in self research. *Trends Neurosci* 40:643-  
800 653.

801 Summerfield JJ, Hassabis D, Maguire EA (2009) Cortical midline involvement in  
802 autobiographical memory. *Neuroimage* 44:1188-1200.

803 Tamir DI, Mitchell JP (2011) The default network distinguishes construals of proximal versus  
804 distal events. *J Cogn Neurosci* 23:2945-2955.

805 Thornton MA, Mitchell JP (2018) Theories of person perception predict patterns of neural  
806 activity during mentalizing. *Cereb Cortex* 28:3505-3520.

807 Wagner DD, Chavez RS, Broom TW (2019) Decoding the neural representation of self and  
808 person knowledge with multivariate pattern analysis and data-driven approaches. *Wiley*  
809 *Interdiscip Rev Cogn Sci* 10:e1482.

810 Wake SJ, Izuma K (2017) A common neural code for social and monetary rewards in the human  
811 striatum. *Soc Cogn Affect Neurosci* 12:1558-1564.

812 Wen T, Mitchell DJ, Duncan J (2020) The functional convergence and heterogeneity of social,  
813 episodic, and self-referential thought in the default mode network. *Cereb Cortex*  
814 30:5915-5929.

815 Woo CW, Koban L, Kross E, Lindquist MA, Banich MT, Ruzic L, Andrews-Hanna JR, Wager TD  
816 (2014) Separate neural representations for physical pain and social rejection. *Nature*  
817 *communications* 5:5380.

818 Yankouskaya A, Humphreys G, Stolte M, Stokes M, Moradi Z, Sui J (2017) An anterior-posterior  
819 axis within the ventromedial prefrontal cortex separates self and reward. *Soc Cogn*  
820 *Affect Neurosci* 12:1859-1868.

821 Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated  
822 synthesis of human functional neuroimaging data. *Nat Methods* 8:665-U695.

823 Yeshurun Y, Nguyen M, Hasson U (2021) The default mode network: where the idiosyncratic  
824 self meets the shared social world. *Nat Rev Neurosci* 22:181-192.

825

826

## Figure legends

827 **Figure 1. Examples of a trial/block for each of the seven conditions across the three**  
828 **tasks.** The self/other trait judgment task consisted of (a) self-reference condition, (b) other-  
829 reference condition, and (c) semantic condition. The introspection task consisted of (d)  
830 introspection task and (e) categorization task. The autobiographical memory task consisted of  
831 (f) memory condition and (g) general knowledge condition.

832

833 **Figure 2. Schematic illustrations of representational similarity analysis (RSA).** (a) For  
834 each participant, we created a neural representational similarity matrix (RSM) by computing  
835 Pearson correlations between activation patterns during two tasks across five runs. (b) Self =  
836 Other model RSM. (c) Self = Introspection model RSM. (d) self = memory model RSM. In each  
837 neural/model RSM, we excluded cells in black from the analysis. In panels b-d, cells in cyan  
838 represent 1 (similar) while cells in white represent 0 (dissimilar). We evaluated fit between the  
839 neural versus each model RSMs through Kendall's tau-a (Nili et al., 2014).

840

841 **Figure 3. Multivariate pattern regression.** Activation patterns of the self > semantic contrast  
842 were a dependent variable, whereas activation patterns of the other three contrast were  
843 independent variables. Independent and dependent variables were always from different runs.

844

845 **Figure 4. Sagittal slices ( $x = -6$ ) showing results of the univariate analyses.** (a) Areas  
846 significantly activated by the self > semantic contrast. (b) Areas significantly activated by the  
847 other > semantic contrast. (c) Areas significantly activated by the introspection > categorization  
848 contrast. (d) Areas significantly activated by the memory > knowledge contrast. (e) Areas  
849 commonly activated by the all four contrasts (1,565 voxels). Only mPFC showed significant 4-  
850 way overlap. For display purposes, we set statistical threshold at  $p < 0.005$  and cluster- $p < 0.05$   
851 (FWE corrected). (f) Parameter estimates of the four contrasts within the mPFC areas  
852 commonly activated by the four contrasts (panel e). (g) Areas significantly activated by the rest  
853 > semantic + categorization + knowledge contrast.

854

855 **Figure 5. (a-c) Sagittal slices (x = -6) showing results from the RSA.** Significant areas  
856 indicate that activation patterns of the two contrasts were similar. **(d-f)** Sagittal slices (x = -6)  
857 showing results from the MVPA testing pattern discriminability. Significant areas indicate that  
858 activation patterns of the two contrasts were distinguishable. For display purposes, we set  
859 statistical threshold at  $p < 0.005$  and cluster- $p < 0.05$  (FWE corrected). **(g)** A sagittal slice (x = -  
860 6) showing the mPFC area that showed 6-way overlap (overlap across areas shown in panel a-  
861 f). **(h)** A sagittal slice (x = -6) showing overlap between univariate and MVPA results. Magenta  
862 represents areas activated commonly by the four univariate contrasts (Figure 4e), and white  
863 represents 6-way overlapped region depicted in panel **g**. **(i)** A sagittal slice (x = -6) showing  
864 overlap (white areas) between areas activated by the rest > semantic + categorization +  
865 knowledge contrast (magenta; Figure 4g) and the 6-way overlapped region depicted in panel **g**  
866 (cyan).

867  
868 **Figure 6. Results of the multivariate pattern regressions.** **(a)** A sagittal slice (x = -6) showing  
869 the mPFC areas used in the ROI analysis. We defined the mPFC ROI with the term “self-  
870 referential” based on Neurosynth term-based meta-analysis. **(b)** Beta values from the  
871 multivariate pattern regression with activation patterns of the self > semantic contrast as a  
872 dependent variable. Colored horizontal lines indicate mean beta values, and lower/upper box  
873 limits represent 95% confidence intervals (CIs). **(c)** Adjusted  $R^2$  from the original regression  
874 model (Figure 3) and the noise ceiling model. Pink circles indicate mean  $R^2$ , and black/grey  
875 circles indicate  $R^2$  of individual subjects. **(d)** Variance in mPFC ROI activation patterns of the  
876 self-reference condition that was explained by activation patterns of the other, introspection, and  
877 memory conditions. In panels **b** and **d**, bell shaped gray areas indicate permutation distribution.

878  
879 **Figure 7. Results of the multivariate pattern regressions and variance partitioning**  
880 **analyses.** We defined the self-related brain regions as “self-referential” based on the  
881 Neurosynth meta-analysis map. Regions within the default mode network were based on  
882 Andrews-Hanna et al. (2010). For the self-related ROIs, we used all voxels within each cluster.  
883 For the ROIs from the default mode network, we used a 9-mm sphere surrounding the center  
884 coordinate (maximum of 123 voxels). \*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , and \*  $p < 0.05$  (uncorrected). All  
885  $p$  values rely on permutation test (1,000 times). *n.s.* non-significant. dmPFC, dorsomedial  
886 prefrontal cortex. PCC, posterior cingulate cortex. TPJ, temporoparietal junction. TempP,



887 temporal pole. vmPFC, ventromedial prefrontal cortex. amPFC, anterior-medial prefrontal  
888 cortex. TC, temporal cortex. PHC, parahippocampal cortex. pIPL, posterior inferior parietal  
889 lobule. Rsp, retrosplenial cortex. TempP, temporal pole.

890

891 **Figure 8.** Schematic illustrating of the proposed integrative function of the mPFC. When  
892 participants engage in the self-reference task, the mPFC integrates information from other brain  
893 regions, each of which performs distinct cognitive processes (listed here), resulting in unique  
894 activation patterns (Figures 5-7). The variance partitioning analysis showed that variances  
895 explained by each of the seven portions were all significantly positive (Figure 6d), suggesting  
896 that the mPFC activation pattern reflects at least seven different cognitive processes at play  
897 simultaneously (plus self-specific processes; see Discussion). The table on the right lists  
898 possible cognitive processes corresponding to each of the seven significant portions of the  
899 variance partitioning analysis plus self-specific processes. Note that the possible candidates we  
900 listed are purely speculative; we do not claim that these processes are responsible for the  
901 results.

902

JNeurosci Accepted Manuscript

**a**

Describe you?

Friendly

Yes No

Self condition

**b**

Describe your friend?

Kind

Yes No

Other condition

**c**

Positive or negative?

Selfish

Positive Negative

Semantic condition

**Self/other trait judgment task****d**

How do you feel?



Positive Negative

Introspection condition

**e**

Where is it taken?



Indoor Outdoor

Categorization condition

**f**

During the summer vacation of my \_\_\_\_th year in junior high school, on the \_\_\_\_th day of the soccer club's training camp, I injured my \_\_\_\_, and afterward, I couldn't play anymore, which made me feel frustrated.

Memory condition

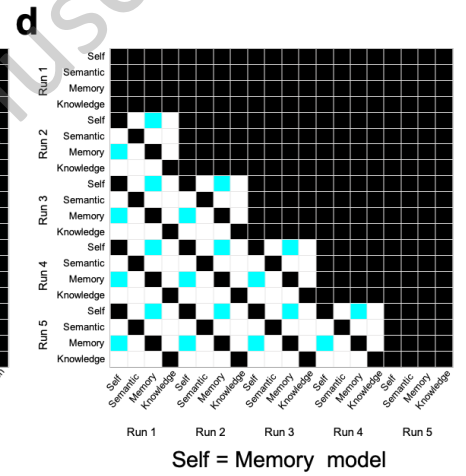
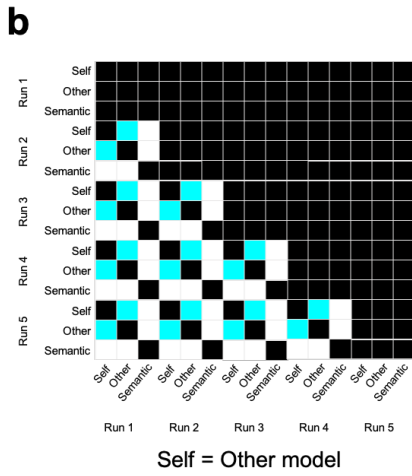
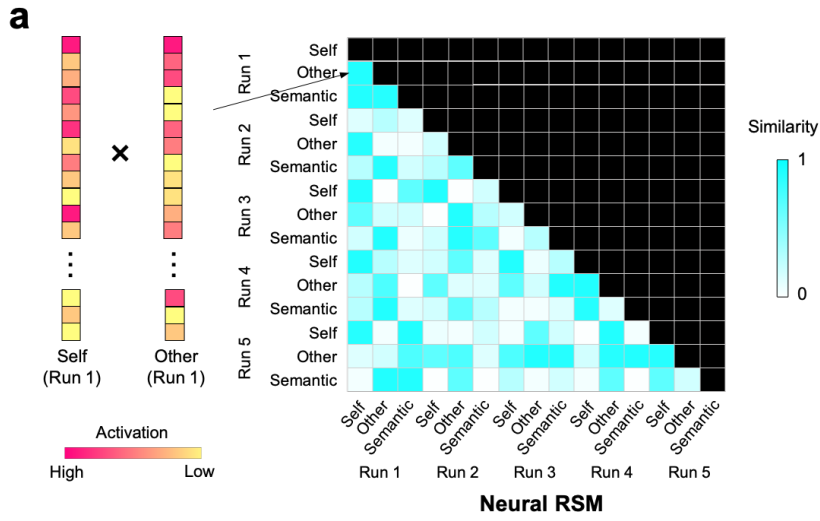
**g**

Mount Fuji is Japan's tallest mountain, spanning \_\_\_\_ and \_\_\_\_ Prefectures, with an elevation of \_\_\_\_ meters. It is registered as a World Cultural Heritage site along with its associated cultural assets.

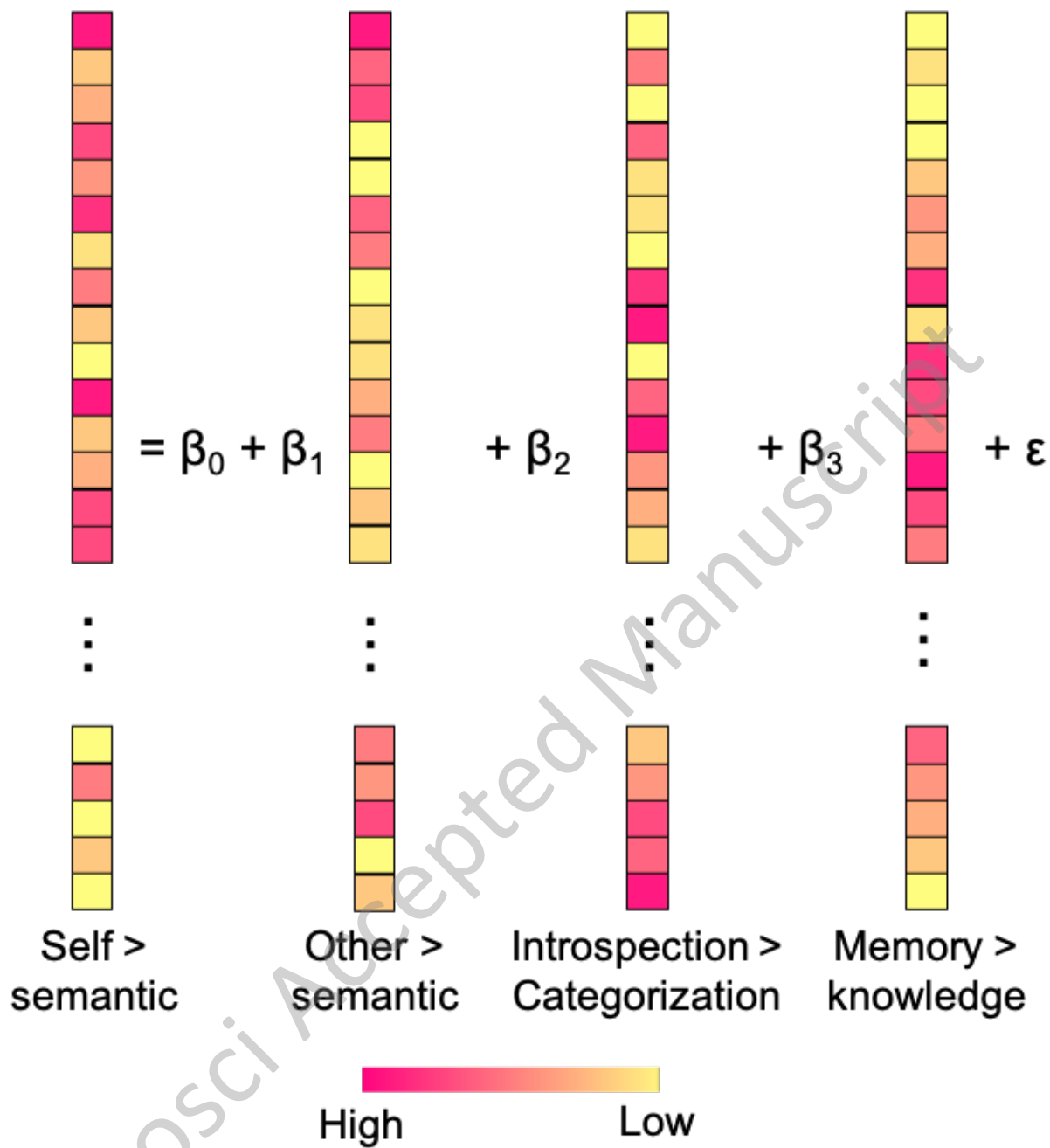
Knowledge condition

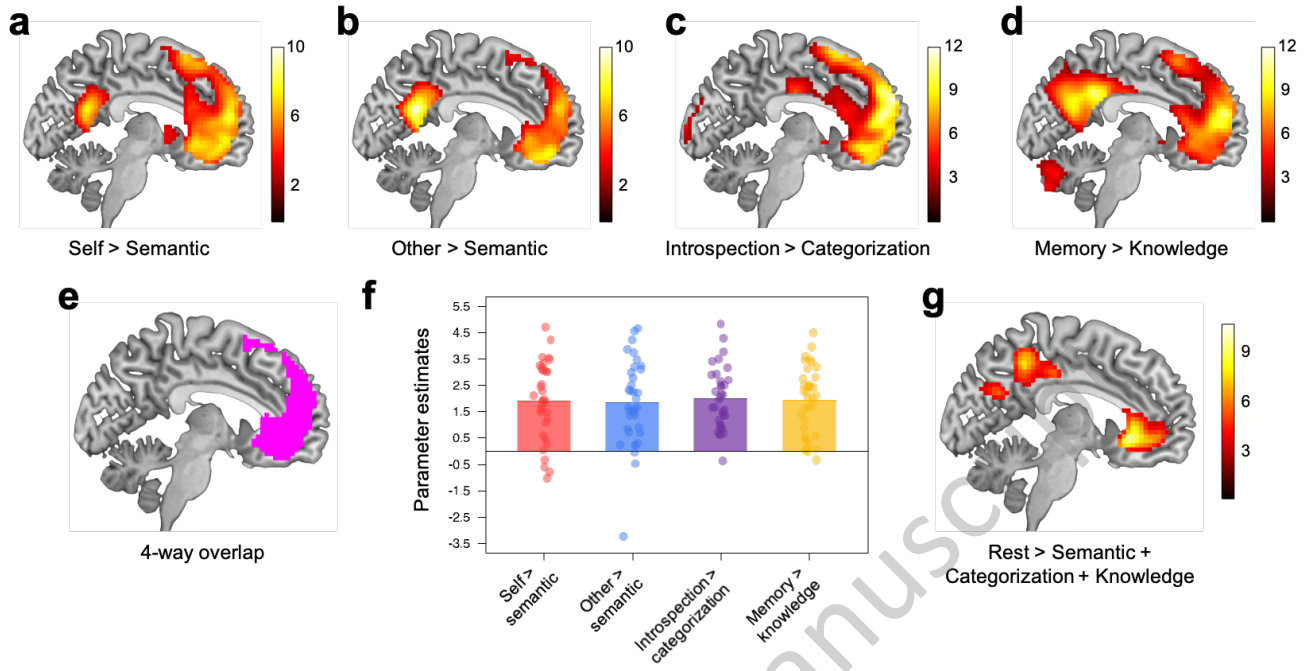
**Introspection task****Autobiographical memory task**

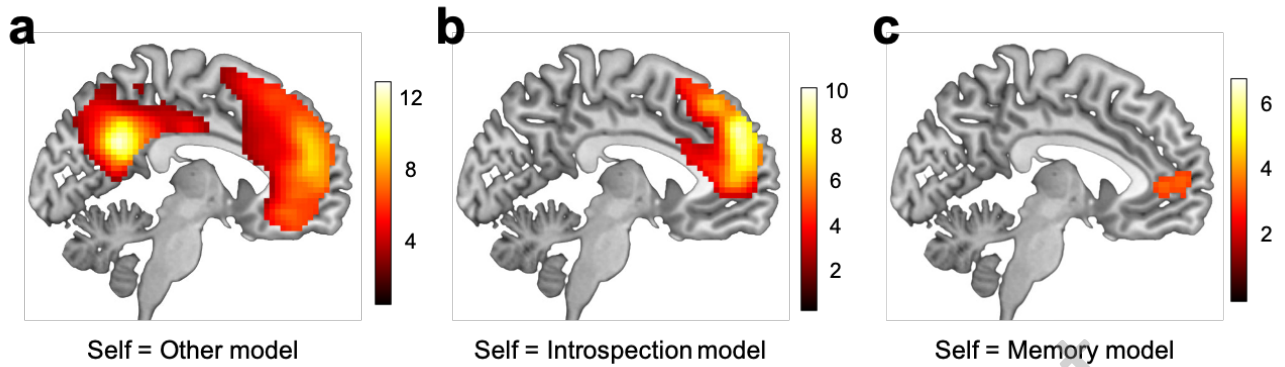
JNeurosci Accepted Manuscript



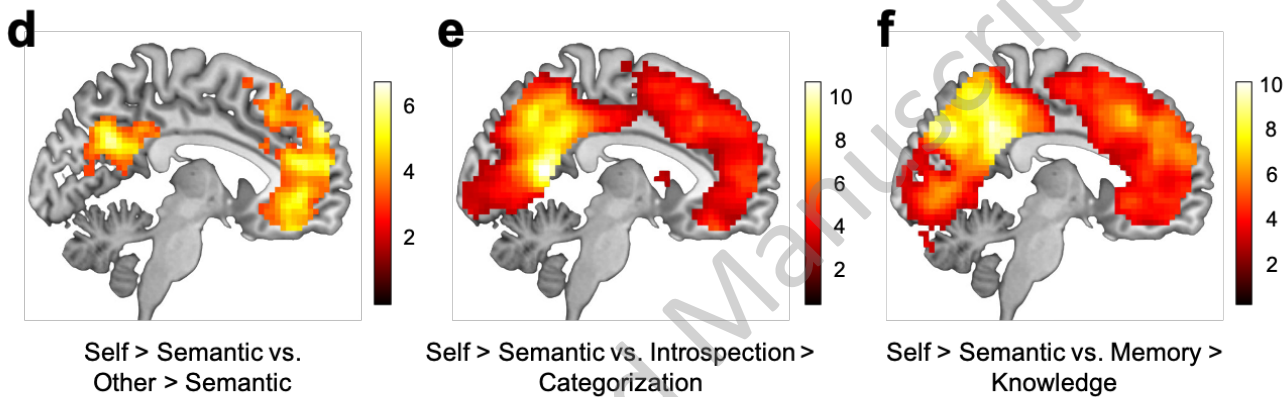
JNeurosci Accepted Manuscript



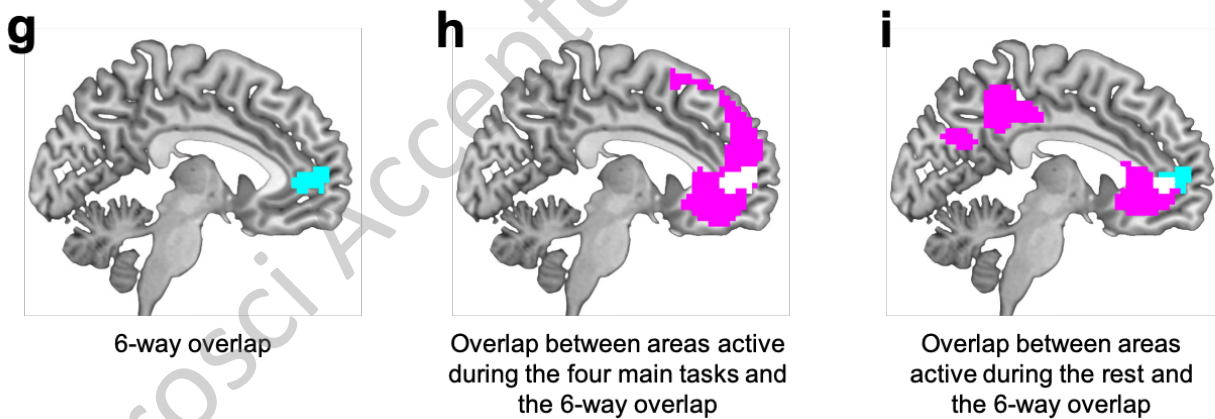


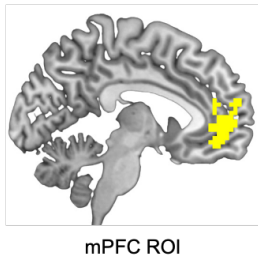
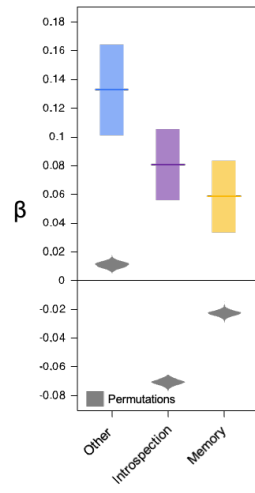
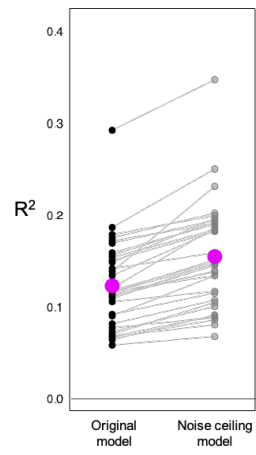
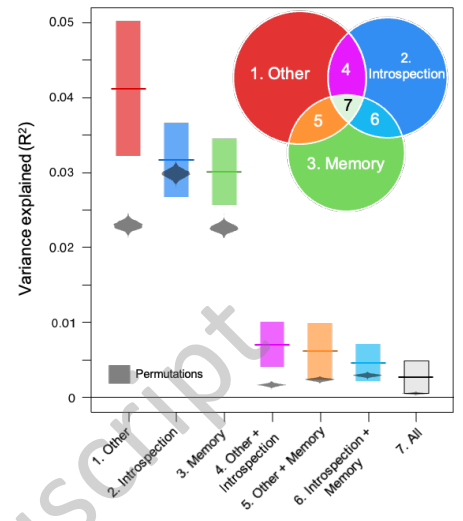


**Representational Similarity Analysis (RSA)**



**MVPA testing pattern discriminability**



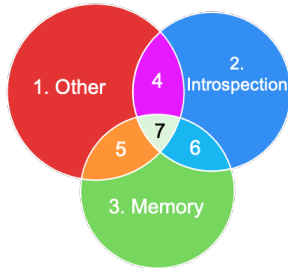
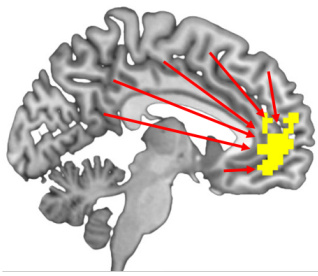
**a****b****c****d**

JNeurosci Accepted Manuscript

	region	MNI coordinate			voxels	Multivariate Regression			Variance Partitioning Analysis							
		x	y	z		Other-ref	Introspection	Memory	1. Other-ref	2. Introspection	3. Memory	4. Other + introspection	5. Other + memory	6. Introspection + memory	7. All	
Self-related regions (Neurosynth)	dmPFC	-6	41	41	56	***	***	***	***	***	***	***	***	n.s.	***	
	PCC	-3	-58	26	131	***	***	**	***	n.s.	***	***	***	***	***	
	Left TPJ	-48	-64	29	66	***	***	n.s.	***	n.s.	n.s.	***	***	n.s.	***	
	Left TP	-60	-13	-22	70	***	***	n.s.	***	***	***	***	***	***	n.s.	
Default mode network (Andrews-Hanna et al., 2010)	vmPFC	0	26	-18	39	***	***	n.s.	*	**	n.s.	n.s.	*	***	n.s.	
	amPFC	-6	52	-2	93	***	***	***	***	**	***	***	***	**	***	
	dmPFC	0	52	26	116	***	***	***	***	***	***	***	***	***	***	
	Left TC	-60	-24	-18	117	***	***	n.s.	***	**	***	***	***	***	***	n.s.
	Left TPJ	-54	-54	28	110	***	***	n.s.	***	n.s.	n.s.	***	***	n.s.	***	
	PCC	-8	-56	26	116	***	***	**	***	n.s.	***	***	***	***	***	
	PHC	-28	-40	-12	79	n.s.	***	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	
	pIPL	-44	-74	32	78	***	***	n.s.	***	n.s.	n.s.	***	***	n.s.	***	
	Rsp	-14	-52	8	83	n.s.	***	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.	
	TempP	-50	14	-40	21	n.s.	***	***	*	**	n.s.	n.s.	n.s.	n.s.	n.s.	

JNeurosci Accepted Manuscript





	Cognitive process shared across self-reference and other tasks	Possible candidates
1	Cognitive process common to self-reference and other-reference tasks	<ul style="list-style-type: none"> <li>Evaluating consistency with internal goal (e.g., self-enhancement and other-enhancement)</li> <li>Evaluating personal/social relevance</li> <li>Mentalizing (Theory of Mind): thinking about traits or attributes for oneself or another person</li> <li>Conceptual and social information processing: Accessing and evaluating social knowledge, including how the self or another person fits within social norms</li> </ul>
2	Cognitive process common to self-reference and introspection tasks	<ul style="list-style-type: none"> <li>Paying attention to one's internal state</li> <li>Self-referential processing or self-related cognition</li> <li>Internally constructing mental representations about oneself</li> <li>Self-awareness and evaluation of one's own thoughts</li> </ul>
3	Cognitive process common to self-reference and memory tasks	<ul style="list-style-type: none"> <li>Retrieval of personal memories</li> <li>Conceptual and semantic integration: Accessing self-related concepts and integrating them with existing knowledge structures</li> </ul>
4	Cognitive process common to self-reference, other-reference, and introspection tasks	<ul style="list-style-type: none"> <li>Inferential processing</li> <li>Affective evaluation</li> </ul>
5	Cognitive process common to self-reference, other-reference, and memory tasks	<ul style="list-style-type: none"> <li>Imagery and mental simulation: Visualizing past experiences or conceptualizing personal traits</li> <li>Social knowledge retrieval: Accessing stored semantic and episodic knowledge about people</li> </ul>
6	Cognitive process common to self-reference, introspection, and memory tasks	<ul style="list-style-type: none"> <li>Affective self-evaluation: Reflecting on one's own emotions and personal significance of memories or traits</li> </ul>
7	Cognitive process common to self-reference, other-reference, introspection, and memory tasks	<ul style="list-style-type: none"> <li>Judgment based on internally constructed representations</li> <li>Emotion regulation</li> <li>Abstract thought and conceptual integration: forming abstract representations about traits, experiences, or mental states rather than processing immediate sensory input</li> </ul>
8	Cognitive process specific to self-reference	<ul style="list-style-type: none"> <li>Access to self-knowledge stored in the mPFC in terms of self-importance</li> <li>Affective processing</li> <li>Evaluating consistency with a self-enhancement goal</li> </ul>

JNeurosci Accepted Manuscript