How does the brain navigate knowledge of social relations? Overlapping but distinct mechanisms for attentional shifts in space and social knowledge

Meng Du[1], Ruby Basyouni[1], & Carolyn Parkinson[1,2*]

[1]UCLA Department of Psychology, 1285 Psychology Building, Los Angeles, CA 90095

[2]UCLA Brain Research Institute, Los Angeles, CA 90095

*Correspondence: cparkinson@ucla.edu (C. Parkinson)

Abstract

How does the human brain support reasoning about social relations (e.g., social status, friendships)? Converging theories suggest that navigating knowledge of social relations may co-opt neural circuitry with evolutionarily older functions (e.g., shifting attention in space). Here, we analyzed multivoxel response patterns of fMRI data to examine the neural mechanisms for shifting attention in knowledge of a social hierarchy. The “directions” in which participants mentally navigated social knowledge were encoded in multivoxel patterns in superior parietal cortex, which also encoded the direction of attentional shifts in space. Exploratory analyses implicated additional regions of posterior parietal and occipital cortex in encoding analogous mental operations in space and social knowledge. However, cross-domain analyses suggested that attentional shifts in space and social knowledge are encoded in functionally independent response patterns. These results elucidate the neural basis for navigating abstract knowledge of social relations, and its connection to more basic mental operations.

Keywords: social cognition, social status, social hierarchy, social relations, social neuroscience, multivoxel pattern analysis, superior parietal lobule, posterior parietal cortex

Effectively navigating the human social world requires tracking, encoding, and reasoning about the bonds, rivalries, and hierarchies that comprise it. Correspondingly, humans and other group-living primates possess sophisticated social cognitive abilities. One such ability involves learning and reasoning about others' relative ranks in social hierarchies, which is important for strategically choosing allies and avoiding potentially harmful conflicts[1]. Rough, approximate assessments of someone's power or status can be gleaned from perceptual cues[2–4], but more precise knowledge is achieved through observing and participating in encounters between pairs of individuals, then using transitive inference to ascertain the relative status of those whom one has not seen interacting with each other[1,5–10]. Indeed, there is evidence that humans and other highly social animals acquire nuanced knowledge of others' relative status in social hierarchies in this way[1,5,8], and use such knowledge strategically, such as when choosing social partners[1]. Past research has examined the neural basis of how people learn about social hierarchies and the neural correlates of social status[8,11,12]. Yet, relatively little is known about the neural mechanisms that support reasoning about social relations, or how they relate to less abstract mental operations.

Theories from neuroscience, psychology, and cognitive linguistics have highlighted widespread parallels in how people speak and think about space and social knowledge, and have inspired suggestions that mentally navigating social structures might have co-opted pre-existing neural architecture that originally supported evolutionarily older functions, such as representing and navigating physical space. For example, Conceptual Metaphor Theory suggests that the language used to describe abstract information can shed light on how the mind processes such information[13]. In particular, the Spatialization of Form Hypothesis[14] suggests that the widespread use of spatial language to describe conceptual relations, such as social relations, reflects the spatial organization of knowledge of such relations. Thus, phrases like "top of the pecking order" and "low status" may not just be figures of speech, but rather, figures of thought that illuminate the structure of underlying mental representations[15]. Research findings and theoretical perspectives from cognitive neuroscience have converged with these ideas, suggesting that many cognitive capacities emerge through the repurposing of existing neural architecture, both over the course of evolution and in the case of culturally learned capacities, through development[16–20]. Such accounts provide potential neural mechanisms for theories from cognitive linguistics[14] and psychology[21] that suggest that representational resources with a precursory role in processing space were repurposed to process more abstract domains, such as time and social relations. For example, it has been suggested that, over the course of human evolution, regions of the posterior parietal cortex (PPC) originally devoted to representing and allocating attention in peripersonal space were co-opted to perform analogous operations on increasingly abstract contents[20]. Consistent with this possibility, findings from neuroimaging studies suggest considerable overlap in the brain regions that are recruited when processing spatial information and abstract social knowledge, particularly in the PPC[18,20].

While past findings suggest that areas in the PPC play a role in processing knowledge of social relations, many questions remain. For example, a previous study found that regions of the inferior PPC were more active when people made more difficult social status comparisons (e.g., determining which of two naval officers had a higher rank, when their ranks were very similar) than for easier ones[22] (e.g., determining which of two naval officers had a higher rank, when their ranks differed greatly). Yet, it is difficult to ascertain if such effects reflect the neural

representation of status knowledge or the more demanding nature of more difficult comparisons, especially given that the same brain region is also implicated in unrelated tasks when they demand attention or are difficult[23,24]. Thus, whereas the neural basis of *acquiring* knowledge about social relations (e.g., social status hierarchies) has been relatively well-characterized[8], further work is needed to understand how the brain supports reasoning about, or mentally "navigating" knowledge of social relations.

Addressing this question is likely to benefit from approaches that consider the information contained in spatial patterns of responses within brain areas, rather than analyses that only consider the overall magnitude of neural responses: Whereas many regions, including regions of the PPC, may be more active when tasks are more difficult or attention-demanding, regions that specifically encode the mental navigation of social knowledge should respond differently when mentally traversing knowledge of social relations in different ways, even when equating for difficulty. Multivoxel pattern analysis (MVPA) is likely to be especially fruitful in this context, given that it is sensitive to differences in the spatial patterning of responses in a brain region (e.g., for different kinds of stimuli or mental operations), even when the overall response magnitudes are equivalent[25].

Here, to elucidate the neural basis of representing and mentally navigating social hierarchy knowledge, we characterized the response patterns evoked when people reasoned about individuals' relative location in a learned social hierarchy. Specifically, we used MVPA to compare neural response patterns on trials that were matched for difficulty, but that differed in the nature of mental navigation within participants' knowledge of a set of social relations. The trials differed in the “directions” of navigation, such that participants navigated towards either a more powerful person or a less powerful person in the social hierarchy. If response patterns within a brain region subserve mentally navigating internal representations of social hierarchies, then in that region, response patterns should be more similar for shifts of attention in the same "direction" than in different "directions" within knowledge of a social hierarchy.

Past research has highlighted the importance of the superior parietal lobule (SPL) in supporting shifts of attention both in the perceptual environment and in internal knowledge representations[26–29]. Thus, we focused on the SPL in our main analyses. In addition, given that relatively little is known about how the brain supports traversing knowledge of social relations, and given that past related research has tended to use mass univariate analyses, which are not sensitive to effects carried in multivoxel patterns, we also performed exploratory analyses across regions spanning all of cerebral cortex and subcortical gray matter structures.

We also sought to test if shifts of attention in knowledge of social relations (exemplified by status hierarchies here) are encoded similarly to shifts of attention in physical space. As described above, considerable overlap has been documented in the brain regions that are recruited when processing knowledge of social hierarchies and other domains of content. However, shared encoding mechanisms cannot be inferred from overlapping fMRI activations alone, as such results can be produced either from common codes for processing different domains or modalities of information[30,31], or by spatially overlapping but functionally independent response patterns[32,33]. Thus, MVPA affords a more stringent test of whether overlapping fMRI activations reflect shared or distinct coding mechanisms[25]. For example, Knops et al.[29] trained a machine learning classifier only to discriminate fMRI response patterns in the SPL during overt leftward and rightward attentional shifts, and found that it correctly generalized to distinguishing

brain responses during mental subtraction and addition. This suggests that mental arithmetic relies on neural mechanisms that also support shifting attention in space, and may involve mental operations akin to shifting attention along a horizontally oriented mental number line. Using MVPA in this manner can provide insight into the representational contents of brain regions, and assess if common encoding mechanisms support analogous mental operations in different domains of information.

In the current study, we sought to elucidate the neural basis of mentally navigating social hierarchy knowledge, and to test if common neural mechanisms are involved in shifting attention in space and internal representations of social hierarchies. Participants first learned, through trial and error, the relative positions of 9 individuals in a fictive social hierarchy (Figs. 1 and 2). During scanning, participants performed a social navigation task, where they determined who in the hierarchy was a given number of steps more or less powerful than a reference person (Fig. 3). Participants later performed an ostensibly unrelated eye movement task during fMRI scanning, which allowed us to characterize activity patterns associated with overt spatial shifts of attention. After scanning, participants performed an open-ended task probing the structure of their mental representations of the social hierarchy (Fig. 4). We tested if distributed response patterns in the SPL encode shifts of attention in knowledge of social relations. We also tested if shifts of attention in space and social relationship knowledge are encoded in the same way in the SPL, or in any region of the brain. Testing these questions seeks to advance our understanding of how the human brain supports the ability to navigate knowledge of social relations, and more generally, how abstract feats of human social cognition relate to more concrete operations, such as processing the space around oneself.

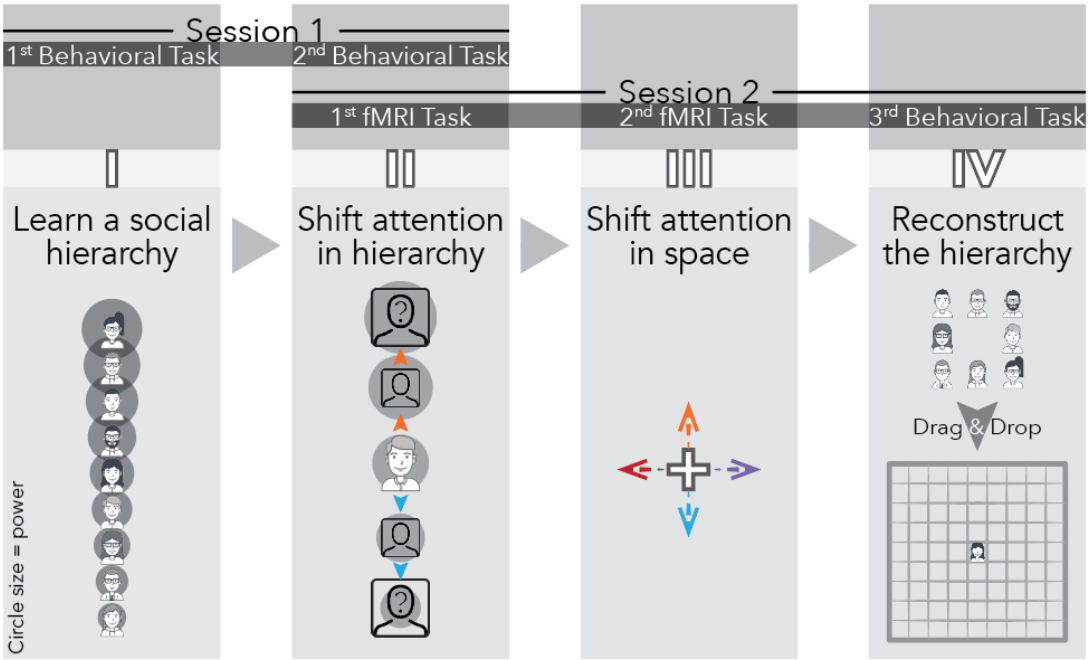

Figure 1. Overview of paradigm. Participants attended a behavioral session (Session 1) and an fMRI session (Session 2). In Session 1, they first learned a social hierarchy of 9 people through trial and error (Task I*). They then completed a social navigation task (Task II*), where they were repeatedly asked to determine who was a given number of steps more or less powerful than another person in the learned social hierarchy. In Session 2, participants underwent fMRI scanning while performing the social navigation task (Task II) again, followed by Task III, which

involved overt shifts of attention in external space (i.e., eye movements). At the end of the experiment, we probed participants' mental representations of the social hierarchy (i.e., mappings between positions in the hierarchy and locations in space) by asking them to arrange the faces in the way that they thought best reflected the people's relative power in the organization (Task IV).

*Note: in schematic illustrations of Tasks I and II, faces are depicted at different locations and within circles of varying sizes to indicate relative positions in the social hierarchy. These cues are included here only for illustration. This vertical arrangement, or any systematic spatial arrangement of the faces, was never shown to participants, nor were the faces depicted in circles of varying sizes. No spatial cues were provided to participants in Tasks I and II.

Results

Behavioral Data

Learning the Social Hierarchy. Participants first learned about the structure of a fictive social hierarchy of 9 people through trial and error, using a task adapted from prior related work[8] (Fig. 2). Participants were told that we were interested in how people learn about social information and that they would be learning about who had more power in an organization. On each trial, participants chose which of two people they thought had more power in the organization. All phases of this task were designed to avoid biasing participants towards any particular mapping between locations in space and in the social hierarchy, as well as more generally, between the social and spatial domains. For example, faces were presented one at a time in the same location on the screen (i.e., faces were never shown above or beside one another). Participants in the fMRI study attained an average accuracy of 98.24% in the last 3 blocks of Session 1, indicating that they had effectively learned the social hierarchy.

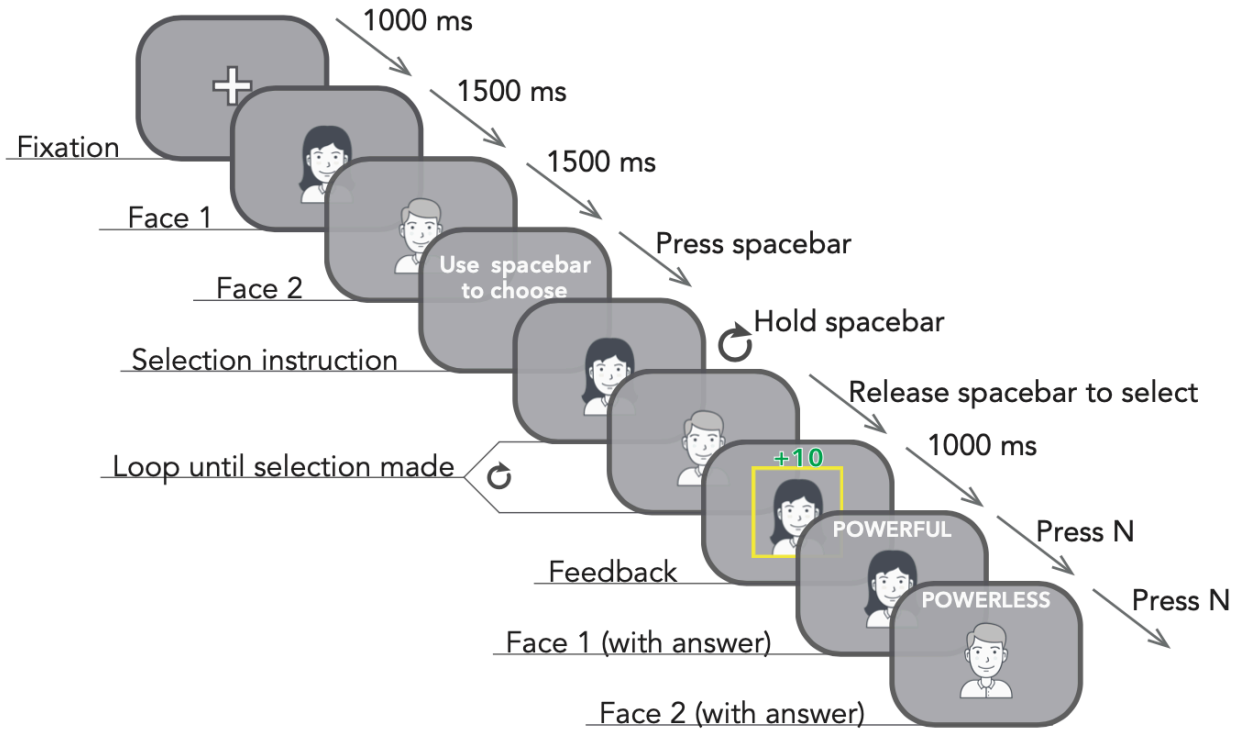

Figure 2. Learning the hierarchy (Task I). In this task, participants learned each person's relative position in the social hierarchy through trial and error. To avoid biasing participants towards thinking of the hierarchy in spatial terms, or using any particular spatial mapping, both presentation of the faces and the response paradigm were

designed to be sequential, rather than spatial (for example, faces and response keys were never arranged side-by-side or vertically). In each trial, two faces were presented on the screen one at a time, in a random order. Participants were then prompted to select the more powerful person in the pair by pressing and holding the space bar. While they held down the spacebar, the two face images were cycled through repeatedly in the same order as earlier in the trial. Participants chose one of the faces by releasing the space bar whenever they saw the person they intended to choose. The selected face was then highlighted together with a score (e.g., "+10 points" for correct responses, or "-10 points" for incorrect responses). Finally, the two faces were presented sequentially again, along with the correct information about the relative power of each face. Participants then pressed "N" to advance to the next trial.

Shifting Attention in Social Hierarchy Knowledge and in External Space. After learning the social hierarchy, participants performed a second task involving mentally navigating (shifting attention within) knowledge of the social hierarchy (Fig. 3; Methods). On each trial, participants saw a reference face, then were instructed to mentally navigate towards a more or less powerful person in the hierarchy. Participants who successfully performed this task in Session 1 (see Methods for more details) were eligible to return for the fMRI study (Session 2; see Fig. 1). During the fMRI study, participants completed several rounds of the social hierarchy navigation task (average accuracy = 94.35%). Participants then performed a second fMRI task (adapted from previous work[29]; involving overt shifts of visual attention upward, downward, leftward, and rightward in space (Fig. 1; Methods). Together, these two tasks allowed us to characterize the neural response patterns associated with shifts of attention in different directions in internally represented knowledge of the social hierarchy and in external space.

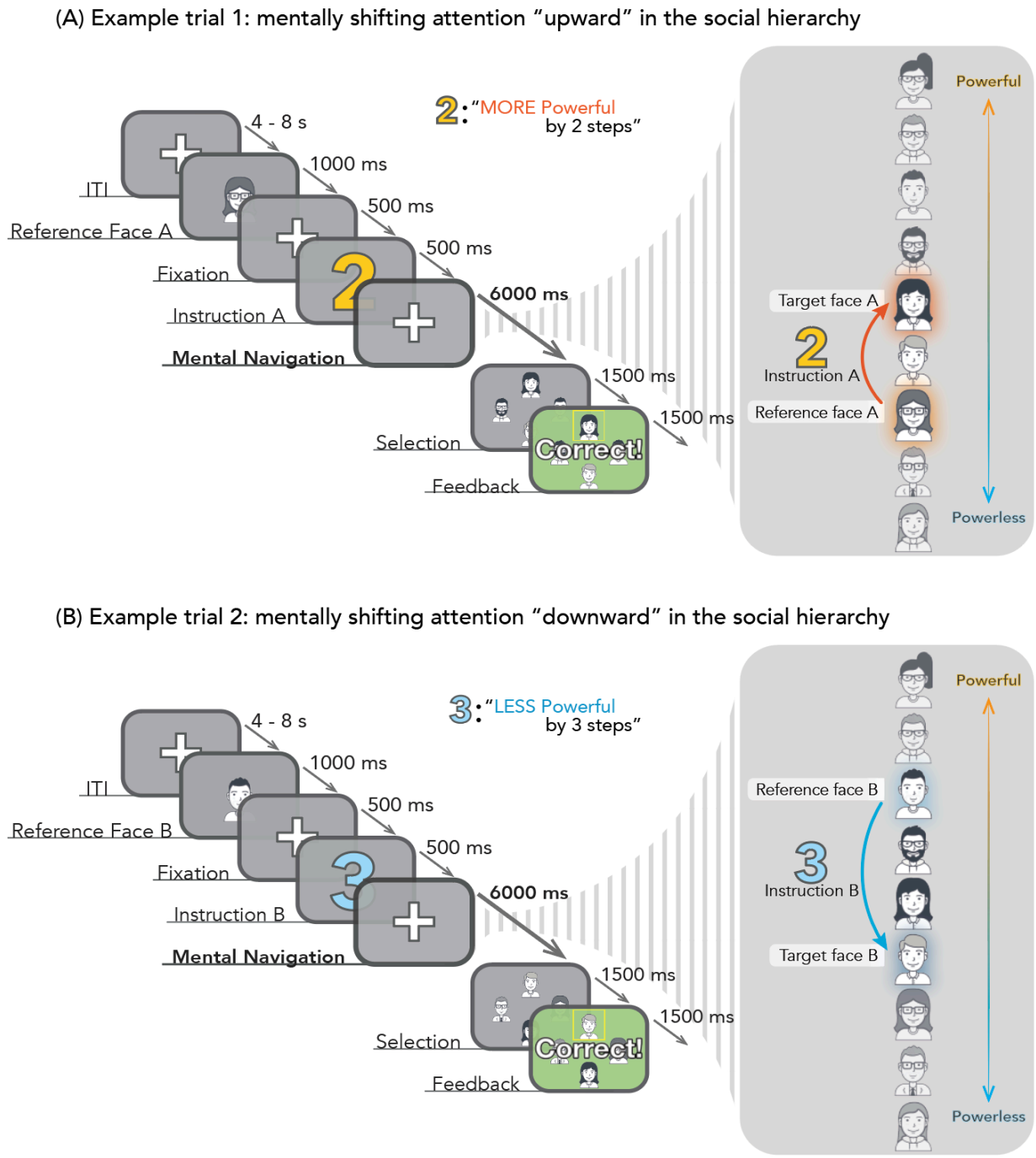

Figure 3. Mentally navigating the hierarchy (Task II). In each trial, participants were shown a reference face, followed by a colored number. The color of the number indicated whether they would need to determine the identity of a person more (**A**) or less (**B**) powerful than the reference person. The number itself indicated the number of steps participants would need to move from the reference person to reach the target person. Next, participants had 6 seconds to mentally "navigate" through the hierarchy and identify the target person. They were then prompted to select an answer from four options, and feedback was presented afterwards. The response options were displayed very briefly to force participants to arrive at their answers during the Mental Navigation portion of the trial.

Probing Mental Representations of the Social Hierarchy. After scanning, participants completed tasks probing their mental representation of the social hierarchy. Participants performed a drag-and-drop task where they arranged hierarchy members' faces into whatever configuration they thought best represented their relative power in the organization (Fig. 4; Methods). The majority of participants arranged faces vertically, with the most powerful people at the top; about one third of participants used alternative configurations (Fig. 4). Each participant's response on this spatial arrangement task informed how correspondences between the neural encoding of shifts of attention in space and in social knowledge were tested in cross-domain fMRI analyses.

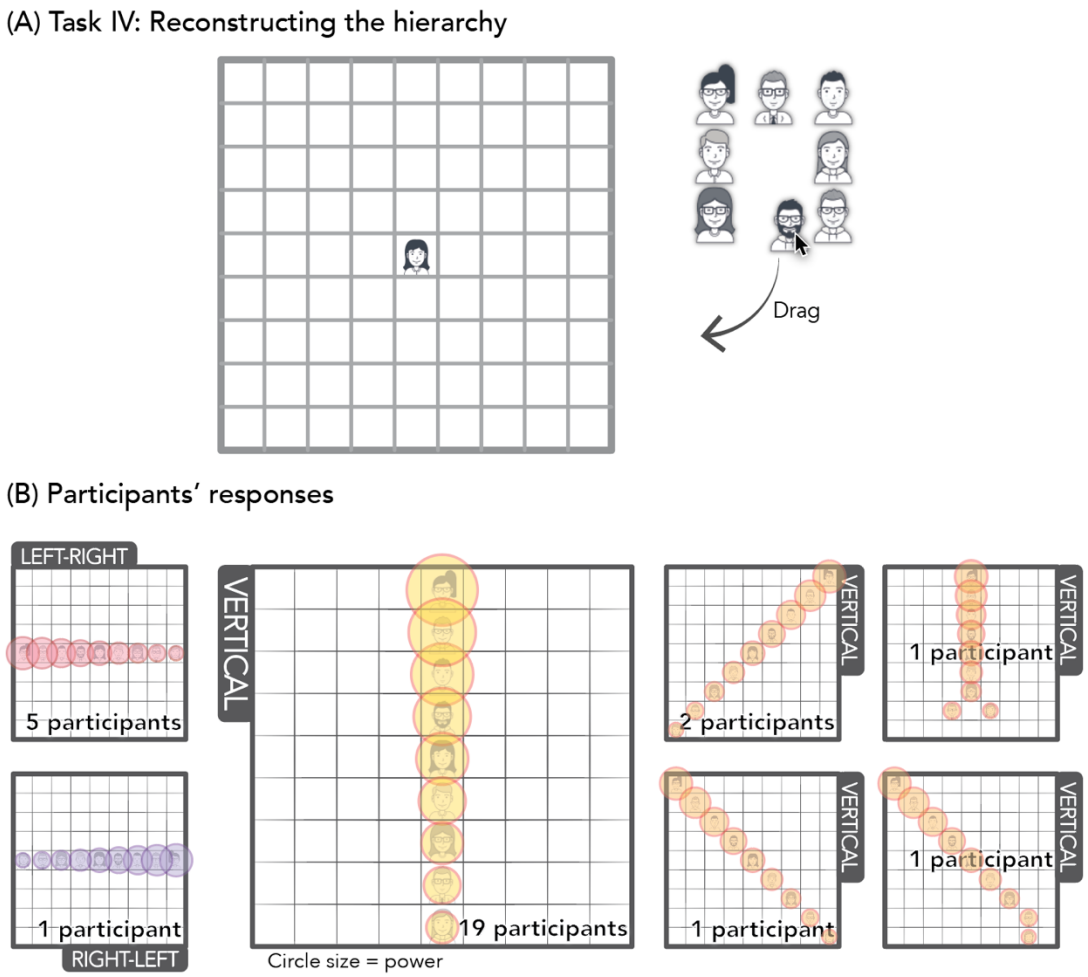

Figure 4. Reconstructing the hierarchy (Task IV). (**A**) In this task, participants arranged the faces on a 9x9 grid, in any configuration that they thought best represented these people's relative power. The middle person (i.e., the 5th person) in the hierarchy was already set at the center, and participants were asked to drag and drop the other 8 faces onto the grid. (**B**) All participants made linear arrangements, or close to linear arrangements, in this task. Based on the overall structure, 20 of the 30 participants arranged the social hierarchy in a top (most powerful) to bottom (least powerful) manner, and 10 participants arranged them left-to-right (N = 5), diagonally (N = 4), or right-to-left (N = 1). In cross-task analyses of fMRI data, we related multi-voxel response patterns during social navigation to vertical shifts of spatial attention if a participant responded with a vertical or diagonal representation here (i.e., more power corresponds to higher locations in space), and to horizontal shifts of spatial attention if they arranged the faces horizontally here (i.e., more power corresponds to more leftward or rightward locations in space, depending on whether a participant arranged faces in a left-right or right-left configuration, respectively).

Neuroimaging Data

SPL Encodes the Direction of Mental Navigation in Social Knowledge. We first tested if the SPL encoded the "direction" of shifts of attention within knowledge of social relations. To do this, we used data from the social hierarchy navigation task (see Fig. 3) to test if multivoxel response patterns evoked by mental navigation of the social hierarchy in the same "direction" (e.g., two trials where the participant had to determine who was "above" a reference person in the social hierarchy) would be more similar than those evoked by mental navigation in opposite "directions" in the social hierarchy. To quantify this, we computed an average pattern similarity score ($\bar{r}$) between multi-voxel response patterns evoked by shifts of attention in the same directions ($\bar{r}_{matching}$) and in opposite directions ($\bar{r}_{mismatching}$), then compared these two values (see Fig. 5a; Methods). As shown in Fig. 5a, multivoxel response patterns in both the left and right SPL signaled directions of attentional shifts in knowledge of social relations. In other words, $\bar{r}_{matching}$ was significantly greater than $\bar{r}_{mismatching}$ in both the left ($t(29) = 5.50$, $p = 3.15 \times 10^{-6}$) and right ($t(29) = 4.46$, $p = 5.68 \times 10^{-5}$) SPL (unless otherwise indicated, all reported *p*-values are one-tailed, given that all hypotheses were directional). That is, when mentally navigating the social hierarchy, the SPL encoded attentional shifts in the same direction with more similar multi-voxel patterns, compared to attentional shifts in different directions.

SPL Encodes the Direction of Attentional Shifts in External Space. We next used analogous data analytic methods to test for the encoding of the directions of overt attentional shifts in space (Fig. 5b). The SPL also encoded shifts of attention in external space: Spatial attentional shifts in matching directions evoked more similar multivoxel response patterns than those in mismatching directions, in both the right ($t(29) = 8.12$, $p = 3.00 \times 10^{-9}$) and left ($t(29) = 9.71$, $p = 6.41 \times 10^{-11}$) SPL.

Does the SPL Encode Directions of Attentional Shifts in Social Knowledge and External Space in the Same Way? Given that the directions of attentional shifts in external space and internal representations of social relations were both encoded in multivoxel response patterns in the SPL, we next tested if the SPL encoded shifts of attention in space and social knowledge in the same way. To determine which direction of spatial attentional shifts (up, down, left, right) corresponded to which direction of mental navigation in social knowledge (more powerful, less powerful) for each participant, we used the mappings between locations in space and the social hierarchy that the participant indicated in the post-scan drag-and-drop task (Fig. 4). For example, if, after scanning, a participant had arranged the faces vertically, with the most powerful people at the top, we tested if neural response patterns evoked when navigating towards more powerful people in the social hierarchy resembled those evoked when shifting attention upward in space. On the other hand, if a participant arranged the faces horizontally, with the most powerful people on the left, we tested if neural response patterns evoked when navigating towards more powerful people resembled those evoked when shifting attention leftward in space. As shown in Fig. 5c, in this cross-domain analysis, attentional shifts in matching directions across domains were not more similar than attentional shifts in mismatching directions in either the left or right SPL. These results suggest that overlapping but distinct codes in the SPL encode shifts of attention in social knowledge and in external space.

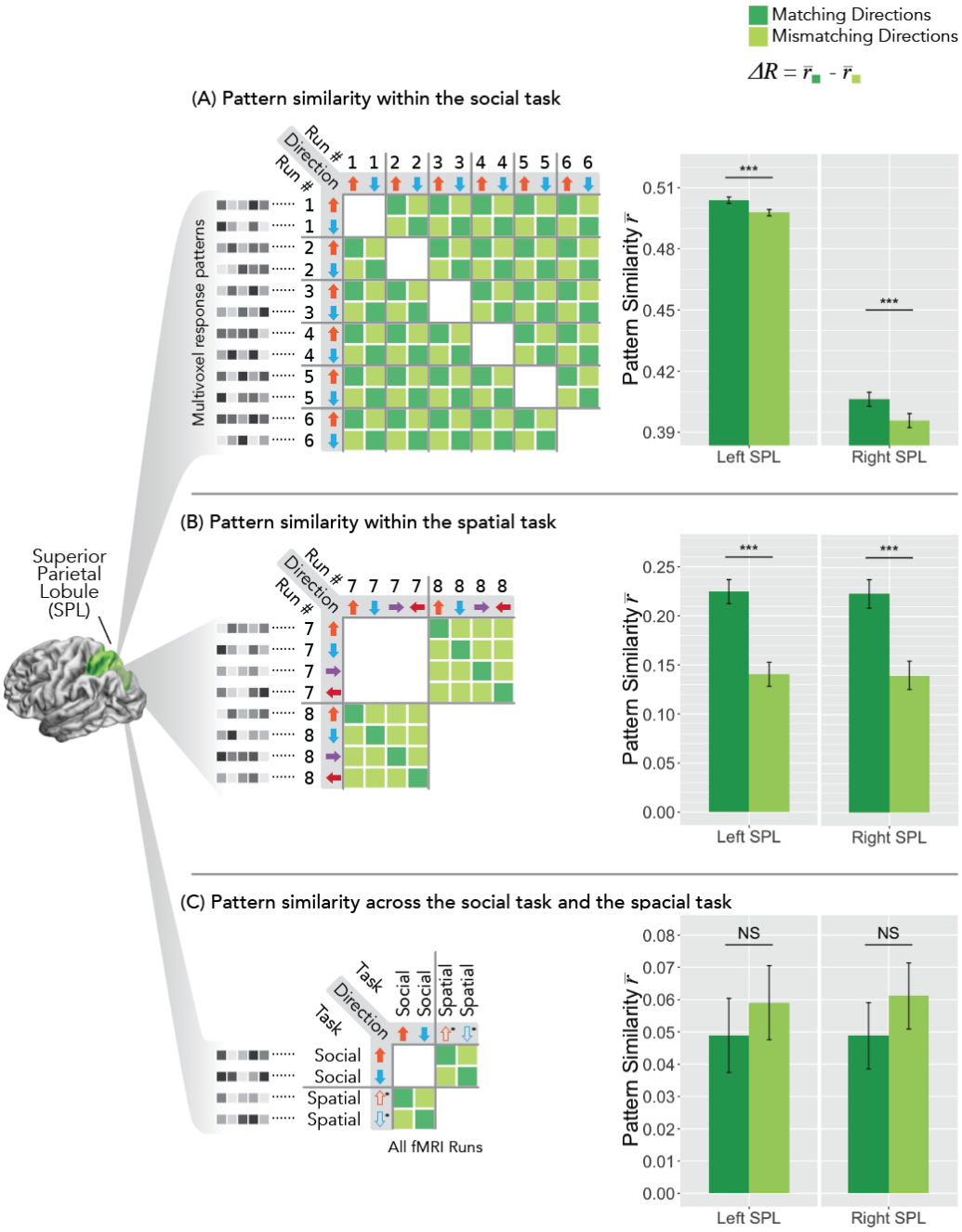

Figure 5. The SPL signals directions of attentional shifts in external space and within internal representations of social relations in overlapping but distinct multi-voxel response patterns. (A) Average multi-voxel response patterns in the SPL were calculated for shifts of attention "upward" and "downward" within knowledge of the social hierarchy for each participant. Mean similarities of multi-voxel response patterns in the SPL across runs ($\bar{r}$) were computed for each participant for matching (dark green) and mismatching (light green) directions of attentional shifts in social hierarchy knowledge. In the social navigation task, multi-voxel response patterns corresponding to matching directions of attentional shifts were significantly more similar to each other than those for mismatching directions, in both the left and right SPL. The same data analytic procedure was repeated to compare cross-run pattern similarities ($\bar{r}$) in the SPL for matching (dark green) and mismatching (light green) directions of attentional shifts in the spatial attention task (**B**) and across social and spatial tasks (**C**). In the spatial attention task (**B**), multi-voxel response patterns for matching directions of attentional shifts were more similar to each other than those for mismatching directions of attentional shifts in both the left and right SPL. When comparing multi-

voxel response patterns for corresponding “directions” across the social and spatial tasks (**C**), response patterns for matching directions (based on each participant’s responses to the social hierarchy reconstruction task) were not more similar than those for mismatching directions in either the left or right SPL. Thus, multi-voxel response patterns in the SPL encoded the direction of attentional shifts in external space (**B**) and within internal representations of social relations (**A**), but attentional shifts in the same direction* were not signified by the same patterns across social and spatial tasks (**C**). Error bars show 95% CIs. ***$p < .0001$

Do Any Brain Regions Encode Shifts of Attention in Social Knowledge and External Space in the Same Way?

We next conducted exploratory analyses testing if any brain regions encode shifts of attention in external space and knowledge of social relations in the same way. These analyses were conducted in two ways. First, we repeated the pattern similarity analyses that had been conducted in the SPL in anatomically defined parcels spanning all of cerebral cortex and subcortical gray matter structures (see Methods; Figs. 6a and 7). Second, we performed a whole-brain searchlight analysis (see Methods; Fig. 6b). As shown in Fig. 6, results from the parcel-based and searchlight approaches were largely consistent with one another.

Exploratory Whole-Brain Pattern Similarity Analyses. Results showed considerable overlap in the brain regions that encoded the direction of shifts of attention in both tasks, particularly in regions of parietal and occipital cortex (Figs. 6 and 7a). Additionally, several brain regions encoded the direction of shifts of attention only in social knowledge (e.g., left lateral prefrontal cortex, bilateral supramarginal gyri; Fig. 7b) or only in external space (e.g., hippocampus, parahippocampal cortex, entorhinal cortex; Fig. 7b). However, these exploratory whole-brain analyses did not identify any brain regions where matching directions of attentional shifts in social knowledge and external space were encoded similarly. Similar to our primary SPL-focused analyses, these results suggested while there is much overlap in the brain regions that encode shifts of attention in space and social knowledge, these mental operations are signaled by distinct neural population codes.

Exploratory Whole-Brain Cross-Domain Classification Analyses. To further test the above conclusion from pattern similarity analyses, we performed cross-domain classification analyses, both within the SPL and across the brain. As pattern similarity analyses are based on correlations of all voxels in a region, they weight values from each voxel in a pattern equally. In contrast, other approaches, such as cross-domain decoding using support vector machine (SVM) learning, which can also assess if encoding mechanisms are shared across domains[31,34], do so in a manner that gives higher weights to more relevant voxels[35]. Thus, within each brain region (in both the parcel-based and searchlight analyses), we performed cross-domain decoding analyses using SVM classification. SVM classifiers were trained to decode directions of spatial attentional shifts based on evoked neural response patterns, then tested on their ability to decode the direction of shifts of attention in social knowledge, and vice versa. As in the pattern similarity analyses, participants' responses on the post-scan drag-and-drop task determined which directions in space were matched to which “directions” in the social hierarchy. In the parcel-based analyses, there were no regions where cross-domain decoding was possible. In the searchlight analyses, a small region in early visual cortex was implicated (Fig. S3). Follow-up analyses indicated that for each task, within-task decoding of the direction of attentional shifts was possible in many of the same regions of parietal and occipital cortex that were implicated in the within-task pattern similarity analyses (see Supplement; Fig. S2). This

suggests that the lack of success in cross-domain decoding in these regions did not result from an inability of the algorithms to learn distinctions within each task, but rather, from the distinctions learned within tasks not generalizing across tasks. Taken together, these results cohere with those of our primary analyses, and suggest that regions of parietal and occipital cortex encode the direction of attentional shifts in both social knowledge and external space, but this information is encoded using largely distinct underlying mechanisms.

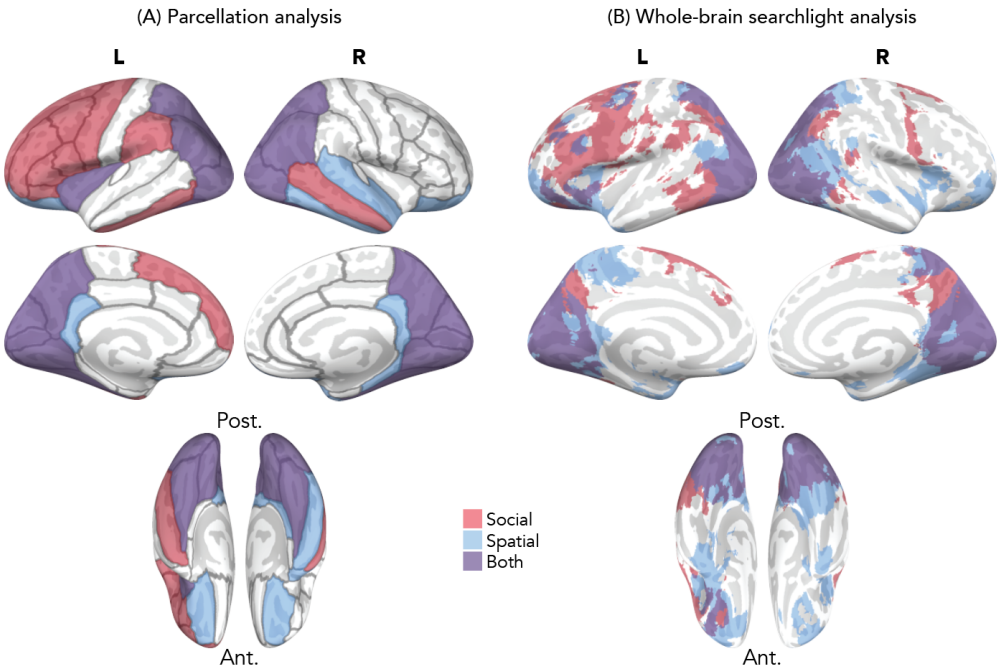

Figure 6. Shifting attention in space and social knowledge is supported by distinct codes carried in largely overlapping brain regions. Both panels show results of exploratory whole-brain pattern similarity analyses. Red indicates areas that encoded "directions" of mental navigation in social knowledge, blue indicates areas that encoded directions of attentional shifts in the spatial attention task, and purple indicates areas that encoded "directions" of attentional shifts in both tasks. A similar pattern of results was observed when analyses were conducted within anatomically defined parcels **(A)** and with a searchlight approach **(B)**. Cross-domain pattern similarity analysis did not reveal any brain regions that encoded attentional shifts similarly across tasks. All visualized results are significant at a corrected threshold of $p < .01$ (one-tailed).

	Social			Spatial	
	t(29)	FDR p		t(29)	FDR p
(A)					
Lateral occipital cortex (L)	6.05	4.1×10^{-5}	***	11.21	2.5×10^{-11}
Lateral occipital cortex (R)	5.88	4.1×10^{-5}	***	11.28	2.5×10^{-11}
Cuneus cortex (R)	5.61	5.8×10^{-5}	***	12.27	6.5×10^{-12}
Superior parietal lobule (L)	5.50	5.8×10^{-5}	***	9.71	5.9×10^{-10}
Fusiform gyrus (L)	5.31	6.6×10^{-5}	***	6.54	1.1×10^{-6}
Precuneus (R)	5.42	5.9×10^{-5}	***	4.46	2.6×10^{-4}
Lingual gyrus (L)	4.62	3.3×10^{-4}	***	9.25	1.6×10^{-9}
Inferior parietal lobule (L)	4.66	3.3×10^{-4}	***	4.81	1.1×10^{-4}
Pericalcarine cortex (R)	4.48	4.2×10^{-4}	***	15.64	4.2×10^{-14}
Superior parietal lobule (R)	4.46	4.2×10^{-4}	***	8.12	2.2×10^{-8}
Lingual gyrus (R)	4.29	6.0×10^{-4}	***	11.44	2.1×10^{-11}
Precuneus (L)	4.19	6.7×10^{-4}	***	4.70	1.4×10^{-4}
Pericalcarine cortex (L)	4.05	8.5×10^{-4}	***	12.85	3.1×10^{-12}
Cuneus cortex (L)	3.79	0.0016	**	11.76	1.4×10^{-11}
Inferior parietal lobule (R)	3.73	0.0016	**	5.03	6.7×10^{-5}
Insula (L)	3.76	0.0016	**	3.28	0.0042
Fusiform gyrus (R)	3.12	0.0056	**	7.51	9.4×10^{-8}
(B)					
Rostral middle frontal gyrus (L)	3.71	0.0016	*	2.39	0.028
Inferior temporal gyrus (L)	3.61	0.0019	*	2.65	0.019
Middle temporal gyrus (R)	3.49	0.0024	*	2.63	0.019
Putamen (R)	3.32	0.0034	*	2.43	0.026
Caudal middle frontal gyrus (R)	2.71	0.014	*	2.22	0.037
Lateral orbitofrontal gyrus (R)	2.66	0.015	*	3.29	0.0042
Superior temporal gyrus (R)	2.61	0.017	*	3.47	0.0029
Isthmus of cingulate cortex (L)	2.58	0.017	*	3.55	0.0025
Caudate nucleus (R)	2.40	0.023	*	2.36	0.029
Paracentral lobule (L)	2.38	0.023	*	2.18	0.039
Lateral orbitofrontal gyrus (L)	2.21	0.029	*	4.38	3.1×10^{-4}
(C)					
Superior frontal gyrus (L)	4.23	6.6×10^{-4}		0.73	0.28
Precentral gyrus (L)	4.06	8.5×10^{-4}		1.47	0.12
Triangular part of inferior frontal gyrus (L)	3.75	0.0016		0.87	0.24
Caudal middle frontal gyrus (L)	3.69	0.0016		1.44	0.12
Opercular part of inferior frontal gyrus (L)	3.52	0.0023		0.68	0.28
Orbital part of inferior frontal gyrus (L)	3.46	0.0025		2.00	0.054
Supramarginal gyrus (L)	2.97	0.0078		-0.61	-
Isthmus of cingulate cortex (R)	1.65	0.080		3.28	0.0042
Parahippocampal gyrus (R)	1.25	0.14		4.03	7.6×10^{-4}
Inferior temporal gyrus (R)	1.12	0.17		3.61	0.0022

Figure 7. Regions that encode the direction of attentional shifts in space, social knowledge, or both: Results of exploratory whole-brain pattern similarity analyses. This table shows brain regions where within each task, multi-voxel response patterns for matching directions of attentional shifts were significantly more similar to each other compared to those for mismatching directions (i.e., this figure provides additional details for the results shown in Fig. 6a). The upper two panels, (**A**) and (**B**), contain brain regions where multivoxel response patterns encoded the direction of attentional shifts (i.e., where $\Delta R > 0$) in both social and spatial tasks at significance levels of $p < .01$ and $p < .05$, respectively (FDR-corrected). Asterisks and shades of purple in the middle column show the significance levels that *p*-values met in both social and spatial tasks: *$p < .05$ (FDR-corrected) on both tasks, **$p < .01$ (FDR-corrected) on both tasks, ***$p < .001$ (FDR-corrected) on both tasks. Areas in occipital and parietal cortex most robustly encoded attentional shifts in both social and spatial tasks. The lower panel (**C**) contains brain regions that encode attentional shift directions in either the social or spatial task, but not both. Brain regions where FDR-corrected *p*-values > 0.01 in both tasks are shown in Fig. S1.

Discussion

Here, we found that the SPL, a region with a long-established role in directing shifts of attention in external space, also encodes shifts of attention in internal representations of social relations. Multivoxel response patterns in this region signaled whether participants were mentally navigating "upward" or "downward" in knowledge of a social hierarchy. Convergent results were obtained using both pattern similarity and decoding analyses. In line with prior work, multivoxel response patterns in the SPL also encoded the direction of shifts of attention in external space. Furthermore, exploratory whole-brain analyses implicated additional regions of the posterior parietal cortex and occipital cortex in encoding direction of attentional shifts in space and social knowledge. Thus, shifts of attention in internal representations of social knowledge and in external space are encoded by partially shared brain regions. At the same time, cross-domain pattern similarity and decoding analyses illustrated that shifts of attention in space and social knowledge were encoded in functionally independent response patterns in the SPL and in most other implicated brain regions. These results shed light on the neural basis of reasoning about social relations, and suggest that shifts of attention in space and in social knowledge are supported by distinct neural mechanisms within largely overlapping brain regions.

Multivoxel Patterns in the SPL Encode Shifts of Attention in Social Knowledge

Pattern similarity analyses were used to characterize the mental operations within the SPL that are involved in reasoning about social relations. This data analytic approach was chosen so that we could rule out the possibility that SPL was merely recruited when processing social hierarchy knowledge without actually encoding mental operations within such knowledge. For example, SPL responses could have been organized only by task difficulty (e.g., with pattern similarity and/or response magnitude only reflecting the similarity of the amount of "distance traveled" on particular trials). That possibility would have been consistent with prior work where more difficult social status comparisons evoked stronger responses in the vicinity of this region[22]. However, we found that the multivoxel pattern similarity in the SPL reflected the "direction" in which people were navigating in their knowledge of the social hierarchy. This suggests that the SPL directly encodes the mental navigation of social knowledge.

Why Would the SPL Encode Shifts of Attention in Space and Social Knowledge?

The SPL has been widely implicated in shifting attention in the external world[28,36]. The fact that it also encoded the "direction" of attentional shifts within social knowledge is consistent with behavioral and linguistic evidence that operating on space and social knowledge may involve common representational resources[37–39]. Furthermore, it has been suggested that over the course of human brain evolution, regions of the PPC expanded[40] and formed new connections[41] as they came to support increasingly abstract aspects of social cognition[18,20]. This is consistent with the substantial overlap we found in the brain regions that encoded spatial and social shifts of attention (Figs. 6 and 7a), which could reflect similarities in the processes involved in these mental operations. In other words, shifts of attention in social knowledge may be supported by brain regions with a long-standing role in spatial attention because characteristics of such brain regions (e.g., their internal structure and connectivity with other

brain regions) afford operations common to directing shifts of attention in space and in knowledge of social relations (e.g., coordinate transformations).

Shifting Attention in Space and Social Knowledge is Supported by Distinct Codes within Overlapping Brain Regions

Relatively little is known about how the brain supports mental navigation of social relations. Further, most relevant past research has used mass univariate analyses, and thus, would be insensitive to effects carried in multivoxel response patterns. Therefore, we performed exploratory whole-brain analyses in addition to our main analyses on the SPL. The cross-domain pattern similarity analyses found several regions encoding shifts of attention in both space and social knowledge, including superior and inferior aspects of lateral PPC, as well as the precuneus, fusiform cortex, and regions of occipital cortex (Figs. 6 and 7a).

Considerable efforts were taken to avoid biasing participants toward thinking about the hierarchy spatially. For example, instructions avoided the use of spatial language and faces were never shown above, below or beside one another during the learning task (Fig. 2). In the response selection phase of the social navigation task, faces (Fig. 3) and response button choices were configured such that there was no systematic relationship between response location–on the screen or response pad–and faces' positions in the hierarchy. Additionally, participants were not informed of the spatial attention task until after all social navigation trials had been completed to avoid inducing spurious similarities between how participants would perform the "social navigation" task and the spatial attention task. Further, participants were instructed to focus on the center of the screen during the social task, and the results of cross-domain analyses and an additional eye-tracking study (see Supplement) suggest that neural activity during the social task did not reflect eye movements. Still, there was considerable overlap in the brain regions encoding social and spatial operations. As discussed earlier, this is consistent with a large body of research linking mental representations of space and social hierarchy knowledge[37–39]. Moreover, participants' post-scan responses indicated that they tended to ascribe a systematic spatial organization to people's relative locations in the status hierarchy (usually a vertical organization, see Fig. 4), suggesting an association between spatial and social contents in their internal representations.

At the same time, brain regions encoding both social and spatial mental operations appeared to encode them in largely independent coding schemes. Using both anatomically defined parcels and searchlights, the cross-domain pattern similarity analyses did not identify any region that encoded attentional shifts in social knowledge and space in the same way. As described in the Results section, we complemented these cross-domain pattern similarity analyses with exploratory decoding analyses to provide sensitivity to distinctions driven by only a subset of voxels within a region. Whole-brain cross-domain decoding analyses conducted within anatomically defined parcels identified no region where directions of attentional shifts in space and social knowledge were encoded similarly, consistent with findings from whole-brain pattern similarity analyses. Only when conducted using a searchlight approach, cross-domain decoding analyses identified a cluster in early visual cortex where classification accuracy was above chance (Fig. S3). As this region encodes the direction of imagined visual motion[42], it is possible that this reflects mental imagery during the social navigation task, such as imagining moving up or down in the social hierarchy. However, given the exploratory nature of these

results, and that significant results were not observed in other exploratory whole-brain analyses (i.e., parcel-based or searchlight-based cross-domain pattern similarity analyses; parcel-based decoding analyses), we interpret these results with caution, particularly relative to the results of other analyses of these data, which showed considerable convergence across diverse data analytic approaches (e.g., parcel-based and searchlight-based within- and cross-domain pattern similarity analyses). We present them here for transparency and to potentially inform future research.

Despite considerable overlap in the brain regions involved in shifting attention in space and social status knowledge, in all but one exploratory analysis, we did not find evidence for a common encoding of these two mental operations. Notably, this was not because these analyses lacked sensitivity to within-domain information: Both pattern similarity analyses and classification analyses consistently identified regions, including regions with long-established roles in spatial attention, that signaled directions of attentional shifts within each domain (see Fig. 6 and Supplement). Nonetheless, these reliable within-domain distinctions did not tend to generalize across domains. Thus, the current results suggest that shifting attention in space and social knowledge relies on overlapping but distinct coding schemes. As mentioned above, encoding shifts of attention in abstract knowledge of social relations may take place in regions that also encode shifts of attention in space because of partially shared processing demands for these mental operations.

Several psychological factors may have contributed to neural response patterns in the social task being distinct from those in the spatial task. For example, it is possible that participants' post-scan arrangements of the social hierarchy, which we used to inform the cross-domain correspondence of directions in our analyses, do not accurately reflect participants' actual mental representations during the social navigation task. Rather, their actual representations may have been hard to draw or may have changed over time, or it may be difficult for participants to accurately introspect about how these mental representations are organized. Alternatively, although the SPL encodes directions of attentional shifts in both social and spatial domains, it may encode the operation of shifting attention upward in space completely differently from shifting attention "upward" abstractly in social hierarchy knowledge. In other words, the two corresponding "directions" of attentional shifts across domains may be represented distinctly by the brain, even though participants mapped "higher" positions in the hierarchy to higher positions in space (Fig. 4), and even though we describe both as "upward" in common parlance. That is to say, "directions" within representational spaces encoded by the SPL (and other brain regions) may not have straightforward correspondence across the two domains. This could be why distinct coding schemes were found for mental operations across the social and spatial domains in most of our analyses.

Limitations and Future Directions

The current study focused on a single kind of social relational knowledge (a status hierarchy). This type of knowledge afforded a clear hypothesis for the nature of cross-domain mappings, which was necessary to examine if any brain regions encode shifts of attention in space and social knowledge in the same way. Thus, in this study, social hierarchy knowledge served as a practical place to start for examining these questions. Future research can extend

these findings by examining how the brain supports mental navigation of other kinds of social relational knowledge, and how this compares to other kinds of mental operations.

We also used very constrained social tasks. Using a highly controlled social navigation task allowed us to ascertain when participants were mentally traveling “up” or “down” in their knowledge of the social hierarchy, which was necessary to extract corresponding neural response patterns. Furthermore, a few considerations suggest the efficacy of the hierarchy learning task in eliciting social processing. First, this task was adapted from a prior study[8], in which providing a social context merely through verbal instructions evoked distinct neural processes for learning a social, compared to non-social, hierarchies. Second, this task resembles real-world social status hierarchy learning in that in both cases, status knowledge largely emerges from observing relations between pairs of individuals and using transitive inference to construct mental representations of social hierarchies[1,5]. However, both tasks differed in many respects from how social relations are learned and navigated in the real-world. Other researchers[43] have developed ecologically valid paradigms to examine how the brain tracks subjective affiliation and power during social interactions. Future research could extend the current work using similarly ecologically valid paradigms, such as naturalistic social interactions or veridical social knowledge.

It would also be interesting to compare the neural response patterns evoked when mentally navigating social and non-social relational knowledge (e.g., shifting attention between memories of different spatial locations, rather than between different spatial locations themselves; other forms of relational knowledge). Past research has shown that shifting attention along the mental number line (as in mental arithmetic) evokes patterns similar to shifting attention in external space[29]. The asymmetry between those results and the results of the current study suggests that shifts of attention in knowledge of social relations and other forms of relations (e.g., numerical magnitude) would be encoded differently, but future studies could more directly tease these phenomena apart. Regardless of the results of such studies, the capacity to perform transitive inferences on relational knowledge may be fundamentally social in origin: This capacity is thought to be particularly important for highly social species living in large groups containing linear status hierarchies[1,44] and may have evolved to support social complexity[45].

Conclusions

Humans have a sophisticated capacity to learn and reason about social relations. The current results demonstrate considerable overlap in the brain regions that encode shifts of attention in the space around oneself and in internal representations of social knowledge. At the same time, within these overlapping brain regions, multivoxel codes within each domain (space, social knowledge) were largely independent. These findings provide insight into the neural basis of reasoning about knowledge of social relations. More generally, these findings also highlight the value of examining distributed response patterns for elucidating the diverse mental operations that brain regions encode, and how they relate to one another.

Acknowledgments

The authors wish to thank Ryan Hyon, Savanna Gharibian, and Georgia Perris for assistance with data collection, as well as those who provided helpful feedback on the initial presentation of these results at the 2018 meeting of the Social and Affective Neuroscience Society. This work was supported by funding from a UCLA Academic Senate Council on Research Faculty Research Grant, a Sloan Foundation Research Fellowship, and the UCLA Center for Cognitive Neuroscience.

Author Contributions

Conceptualization, C.P.; Methodology, C.P. and M.D.; Software, M.D.; Formal Analysis, M.D.; Investigation, M.D. and R.B.; Resources, C.P.; Data curation, M.D.; Writing - Original Draft, C.P. and M.D.; Writing - Review & Editing, C.P., M.D., and R.B., Visualization, M.D.; Supervision, C.P.; Project Administration, C.P.; Funding Acquisition, C.P.

Declaration of Interests

The authors declare no competing interests.

Data and Code Availability

The data and code in this study are available at https://github.com/CSNLab/hierarchy-study.

Methods

Overview of Paradigm and Analysis

The overall procedure of this study is illustrated in Fig. 1, and is outlined below and in the Results section. Detailed descriptions of each component of the study are provided in the Procedure section. Participants attended a behavioral session (Session 1) and a subsequent fMRI session (Session 2). Qualifying participants completed the fMRI session 2.5 hours - 5.85 days after the initial behavioral session (*M* = 2.71 days).

In Session 1, participants first learned a linear social hierarchy of 9 people through trial and error (Task I, adapted from a prior study[8]), and then performed a social navigation task (Task II), where they were repeatedly asked to shift attention within their knowledge of the social hierarchy (i.e., to find a target person in the hierarchy, given another reference person, a distance and "direction"). Participants who had satisfactory accuracy (>70%) in Task II of Session 1 were invited to participate in Session 2, which involved performing the social navigation task (Task II) again in the MRI scanner.

Upon arriving for Session 2, participants first completed short versions of Tasks I and II again outside of the scanner to refresh their memories of the social hierarchy and of the social navigation task, respectively. Then, while being scanned, participants performed the same social navigation task (Task II), followed by a task involving overt shifts of attention (i.e., eye movements) in different directions in external space (Task III, adapted from a prior study[29]). After scanning, participants' mental representations of the social hierarchy were probed by asking them to rearrange the faces in whatever spatial configuration they thought best represented the people's relative power (Task IV; Fig. 4). These responses were used to determine which directions of a participant's attentional shifts should be matched across domains for each participant in pattern similarity analyses and cross-domain decoding.

Importantly, in the experimental materials, we avoided instructions, stimulus arrangements, and response formats that might lead participants to think about the social hierarchy in any particular spatial terms, or to visualize the hierarchy in any particular way. In the instructions, for example, we described the hierarchy as an "organization" where one person is a certain number of "degrees" more or less "powerful" than another person. During the entire experiment, faces from the hierarchy were displayed in one of three ways: a) alone at the center of the screen, or randomly arranged b) in a group of four shaped like a diamond (see the last two screens in Fig. 3), or c) in a group of 8 shaped like a square (a 3 x 3 grid with an empty central cell, see Fig. 4a). Thus, participants never saw the social hierarchy arranged in a linear configuration during the experiment; therefore, any linear arrangement of the hierarchy in their mental representations would have been derived on their own.

Participants

Participants were recruited at the University of California, Los Angeles. 31 participants qualified for the fMRI study and completed all tasks after passing the accuracy criterion (70%) in Session 1, as described later in this section. Data from one participant were discarded from analysis due to excessive movements during the fMRI scan (>3 mm in all runs). All participants in the final sample were right-handed, native English speakers, and between the ages of 18 and 31 (M_{age} = 20.17, 11 males, 19 females). Participants met standard MRI safety criteria, and

provided informed consent in accordance with the policies of the UCLA Institutional Review Board.

Stimuli

We obtained 18 face images (9 white males and 9 white females) from the Chicago Face Database[46] with approximately matched ratings of perceived age, affect, dominance, and attractiveness. The gender of the social hierarchy was matched to the gender of the participant. To facilitate participants' ability to individuate and remember the people about whom they learned, and thus, reduce the difficulty of tasks, images were edited to make the T-shirts different colors (rather than all gray) and names were added below the faces. The names were selected from the most popular given names for babies born during the 1990s, as published by the US Social Security Administration (https://www.ssa.gov/oact/babynames/decades/names1990s.html). Assignment of faces to positions in the social hierarchy was randomized across participants.

Procedure

All paradigms described below, except the social hierarchy reconstruction task (Task IV), were created using Python 2.7 with PsychoPy 1.85[47]. The social hierarchy reconstruction task was created with HTML, CSS, and JavaScript.

Task I: Learning the social hierarchy. In the learning task (see Fig. 2, adapted from a prior study[8]), participants learned each person's relative position in the social hierarchy through trial and error. The following instructions were used to verbally inform the participants of the content of this task:

"Our experiment looks at how people learn social information. There will be two parts of the experiment in today's session, and your performance in both of them will determine whether you are eligible for the second fMRI session. In the first part, you will be looking at pictures of nine people who all belong to an organization, and your goal will be to figure out who has more power in this organization and who has less power. You will see pictures of two people in a trial, and all you have to do is choose the person who has more power in the organization. You will be using trial and error... So you'll have to guess a little bit in the beginning, and you will get a score that tells you if you chose right or wrong. It might seem a little difficult at first, but most people who keep trying do figure it out… So just try your best."

Participants were also informed that they would use the information they learned from this task in later parts of the experiment: *"When you're done with this part of the experiment, you will use what you learn about the people from this part to complete the second part today and the rest of the experiment. That means your performance in the rest of the experiment depends on how well you learn the information from this part… So try to get the highest score possible in this part."*

On each trial, participants saw 2 people and selected who had more power in the organization. Each trial started with a 1000 ms fixation cross, and then two faces were presented sequentially on the screen in a random order, with the presentation of each face lasting 1500 ms, with a 50 ms interval in between. Participants were then prompted to select the more powerful person in the pair by pressing and holding down the space bar, then releasing it when their preferred response choice appeared. Once the participant started to press the

spacebar, the two face images were cycled through again in the same order in which they had been displayed at the start of the trial. While participants held down the spacebar, each face image was presented for 1000 ms at a time on the screen. An expanding circle was displayed behind each face; the size of this circle indicated how long the current face would remain on the screen. Participants then selected one of the faces by releasing the space bar when they saw the person they intended to choose. After a choice was made, the selected face was highlighted together with a score (e.g. “+10 points” after a correct answer or “-10 points” after an incorrect answer) displayed on the screen for 1000 ms. Afterwards, the two faces were presented again with the correct answers (i.e., who was relatively powerful and who was relatively powerless in the organization; see Fig. 2), and participants pressed “N” to advance to the next trial. During the entire task, stimuli and response options were not mapped to any spatial configuration (i.e., only one face was shown at a time at the center of the screen, instead of two faces appearing in a horizontal or vertical configuration; response choices between two options were made by releasing the spacebar when the preferred option appeared, rather than by pressing two response keys which would form a spatial configuration on the keyboard). This was to avoid inducing any extraneous mappings between positions in space and in the social hierarchy, and to avoid biasing participants towards thinking about the hierarchy in spatial terms any more than they typically would.

The task contained 2 types of trials, adjacent trials and non-adjacent trials. Each block in this task began with 16 adjacent trials in which the pairs of people presented were adjacent to each other in the social hierarchy (e.g., the most powerful person and the second most powerful person). Every possible adjacent pair of people was presented twice at the start of the block in a randomized order. The next portion of the block contained non-adjacent trials, in which the pairs of people presented were farther away from each other in the hierarchy (specifically, either 2 or 3 steps away from each other, e.g., the most powerful person and the third most powerful person). Thus, responding correctly on non-adjacent trials required participants to perform transitive inference on the social hierarchy knowledge that they had acquired during the first part of the block. During these non-adjacent trials, participants were also asked to rate how confident they were in their answer before receiving feedback about its accuracy. Each adjacent trial of a block was worth 10 points, each non-adjacent trial was worth 100 points, and participants were asked to try to get the highest score possible.

Participants first performed this task at the beginning of Session 1, in which a minimum of 12 blocks were displayed. If their accuracy was unsatisfactory after 12 blocks – i.e., they had either more than 3 errors in the adjacent trials of the last 3 blocks, or more than 2 errors in the non-adjacent portions of the last 3 blocks – the task would keep going until a maximum of 14 blocks had been completed.

At the beginning of Session 2, participants performed this task again outside of the scanner to refresh their memories of the social hierarchy. A minimum of 4 blocks were displayed during Session 2, and the task continued until either they had reached a perfect score in all of the last 3 blocks, or it had taken over 30 minutes.

Task II: Mentally navigating the social hierarchy. In this task, participants mentally shifted attention within the social hierarchy that they had learned in Task I in order to answer a series of questions (see Fig. 3). Each question asked the participant to find a target person in the hierarchy, given another person who served as a point of reference, a distance, and a

direction. In each trial, the participant first saw the reference face from the hierarchy (1000 ms), and then a fixation cross (500 ms), followed by a colored number (500 ms). The number's magnitude indicated the distance between the reference and the target person, and it was colored in either orange or blue, indicating the target person was either more or less powerful than the reference person. This mapping between color and direction was randomized across participants (i.e., for some participants, orange meant more powerful and blue meant less powerful; for the others, the opposite was true). Next, participants had 6000 ms to mentally navigate through their knowledge of the social hierarchy to determine their answers while a fixation cross was displayed. Afterwards, four faces were presented as response options, and the participants had 1500 ms to choose an answer. These faces were presented in a diamond configuration (Fig. 3) and while being scanned, participants responded using a response box with buttons configured in the same diamond shape (Current Designs 4-Button Diamond Fiber Optic Response Pad). Rather than using linear configurations of response choices and/or response buttons where faces and/or button options appeared side-by-side, we used this diamond configuration to avoid experimentally biasing participants to think more about the social hierarchy navigation task in any particular spatial manner. The correct answer was equally likely to appear in each of the four possible positions on the screen, and the incorrect response options were three other faces that were as close as possible to the correct answer in the hierarchy. After the response option screen had been presented, feedback was displayed for 1500 ms. If the participant did not make a choice in time, they were prompted to respond faster next time; otherwise, they were informed whether they were right or wrong while their selected answer was highlighted. Finally, a fixation cross was displayed between trials. During the behavioral session, each inter-trial interval (ITI) was 2 or 3 seconds. During the fMRI session, ITIs ranged from 4-8 seconds and were determined using NeuroDesign[48], as described in more detail below.

Participants completed 3 blocks of this task at the end of Session 1, one block at the beginning of Session 2 right before they were scanned, and 6 runs inside the scanner. Each block contained 24 unique trials. For the blocks completed outside of the scanner, trial order was randomized. For those runs completed inside of the scanner, the stimuli order and ITI durations were generated using the Python package NeuroDesign[48] to optimize our power to detect differences between attentional shifts in contrasting directions within each subject. The total duration was 402 seconds for 5 out of the 6 runs and 404 seconds for one run; run order was counterbalanced across participants using a Latin square design.

To ensure that participants were engaged in mental navigation during the 6000-ms period (which thus ensures the sensitivity of our fMRI analyses), we took efforts to minimize the possibility that they would be able to 1) predict the direction of the correct answer immediately upon seeing the reference face (and thus begin mental navigation away from the anchor face before the designated part of the trial), or 2) wait until after the response options appeared and then start figuring out the answer, or 3) reach an answer easily without navigating through the hierarchy. First, all of the reference faces and target faces were selected among the third to seventh people in the 9-person hierarchy, and the only possible distances between the two relevant faces were 2, 3, or 4 steps. Each of the 24 possible trials was presented once in each run. This avoided situations where a face from either extreme end of the hierarchy was used as the reference, in which case participants would have been able to ascertain the direction of

navigation before seeing the instruction regarding direction (and thus begin mentally navigating away from the anchor face in that direction); it also avoided situations where participants could directly "jump" to an answer without needing to mentally navigate their knowledge of the hierarchy, e.g., when the answer was either around the top or bottom of the hierarchy, or adjacent to the reference face. Second, the four candidate response options were presented very briefly, so that participants were merely able to quickly skim the options and indicate the answer they had already figured out before seeing these options. This ensured that participants would have mentally navigated through their knowledge of the hierarchy before the options appeared and not after; if they did not already have an answer in mind before seeing the options, they would not have enough time to figure it out when the options appeared. This aspect of the experimental design facilitated our ability to ascertain which time points reflected shifts of attention in particular directions in social hierarchy knowledge.

Task III: Shifting attention spatially. This task was adapted from previous related research[29]. Participants were asked to follow a white target cross on the screen with their eyes. The target cross appeared at the center of the screen at the beginning of each trial, then moved up, down, left or right in 4 steps, and finally went back to the center during the ITI. Each step within a trial lasted for the same amount of time (500, 750 or 1000 ms), during which the target cross moved approximately 1.25° of visual angle, with a random jitter in the range of ± 0.42º in the direction of movement, and a random jitter of ± 0.26° perpendicular to the direction of movement. At the end of each trial, the target cross stayed in its last position for the first 3 seconds of the ITI to prevent inducing eye movements back to the center of the screen immediately after attentional shifts in the direction of interest. Then the target cross returned to the center and remained there for the second half of the ITI. The ITI durations were 6, 7 or 8 seconds. As in Task II, the order and timing of trials was determined using NeuroDesign[48] in order to optimize power to detect differences between shifts of attention in different directions with each participant. Each run contained 40 trials and participants completed 2 runs, which lasted 409 seconds and 396 seconds.

Task IV: Reconstructing the social hierarchy. At the end of the experiment, we probed participants' mental representations of the social hierarchy. Responses to the hierarchy reconstruction task (Task IV) were used to determine, for each individual participant, which directions of attentional shifts in the spatial task (Task III) corresponded to which directions of attentional shifts in the social task (Task II). In Task IV, we showed participants a 9 x 9 grid that was empty except for the central cell (see Fig. 4a). The middle person (i.e., the 5th person) in each participant's social hierarchy appeared at the center of the grid, and the other 8 people appeared outside the grid in a random configuration. Participants were asked to drag the 8 people to the grid and arrange them in any configuration that they thought best represented these people's relative power. Each participant's answer was coded according to the way that they arranged the faces (see Fig. 4b).

fMRI Data Acquisition

MRI data were collected using a Siemens 3 Tesla Prisma Fit MRI Scanner with a 32-channel head coil. Functional scans were obtained using a gradient echo sequence with 60 transverse interleaved slices (TR = 1000 ms, TE = 29 ms, flip angle = 52°, FOV = 198 mm, 2.4 mm isotropic voxels). For each subject, two field map scans were obtained before functional

scans began in order to correct for the effects of field inhomogeneity. Participants used a 4-button diamond-shaped response box (Current Designs 4-Button Diamond Fiber Optic Response Pad) to make choices during the social navigation task. Finally, a T1-weighted (T1w) MPRAGE sequence was acquired after the functional runs (192 slices, TR = 2000 ms, TE = 2.52 ms, flip angle = 12°, FOV = 256 mm, 1 mm isotropic voxels).

Data Analysis

Image preprocessing. Preprocessing was performed using fMRIPrep 1.1.8[49], which is based on Nipype 1.1.3[50] and Nilearn 0.4.2[51]. Internal operations of fMRIPrep also use ANTs 2.1.0[52], AFNI 17.2.07[53], FSL 5.0.10[54], and FreeSurfer 6.0.0[55]. The descriptions of anatomical and functional data preprocessing provided in the following two sections are based on the recommended citation boilerplate text that is generated by fMRIPrep and released under a CC0 license, with the intention that researchers reuse the text to facilitate clear and consistent descriptions of preprocessing steps, thereby enhancing the reproducibility of studies.

Anatomical data preprocessing. The T1w image was corrected for intensity non-uniformity using N4BiasFieldCorrection in ANTs 2.1.0, and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped using antsBrainExtraction.sh (ANTs 2.1.0), using OASIS as target template. Brain surfaces were reconstructed using recon-all from FreeSurfer 6.0.0, and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle[56]. Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.1.0), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.10).

Brain parcellation. Automated brain parcellation methods were used to identify our key regions of interest (right and left SPL) and to delineate other regions to be used in exploratory analyses. As mentioned in the preceding section, cortical surfaces for each participant were reconstructed by applying FreeSurfer's anatomical parcellation algorithm to each participant's T1w image. This process entails removal of non-brain tissue, automated segmentation of the cerebral cortex, subcortical white matter, brainstem, cerebellum, and deep gray matter structures, creating of a model of the participant's cortical surface, and automatically parcellating this cortical surface model into discrete regions based on the participant's cortical folding patterns. The Desikan-Killiany-Tourville (DKT) atlas[57], implemented in FreeSurfer 6.0.0, was used to assign labels to cortical brain regions in each participant's native space. This gyral-based atlas largely uses sulci to define the boundaries of adjacent gyri; as such, a particular gyral label in this atlas corresponds to both the gyrus itself and the adjacent banks of its limiting sulci. Using this procedure, each participant's cerebral cortex was parcellated into 31 regions in each hemisphere. Six subcortical gray matter structures delineated using the tissue segmentation procedure described above were also retained for exploratory analyses (i.e., hippocampus, amygdala, caudate nucleus, palladium, putamen, nucleus accumbens). Thus, while our primary analyses concerned the right and left SPL, exploratory analyses spanned 74 anatomically defined brain regions.

Functional data preprocessing. For each of the 9 blood oxygen level-dependent (BOLD) runs per subject (across all tasks), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A deformation field to correct for susceptibility distortions was estimated based on two echo-planar imaging (EPI) references with opposing phase-encoding directions, using 3dQwarp (AFNI 17.2.07). Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer 6.0.0), which implements boundary-based registration. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) were estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.10). The BOLD time-series were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as preprocessed BOLD. Automatic removal of motion artifacts using independent component analysis (ICA-AROMA) was performed on the preprocessed BOLD in MNI space time-series after spatial smoothing with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum). Several confounding time-series were calculated based on the preprocessed BOLD, including framewise displacement (FD), and three region-wise global signals extracted within the CSF, the WM, and the whole-brain masks. FD was calculated for each functional run using implementations in Nipype. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (CompCor). Principal components were estimated after high-pass filtering the preprocessed BOLD time-series (using a discrete cosine filter with 128s cut-off) for anatomical CompCor (aCompCor). A subcortical mask was obtained by heavily eroding the brain mask, which ensured it does not include cortical GM regions. Six aCompCor components were calculated within the intersection of the aforementioned mask and the union of CSF and WM masks calculated in T1w space, after their projection to the native space of each functional run (using the inverse BOLD-to-T1w transformation). In addition, non-steady volumes identified by fMRIPrep were removed from data prior to subsequent analyses.

First-level analysis. To conduct run-wise analyses, a general linear model (GLM) was fit to each run of the fMRI data using Nipype[50] to estimate the BOLD response evoked for each direction of attentional shift in social knowledge (Task II) and external space (Task III). The following confounding variables (estimated during the preprocessing steps described in the preceding section) were included in the model as nuisance regressors: three translational motion parameters, three rotational motion parameters, three global signals extracted within the CSF, WM, and whole-brain masks, the set of aCompCor regressors, the set of discrete cosine basis functions that were used when extracting the aCompCor regressors, and the motion-related components identified by ICA-AROMA. All regressors of interest were convolved with a double gamma hemodynamic response function (HRF); the temporal derivative of the HRF was also included in the model. The *t*-statistic maps (i.e., maps of beta coefficients divided by their standard error estimates) resulting from these run-wise analyses were used for pattern similarity analysis (our primary analyses), as described in more detail below. In addition, we conducted trial-wise analyses prior to conducting classification analyses, in order to provide more training

data to the SVM algorithm. Specifically, we repeated the above-described procedure modeling the response for each trial (rather than for each condition/direction of attentional shift) separately.

Pattern similarity analysis. We conducted a pattern similarity analysis in the SPL as our primary analysis. For each of Task II (the social navigation task) and Task III (the spatial task), we first calculated the similarity of the neural response patterns evoked when shifting attention in the same (matching) direction ($\bar{r}_{matching}$) and in opposite (mismatching) directions ($\bar{r}_{mismatching}$) across runs. Specifically, $\bar{r}_{matching}$ was calculated as the mean correlation between patterns evoked during attentional shifts in the same direction across different runs (for example, between patterns evoked during "upward" shifts of attention in social knowledge in run 1 and "upward" shifts of attention in social knowledge in run 2). Contrastingly, $\bar{r}_{mismatching}$ was calculated as the mean correlation between patterns evoked during attentional shifts in opposing directions across different runs (for example, between patterns evoked during "upward" shifts of attention in social knowledge in run 1 and "downward" shifts of attention in social knowledge in run 2). If response patterns in a given brain region encode the direction of attentional shifts in social knowledge, then response patterns evoked during shifts of attention in the same direction should be more similar than those evoked during shifts of attention in opposing directions (i.e., $\bar{r}_{matching}$ should be greater than $\bar{r}_{mismatching}$). Thus, we compared $\bar{r}_{matching}$ to $\bar{r}_{mismatching}$ by defining ΔR as $\bar{r}_{matching} - \bar{r}_{mismatching}$, and testing whether ΔR was significantly greater than zero (i.e., whether $\bar{r}_{matching}$ was significantly greater than $\bar{r}_{mismatching}$). The ΔR values were transformed with Fisher's *z*-transformation, then tested against zero with a one-sample *t*-test.

When conducting pattern similarity analyses across the social navigation and spatial tasks, $\bar{r}_{matching}$ was calculated as the mean correlation between response patterns evoked during attentional shifts in matching directions across runs from different tasks (for example, between "upward" shifts of attention in social knowledge in run 1 of the social navigation task, and upward shifts of visual attention in run 7 of the spatial task). $\bar{r}_{mismatching}$ was calculated as the mean correlation between response patterns evoked during attentional shifts in mismatching directions across runs from different tasks (for example, between patterns evoked during "upward" shifts of attention in social knowledge in run 1 of the social navigation task, and downward shifts of visual attention in run 7 of the spatial task). As in the within-task pattern similarity analyses described in the preceding paragraph, ΔR was defined as $\bar{r}_{matching} - \bar{r}_{mismatching}$ and was again tested against zero with a one-sample *t*-test.

After conducting our primary analyses in the left and right SPL, we repeated these pattern similarity analyses within each region of the DKT cortical parcellation, and in subcortical grey matter structures, using FDR-correction to account for multiple comparisons across regions. Additional exploratory whole-brain analyses were performed with searchlights, as described in the subsequent section.

Searchlight analyses. Whole-brain searchlights were run on the social navigation task and the spatial task, both separately and across the two tasks. The searchlights were run in each subject's native space, within a dilated cortical ribbon mask. The cortical ribbon masks for each subject were obtained with FreeSurfer 6.0.0 based on their structural scans, downsampled to the same resolution as the functional scans, and then dilated with a spherical kernel of 3 mm.

The searchlights were implemented in PyMVPA[58] as spheres with radii of 4 voxels. Both the searchlight-based pattern similarity analyses (described in the preceding section) and additional exploratory classification analyses (described in the subsequent section) were run within each searchlight sphere. Maps of results from these analyses were smoothed with a gaussian kernel of 3 mm FWHM and transformed to MNI space prior to group analyses. The results of searchlight analyses were tested against either 0 (for pattern similarity analyses) or 0.5 (chance accuracy for classifications). FSL "randomise" was used to perform permutation tests and to generate null distributions of cluster masses for multiple comparisons correction (cluster-forming threshold: $p < .01$, one-tailed; 10,000 permutations).

Classification analyses. Searchlight-based classifications were implemented in PyMVPA. Within each searchlight sphere, a linear support vector machine (SVM) classifier was trained on the social navigation task (Task II) to classify the directions of attentional shifts ("upward" or "downward"), and its classification accuracy was tested, without further training, on the corresponding directions in the spatial task (Task III); another SVM classifier was trained on the spatial task and tested on the social navigation task. Here, the correspondence between the attentional shift directions across the two tasks was determined individually, based on how each participant reconstructed the social hierarchy (see Fig. 4).

To better understand the results of this classification analysis, we conducted a separate analysis of within-task decoding (see Supplement).

Visualization. Visualizations of neuroimaging data throughout this paper were implemented with PySurfer[59].

Supplemental Analyses and Results

Within-Domain Decoding

As described in the main text, when using a parcel-based approach, cross-domain decoding analyses identified no regions that encoded shifts of attention in space and social knowledge similarly, mirroring the results of the cross-domain pattern similarity analyses. Using a searchlight approach, cross-domain decoding resulted in above-chance accuracy only in a small cluster in early visual cortex (Fig. S3). Unsuccessful cross-domain decoding could reflect either attentional shift directions being encoded in different ways across domains, or our classification method having insufficient power to detect systematic distinctions between shifts of attention in different directions within domains. In other words, chance-level cross-domain decoding could stem from classifiers being unable to learn reliable distinctions within domains, or from classifiers learning reliable distinctions within domains that do not generalize across domains. To arbitrate between these possibilities, we conducted a within-domain decoding analysis.

Data Analysis. This within-domain analysis was implemented similarly to the cross-domain decoding analyses described in the main text, with whole-brain searchlights and linear support vector machine (SVM) classifiers (see Methods for more details). Instead of training classifiers on either the social navigation task or the spatial task and then testing classifiers on the left-out task, this analysis was carried out on each of the two tasks separately. For each task, specifically, we used leave-one-run-out cross-validation within each searchlight sphere, such that classifiers were trained on data from all fMRI runs except one, and then tested for their performance on the left-out run. Six-fold cross-validation was used for the social navigation task (6 runs) and 2-fold cross-validation was used for the spatial task (2 runs), so that the data from each run was left out exactly once; classifier accuracies were then averaged across folds.

Results. Results from this within-task decoding analysis showed significantly above-chance accuracy for both the social navigation task and the spatial task (Fig. S2a), in the superior parietal cortex, as well as in regions of the occipital cortex.

These results resembled the results we obtained through pattern similarity analyses (Fig. 6b/Fig. S2b). This suggests that this classification method was able to detect information signaled in multivoxel patterns about the direction of spatial and social attentional shifts in many of the same brain regions identified using the methods employed in our main analysis. This analysis generally resulted in smaller significant areas compared to the pattern similarity analysis, which may be due to the relatively small amount of training data available to the machine learning classifiers.

Using the same classification method, however, the performance of classifiers was considerably lower in the cross-domain decoding analysis described in the main text (Fig. S3) compared to the within-domain analysis described here (Fig. S2a); this difference was particularly notable in the superior parietal lobule (SPL). This indicates that the lack of cross-domain decoding success in the SPL was unlikely due to an inability of the current statistical methods to detect within-domain distinctions. Instead, SPL may not encode attentional shift directions in social and the spatial contents in the same way.

Eye-Tracking Study

To confirm that the results observed in the social navigation task did not merely result from distinctive eye movements during mental navigation of social knowledge, we ran a separate study with the same tasks, while participants' eye movements were tracked during the social navigation task.

In this eye-tracking study, participants first completed the trial-and-error learning task, followed by 6 blocks of the social navigation task if they passed an accuracy criterion in the learning task (i.e., no more than 3 errors in adjacent trials of at least 3 consecutive blocks, and no more than 2 errors in non-adjacent trials of at least 3 consecutive blocks). We only analyzed data from participants who had satisfactory accuracy in the social navigation task (>70%). Both paradigms are exactly the same as in the behavioral portion of the fMRI study.

Participants. Sixty-four participants were recruited from the UCLA community; 46 people passed the accuracy criterion in the learning task and thus continued to the social navigation task (i.e., the eye-tracking portion), but 4 people were unable to complete the social navigation task due to equipment malfunction. Out of the 42 people who completed the study, 12 were excluded from data analysis due to unsatisfactory accuracy (11 people) or being non-native English speakers (1 person; only native English speakers were included in the eye-tracking study to mirror the selection criteria for the fMRI study as closely as possible). The remaining 30 participants (10 males, 20 females) were native English speakers between the ages of 19 and 28 (M = 20.5)[1].

Data acquisition. Eye movements were recorded using an EyeTribe tracker at a sampling rate of 60 Hz. Participants were seated with their chin on a chinrest, which was placed approximately 50 cm from the screen. The social navigation task was divided into two parts of 3 blocks each; participants had the chance to take a break in between the first and second part of the task. The eye tracker was calibrated at the beginning of each part using the EyeTribe UI and 16-point grid.

Data analysis. We extracted the raw coordinates of eye gaze positions during the 6000-ms mental navigation period of each trial, and calculated the median x and y coordinates for each trial. As described in the Methods section, the social navigation task had two types of trials, where participants either navigated "up" (to find a more powerful person) or "down" (to find a less powerful person) in their knowledge of the learned social hierarchy. The median coordinates of eye gaze positions were then averaged across the "up" and "down" mental navigation trials for each participant. A paired t-test was used to test for differences between eye gaze coordinates in trials involving mental navigation of social knowledge in different "directions."

Results. On the x-axis, the mean difference of gaze positions between trials involving mental navigation in different "directions" (i.e., navigating "up" vs. "down" in social knowledge) was 1.97 pixels ($t(29) = 0.93$, $p = 0.36$). On the y-axis, the mean difference was -1.95 pixels

[1] The same pattern of results that is reported here was observed without any participant exclusions (on the x-axis, $t(41) = 0.26$, $p = 0.80$; on the y-axis, $t(41) = -0.88$, $p = 0.39$). However, to provide a more stringent test of systematic associations between eye movements and mental navigation of social knowledge in our fMRI task, we focus on the results obtained from a sample with characteristics that mirror our fMRI sample (i.e., native English speakers who performed the social navigation task with high accuracy).

($t(29) = -1.18$, $p = 0.25$). These results suggest no significant systematic differences in eye movements between trials corresponding to different navigation directions in the social navigation task.

Supplemental Figures

	Social		Spatial	
	t(29)	FDR p	t(29)	FDR p
Superior frontal gyrus (L)	4.23	6.6×10^{-4}	0.73	0.28
Precentral gyrus (L)	4.06	8.5×10^{-4}	1.47	0.12
Triangular part of inferior frontal gyrus (L)	3.75	0.0016	0.87	0.24
Caudal middle frontal gyrus (L)	3.69	0.0016	1.44	0.12
Opercular part of inferior frontal gyrus (L)	3.52	0.0023	0.68	0.28
Orbital part of inferior frontal gyrus (L)	3.46	0.0025	2.00	0.054
Supramarginal gyrus (L)	2.97	0.0078	-0.61	-
Middle temporal gyrus (L)	2.72	0.014	1.30	0.14
Superior frontal gyrus (R)	2.56	0.017	1.45	0.12
Postcentral gyrus (L)	2.51	0.019	0.68	0.28
Opercular part of inferior frontal gyrus (R)	2.45	0.021	1.65	0.094
Supramarginal gyrus (R)	2.37	0.023	1.39	0.13
Posterior cingulate cortex (L)	2.31	0.026	0.70	0.28
Orbital part of inferior frontal gyrus (R)	2.26	0.028	1.80	0.074
Pallidum (L)	2.27	0.028	0.19	0.45
Posterior cingulate cortex (R)	2.22	0.029	1.29	0.14
Precentral gyrus (R)	2.17	0.032	1.07	0.19
Superior temporal gyrus (L)	2.06	0.039	1.48	0.12
Triangular part of inferior frontal gyrus (R)	1.89	0.054	1.94	0.060
Paracentral lobule (R)	1.84	0.058	1.01	0.20
Postcentral gyrus (R)	1.73	0.071	2.23	0.036
Insula (R)	1.70	0.074	1.92	0.061
Isthmus of cingulate cortex (R)	1.65	0.080	3.28	0.0042
Medial orbitofrontal gyrus (L)	1.59	0.088	2.61	0.019
Medial orbitofrontal gyrus (R)	1.57	0.088	1.22	0.15
Caudal anterior cingulate cortex (L)	1.52	0.096	-0.48	-
Putamen (L)	1.49	0.099	0.54	0.32
Rostral middle frontal gyrus (R)	1.41	0.11	1.57	0.11
Amygdala (L)	1.38	0.11	2.45	0.026
Parahippocampal gyrus (L)	1.28	0.13	1.80	0.074
Parahippocampal gyrus (R)	1.25	0.14	4.03	7.6×10^{-4}
Transverse temporal gyrus (L)	1.20	0.15	1.74	0.082
Inferior temporal gyrus (R)	1.12	0.17	3.61	0.0022
Entorhinal cortex (L)	0.92	0.22	2.57	0.021
Nucleus accumbens (L)	0.73	0.28	-0.028	-
Amygdala (R)	0.71	0.28	-0.084	-
Hippocampus (R)	0.60	0.31	1.34	0.14
Transverse temporal gyrus (R)	0.58	0.32	0.68	0.28
Rostral anterior cingulate cortex (L)	0.46	0.36	0.037	0.49
Caudate nucleus (L)	-0.38	-	0.99	0.21
Entorhinal cortex (R)	0.38	0.38	0.11	0.48
Nucleus accumbens (R)	-0.23	-	1.43	0.12
Rostral anterior cingulate cortex (R)	0.23	0.43	1.31	0.14
Hippocampus (L)	-0.14	-	2.26	0.035
Pallidum (R)	-0.082	-	0.70	0.28
Caudal anterior cingulate cortex (R)	0.061	0.48	1.46	0.12

Figure S1. Additional statistical results of parcel-based pattern-similarity analyses. This table is a continuation of Fig. 7b. Both Fig. 7 and this figure contain results from within-domain pattern similarity analyses (analysis of ΔR), which tested if the multi-voxel response patterns for attentional shifts in matching directions were significantly more similar to each other compared to those in mismatching directions. Fig. 7b only shows the brain regions that were significant at an FDR-corrected threshold of $p < 0.01$, one-tailed, in either the social navigation task or the spatial task. This figure shows results for brain regions not included in Fig. 7a (i.e., regions that did not surpass a one-tailed, FDR-corrected threshold of $p < .01$ on one or both tasks).

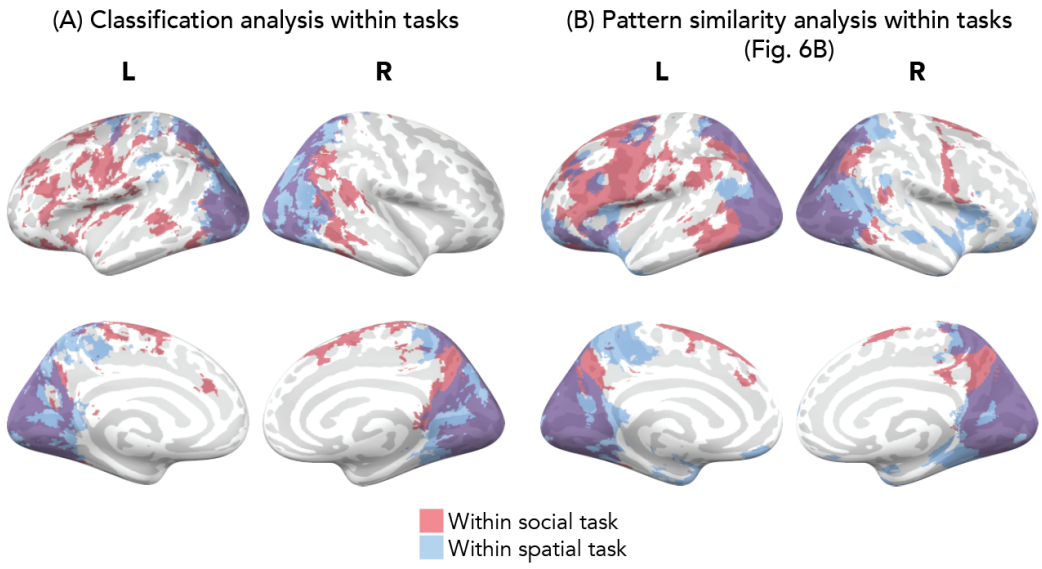

Figure S2. Within-task classification analyses yielded similar results to within-task pattern similarity analyses. We conducted a decoding analysis within each of the social and spatial domains in order to examine if the lack of significant results in the cross-domain decoding analysis (particularly in the superior parietal lobule; Fig. S3) was due to an inability of classifiers to learn distinctions within tasks, or to within-task decision boundaries not generalizing across tasks. The results from this within-domain analysis (**A**) showed significantly above-chance accuracy in brain areas in much of the occipital and parietal lobule, largely replicating the results from the within-domain pattern similarity analysis (**B**). Panel B shows the results from the pattern similarity analysis, i.e., Fig. 6b, to provide a side-by-side comparison. All visualized results are significant at a corrected threshold of $p < .01$ (one-tailed).

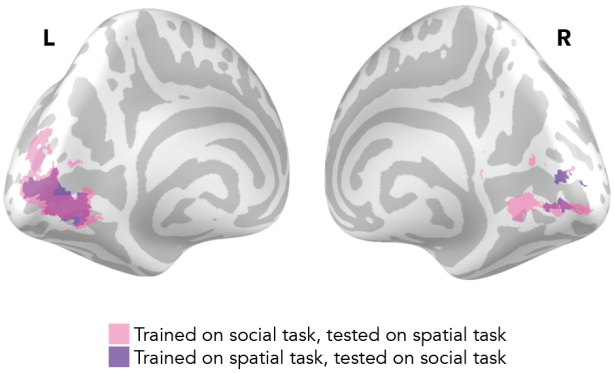

Figure S3. Exploratory whole-brain cross-task classification analyses revealed a significant cluster in early visual cortex. At each searchlight center, a support vector machine classifier was trained to classify the two directions of attentional shifts based on neural response patterns in the social navigation task, and then tested on its ability to decode the corresponding directions of attentional shifts in the spatial task. Another classifier was analogously trained on data from the spatial task, and tested on data from the social navigation task. The correspondence of directions across two tasks was determined based on participants' answers in the post-scan hierarchy reconstruction task (see Fig. 4). In the colored areas of this figure, the two classifiers achieved accuracies that were significantly above chance. Although there was substantial overlap in the brain regions that encoded directions of attentional shifts within each domain (see Figs. 6 and 7), this cross-task classification showed less robust results, suggesting that directions of attentional shifts in space and social knowledge are encoded in largely distinct multi-voxel response patterns. We note that significant results were not observed in other cross-domain analyses (i.e., parcel-based cross-domain pattern similarity analyses; searchlight-based cross-domain pattern similarity analyses; parcel-based decoding analyses); thus, these results should be interpreted with caution. Depicted results are thresholded at $p < 0.05$, corrected (one-tailed; note that this is a more liberal threshold than the results visualized in the main text). Specifically, $p = 0.015$ (one-tailed) in the pink cluster (trained on the social task and tested on the spatial task), and $p = 0.036$ (one-tailed) in the purple cluster (trained on the spatial task and tested on the social task).

Reference

1. Cheney, D. L. & Seyfarth, R. M. *How monkeys see the world: Inside the mind of another species*. (University of Chicago Press, 1990).
2. Hall, J. A., Coats, E. J. & LeBeau, L. S. Nonverbal behavior and the vertical dimension of social relations: a meta-analysis. *Psychol. Bull.* **131**, 898–924 (2005).
3. Marsh, A. A., Blair, K. S., Jones, M. M., Soliman, N. & Blair, R. J. R. Dominance and submission: the ventrolateral prefrontal cortex and responses to status cues. *J. Cogn. Neurosci.* **21**, 713–724 (2009).
4. Mattan, B. D., Kubota, J. T. & Cloutier, J. How Social Status Shapes Person Perception and Evaluation: A Social Neuroscience Perspective. *Perspect. Psychol. Sci.* **12**, 468–507 (2017).
5. Cheney, D. L. & Seyfarth, R. M. The representation of social relations by monkeys. *Cognition* **37**, 167–196 (1990).
6. Gazes, R. P., Hampton, R. R. & Lourenco, S. F. Transitive inference of social dominance by human infants. *Dev. Sci.* **20**, (2017).
7. Grosenick, L., Clement, T. S. & Fernald, R. D. Fish can infer social rank by observation alone. *Nature* **445**, 429–432 (2007).
8. Kumaran, D., Melo, H. L. & Duzel, E. The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron* **76**, 653–666 (2012).
9. Paz-y-Miño, G. C., Bond, A. B., Kamil, A. C. & Balda, R. P. Pinyon jays use transitive inference to predict social dominance. *Nature* **430**, 778–781 (2004).
10. White, S. L. & Gowan, C. Brook trout use individual recognition and transitive inference to determine social rank. *Behav. Ecol.* **24**, 63–69 (2013).
11. Muscatell, K. A. *et al.* Social status modulates neural activity in the mentalizing network. *Neuroimage* **60**, 1771–1777 (2012).
12. Zink, C. F. *et al.* Know your place: neural processing of social hierarchy in humans. *Neuron*

58, 273–283 (2008).

13. Lakoff, G. & Johnson, M. *Metaphors We Live By*. (University of Chicago Press, 1980).
14. Lakoff, G. *Women, Fire, and Dangerous Things*. (University of Chicago Press, 1987).
15. Lakoff, G. A Figure of Thought. *Metaphor and Symbolic Activity* **1**, 215–225 (1986).
16. Anderson, M. L. Neural reuse: a fundamental organizational principle of the brain. *Behav. Brain Sci.* **33**, 245–66; discussion 266–313 (2010).
17. Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).
18. Parkinson, C. & Wheatley, T. Old cortex, new contexts: re-purposing spatial perception for social cognition. *Front. Hum. Neurosci.* **7**, 645 (2013).
19. Parkinson, C. & Wheatley, T. The repurposed social brain. *Trends Cogn. Sci.* **19**, 133–141 (2015).
20. Yamazaki, Y., Hashimoto, T. & Iriki, A. The posterior parietal cortex and non-spatial cognition. *F1000 Biol. Rep.* **1**, 74 (2009).
21. Williams, L. E., Huang, J. Y. & Bargh, J. A. The Scaffolded Mind: Higher mental processes are grounded in early experience of the physical world. *Eur. J. Soc. Psychol.* **39**, 1257–1267 (2009).
22. Chiao, J. Y. *et al.* Neural representations of social status hierarchy in human inferior parietal cortex. *Neuropsychologia* **47**, 354–363 (2009).
23. Göbel, S. M., Johansen-Berg, H., Behrens, T. & Rushworth, M. F. S. Response-selection-related parietal activation during number comparison. *J. Cogn. Neurosci.* **16**, 1536–1551 (2004).
24. Shuman, M. & Kanwisher, N. Numerical magnitude in the human parietal lobe; tests of representational generality and domain specificity. *Neuron* **44**, 557–569 (2004).
25. Peelen, M. V. & Downing, P. E. Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends in cognitive sciences* vol. 11 4–5 (2007).
26. Cabeza, R., Ciaramelli, E., Olson, I. R. & Moscovitch, M. The parietal cortex and episodic

memory: an attentional account. *Nat. Rev. Neurosci.* **9**, 613–625 (2008).

27. Hutchinson, J. B., Uncapher, M. R. & Wagner, A. D. Posterior parietal cortex and episodic retrieval: convergent and divergent effects of attention and memory. *Learn. Mem.* **16**, 343–356 (2009).
28. Serences, J. T., Schwarzbach, J., Courtney, S. M., Golay, X. & Yantis, S. Control of object-based attention in human cortex. *Cereb. Cortex* **14**, 1346–1357 (2004).
29. Knops, A., Thirion, B., Hubbard, E. M., Michel, V. & Dehaene, S. Recruitment of an area involved in eye movements during mental arithmetic. *Science* **324**, 1583–1585 (2009).
30. Peelen, M. V., Atkinson, A. P. & Vuilleumier, P. Supramodal representations of perceived emotions in the human brain. *J. Neurosci.* **30**, 10127–10134 (2010).
31. Parkinson, C., Liu, S. & Wheatley, T. A common cortical metric for spatial, temporal, and social distance. *J. Neurosci.* **34**, 1979–1987 (2014).
32. Downing, P. E., Wiggett, A. J. & Peelen, M. V. Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *J. Neurosci.* **27**, 226–233 (2007).
33. Peelen, M. V., Wiggett, A. J. & Downing, P. E. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron* **49**, 815–822 (2006).
34. Chavez, R. S., Heatherton, T. F. & Wagner, D. D. Neural Population Decoding Reveals the Intrinsic Positivity of the Self. *Cereb. Cortex* **27**, 5222–5229 (2017).
35. Haynes, J.-D. A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron* **87**, 257–270 (2015).
36. Molenberghs, P., Mesulam, M. M., Peeters, R. & Vandenberghe, R. R. C. Remapping attentional priorities: differential contribution of superior parietal lobule and intraparietal sulcus. *Cereb. Cortex* **17**, 2703–2712 (2007).
37. Dai, Q. & Zhu, L. Verbal-spatial and visuospatial coding of power-space interactions.

Conscious. Cogn. **63**, 151–160 (2018).

38. Schubert, T. W. Your highness: vertical positions as perceptual symbols of power. *J. Pers. Soc. Psychol.* **89**, 1–21 (2005).
39. Zanolie, K. *et al.* Mighty metaphors: behavioral and ERP evidence that power shifts attention on a vertical dimension. *Brain Cogn.* **78**, 50–58 (2012).
40. Van Essen, D. C. *et al.* Mapping visual cortex in monkeys and humans using surface-based atlases. *Vision Res.* **41**, 1359–1378 (2001).
41. Mantini, D., Corbetta, M., Romani, G. L., Orban, G. A. & Vanduffel, W. Evolutionarily novel functional networks in the human brain? *J. Neurosci.* **33**, 3259–3275 (2013).
42. Emmerling, T. C., Zimmermann, J., Sorger, B., Frost, M. A. & Goebel, R. Decoding the direction of imagined visual motion using 7T ultra-high field fMRI. *Neuroimage* **125**, 61–73 (2016).
43. Tavares, R. M. *et al.* A Map for Social Navigation in the Human Brain. *Neuron* **87**, 231–243 (2015).
44. Hogue, M. E., Beaugrand, J. P. & Laguë, P. C. Coherent use of information by hens observing their former dominant defeating or being defeated by a stranger. *Behav. Processes* **38**, 241–252 (1996).
45. Maclean, E. L., Merritt, D. J. & Brannon, E. M. Social Complexity Predicts Transitive Reasoning in Prosimian Primates. *Anim. Behav.* **76**, 479–486 (2008).
46. Ma, D. S., Correll, J. & Wittenbrink, B. The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. Methods* **47**, 1122–1135 (2015).
47. Peirce, J. W. PsychoPy--Psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).
48. Durnez, J., Blair, R. & Poldrack, R. A. Neurodesign: Optimal Experimental Designs for Task fMRI. *bioRxiv* 119594 (2018) doi:10.1101/119594.
49. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat.*

Methods **16**, 111–116 (2019).

50. Gorgolewski, K. *et al.* Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Front. Neuroinform.* **5**, 13 (2011).
51. Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Front. Neuroinform.* **8**, 14 (2014).
52. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med. Image Anal.* **12**, 26–41 (2008).
53. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).
54. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23 Suppl 1**, S208–19 (2004).
55. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).
56. Klein, A. *et al.* Mindboggling morphometry of human brains. *PLoS Comput. Biol.* **13**, e1005350 (2017).
57. Klein, A. & Tourville, J. 101 labeled brain images and a consistent human cortical labeling protocol. *Front. Neurosci.* **6**, 171 (2012).
58. Hanke, M. *et al.* PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics* **7**, 37–53 (2009).
59. Waskom, M. *et al. nipy/PySurfer: 0.10.0*. (Zenodo, 2020). doi:10.5281/ZENODO.592515.