OXFORD

Social Cognitive and Affective Neuroscience, 2022, **17**, 939–955

DOI: https://doi.org/10.1093/scan/nsac017
Advance Access Publication Date: 7 March 2022
Original Manuscript

Transforming social perspectives with cognitive maps

Shahar Arzy[1,2] and Raphael Kaplan[3]

[1]Faculty of Medicine and the Department of Cognitive Sciences, Hebrew University of Jerusalem, Jerusalem 91120, Israel
[2]Department of Neurology, Hadassah Hebrew University Medical School, Jerusalem 91120, Israel
[3]Department of Basic Psychology, Clinical Psychology, and Psychobiology, Universitat Jaume I, Castelló de la Plana 12071, Spain
Correspondence should be addressed to Raphael Kaplan, Department of Basic Psychology, Clinical Psychology, and Psychobiology, Universitat Jaume I, Avinguda de Vicent Sos Baynat, Castelló de la Plana, Spain. E-mail: kaplan@uji.es.

Abstract

Growing evidence suggests that cognitive maps represent relations between social knowledge similar to how spatial locations are represented in an environment. Notably, the extant human medial temporal lobe literature assumes associations between social stimuli follow a linear associative mapping from an egocentric viewpoint to a cognitive map. Yet, this form of associative social memory does not account for a core phenomenon of social interactions in which social knowledge learned via comparisons to the self, other individuals or social networks are assimilated within a single frame of reference. We argue that hippocampal–entorhinal coordinate transformations, known to integrate egocentric and allocentric spatial cues, inform social perspective switching between the self and others. We present evidence that the hippocampal formation helps inform social interactions by relating self *vs* other social attribute comparisons to society in general, which can afford rapid and flexible assimilation of knowledge about the relationship between the self and social networks of varying proximities. We conclude by discussing the ramifications of cognitive maps in aiding this social perspective transformation process in states of health and disease.

Key words: cognitive map; perspective-taking; hippocampus; episodic memory; spatial navigation; social learning

Introduction

Allocentric dimensions are commonly used in both psychology and neuroscience, but have significantly different meanings. In experimental psychology and neuroscience, allocentric processing stems from Tolman's cognitive map, where the relative locations of phenomena in their everyday environment are stored in a mental model (Tolman, 1948; O'Keefe and Nadel, 1978; Gallistel, 1990). The cognitive map has been applied with great success in the neurobiology of spatial cognition, where allocentric processing in spatial cognition entails maintaining a mapping of the relative locations of different stimuli in a physical environment (O'Keefe and Nadel, 1978). Consequently, spatially tuned neurons, mainly in the hippocampal–entorhinal system, transform different spatial cues from relative, egocentric spatial coordinates to absolute, allocentric coordinates within an environment (Hartley *et al.*, 2013). In contrast, allocentrism in social psychology is thought to signify having one's attention and actions centered on others over themselves (Triandis *et al.*, 1985). Allocentric coding in social cognition may be further developed in the context of social networks and constituent psychological distances based on aspects like intimacy, affiliation, power or preferences (Trope and Liberman, 2010; Parkinson *et al.*, 2014; Dunbar, 2018). In this context, the relation between the distance of the self and others would be egocentric, while the distances reflecting how other people relate to both the self, social networks and society would be allocentric (Peer *et al.*, 2021a). Notably, this type of social allocentrism remains little explored at the neural level and unlike spatial coding it is unclear whether allocentric coding in the social domain is completely detached from a first-person reference frame. Despite the lack of contemporary findings, Tolman's original conception of cognitive maps also incorporated the social psychology view of prioritizing others over the self (Tolman, 1948). Given the complexity of human social networks and storing related knowledge, maintaining an internal model of the social world would be advantageous. We argue that transformation of egocentric and allocentric spatial cues primarily by the hippocampal–entorhinal system makes it ideally placed to help assimilate knowledge about the preferences of the self and others to flexibly make social comparisons in the prosocial world (Figure 1).

Following the award of the 2014 Nobel Prize in Physiology and Medicine for the discovery of allocentric mapping of physical space by grid cells in the medial entorhinal cortex and hippocampal place cells, recent cognitive neuroscience findings have linked cognitive maps in the hippocampal formation (including the entorhinal cortex) to more abstract domains in the brains of model organisms (Aronov *et al.*, 2017; Sarel *et al.*, 2017) and humans (Constantinescu *et al.*, 2016). This emerging research line

Received: 30 August 2021; **Revised:** 17 December 2021; **Accepted:** 7 March 2022

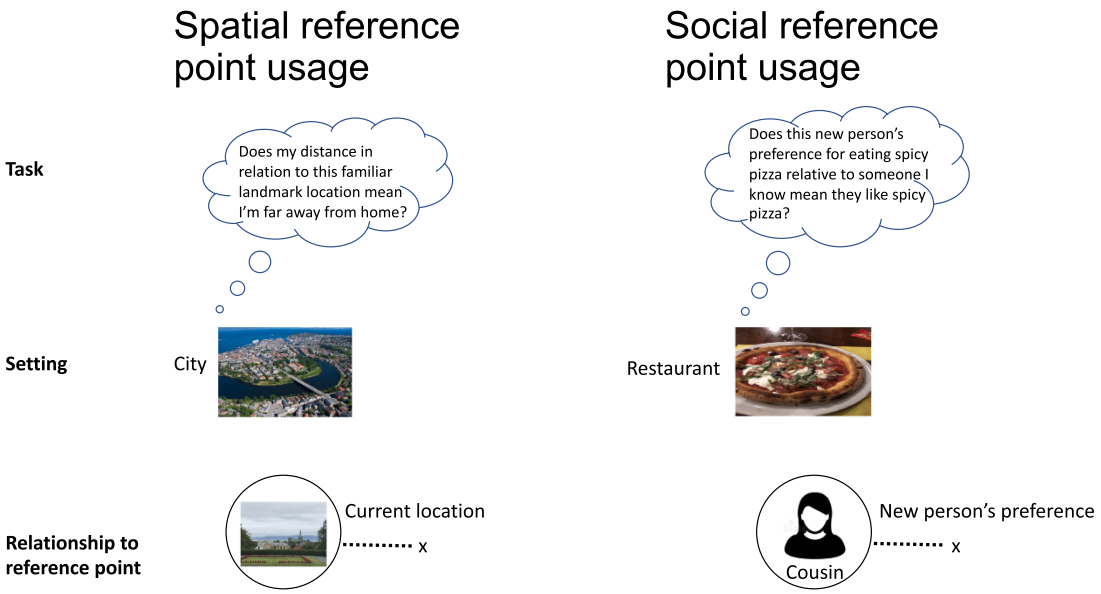

Fig. 1. Common reference point usage along one dimension in spatial and social tasks. Illustration highlights similar alignment to spatial and social points of reference for one dimension during spatial navigation in a city and a social decision in a restaurant (Overhead city image taken by Åge Hojem/Trondheim Havn).

has now expanded to the social domain, where hippocampal coding of social variables like the location of others is investigated at the level of single neurons in model organisms (Danjo *et al.*, 2018; Omer *et al.*, 2018). In parallel, research investigating the role of the human hippocampal–entorhinal system in social cognition has found abstract hippocampal mapping of power and affiliation for different individuals (Tavares *et al.*, 2015), hippocampal–entorhinal transformation of relatively learned social attributes to an absolute scale (Kaplan and Friston, 2019), along with hippocampal distance and entorhinal grid-like coding of a 2D social hierarchy (popularity and competence) structure (Park *et al.*, 2020, 2021). Taken together, this mounting evidence implicates the hippocampal formation, a region putatively associated with declarative memory, in playing a key role in social cognition (Hitti and Siegelbaum, 2014; Montagrin *et al.*, 2018; Schafer and Schiller, 2018a; Leblanc and Ramirez, 2020).

The human hippocampal formation in social cognition

Despite advances linking the hippocampus and surrounding brain regions to social cognition, the specific contribution of this region to social behavior is not well understood. Building on the role of the hippocampus in memory and spatial navigation, the hippocampal formation has been linked to remembering interactions with others (Moscovitch *et al.*, 2016), recalling social hierarchies (Kumaran *et al.*, 2012, 2016; FeldmanHall *et al.*, 2021), investigating the influence of episodic memory on recalling and maintaining social networks (Stiller and Dunbar, 2007; Davidson *et al.*, 2012), maintaining shared group knowledge (Hirst *et al.*, 2018) and maintaining the spatial location of other individuals (Stangl *et al.*, 2021). Further neuroimaging work has built on findings about memory for social networks by focusing on hippocampal-dependent memory for social hierarchies (Kumaran *et al.*, 2012, 2016). In these studies, Kumaran *et al.* (2012), (2016) found that the hippocampus maintains the rank of different individuals in a social hierarchy. Yet, the role of the human hippocampus in 'social memory' is not limited to remembering who is important in social hierarchies. Rather, it can be extended to 'person memory' which is general information about a person such as whether an individual treated you well or poorly in the past (Srull and Wyer, 1989). Likewise, in a social decision-making task where participants needed to make decisions about whether specific individuals would treat them well or not, FeldmanHall and colleagues(2021) found that the hippocampus was sensitive to adaptive *vs* maladaptive social choices. Notably, this work highlights a hippocampal role in social memory, similar to what is been observed in rodents (Hitti and Siegelbaum, 2014; Leblanc and Ramirez, 2020). Still, the aforementioned findings do not appear to directly relate to the hippocampus' putative role in spatial navigation. Linking memory for social networks to navigation tasks, a study by Tavares *et al.* (2015) extended these ideas further by uncovering a hippocampal role in learning multidimensional (2D) social hierarchies/rankings for which the two dimensions were power and affiliation. However, how this data related to construction of cognitive maps and map-like distance coding in a more general manner was unclear. Addressing this issue, Park *et al.* (2020) trained participants to learn a 2D social hierarchy defined by two independent dimensions of popularity and competence that could be reconstructed after learning the outcomes from a series of binary decisions along a single dimension, where two dimensions were learned on different days. Crucially, the true 2D hierarchies were never shown to participants, but could be reconstructed via transitive inference across one dimension after a session, and between two dimensions across days. Using functional magnetic resonance imaging (fMRI), the authors demonstrated that the hippocampus maintains and updates an integrated map of a social hierarchy structure across the two dimensions with an entorhinal contribution coding the distance between key individuals, which was also observed in medial prefrontal cortex (mPFC). Furthermore, the hippocampus reinstated the social hierarchy structure coded by the entorhinal cortex, pointing to the integrative function of the hippocampal–entorhinal circuit in assimilating social knowledge. A subsequent fMRI study by Park *et al.* (2021) using the same

2D social hierarchy task found evidence for grid cell-like coding in entorhinal cortex, mPFC and the posterior cingulate cortex for inferred direct trajectories over the implicitly learned 2D social space during decision-making (see reviews by Boorman *et al.*, 2021; Russin *et al.*, 2021 for more information). These two studies by Park and co-authors highlight how hippocampal–entorhinal cognitive map-like coding could help guide social learning in coordination with more traditionally studied social brain networks. By incorporating a learning component and flexible comparisons between individuals in their tasks, Park *et al.* (2020) observed hippocampal–entorhinal involvement in encoding relative distances between individuals in social networks where previous studies had primarily found parietal midline, temporoparietal junction (TPJ) and prefrontal cortex involvement in similar tasks (Parkinson *et al.*, 2014; Hayman and Arzy, 2021; Peer *et al.*, 2021a). Still, it remains to be seen whether these sorts of mechanisms are domain-general or if spatial neural coding mechanisms uniquely inform social cognition relative to other cognitive domains.

Paralleling fMRI findings in conceptual knowledge (Constantinescu *et al.*, 2016; Theves *et al.*, 2019), odor (Bao *et al.*, 2019) and semantic(Viganò and Piazza, 2020) 2D spaces, the few studies examining hippocampal–entorhinal coding of social hierarchies in 2D (Tavares *et al.*, 2015; Park *et al.*, 2020) have focused on purely associative spaces. In associative social spaces, an individual with two social characteristics (e.g. power and affiliation) is associated with coordinates in a continuous map of a social hierarchy, which builds on a rich literature of power and affiliation in social learning (see Magee and Galinsky, 2008; Fiske, 2010 for reviews). Research focusing on hippocampal maintenance of a one-to-one linear relationship between an individual and their 2D position in a social hierarchy map, extends the well-known role of the hippocampus in associative memory encoding along continuous dimensions (Eichenbaum and Cohen, 2014). These studies demonstrate that distance can be represented as a domain-general metric in the brain (Parkinson *et al.*, 2014; Peer *et al.*, 2015), but the hippocampus appears to only be engaged when there is an associative memory component (Schiller *et al.*, 2015). The necessity of a hippocampal-dependent mnemonic component to the task is further supported by recent findings from Lau and colleagues (2020) which showed that structure learning using category learning and overtraining, task demands typically not associated with hippocampal learning processes, primarily engages insula instead of the hippocampus. Comparative evidence showing structure learning of novel individuals' preferences within social group structures that could be reflected in hippocampal cognitive maps is still missing. Indeed, relational spaces of social knowledge can be helpful in guiding prosocial behavior, but they lack a key capacity of cognitive maps that would afford flexible social memory and decision-making. In cognitive maps of physical space, spatial information coded within an egocentric reference frame is typically mapped in a non-linear fashion to their location in an environmental/allocentric reference frame (Burgess, 2006). Non-linear transformation of spatial cues involves translating spatial coordinates that could be in the top left of an egocentric viewpoint and in the center of an allocentric map, where the spatial coordinates are not a one-to-one linear mapping between reference frames (e.g. coordinates multiplied by two). Rather, the mapping between reference frames could be a logarithmic or more complex transformation of coordinates. Consequently, a non-linear transformation process affords efficient and flexible remapping in response to environmental changes (Burgess, 2006; Hartley *et al.*, 2013). Given the emerging role of cognitive maps in representing non-spatial knowledge (Schiller *et al.*, 2015; Kaplan *et al.*, 2017a; Behrens *et al.*, 2018; Bellmund *et al.*, 2018), might similar flexible non-linear remapping processes in the human hippocampal–entorhinal system support transformations of abstract knowledge learned within one reference frame to another frame of reference? If these sorts of computations are domain-general, it could allow for the rapid and flexible relation of different social perspectives like value systems between one's own egocentric perspective and many others at different levels of familiarity.

Reference frame transformations in abstract domains

Despite the capacity of cognitive maps to flexibly translate multidimensional information between reference frames, there is limited evidence of non-linear transformations in human abstract cognitive mapping studies (however see Aronov *et al.*, 2017; Omer *et al.*, 2018; Danjo *et al.*, 2018 for evidence in model organisms). One human neuroimaging study investigated these non-linear transformations in a 1D space of different individuals' social preferences by having participants learn a stranger's preference for an everyday activity relative to one of three personally known individuals and subsequently decided how the stranger's preference related to the other two individuals' preferences (Figure 2; Kaplan and Friston, 2019). Making the non-linear transformation from social preferences presented in a relative reference frame (themselves or a familiar reference individual) to an absolute (allocentric) one, entorhinal and hippocampal subiculum signals reflected the absolute distance between the ratings of the stranger and the familiar choice options, particularly when preferences learned relative to other individuals need to be related back to the self (Figure 2). Kaplan and Friston (2019) findings suggest that human entorhinal cortex and hippocampal subiculum, which are known to integrate egocentric and environmental spatial cues during navigation (Hartley *et al.*, 2013), also help people relate others' social preferences from multiple reference points to their own. It is important to note that the authors did not test whether this one-dimensional transformation of abstract knowledge was specific to social knowledge, or how it might extend to multiple dimensions. Notably, there is some behavioral evidence that non-linear transformations of abstract knowledge from an egocentric reference frame to an absolute one can be made along two dimensions. Kuhrt and colleagues (2021) used immersive virtual reality technology to test whether people can navigate abstract knowledge in the same way that they navigate the physical world. Crucially, the abstract knowledge was organized in a collection of different shapes of varying shades that were systematically controlled for visual input (e.g. luminance, optic flow). The authors found that participants learned to navigate using a first-person perspective and formed accurate representations of the abstract space. Navigation in the quantity space resembled behavioral patterns observed in navigation studies using environments with natural visuospatial cues, where non-linear transformations between the participant's first-person viewpoint and a cognitive map of the abstract environment were made. Taken together, these studies represent promising first steps in determining how hippocampal–entorhinal non-linear transformations between egocentric and other reference frames could allow humans to flexibly and rapidly switch between using different points of reference when gathering various types of abstract knowledge.

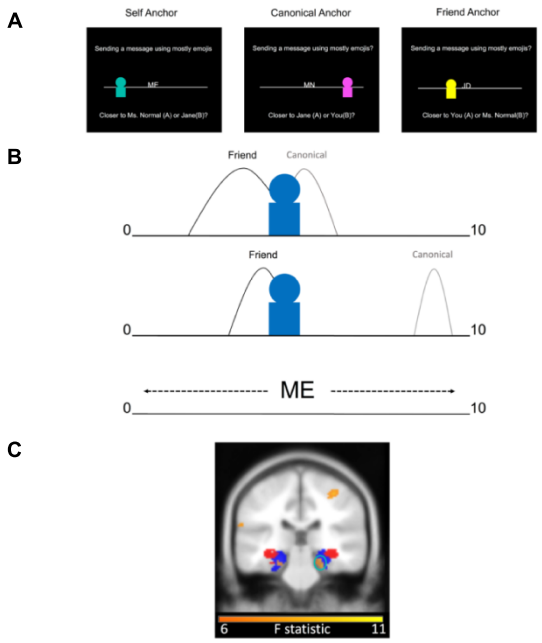

Fig. 2. Social coordinate transformations by the human hippocampal–entorhinal system. A) In (Kaplan and Friston, 2019), subjects were instructed to choose a friend. Subsequently, subjects rated from 1–9, on a 0–10 scale, how likely (likelihood) they (self), a close friend (friend) and the typical person (canonical) were to partake in a variety of everyday scenarios (e.g. eat spicy food, read a book, cycle to work), as well as subjective confidence ratings on a 1–5 scale. To allow for strangers with more extreme ratings than the familiar individuals in the fMRI paradigm, subjects were restricted to rating between 1 and 9. During a forced-choice fMRI task, subjects made a decision on the relative proximity of a stranger's likelihood rating for an everyday scenario relative to the likelihood ratings for the self, friend and canonical individuals for that same scenario. On each self-paced trial, subjects viewed a personal preference for a new stranger presented relative to one of the known individuals' initials on a number line (the anchor individual). Subjects had to determine which one of the two remaining familiar individuals was closer to the stranger's rating. Crucially, the anchor individual (e.g. ME = self; MN = canonical 'Mr./Ms. Normal;' JD = friend's initial) was always placed in the middle of the scale, ensuring that subjects had to use memory of their ratings made before fMRI scanning to infer the stranger's absolute preference relative to the anchor's true rating and the ends of the number line. Additionally, subjects were instructed that the number line ranged from 0 to 10. For example, if the stranger was three-fourths to the right of the anchor with a rating of 9, the participant would infer that the stranger's rating was three-fourths of the way between 9 and 10 (the right boundary of the scale). Notably, the numbers on the scale were not visible during the task. B) Top illustration shows an ambiguous, less discriminable choice, while the bottom illustration shows a straightforward, highly discriminable choice. We quantified the difficulty of discriminating a particular choice by fitting a formal signal detection model, based on the absolute distance between the two choice individuals on the scale and how confident subjects were in their ratings (e.g. comparing the stranger's rating, represented by the blue avatar, with their rating for their friend and the canonical individual). Subjective confidence was represented by the standard deviation for each rating (e.g. curve width in the illustrations), where lower confidence entails higher standard deviations, and helped account for the influence of memory on choice behavior. Center. Illustration of how the stranger's rating is inferred by mentally rescaling the anchor individual's rating from its perceived relative position on the screen (5), to its absolute position (participant's rating) on the preference scale. Bottom. Entorhinal/subicular region exhibiting an effect of absolute distance between the preferences of different familiar individuals circled in turquoise. For visualization purposes, entorhinal/subicular (blue) and hippocampal body (red) probabilistic masks are presented. Figure adapted from (Kaplan and Friston, 2019).

Transformations in social frames of reference

It is now becoming increasingly recognized that non-linear reference frame transformation processes are not limited to spatial cognition. Yet, it is unknown which other cognitive domains might also benefit from this type of coding scheme and along which dimensions this may occur. One promising application for these transformations could be in the domain of social cognition, where individuals could infer the preferences of others and rapidly relate them to one's own preferences and social networks of varying proximities. Furthermore, non-linear transformation of abstract knowledge would enable better differentiation between similar options during decision-making such as the flexible implementation of abstract decision boundaries (e.g. what is the threshold for deciding when a food is too spicy). Previous accounts have highlighted the importance of neocortical coordinate transformation in guiding self *vs* other comparisons (Chang, 2013), particularly reward outcome assignment during social exchanges between two individuals. However, a theoretical role for hippocampal–entorhinal coordinate transformations involving the integration of multiple points of reference (e.g. >3 individuals) or cognitive collage (Tversky, 2014) in a cognitive map of the social world has been missing in these accounts (Kaplan and Friston, 2019). The lack of a role for the hippocampal formation in social coordinate transformations might be due to the perceived limited role of allocentricity beyond the spatial domain.

The lack of literature investigating the use of abstract cognitive maps beyond simple relations is emblematic of the lack of translation in psychological research to cognitive neuroscience, especially when investigating spatial primacy in other cognitive domains (Boroditsky, 2000; Tversky, 2000; Boroditsky and Ramscar, 2002). One example of the dearth of translation between spatial and social cognition is the lack of research on the use of allocentric processing in social neuroscience. For instance, it is notable that most of the aforementioned findings provide insights into how an individual can store knowledge about the social world, but does not actually involve *navigating* the social world (e.g. flexibly switching between the perspectives of others, using familiar individuals as a reference to learn about others). Putative neural computations from spatial navigation research that could be used to explain these social behaviors include allocentric coding and non-linear transformation of information learning via multiple points of reference. In particular, non-linear transformation of social attributes learned via multiple social points of reference could be used to relate self *vs* other comparisons to social networks or society in general. Through a clearer theoretical incorporation of these spatial concepts in the social domain, we believe that the role of the hippocampal–entorhinal region in social neuroscience can be better understood. In what follows, we draw upon neuroscientific evidence from animals and social psychology studies in humans to highlight potential roles for the hippocampal–entorhinal system in social cognition, particularly social perspective switching between more than two individuals.

Are there similar hippocampal–entorhinal cell types for spatial and social perspective-taking?

Hippocampal-entorhinal spatial coding principles can inform human social behavior and we'll detail in this section how specific subregions in the hippocampal–entorhinal circuit inform self *vs*

other processing. We'll then speculate how that circuit could work in concert with other brain regions like the retrosplenial cortex (RSc) and the TPJ during social perspective switches and mentalizing. Taking inspiration from neural models of spatial imagery (Byrne *et al.*, 2007; Bicanski and Burgess, 2018; also see Gaesser *et al.*, 2018; Gaesser, 2020), we relate spatial perspective-taking models to coordinate transformations in the social domain. The hippocampal formation plays an important role in assimilating egocentric and allocentric spatial coordinates during navigation. Furthermore, these sorts of coordinate transformations are also thought to be important in the social domain since stimuli often needs to be translated from relative to absolute frames of reference during social cognition, where there is not typically a one-to-one linear mapping between reference frames (Chang, 2013). Similar to how different types of spatially-modulated neurons guide coordinate transformations during spatial navigation, social coordinate transformations involving an allocentric reference frame likely rely upon a scope of different neuron types in the hippocampal–entorhinal circuit.

There are different hippocampal–entorhinal cell types that afford the transformation of egocentric and allocentric spatial coordinates during navigation. One foundational example is hippocampal place cells that are active when an animal visits a particular location (O'Keefe and Nadel, 1978) and medial entorhinal/subicular grid cells that fire when an animal visits place fields equally placed on a triangular grid (Hafting *et al.*, 2005; Boccara *et al.*, 2010). Together, place and grid cells help dynamically encode the relations and distances between different locations in an allocentric reference frame (Bush *et al.*, 2015; Stemmler *et al.*, 2015). In parallel, head-direction neurons first found in hippocampal–entorhinal subregions fire when an animal is facing a specific direction (Taube *et al.*, 1990). In terms of coordinate transformations, head-direction cells maintain where the direction of an animal's head is in an allocentric reference frame. Boundary and object vector cells in the entorhinal/subicular region fire when an animal is a given distance and/or direction from either environmental boundaries (Lever *et al.*, 2009) or objects (Høydal *et al.*, 2019), respectively. These cell types help transform the relative position between an environmental cue and the animal into absolute, allocentric distances during spatial coordinate transformations. Recent work has begun to uncover how social variables might influence these spatially tuned neurons. Investigating whether hippocampal place cell represent the spatial position of others in an animal's environment, Danjo *et al.* (2018) found that rat hippocampal CA1 subfield neurons coded the spatial locations of others and the animal itself, while Omer *et al.* (2018) observed a similar place coding scheme in bat hippocampal CA1. These social place cell studies found cells conjunctively coding the spatial position and relationship between the self and other animals. Further work is necessary to link existing spatially tuned cells (e.g. grid, HD, boundary vector, object vector cells) to other social variables (e.g. importance, group affiliation, danger) in model organisms. One promising direction may come through investigating spatially tuned hippocampal neurons in model organisms like marmosets that form colonies with clear social relationships such as hierarchies and group roles (Miller *et al.*, 2016).

Beyond spatially tuned neurons, an emerging body of literature in rodents has explored the importance of the hippocampal CA2 subfield in storing memory for other individuals/conspecifics (Hitti and Siegelbaum, 2014; Tzakis and Holahan, 2019). Hitti and Siegelbaum found that genetically targeted inactivation of CA2 neurons in mice caused a dramatic impairment in the animal's ability to remember a conspecific, which the authors define as 'social memory' in their paradigms. Subsequent work in this research line has uncovered an important role for CA2 in encoding, consolidating and retrieving social memories (Meira *et al.*, 2018; Oliva *et al.*, 2020). Beyond social memory, this hippocampal subregion has also been shown to play a key role in social behavior, where a lateral septum-CA2 circuit helps disinhibit social aggression (Leroy *et al.*, 2018). Differing from findings in model organisms, parallel evidence investigating social memory coding in human hippocampal CA2, or any subregion, is lacking.

There has not been evidence of human hippocampal place coding of the spatial location of others. However, there does appear to be some indices of human hippocampal coding of other individuals in a physical environment. The temporal and functional integration of all spatial cell types like place and grid cells are afforded through phase locking with the hippocampal theta rhythm (Barry *et al.*, 2012), which is approx. ∼3–8 Hz in humans and 6–10 Hz in rodents. Linking the human hippocampal theta rhythm to signaling the location of others, Stangl and colleagues found that freely exploring epilepsy patients with chronically implanted wireless hippocampal electrodes exhibited increases in hippocampal theta power for the distance of themselves and other individuals relative to environmental boundaries during navigation (Stangl *et al.*, 2021).

The spatial coding studies in this section raise an important issue about how cognitive factors could play a key role in distorting how the brain represents physical and more abstract spaces. There is now an emerging literature showing spatially tuned cells in the hippocampal–entorhinal circuit can be distorted by different cognitive variables, particularly reward. Building on previous evidence showing that hippocampal replay, the reactivation of place cells related to previously visited spatial trajectories, is stronger for cell representation of rewarding locations (Singer and Frank, 2009; Dupret *et al.*, 2010), Gauthier and Tank (2018) found that mice in a navigation with shifting reward contingencies to distinguish place *vs* reward encoding had specific place cell populations in hippocampal CA1 and the subiculum that only fired near reward locations. In parallel, Butler *et al.* (2019) found that spatial cognitive maps in the rodent medial entorhinal cortex incorporated a learned reward location, and that learning of reward locations across sessions afforded better positional decoding in a cognitive map. These findings in model organisms are not limited to reward either, where a vectorial relationship to navigational goal locations has been observed (Poucet and Hok, 2017; Sarel *et al.*, 2017). Altogether, findings on spatially-modulated neurons with firing fields that are distorted by the presence of a reward or goal gives us an idea of how the hippocampal–entorhinal circuit can incorporate the flexible coding of abstract variables like rewards and social knowledge over time. Moreover, these data provide clues on how lingering cognitive biases can affect spatially tuned neurons and how this might hold even greater importance in humans.

Potential roles for canonical spatially-modulated hippocampal neurons in social cognition

Given the outstanding role for canonical spatially tuned and mnemonic hippocampal–entorhinal cell types in social neuroscience beyond social place cells, we highlight potential repurposed roles for these cells in social cognition (Figure 3, Table 1). Allocentric reference frames and boundaries can be important constructs beyond spatial cognition, especially in the social domain. For instance, humans often use decision boundaries (Kaplan *et al.*, 2017a) or recognize limits of expected social norms

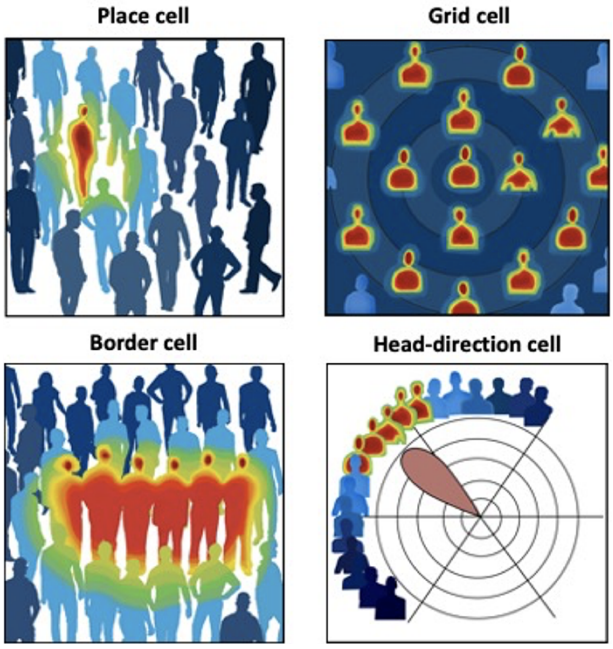

Fig. 3. Implication of major spatially tuned neurons in social cognition. Theoretical potential representation of spatially tuned neurons is illustrated in the social domain. Place cells (upper left) may indicate a specific person within a given social situation; Grid cell (upper right) may indicate a hexagonal organization of a certain social network according to a given 2D coordinate system; Border cell (lower left) may point on a given social boundary (defined group, e.g. close family); Head-direction cell (lower right) may point on a certain 'angle' within the same social layer (e.g. among all cousins, those in my age). Colors indicate level of proximity to the specific person (for further explanations and more cell types see Table 1; Courtesy: Hagar Segev).

and preferences (Kahneman and Miller, 1986). Moreover, allocentric coding as defined by Triandis, having one's interests centered on others over themselves (Triandis *et al.*, 1985), would be expected to play a key role in empathy and theory of mind (i.e. putting yourself in someone else's shoes Frith and Frith, 2003) by contributing a common reference frame that could facilitate social behaviors that are not putatively associated with the medial temporal lobe (MTL; see subsequent section for information about functional interactions between the hippocampal formation and brain regions responsible for theory of mind). In terms of boundary coding, one could envision a role for boundary vector cells and border cells in coding the abstract distance from figurative social boundaries/norms. Such a role could portend social boundaries influencing underlying computations by grid cells in a similar manner as irregular geometric boundaries (Barry *et al.*, 2007; Krupic *et al.*, 2012). Consequently, grid and place signals could potentially fragment and reset due to 'boundaries' between social groups (Figure 3). In parallel, other types of vectorial coding associated with object vector and grid cells could then afford the non-linear entorhinal/subicular transformations between relative and absolute frames of reference observed in human social decision-making tasks (Kaplan and Friston, 2019). However, one caveat is that social boundaries and coordinate transformations could be even more familiarity-dependent than the coding of physical boundaries and transformations, which may result in different neural coding mechanisms.

Notably, grid cells are typically associated with 2D (and sometimes 3D, see Stella and Treves, 2015) spaces. The potential for grid cells to incorporate multiple elements of social knowledge could allow the brain to flexibly compare different individuals based on different social factors (e.g. preferences, status, and threat level). How might grid cells contribute to social perspective switches in continuous spaces? Since grid cells simultaneously code multiple locations, these entorhinal neurons might be able to encode complex relations between members of large-scale social networks or how different social characteristics compose the personality or familiarity of a single individual (Bellmund *et al.*, 2018; Schafer and Schiller, 2018a). Within this framework, head-direction cells could signal the abstract direction for social variables like a rank in a hierarchy (e.g. most popular, generous, etc.; Table 1). Furthermore, reward-driven distortions in hippocampal and entorhinal spatial coding can be expected to translate to social dominance hierarchies, but whether hierarchies might maintain more of a graph- or map-like structure is unclear (Peer *et al.*, 2021b). Formulating social reference points like familiar others (i.e. does person A cook better than person B?, is person A nicer than my best friend?) could therefore be envisioned as a landmark or border within an abstract environment similar to how objects and boundaries are used to test for spatially tuned neural responses (Figure 1, Table 1). We'll detail in a subsequent section how neocortical regions forming the default network might contribute to this process.

How might these social reference frame transformations be computed by the hippocampal–entorhinal system? Recent advances have been made in understanding how hippocampal–entorhinal cognitive maps of physical space and their constituent non-linear transformations can translate to learning and memory more generally. Specifically, this overlap between spatial and social perspective switching could be facilitated by recent computational models. One such model is the Tolman–Eichenbaum Machine (Whittington *et al.*, 2020), which uses factorization (simplification) and conjunctions of representations to simulate how hippocampal–entorhinal knowledge, like social knowledge, can be flexibly (and logarithmically/non-linearly) abstracted. Other potential models that would afford the non-linear transformation of non-spatial social knowledge in the hippocampal–entorhinal system could also be neurally implemented by using neural manifold models (Low *et al.*, 2018). Either of these computational models is well-positioned to test and connect any empirical findings across different species and cognitive domains. Focusing on human cognitive computations, a model that directly addresses how non-linear transformations are implemented in the hippocampal formation and associated cortical midline regions is the agency–ownership (A–O) model (see Box 1 for an in-depth summary). In the A–O model, the brain uses prior information (e.g. beliefs, predictions) and world-information applied on an existing cognitive map to predict a cognitive map of social space via a forward model (Arzy and Schacter, 2019). The model accounts for egocentric-allocentric transformations to correct the predicted map following iterative steps (see Box 1 and Arzy and Schacter, 2019 for details). Further developments and elaborations of such models may better help to understand non-linear transformations in representations of social information within the cognitive mapping system.

Hippocampal formation-neocortical interactions during spatial and social perspective-taking

It is important to examine how the hippocampal–entorhinal region relates to brain areas putatively linked with social interactions, particularly perspective switches. Notably, neocortical

Table 1.

Category Type	Spatial location	Spatio-social conspecific's location	Social network location over structure	Social characters e.g. power-affiliation
Place	Self-location	Others' location	Self-location in social hierarchy	Combined characters (2D)
Grid	Spatial plane	Spatial plane	Social circles	Abstract plane
Boundary	Spatial border	Spatial border for a conspecific	Border between social circles	Qualitative difference in social characteristics
Head-direction	Head-direction	Head-direction for conspecific	People with a similar social role	'Direction' in the abstract space
Reward	Place with reward	Reward for a conspecific	More valued individual	More valued attributes of an individual

regions linked to spatial perspective switches are also implicated in social perspective-taking and form part of the default network (Vogeley and Fink, 2003; Arzy *et al.*, 2006; Yeshurun *et al.*, 2021). These default network regions including the RSc, posterior cingulate cortex (PCC), mPFC and TPJ play key roles in social cognition, memory processes, and have tightly linked functional roles and anatomical connections to the MTL/hippocampal–entorhinal system (Buckner and Carroll, 2007; Buckner *et al.*, 2008, for meta-analyses see Spreng *et al.*, 2009; Dafni-Merom and Arzy, 2020). Accordingly, the default mode network (DMN) has been recently suggested to be the functional brain network that serves as an interface between the self and the social world (Yeshurun *et al.*, 2021). Therefore, it follows that hippocampal–entorhinal components would make a strong contribution to this network in maintaining knowledge and transforming perspectives through the assimilation of new social knowledge with what is already known about the prosocial world (Arzy and Schacter, 2019). More specifically, by examining the well-described role of cortical midline regions and the TPJ in social behaviors like mentalizing, perspective-taking and social learning, we can gain insights into how interactions between the MTL and aforementioned other DMN regions can help facilitate these behaviors.

mPFC interactions

We have highlighted how the hippocampus and entorhinal cortex could help assimilate information from different reference frames during social behavior and nowhere is the presence of different social points of reference clearer in social cognition than with self *vs* other comparisons. Self *vs* other comparisons are commonly studied in the investigation of mentalization, the ability to understand oneself and others (Frith and Frith, 2006). Furthermore, self *vs* other comparisons are primarily linked with the mPFC and right TPJ (Frith and Frith, 2006; Koster-Hale and Saxe, 2013), where the mPFC linearly signals the discrepancy between preference ratings for the self and other (Tamir and Mitchell, 2010). Moreover, functional interactions between the mPFC and MTL, particularly the hippocampal formation, are thought to play a key role in memory-guided decision-making (Biderman *et al.*, 2020), including the social domain. Yet, the relationship between the mPFC and hippocampal formation in guiding social cognition has its limits, where a key cognitive ability derived from mentalizing, theory of mind, is not impaired in patients with focal MTL damage (Rosenbaum *et al.*, 2007). Why might this be? We speculate that the complex structural inferences afforded by hippocampal cognitive maps likely are not necessary during theory of mind tasks, where there is a simple comparison between the self and other (or switching between two first-person perspectives). However, if more than two perspectives need to be integrated then the MTL could be required. Further work testing hippocampal-dependent autobiographical memory in comparison with theory of mind, shows that the right TPJ, but not mPFC, is selectively engaged by theory of mind demands (Rabin *et al.*, 2010). In the same vein, considering theory of mind as a self-projection in the social domain (Buckner and Carroll, 2007), a recent study found implications of the mPFC, anterior cingulate cortex (ACC), right TPJ and medial parietal cortex, but not the MTL (Hayman and Arzy, 2021).

Another insight may derive from an in-depth dissection of the mPFC itself. Although the mPFC is associated with theory of mind, there appears to be functional dissociations between dorsal and ventral areas of rostral mPFC during social decision-making tasks. Namely, dorsal portions of mPFC were engaged when making value-guided choices for others or when someone else decided for the participants (Nicolle *et al.*, 2012). In contrast, ventral portions of mPFC, putatively associated with autobiographical memory, are engaged when making choices for oneself. This functional gradient in mPFC might not be limited to the social domain either. In a multi-step spatial planning task, Kaplan *et al.* (2017b) found that more dorsal areas of rostral mPFC were engaged when planning next-step choices, while ventral mPFC (vmPFC) was engaged by easier initial and next-step choices. Notably, functional connectivity between the dorsal mPFC region and hippocampus reflected the subsequent accuracy of next-step choices. Taken together, the findings by (Kaplan *et al.*, 2017b) and (Nicolle *et al.*, 2012), highlight a potential domain-general role for dorsal regions of mPFC in thinking beyond a current subjective state to facilitate prospective thinking or adopting another person's perspective. Furthermore, the mPFC might hold a more similar functional role to the MTL in learning and decision-making than other neocortical default network regions (Kaplan *et al.*, 2017b), where vmPFC (and rodent orbitofrontal cortex; OFC) is thought to maintain a cognitive map of task space (Wilson *et al.*, 2014). In this formulation by Niv (2019), the vmPFC maintains the current location within a multi-faceted task and serves as a pointer to where specific knowledge is stored in the hippocampus during decision-making. In this context, one could envision the mPFC/OFC as a pointer to hippocampal–entorhinal social knowledge of both the self and other individuals during social decision-making, which could be enhanced by the mPFC's known role in enhancing spatial imagery (Kaplan *et al.*, 2017c; Gaesser *et al.*, 2020).

Parietal midline interactions

The parietal midline, including PCC, precuneus and RSc, is a key region in the default network. In parallel, parietal midline regions like the RSc are known to play a key role in spatial coordinate transformation (Epstein, 2008). Specifically, RSc is thought to translate between different perspectives of the external world such as viewpoint-dependent and viewpoint-independent reference frames during navigation and other spatial behaviors

Box 1. Computational models for nonlinear transformation in cognitive mapping

The Agency–Ownership (A-O) model is an example of the cross-talk between the entorhinal–hippocampal system and the cortex (mostly the default mode network, DMN). In the A–O model (Figure 6), the agency-branch (bottom-up) uses self-originated (top-down) agency-derived assumptions and the current generated cognitive map to predict the next version of the map using a forward model. The ownership-branch (bottom) uses one's personal schema of ownership and egocentric-allocentric translations (difference between the predicted allocentric map and the actual egocentric feedback perceived by path-integration and perceived world-information) to correct the predicted map following iterative steps. The next generated map is dependent on these two processes (with different weighting; see Arzy and Schacter, 2019 for details).

In each step, the previously stored map is read out, together with new information from the real-world ('world-information', w_{real}). The measured information ('measurement', w_m) is then used, together with the previously stored map, to generate a new map which estimates the real-world using a stochastic model. Note that there is always a gap (error) between the estimation (in the social world: 'what a person expect' (FeldmanHall and Nassar, 2021) represented on the map (w_{est}) and the real-world ('what a person gets'), due to the combination of noise and stochasticity in the measurement, read out and model. The estimated map is compared with the real-world measurement to generate an error correction (e.g. a gain factor G(Δ); Kalman, 1960), which determines the relative weight of the measurement and the model estimate in the updated cognitive map. G(Δ) is a monotonically increasing function of the prediction error Δ, such that when Δ is large, more weight is given to the world measurement, and when Δ is small, the output is closer to the model's estimate, a process that involves non-linear changes. When new world-information is encountered, the process reiterates. The stored cognitive map after n updates is denoted by w^n. A probability distribution $P_m(w_m|w_{real})$ accounts for noise introduced by both the measurement and the readout. Similarly, there is a probability distribution for the estimation $P_{est}(w_{est}|w^n)$. Note that according to the A–O model the measured information may be egocentric (encountered by the subject) or allocentric (derived from an external source).

The simplest realization of the A–O model handling a social network data is a linear model in which each connection is mapped using a deterministic function:

$$w_{real} = w_m, w_{est} = w^n$$

This model 'copy-pastes' learned information about one's social network into the data in a one-to-one manner. In a more realistic, yet still simple version of the A–O model, we assume that the measurement and the readout of the encountered social information are Gaussian variables:

$$W_m \sim N(w_{real}, \sigma_m),\ W_{est} \sim N(w^n, \sigma_e)$$

where σ is the standard deviation and N indicates the number of connections within one's social network. Another potential non-linear realization of one's social network is a neural network model. In this realization, each individual in the social network is represented by a single unit i. An existing social connection between two individuals in the social network is represented as an interconnection between two units W_{ij}. Therefore, one's social map may be represented by the connectivity matrix of the participant's 'neural population' (Each neuron is assigned with a total synaptic current, which is the total input to the neuron, filtered by an exponential filter representing the synaptic dynamics; several characteristics of neural networks such as synaptic current, firing rate and noise should be further introduced to create the network dynamics).

(continued)

Box 1. (Continued)

The linear model yields a monotonic decrease in the error-per-connection. Note that such a model assumes a perfect memory and zero measurement noise, which are not realistic from a neuro-cognitive POV. In the logarithmic model, the error monotonically decreases, and the error-per-connection saturates (after number of iterations going to N^2). The model performance depends on the real-world measurements and readouts and the decay rate of the gain function. Finally, the neural network model is dependent on the measured social network, and specific values of the parameters involved in the learning (the learning rate and the weight decay rate) lead to different learning performance for each participant. Since agency, ownership, and cognitive mapping are managed in the hippocampal–entorhinal system and DMN, we speculate that such a model implements non-linear transformations in these systems.

(Vann *et al.*, 2009). One potential mechanism for RSc's translation of spatial cues comes from RSc's perceived role in integrating different viewpoints, where the RSc uses spatial mental imagery to merge different viewpoints into a cohesive allocentric representation of the environment (Byrne *et al.*, 2007; Bicanski and Burgess, 2018). Furthermore, recent evidence portends a role for parietal midline structures in similar types of viewpoint changes in social interaction tasks (Lombardo *et al.*, 2010). This role is not limited to social viewpoints and is thought to extend to maintaining social knowledge. Investigating how participants characterize the relationship between individuals in their real-life social networks with fMRI, Peer *et al.* (2021a) found that the RSc maintains knowledge of the higher-order organization of different individuals comprising a social network, while other parietal midline structures maintains knowledge of the affiliation of others in relation to the self. This ensemble of activity was suggested (Peer *et al.*, 2021b) to obey the form of a 'cognitive-graph', which is a representation of topological connection in the environment (Chrastil and Warren, 2014). Yet, in contrast with the hippocampal–entorhinal circuit, medial parietal areas are capable of making self-other perspective shifts without needing to relate to the world at large. Furthermore, medial parietal areas can maintain and translate social information during tasks without a strong mnemonic component unlike hippocampal-dependent social cognition tasks. Still, further work is necessary to differentiate medial parietal *vs* hippocampal formation contributions to translating spatial and social reference frames.

TPJ interactions

The right TPJ is known to play a key role toward integrating self *vs* other perspective-taking in theory of mind (Saxe *et al.*, 2006), where right TPJ signals perspective changes (Hayman and Arzy, 2021). Right TPJ was found to be involved in spontaneous theory of mind (Bardi *et al.*, 2017). Moreover, activation of this region decreased as a function of the psychological closeness between participants and others (Ionta *et al.*, 2020). Notably, functional links between the hippocampal–entorhinal system and TPJ in either spatial or social perspective-taking remain unclear. One possibility might be that theta phase coupling between the hippocampus and neocortical areas like TPJ might guide the integration of different perspectives (Byrne *et al.*, 2007). Indeed, there is preliminary evidence supporting this notion. A magnetoencephalography-transcranial magnetic stimulation (MEG-TMS) study found that theta oscillations in the right TPJ

help transform the embodied self into another's viewpoint (Wang *et al.*, 2016), which is a process that can be selectively interrupted with TMS stimulation (Wang *et al.*, 2016; Martin *et al.*, 2020). Notably, disrupting the right TPJ, but not MTL, seems to conserve active recruitment of prosocial intentions in mentalizing (Gaesser *et al.*, 2019). Finally, the TPJ is involved in logarithmic transformations along 'mental-lines' (Dehaene *et al.*, 1999; Harvey *et al.*, 2013; van Dijk *et al.*, 2021; For discussion of the mental lines see in the next section).

The default network initially gained prominence as one of the most robustly observed resting-state networks in brain imaging (Raichle *et al.*, 2001). Subsequently, Buckner and colleagues highlighted a role for the default network in internal mentation which includes cognitive processes like autonoesis, episodic memory, mentalizing, imagination and fictive planning (Buckner *et al.*, 2008; Dafni-Merom and Arzy, 2020). Earlier work highlighted a role for the default network, including the hippocampus, in self-projection and episodic future thinking (Schacter, Buckner *et al.*, 2008). These ideas have led to the default network being characterized as a key center for maintaining endogenously driven cognitive processes that occur on longer timescales (Hasson *et al.*, 2015; Yeshurun *et al.*, 2021), even during offline 'resting' states like sleep (Genzel, 2020). At the circuit level, hippocampal sharp-wave ripples are oscillatory events that occur during resting periods like sleep, as well as awake deliberative periods. Sharp-wave ripples are associated with memory replay since reactivation of hippocampal place cell ensembles co-occurs with ripples and have been related to successful memory consolidation. Hippocampal sharp-wave ripples have been shown to influence ongoing neocortical fMRI signals (Logothetis *et al.*, 2012), which suggests that ripples may influence default network function. Indeed, this idea is supported by findings showing that hippocampal ripple events modulate the ongoing signal in the default network (Kaplan *et al.*, 2016), which has been shown both invasively (Mitra *et al.*, 2016) and non-invasively (Higgins *et al.*, 2021) in humans. Taken together, these studies support the idea that hippocampal replay of past experience may help the default network assimilate knowledge and explore potential outcomes of future decisions from different social perspectives (Dafni-Merom and Arzy, 2020).

An alternative theory for the potential role of the hippocampal formation in social perspective-taking is based on vivid spatial imagery from the hippocampus having an explicit, additive role on mentalizing and social perspective-taking (Gaesser, 2020), similar to the effect method of loci and memory palaces have on episodic memory performance (Spence, 1984). In other words, the Gaesser 'episodic mindreading hypothesis' (2020) postulates that as an episode is more vividly experienced, the mental states of the target in that episode will become more accessible in memory and more readily utilized. Notably, our more implicit framework and the more explicit episodic mindreading hypothesis are not mutually exclusive. The hippocampal–entorhinal circuit facilitates spatial coordinate transformations and imagery, so there remains the possibility that both these capabilities are conserved in the social domain. Still, future research is necessary to formally compare these implicit and explicit additive hippocampal effects on prosocial interactions.

The aforementioned findings in this section imply potential non-linear transformation at the different regions of the default network—logarithmic in the TPJ (see Table 2 below), graph-like in RSC and cell-population-dependent in the MTL. Barbara Tversky (1993) highlighted the significance of all these three by pointing to the importance of different point of views leading to a coherent framework of mental navigation (Tversky, 1993, 2011). According to Tversky, this should be called a 'mental collage' as different points of view are integrated altogether, which meshes nicely with ideas about the default network as an integrative social interface between the self and the external world (Yeshurun *et al.*, 2021).

Unifying spatial and social perspective-taking with the concept of distance

The idea that real-world entities in different domains such as space, time and social networks, are geometrically represented within the brain enabled the inclusion of all these domains into a coherent cognitive framework (Tversky, 1993, 2011; Tversky and Hard, 2009; Arzy and Schacter, 2019). Likewise, spatial and social cognition may use similar concepts and cognitive constructs (Tversky *et al.*, 2013).

One of these potential fundamental concepts is cognitive distance. Cognitive distance is defined as 'mental representations of large-scale environmental distances that cannot be perceived from a single vantage point but require movement through the environment for their apprehension' (Montello, 1991; see also Golledge, 1969, 1978). In other words, cognitive distance is *traversed* by the experiencing self. Accordingly, psychological distance is by definition egocentric, as it is measured or evaluated by the traversing self. Psychological distance is not limited to the spatial domain. It may be found in time: the difference between the present time and that of a past experience or future plan, in the social domain: the estimated 'distance' between the self and a person in one's social world (i.e. how close one feels to another person), and actually in each domain in which a distance dimension may be constituted (Trope and Liberman, 2010; Parkinson *et al.*, 2014; Peer *et al.*, 2015; Dafni-Merom and Arzy, 2020). According to the construal level theory (CLT), the comparison of the self to the imagined target (in space, time or person) involves mental construal, and there is a direct relation between the psychological distance traversed and the mental construal required. Moreover, according to CLT, the psychological mechanisms used to both mental construal and psychological distance traversal in the three domains are similar. Interestingly, Trope and Liberman (2010) link first- and third-person perspective-taking to psychological distance and mental construal: a third-person perspective imposes more distance than a first-person perspective, and as such induces a higher level of construal. Social distance underlies the effect of social discounting, which is the observation that one's generosity toward others decreased as a non-linear (hyperbolic) function of the perceived social distance between them (Jones and Rachlin, 2006). Neuroimaging studies found social discounting to be crucially related to brain activity at the TPJ, as based on the social distance (Strombach *et al.*, 2015; Soutschek *et al.*, 2016; Sellitto *et al.*, 2021).

Here, we highlight the importance of three different cognitive constructs: mental travel/self-projection [from one point-of-view (POV) to another], self-reference (the cognitive distance between the self to each of the represented points) and non-linear transformations (of distances between the POVs, self-to-point or the different points), which all create 'mental lines' of different sorts. These three aspects are using distances to represent important information, through a geometrically-based system that may underlie the different domains of spatial, temporal and social cognition. In the following paragraph, we introduce mental lines, then decipher their underlying governing principles and

Table 2.

Brain region	Spatial perspective changes	Social perspective changes
Medial Prefrontal Cortex/Medial Orbitofrontal Cortex (mPFC/mOFC)	Directing mental imagery and integrating imagined viewpoints from different perspectives	Mentalizing/assuming the perspective of another individual in dorsal regions, while ventral regions maintain a cognitive map of a social task space and pointer for social information stored in the hippocampal–entorhinal system
Retrosplenial Cortex (RSc)	In concert with the hippocampal–entorhinal circuit helps integrate different perspectives of the external world such as viewpoint-dependent and viewpoint-independent reference frames during navigation	Potentially helps integrate different social perspectives in a graph-like absolute reference frame
Posterior Cingulate Cortex (PCC)	Mental imagery and memory retrieval	Infers direct linear trajectories over implicitly learned social spaces
Temporoparietal Junction (TPJ)	Mental rotation/logarithmic coordinate transformations	Logarithmic social perspective change

subsequently focus on discussing distance in social and spatial domains.

One of the main examples of non-linear (logarithmic) transformations in the human brain is mental lines. The archetype of mental lines is the mental number line (MNL; Banks and Hill, 1974; Dehaene *et al.*, 1993). The MNL is characterized by two main features. First—directionality: numbers are represented by humans (and perhaps other species as well; e.g. Rugani *et al.*, 2015) on a horizontal spatial axis running from left to right (Dehaene *et al.*, 1999); second—logarithmic transformation: experiments on mental number scaling in archaic cultures or children revealed the mapping of ordinal numbers along a logarithmic scale rather than a linear one (Dehaene and Cohen, 1995; Siegler and Booth, 2004; Dehaene *et al.*, 2008). Neuroimaging studies localized representation of the MNL to the right TPJ (Dehaene *et al.*, 1999; Harvey *et al.*, 2013; van Dijk *et al.*, 2021). Interestingly, these two features characterize not only the MNL but also other mental lines. The mental time line represents the arrow of time from past to future running from left to right, and response time to events are distributed in a logarithmic manner (Bueti and Walsh, 2009; Arzy *et al.*, 2009a), and the mental emotion line—from negative to positive (Holmes and Lourenco, 2011), localized as well to the TPJ. The original depiction of mental lines refers to the characteristics of the line itself, namely its directionality and logarithmic arrangement. There also is the relation of the different points represented on the line to the location of the experiencer ('self') on this line. To reiterate, we hypothesize that the logarithmic character of the line is derived from the localization of the human self-representation along such a line. This enables prioritization of certain items (such as events, people or places) closer to the self, where two main metrics are considered. The first is the ability to project oneself to a specific 'self-location' along the mental line, which is termed 'self-projection.' In the time domain, for instance, one might 'project' herself to different time-points in the past in order to re-experience an event, an ability referred to as 'mental time travel' (Addis *et al.*, 2007; Arzy *et al.*, 2009b, 2008; see also Schacter *et al.*, 2012). Such 'mental travel' or 'self-projection' is also implacable in other domains. In the emotion domain, one might imagine themselves in different emotional states (Holmes and Lourenco, 2011), and in the spatial domain—at different places (Gauthier and van Wassenhove, 2016). In the social domain, people might project themselves to be 'in the shoes' of other people, a crucial ability for the aforementioned theory of mind (Frith and Frith, 2003, 2006; Buckner and Carroll, 2007; Hayman and Arzy, 2021). Experiments of self-projection to different points on the mental lines (e.g. Arzy *et al.*, 2008; Gauthier and van Wassenhove, 2016; Hayman and Arzy, 2021) show that such a projection requires more mental load (as is expressed by behavioral measures as well as functional neuroimaging ones in specific brain regions) with respect to one's own self-location. However, study of the distribution of 'projection' to different 'locations' on a mental line is still needed to investigate the systematic mathematical definition of projections' distribution.

Self-reference is defined here as the ability to refer to different items on the mental line (e.g. events, places, people, emotions) from one's projected or habitual self-location (but see Rogers *et al.*, 1977). In the time domain, for instance, self-reference is the cognitive distance in between one's self-location in time (the present time or an imagined point in the past or future) and an event they recall (see also Liberman *et al.*, 2002; Herzog *et al.*, 2007). Plotting of reaction times to events in different cognitive distances from one's self-location has been shown to follow a logarithmic distribution (Arzy *et al.*, 2009a). Likewise, ease of retrieval mediated the effect of temporal distance on attitudes (Herzog *et al.*, 2007). Taken together, this suggests that similar logarithmic organization may also be found over other mental lines, such as the social one (Figure 4).

Clinical implications of the overlap between social and spatial cognition

Tolman (1948) once posited that 'social maladjustments can be interpreted as narrowings of our cognitive maps due to too strong motivations or to too intense frustrations.' While this hypothesis is somewhat speculative, the idea of looking at social disturbances through the construct of cognitive mapping remains a promising potential avenue of exploration. In the intervening decades since Tolman formed this idea, the relationship between cognitive maps and neurodegenerative/psychiatric dysfunction is still largely unexplored. Tolman's idea might be better formalized as 'social disorientation' instead of 'maladjustment', as orientation is what one is performing over a map (Peer *et al.*, 2014; Schafer and Schiller, 2018b). Disorientation in the social domain may affect one's own self-location (depersonalization), self-projection to the perspective of others (as in theory of mind), reduplication of self and others or confusion between self *vs* other. In the following paragraph, several exemplary social disorientations are presented and the role of the entorhinal–hippocampal system in these disorders is discussed.

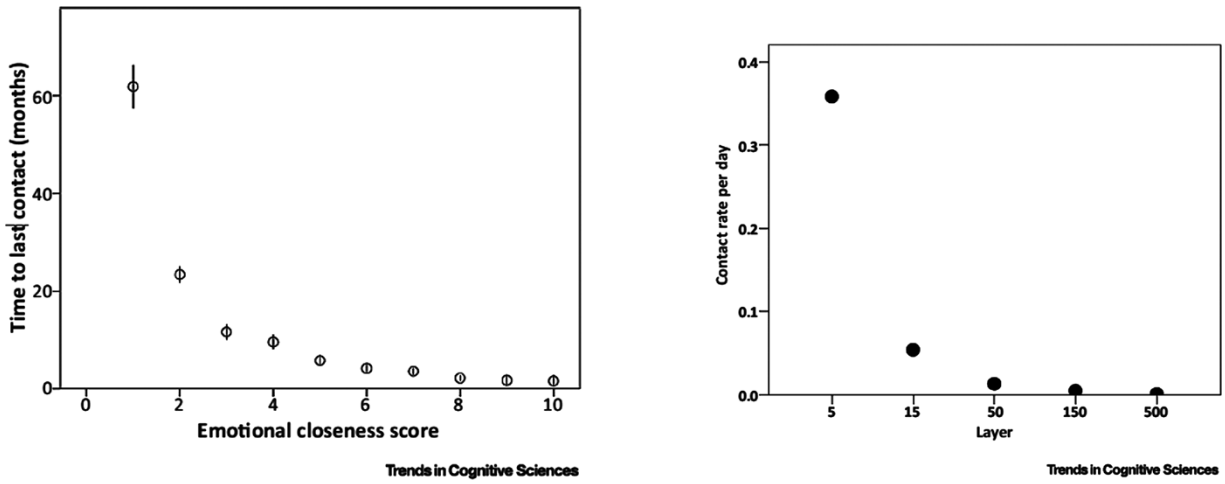

Fig. 4. Logarithmic distributions in the social domain. A. Emotional closeness and the temporal distance from last contact. Mean (±standard errors) temporal distance (months) from last contact is plotted against emotional closeness scale score (1 = low, 10 = high; $n = 20\,249$). Note the decrease of emotional closeness with prolongation of the temporal distance (from Roberts *et al.*, 2009). B. Frequencies of contact. Frequency of contacts one maintains (per day) with others in each layer of one's social network. Note the decrease in contact ray with the distance of the social layer (from Sutcliffe *et al.*, 2012).

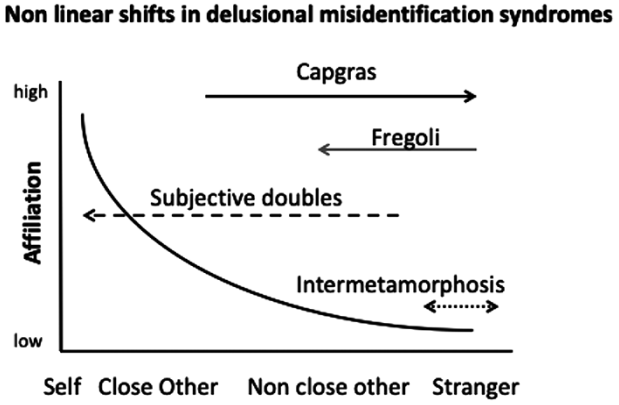

Fig. 5. Shifts along the mental social line in delusional misidentification syndromes. A schematic description of the mental social line is shown, with logarithmic distribution of self, close other, non-close other, and stranger according to their level of affiliation to the self (see Hayman and Arzy, 2021). Arrows depict shifts in proximity in four syndromes (Capgras, Fregoli, subjective doubles and intermetamorphosis), arrowheads depict direction of shift.

Disorientation in the person domain plays a major role in several neuropsychiatric disorders (Peer *et al.*, 2014). Self-referenced personal disorientation can be paraphrased as a detachment of one's self from their own world. It mainly involves the symptoms of derealization and depersonalization. Depersonalization is defined as 'feeling detached from, and as if one is an outside observer of, one's mental processes or body (e.g. feeling like one is in a dream); sense of unreality of self, perceptual alterations; emotional and/or physical numbing; temporal distortions' (Spiegel, 1997). Patients with depersonalization may describe that 'I am not myself anymore, I mean, I know who I am but I just feel like I may be somebody else' (Arzy *et al.*, 2011). In contrast, derealization is defined as the feeling that one's surroundings are not real. Both depersonalization and derealization are hallmarks of dissociative states (Spiegel, 1997; Kihlstrom, 2005; Spiegel *et al.*, 2013), and are frequently observed together. One such entity is 'dissociative fugue' that may be seen in temporal lobe epilepsy (Staniloiu and Markowitsch, 2014) in which patients may travel around while disappearing from their usual environment with no memory of this period after recovery. Functional neuroimaging studies identified such phenomena with alterations (hyperactivation) in mPFC and the TPJ (Reinders *et al.*, 2003; Arzy *et al.*, 2011) hinting at the role of perspective-taking in such disorders. Involvement of the MTL was found when mnemonic functions were implicated in the form of hypoactivation (Staniloiu and Markowitsch, 2014). It is possible that the balance in between the hippocampal–entorhinal mapping system and default network regions is impaired in dissociative disorders, where the latter compensates or governs over the first. Further neuroimaging and computational works (as based on platforms such as the BB or A–O models), may help to further understand these intriguing disorders. This is similar to a proposal regarding the very nature of post-traumatic stress disorder (PTSD): according to this proposal PTSD may be subdivided into dissociative and non-dissociative subtypes. The dissociative subtype is characterized by overmodulation of affect, while the (more common) non-dissociative subtype involves the predominance of re-experiencing the traumatic event itself and hyperarousal symptoms (Lanius *et al.*, 2010). These proposals may be supported by a recent neuroimaging finding, where patients with PTSD related to childhood traumatization showed a smaller hippocampal volume with respect to controls (Chalavi *et al.*, 2015). Additionally, patients who also showed dissociative symptoms had abnormal shape and significantly smaller volume in the CA2-3, CA4-DG and (pre)subiculum parts of the hippocampus as compared with healthy control subjects (Chalavi *et al.*, 2015). Further clinical, experimental and computational work may contribute to the understanding of this common and significant disorder of PTSD.

Social disorientation may be regarded as disorientation in person which involves other people (non-self-referenced), and is generally termed delusional misidentification syndromes (DMS; Feinberg and Roane, 2005; Devinsky, 2009). In DMS, family members and friends are wrongly identified as different people (Christodoulou, 1978; Feinberg and Roane, 2005; Devinsky, 2009; Peer *et al.*, 2014). Specific syndromes may be conceived as a 'shift' of people's 'location' along the 'mental social line' (Figure 5).

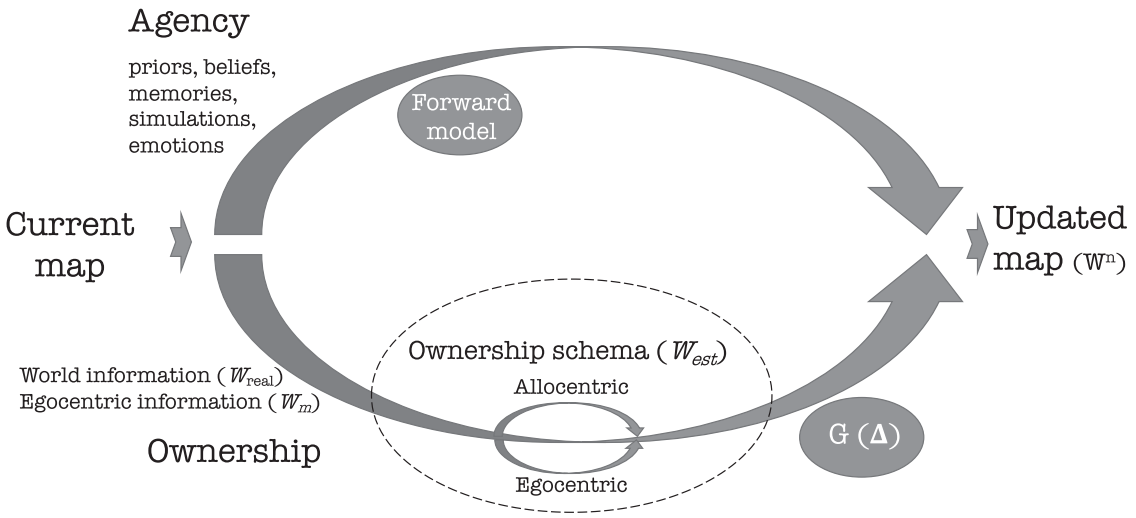

Fig. 6. The Agency–Ownership (A–O) model. In the A–O model, the agency-branch (upper arm), is carrying information related to the agent, that is, it uses self-originated (top-down) agency-derived assumptions (e.g. beliefs, memories, simulations) and the current generated cognitive map to predict the next version of the map implying a forward model . The ownership-branch (bottom arm), carries information 'owned' by the agent while gathering world-information (W_{real}, allocentric). Information may also be measurements 'met' by the individual while navigating the world (W_m, egocentric). The ownership schema is a mechanism that mediates in between the egocentric and allocentric information to create an estimation of the world (W_{est}) by the navigator, which is further subjected to an error correction (Adapted from Arzy, 2021).

Likewise, in *Capgras syndrome* there is a shift of familiar persons (usually close others) to strangers, as they are taken to be impostors (hypofamiliarity; Capgras and Reboul-Lachaux, 1923; Blanke *et al.*, 2008). In contrast, in the syndrome of *Fregoli* the shift is done in the other direction, where a stranger is perceived as a familiar person (hyperfamiliarity; Hudson and Grace, 2000). In the syndrome of *subjective doubles*, the shift is different: the patient believes that a 'double' of themselves exists, and this double may be in relation with a close other (e.g. pretends to be her husband's wife), a non-close other or a famous person (contemporary or revelations of a mythic one; Christodoulou, 1978); Finally, shifts may be seen in the same level of proximity in the condition of *intermetamorphosis* in which another person is conceived as a different person (Malliaras *et al.*, 1978; Bick, 1984).

The neuroanatomy of DMS varies, and case studies highlight lesions in the frontal, temporal and parietal lobes, mostly on the right hemisphere (for review see Feinberg and Roane, 2005; Devinsky, 2009; Peer *et al.*, 2014). The main area for DMS is the right frontal lobe (Capgras and Reboul-Lachaux, 1923; Blanke *et al.*, 2008; Darby and Prasad, 2016). Dysfunction of the dorsal visual pathway and disconnection in the temporal lobe between the fusiform face area and the amygdala have been proposed to account for some of the syndromes (Ellis and Young, 1990; Hirstein and Ramachandran, 1997; Ramachandran and Hirstein, 1998). Disconnection between the frontal lobe (favoring delusional states) and the right temporal lobe (disturbance in face processing, and specifically impairing the association between familiar faces and emotional values) was also suggested (Nuara *et al.*, 2020). Fregoli syndrome was also associated with right frontal and left temporoparietal lesions (Feinberg *et al.*, 1999) as well as lesions in the FFA and the MTL (Hudson and Grace, 2000) in a potentially similar mechanism to Capgras syndrome. Capgras and Reboul-Lachaux proposed in their seminal paper that dysfunction of brain regions involved in the experience of familiarity are underlying the disorder (Capgras and Reboul-Lachaux, 1923). Using a novel method of 'lesion network mapping' applied on 17 cases, Darby and colleagues (Darby *et al.*, 2017) found all lesions underlying these cases to be connected to a default network region, the left RSc. As a result, it is thought that to suffer from a DMS, a 'double hit' is needed including a specific lesion and a disconnection to the RSc. Given the RSc's putative role in maintaining abstract distance information about relations in social networks (Peer *et al.*, 2021a), these findings suggest that DMS may be considered as a combination of a person-specific and domain-general orientation disorder. The latter may involve either an orientation-related mechanism or a specific module-related to egocentric-allocentric transformation, while the first implicates the person domain. Considering the findings by Darby and colleagues, it is possible that DMS can be interpreted as disturbances in the cross-talk between domain-general and domain-specific mechanisms. The domain-general one involves the mechanism of cognitive mapping and/or ego-allocentric representation of such a map, resulting in disorganization, duplication and incorrect placement of locations, events and people, while the domain-specific one in DMS is the person domain. Future research is still needed to investigate this hypothesis.

In spite of the aforementioned syndromes' insights into social cognition and the underlying neuroanatomy, they are primarily anecdotal both by their frequency and their usually transient effect on patients' health. In contrast, a significant and common disorder in which such shifts may play an important role is Alzheimer's disease (AD; Cipriani *et al.*, 2013). Although AD is frequent and people recognition is a hallmark of the disorder (Cipriani *et al.*, 2013), not much is known about the clinical characteristics and their underlying mechanisms of people recognition in AD (for face recognition see Nagaratnam *et al.*, 2003; Donix *et al.*, 2013; Kurth *et al.*, 2015). Clinical observations suggest that in AD people recognition is based on the time first encountered and frequency of meetings more than affiliation. Likewise, people who are best recognized are people who are known for a long time, lived with and/or highly affiliated (e.g. spouse). Grandchildren, supposedly with high affiliation are much less recognized, unless both highly affiliated and frequently met. Professional caregivers, who are encountered with high frequency, but with

a lower affiliation are recognized well unless in the late stages of the disorder). Social orientation is relatively preserved in the early stages of AD, and is apparent only after disorientation in time (e.g. memory disorder) and space (e.g. losing one's way while a usually used road is blocked; Peer *et al.*, 2014). Systematic research is needed to quantify these observations and to investigate their underlying mechanisms, in which non-linear transformations by the hippocampal–entorhinal system may play a central role.

Conclusion

We highlight how hippocampal–entorhinal spatial coding principles can guide social interactions by flexibly relating social attributes to different individuals, social networks and the self. Consequently, cognitive mapping by the hippocampal formation can help the social brain adopt different social perspectives. By integrating models of the hippocampal–entorhinal system's role in spatial coding and memory with the circuit's anatomical links with brain regions guiding social cognition, future work can uncover how the human brain effectively processes social knowledge and maintains intricate details about different individuals in a person's social life.

Acknowledgements

We thank Dr. Nimrod Shaham for his help in creating the computational framework for the A–O model (Box 1) and Hagar Segev for the design of Figure 3.

Funding

S.A. is supported by the ISF (1306/18, 3213/19) and the NIH (AG070877). R.K. is supported by a CIDEGENT talent attraction grant from the Generalitat Valenciana (CIDEGENT/2021/027).

Conflict of interest

The authors declared that they had no conflict of interest with respect to their authorship or the publication of this article.

References

Addis, D.R., Wong, A.T., Schacter, D.L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during even construction and elaboration. *Neuropsychologia*, **45**(7), 1363–77.

Aronov, D., Nevers, R., Tank, D.W. (2017). Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. *Nature*, **543**(7647), 719–22.

Arzy, S., Seeck, M., Ortigue, S., Spinelli, L., Blanke, O. (2006). Induction of an illusory shadow person. *Nature*, **443**(7109), 287.

Arzy, S., Molnar-Szakacs, I., Blanke, O. (2008). Self in time: imagined self-location influences neural activity related to mental time travel. *Journal of Neuroscience*, **28**, 6502–7.

Arzy, S., Adi-Japha, E., Blanke, O. (2009a). The mental time line: an analogue of the mental number line in the mapping of life events. *Consciousness and Cognition*, **18**, 2009–17.

Arzy, S., Collette, S., Ionta, S., Fornari, E., Blanke, O. (2009b). Subjective mental time: the functional architecture of projecting the self to past and future. *European Journal of Neuroscience*, **30**, 2009–17.

Arzy, S., Collette, S., Wissmeyer, M., Lazeyras, F., Kaplan, P.W., Blanke, O. (2011). Psychogenic amnesia and self-identity: a multimodal functional investigation. *European Journal of Neurology*, **18**, 1422–5.

Arzy, S., Schacter, D.L. (2019). Self-agency and self-ownership in cognitive mapping. *Trends in Cognitive Sciences*, **23**, 476–87.

Arzy, S. (2021). Agency, ownership and the potential space. *Brain Sciences*, **11**, 460.

Banks, W.P., Hill, D.K. (1974). The apparent magnitude of number scaled by random production. *Journal of Experimental Psychology*, **102**(2), 353–76.

Bao, X., Gjorgieva, E., Shanahan, L.K., Howard, J.D., Kahnt, T., Gottfried, J.A. (2019). Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron*, **102**(5), 1066–75.

Bardi, L., Six, P., Brass, M. (2017). Repetitive TMS of the temporo-parietal junction disrupts participant's expectations in a spontaneous Theory of Mind task. *Social Cognitive and Affective Neuroscience*, **12**(11), 1775–82.

Barry, C., Hayman, R., Burgess, N., Jeffery, K.J. (2007). Experience-dependent rescaling of entorhinal grids. *Nature Neuroscience*, **10**, 682–4.

Barry, C., Ginzberg, L.L., O'Keefe, J., Burgess, N. (2012). Grid cell firing patterns signal environmental novelty by expansion. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(43), 17687–92.

Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., *et al.* (2018). What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, **100**, 490–509.

Bellmund, J.L.S., Gärdenfors, P., Moser, E.I., Doeller, C.F. (2018). Navigating cognition: spatial codes for human thinking. *Science*, **362**, eaat6766.

Bicanski, A., Burgess, N. (2018). A neural-level model of spatial memory and imagery. *Elife*, **7**, ee33752.

Bick, P.A. (1984). The syndrome of intermetamorphosis. *The American Journal of Psychiatry*, **141**, 588–9.

Biderman, N., Bakkour, A., Shohamy, D. (2020). What are memories for? The hippocampus bridges past experience with future decisions. *Trends in Cognitive Sciences*, **24**(7), 542–56.

Blanke, O., Arzy, S., Landis, T. (2008). Illusory perceptions of the human body and self. *Neuropsychology*, **88**, 429–58.

Boccara, C.N., Sargolini, F., Thoresen, V.H., *et al.* (2010). Grid cells in pre- and parasubiculum. *Nature Neuroscience*, **13**(8), 987–94.

Boorman, E.D., Sweigart, S.C., Park, S.A. (2021). Cognitive maps and novel inferences: a flexibility hierarchy. *Current Opinion in Behavioral Sciences*, **38**, 141–9.

Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition*, **75**(1), 1–28.

Boroditsky, L., Ramscar, M. (2002). The roles of body and mind in abstract thought. *Psychological Science*, **13**, 185–9.

Buckner, R.L., Carroll, D.C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, **11**(2), 49–57.

Buckner, R.L., Andrews-Hanna, J.R., Schacter, D.L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, **1124**, 1–38.

Bueti, D., Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1831–40.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in Cognitive Sciences*, **10**, 551–7.

Bush, D., Barry, C., Manson, D., Burgess, N. (2015). Using grid cells for navigation. *Neuron*, **87**(3), 507–20.

Butler, W.N., Hardcastle, K., Giocomo, L.M. (2019). Remembered reward locations restructure entorhinal spatial maps. *Science*, **363**(6434), 1447–52.

Byrne, P., Becker, S., Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review*, **114**(2), 340–75.

Capgras, J., Reboul-Lachaux, J. (1923). Illusion des sosies dans un delire systematise chronique. *Bulletin de la Société Clinique de Médecine Mentale*, **2**, 6–16.

Chalavi, S., Vissia, E.M., Giesen, M.E., *et al.* (2015). Abnormal hippocampal morphology in dissociative identity disorder and post-traumatic stress disorder correlates with childhood trauma and dissociative symptoms. *Human Brain Mapping*, **36**, 1692–704.

Chang, S.W. (2013). Coordinate transformation approach to social interactions. *Frontiers in Neuroscience*, **7**, 147.

Chrastil, E.R., Warren, W.H. (2014). From cognitive maps to cognitive graphs. *PLoS One*, **9**(11), e112544.

Christodoulou, G. (1978). Syndrome of subjective doubles. *The American Journal of Psychiatry*, **135**, 249–51.

Cipriani, G., Vedovello, M., Ulivi, M., Lucetti, C., Di Fiorino, A., Nuti, A. (2013). Delusional misidentification syndromes and dementia. *American Journal of Alzheimer's Disease and other Dementias*, **28**, 671–8.

Constantinescu, A., O'Reilly, J.X., Behrens, T.E.J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, **352**(6292), 1464–8.

Dafni-Merom, A., Arzy, S. (2020). The radiation of autonoetic consciousness in cognitive neuroscience: a functional neuroanatomy perspective. *Neuropsychologia*, **43**, 107477.

Danjo, T., Toyoizumi, T., Fujisawa, S. (2018). Spatial representations of self and other in the hippocampus. *Science*, **359**, 213–8.

Darby, R., Prasad, S. (2016). Lesion-related delusional misidentification syndromes: a comprehensive review of reported cases. *The Journal of Neuropsychiatry and Clinical Neurosciences*, **28**, 217–22.

Darby, R.R., Laganiere, S., Pascual-Leone, A., Prasad, S., Fox, M.D. (2017). Finding the imposter: brain connectivity of lesions causing delusional misidentifications. *Brain*, **140**, 497–507.

Davidson, P.S.R., Drouin, H., Kwan, D., Moscovitch, M., Rosenbaum, R.S. (2012). Memory as social glue: close interpersonal relationships in amnesic patients. *Frontiers in Psychology*, **3**, 1–9.

Dehaene, S., Changeux, J.P. (1993). Development of elementary numerical abilities: a neuronal model. *Journal of Cognitive Neuroscience*, **5**(4), 390–407.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science*, **284**(5416), 970–4.

Dehaene, S., Izard, V., Spelke, E., Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, **320**(5880), 1217–20.

Dehaene, S., Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, **1**, 83–120.

Devinsky, O. (2009). Delusional misidentifications and duplications: right brain lesions, left brain delusions. *Neurology*, **72**, 80–7.

Donix, M., Jurjanz, L., Meyer, S., *et al.* (2013). Functional imaging during recognition of personally familiar faces and places in Alzheimer's disease. *Archives of Clinical Neuropsychology*, **28**, 72–80.

Dunbar, R.I.M. (2018). The anatomy of friendship. *Trends in Cognitive Sciences*, **22**(1), 32–51.

Dupret, D., O'Neill, J., Pleydell-Bouverie, B., Csicsvari, J. (2010). The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nature Neuroscience*, **13**, 995–1002.

Eichenbaum, H., Cohen, N.J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron*, **83**(4), 764–70.

Ellis, H.D., Young, A.W. (1990). Accounting for delusional misidentifications. *British Journal of Psychiatry*, **157**, 239–48.

Epstein, R.A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, **12**(10), 388–96.

Feinberg, T.E., Eaton, L.A., Roane, D.M., Giacino, J.T. (1999). Multiple fregoli delusions after traumatic brain injury. *Cortex*, **35**(3), 373–87.

Feinberg, T.E., Roane, D.M. (2005). Delusional misidentification. *Psychiatric Clinics of North America*, **28**, 665–683, 678–679.

FeldmanHall, O., Montez, D.F., Phelps, E.A., Davachi, L., Murthy, V. (2021). Hippocampus guides adaptive learning during dynamic social interactions. *The Journal of Neuroscience*, **41**(6), 1340–8.

FeldmanHall, O., Nassar, M.R. (2021). The computational challenge of social learning. *Trends in Cognitive Sciences*, **25**(12), 1045–57.

Fiske, S.T. (2010). Interpersonal stratification: status, power, and subordination. In: Fiske, S.T., Gilbert, D.T., Lindzey, G., editors. *Handbook of Social Psychology*, New York: Wiley. 941–82.

Frith, C.D., Frith, U. (2006). The neural basis of mentalizing. *Neuron*, **50**(4), 531–4.

Frith, U., Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **358**(1431), 459–73.

Gaesser, B., Keeler, K., Young, L. (2018). Moral imagination: facilitating prosocial decision-making through scene imagery and theory of mind. *Cognition*, **171**, 180–93.

Gaesser, B., Hirschfeld-Kroen, J., Wasserman, E.A., Horn, M., Young, L. (2019). A role for the medial temporal lobe subsystem in guiding prosociality: the effect of episodic processes on willingness to help others. *Social Cognitive and Affective Neuroscience*, **14**(4), 397–410.

Gaesser, B. (2020). Episodic mindreading: mentalizing guided by scene construction of imagined and remembered events. *Cognition*, **203**, 104325.

Gallistel, C.R. (1990). *The Organization of Learning*. Cambridge: MIT Press.

Gauthier, B., van Wassenhove, V. (2016). Cognitive mapping in mental time travel and mental space navigation. *Cognition*, **154**, 55–68.

Gauthier, J.L., Tank, D.W. (2018). A dedicated population for reward coding in the hippocampus. *Neuron*, **99**(1), 179–93.e7.

Genzel, L. (2020). Memory and sleep: brain networks, cell dynamics and global states. *Current Opinion in Behavioral Sciences*, **32**, 72–9.

Golledge, R.G., Briggs, R., Demko, D. (1969). The configuration of distances of distances in intraurban space. *Proceedings of the Association of American Geographers*, **1**, 60–5.

Golledge, R.G. (1978). Representing, interpreting, and using cognized environments. *Papers in Regional Science*, **41**(1), 169–204.

Hafting, T., Fyhn, M., Molden, S., Moser, M.B., Moser, E.I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, **436**(7052), 801–6.

Hartley, T., Lever, C., Burgess, N., O'Keefe, J. (2013). Space in the brain: how the hippocampal formation supports spatial cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **369**(1635), 20120510.

Harvey, B.M., Klein, B.P., Petridou, N., Dumoulin, S.O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, **341**(6150), 1123–6.

Hasson, U., Chen, J., Honey, C.J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences*, **19**(6), 304–13.

Hayman, M., Arzy, S. (2021). Mental travel in the person domain. *Journal of Neurophysiology*, **126**(2), 464–76.

Herzog, S.M., Hansen, J., Wänke, M. (2007). Temporal distance and ease of retrieval. *Journal of Experimental Social Psychology*, **43**, 483–8.

Higgins, C., Liu, Y., Vidaurre, D., *et al.* (2021). Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron*, **109**(5), 882–93.

Hirstein, W., Ramachandran, V.S. (1997). Capgras syndrome: a novel probe for understanding the neural representation of the identity and familiarity of persons. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **264**, 437–44.

Hirst, W., Yamashiro, J.K., Coman, A. (2018). Collective memory from a psychological perspective. *Trends in Cognitive Sciences*, **22**(5), 438–51.

Hitti, F.L., Siegelbaum, S.A. (2014). The hippocampal CA2 region is essential for social memory. *Nature*, **508**(7494), 88–92.

Holmes, K.J., Lourenco, S.F. (2011). Common spatial organization of number and emotional expression: a mental magnitude line. *Brain and Cognition*, **77**(2), 315–23.

Høydal, Ø.A., Skytøen, E.R., Andersson, S.O., Moser, M.B., Moser, E.I. (2019). Object-vector coding in the medial entorhinal cortex. *Nature*, **568**(7752), 400–4.

Hudson, A.J., Grace, G.M. (2000). Misidentification syndromes related to face specific area in the fusiform gyrus. *Journal of Neurology, Neurosurgery, and Psychiatry*, **69**(5), 645–8.

Ionta, S., Costantini, M., Ferretti, A., Galati, G., Romani, G.L., Aglioti, S.M. (2020). Visual similarity and psychological closeness are neurally dissociable in the brain response to vicarious pain. *Cortex*, **133**, 295–308.

Jones, B., Rachlin, H. (2006). Social discounting. *Psychological Science*, **17**, 283–6.

Kahneman, D., Miller, D.T. (1986). Norm theory: comparing reality to its alternatives. *Psychological Review*, **93**, 136–53.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *ASME Journal of Basic Engineering*, **82**, 35–45.

Kaplan, R., Adhikari, M.H., Hindriks, R., *et al.* (2016). Hippocampal sharp-wave ripples influence selective activation of the default mode network. *Current Biology*, **26**(5), 686–91.

Kaplan, R., Schuck, N.W., Doeller, C.F. (2017a). The role of mental maps in decision-making. *Trends in Neurosciences*, **40**, 256–9.

Kaplan, R., King, J., Koster, R., Penny, W.D., Burgess, N., Friston, K.J. (2017b). The neural representation of prospective choice during spatial planning and decisions. *PLoS Biology*, **15**, e1002588.

Kaplan, R., Bush, D., Bisby, J., Horner, A.J., Meyer, S.S., Burgess, N. (2017c). Medial prefrontal-medial temporal theta phase coupling in dynamic spatial imagery. *Journal of Cognitive Neuroscience*, **29**(3), 507–19.

Kaplan, R., Friston, K.J. (2019). Entorhinal transformations in abstract frames of reference. *PLoS Biology*, **17**, e3000230.

Kihlstrom, J.F. (2005). Dissociative disorders. *Annual Review of Clinical Psychology*, **1**, 227–53.

Koster-Hale, J., Saxe, R. (2013). Theory of mind: a neural prediction problem. *Neuron*, **79**(5), 836–48.

Krupic, J., Burgess, N., O'Keefe, J. (2012). Neural representations of location composed of spatially periodic bands. *Science*, **337**, 853–7.

Kuhrt, D., St John, N.R., Bellmund, J.L.S., Kaplan, R., Doeller, C.F. (2021). An immersive first-person navigation task for knowledge acquisition. *Scientific Reports*, **11**(1), 5612.

Kumaran, D., Melo, H.L., Duzel, E. (2012). The emergence and representation of knowledge about social and nonsocial hierarchies. *Neuron*, **76**(3), 653–66.

Kumaran, D., Banino, A., Blundell, C., Hassabis, D., Dayan, P. (2016). Computations underlying social hierarchy learning: distinct neural mechanisms for updating and representing self-relevant information. *Neuron*, **92**(5), 1135–47.

Kurth, S., Moyse, E., Bahri, M.A., Salmon, E., Bastin, C. (2015). Recognition of personally familiar faces and functional connectivity in Alzheimer's disease. *Cortex*, **67**, 59–73.

Lanius, R.A., Vermetten, E., Loewenstein, R.J., *et al.* (2010). Emotion modulation in PTSD: clinical and neurobiological evidence for a dissociative subtype. *American Journal of Psychiatry*, **167**, 640–7.

Lau, T., Gershman, S.J., Cikara, M. (2020). Social structure learning in human anterior insula. *Elife*, **9**, e53162.

Leblanc, H., Ramirez, S. (2020). Linking social cognition to learning and memory. *The Journal of Neuroscience*, **40**(46), 8782–98.

Leroy, F., Park, J., Asok, A., *et al.* (2018). A circuit from hippocampal CA2 to lateral septum disinhibits social aggression. *Nature*, **564**(7735), 213–8.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. *Journal of Neuroscience*, **29**(31), 9771–7.

Liberman, N., Sagristano, M.D., Trope, Y. (2002). The effect of temporal distance on level of mental construal. *Journal of Experimental Social Psychology*, **38**, 523–34.

Logothetis, N.K., Eschenko, O., Murayama, Y., *et al.* (2012). Hippocampal-cortical interaction during periods of subcortical silence. *Nature*, **491**(7425), 547–53.

Lombardo, M.V., Chakrabarti, B., Bullmore, E.T., *et al.*, (2010). Shared neural circuits for mentalizing about the self and others. *Journal of Cognitive Neuroscience*, **22**(7), 1623–35.

Low, R.J., Lewallen, S., Aronov, D., Nevers, R., Tank, D.W. (2018). Probing variability in a cognitive map using manifold inference from neural dynamics. bioRxiv.

Magee, J.C., Galinsky, A.D. (2008). Social hierarchy: the self-reinforcing nature of power and status. *Academy of Management Annals*, **2**(1), 351–98.

Malliaras, D.E., Kossovitsa, Y.T., Christodoulou, G.N. (1978). Organic contributors to the intermetamorphosis syndrome. *The American Journal of Psychiatry*, **135**, 985–7.

Martin, A.K., Kessler, K., Cooke, S., Huang, J., Meinzer, M. (2020). The right temporoparietal junction is causally associated with embodied perspective-taking. *The Journal of Neuroscience*, **40**(15), 3089=95.

Meira, T., Leroy, F., Buss, E.W., Oliva, A., Park, J., Siegelbaum, S.A. (2018). A hippocampal circuit linking dorsal CA2 to ventral CA1 critical for social memory dynamics. *Nature Communications*, **9**(1), 4163.

Miller, C.T., Freiwald, W.A., Leopold, D.A., Mitchell, J.F., Silva, A.C., Wang, X. (2016). Marmosets: a neuroscientific model of human social behavior. *Neuron*, **90**(2), 219–33.

Mitra, A., Snyder, A.Z., Hacker, C.D., *et al.* (2016). Human cortical-hippocampal dialogue in wake and slow-wave sleep. *Proceedings of the National Academy of Sciences of the United States of America*, **113**(44), E6868–76.

Montagrin, A., Saiote, C., Schiller, D. (2018). The social hippocampus. *Hippocampus*, **28**(9), 672–9.

Montello, D.R. (1991). Spatial orientation and the angularity of urban routes: a field study. *Environment and Behavior*, **23**(1), 47–56.

Moscovitch, M., Cabeza, R., Winocur, G., Nadel, L. (2016). Episodic memory and beyond: the hippocampus and neocortex in transformation. *Annual Review of Psychology*, **67**, 105–34.

Nagaratnam, N., Irving, J., Kalouche, H. (2003). Misidentification in patients with dementia. *Archives of Gerontology and Geriatrics*, **37**(3), 195–202.

Nicolle, A., Klein-Flügge, M.C., Hunt, L.T., Vlaev, I., Dolan, R.J., Behrens, T.E.J. (2012). An agent independent axis for executed and modeled choice in medial prefrontal cortex. *Neuron*, **75**(6), 1114–21.

Niv, Y. (2019). Learning task-state representations. *Nature Neuroscience*, **22**(10), 1544–53.

Nuara, A., Nicolini, Y., D'Orio, P., *et al.* (2020). Catching the imposter in the brain: the case of Capgras delusion. *Cortex*, **131**, 295–304.

O'keefe, J., Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.

Oliva, A., Fernández-Ruiz, A., Leroy, F., Siegelbaum, S.A. (2020). Hippocampal CA2 sharp-wave ripples reactivate and promote social memory. *Nature*, **587**(7833), 264–9.

Omer, D.B., Maimon, S.R., Las, L., Ulanovsky, N. (2018). Social place-cells in the bat hippocampus. *Science*, **359**, 218–24.

Park, S.A., Miller, D.S., Nili, H., Ranganath, C., Boorman, E.D. (2020). Map making: constructing, combining, and inferring on abstract cognitive maps. *Neuron*, **107**(6), 1226–38.e8.

Park, S.A., Miller, D.S., Boorman, E.D., Park, S.A., Miller, D.S., Boorman, E.D. (2021). Novel inferences in a multidimensional social network use a grid-like code. *Nature Neuroscience*, **24**(9), 1292–301.

Parkinson, C., Liu, S., Wheatley, T. (2014). A common cortical metric for spatial, temporal, and social distance. *Journal of Neuroscience*, **34**(5), 1979–87.

Peer, M., Lyon, R., Arzy, S. (2014). Orientation and disorientation: lessons from patients with epilepsy. *Epilepsy & Behavior*, **41C**, 149–57.

Peer, M., Salomon, R., Goldberg, I., Blanke, O., Arzy, S. (2015). Brain system for mental orientation in space, time, and person. *Proceedings of the National Academy of Sciences*, **112**, 11072–7.

Peer, M., Hayman, M., Tamir, B., Arzy, S. (2021a). Brain coding of social network structure. *The Journal of Neuroscience*, **41**, 4897–909.

Peer, M., Brunec, I.K., Newcombe, N.S., Epstein, R.A. (2021b). Structuring knowledge with cognitive maps and cognitive graphs. *Trends in Cognitive Sciences*, **25**(1), 37–54.

Poucet, B., Hok, V. (2017). Remembering goal locations. *Current Opinion in Behavioral Sciences*, **17**, 51–6.

Rabin, J.S., Gilboa, A., Stuss, D.T., Mar, R.A., Rosenbaum, R.S. (2010). Common and unique neural correlates of autobiographical memory and theory of mind. *Journal of Cognitive Neuroscience*, **22**(6), 1095–111.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, **98**(2), 676–82.

Ramachandran, V.S., Hirstein, W. (1998). The perception of phantom limbs. The D. O. Hebb lecture. *Brain*, **121**(Pt 9), 1603–30.

Reinders, A.A., Nijenhuis, E.R., Paans, A.M., Korf, J., Willemsen, A.T., den Boer, J.A. (2003). One brain, two selves. *Neuroimage*, **20**, 2119–25.

Roberts, S.B.G., Dunbar, R.I.M., Pollet, T., Kuppens, T. (2009). Exploring variations in active network size: constraints and ego characteristics. *Social Networks*, **31**, 138–46.

Rogers, T.B., Kuiper, N.A., Kirker, W.S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, **35**(9), 677–88.

Rosenbaum, R.S., Stuss, D.T., Levine, B., Tulving, E. (2007). Theory of mind is independent of episodic memory. *Science*, **318**(5854), 1257.

Rugani, R., Vallortigara, G., Priftis, K., Regolin, L. (2015). Animal cognition. Number-space mapping in the newborn chick resembles humans' mental number line. *Science*, **347**(6221), 534–6.

Russin, J., Zofalghar, M., Park, S.A., Boorman, E., O'Reilly, R. (2021). Complementary neural networks for relational reasoning. *Cognitive science*, **2021**, 1560–6.

Sarel, A., Finkelstein, A., Las, L., Ulanovsky, N. (2017). Vectorial representation of spatial goals in the hippocampus of bats. *Science*, **355**(6321), 176–80.

Saxe, R., Moran, J.M., Scholz, J., Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. *Social Cognitive and Affective Neuroscience*, **1**(3), 229–34.

Schacter, D.L., Addis, D.R., Hassabis, D., Martin, V.C., Spreng, R.N., Szpunar, K.K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron*, **76**(4), 677–94.

Schafer, M., Schiller, D. (2018a). Navigating social space. *Neuron*, **100**(2), 476–89.

Schafer, M., Schiller, D. (2018b). The hippocampus and social impairment in psychiatric disorders. *Cold Spring Harbor Symposia on Quantitative Biology*, **83**, 105–18.

Schiller, D., Eichenbaum, H., Buffalo, E.A., *et al.* (2015). Memory and space: towards an understanding of the cognitive map. *Journal of Neuroscience*, **35**(41), 13904–11.

Sellitto, M., Neufang, S., Schwed, A., Weber, B., Kalenscher, T. (2021). Arbitration between insula and temporoparietal junction subserves framing-induced boosts in generosity during social discounting. *Neuroimage*, **238**, 118211.

Siegler, R.S., Booth, J.L. (2004). Development of numerical estimation in young children. *Child Development*, **75**(2), 428–44.

Singer, A.C., Frank, L.M. (2009). Rewarded outcomes enhance reactivation of experience in the hippocampus. *Neuron*, **64**, 910–21.

Soutschek, A., Ruff, C.C., Strombach, T., Kalenscher, T., Tobler, P.N. (2016). Brain stimulation reveals crucial role of overcoming self-centeredness in self-control. *Science Advances*, **2**, 2–10.

Spence, J.D. (1984). *The Memory Palace of Matteo Ricci*. New York, NY: Viking Penguin.

Spiegel, D. (1997). Trauma, dissociation, and memory, in psychobiology of posttraumatic stress disorder. In: R. Yehuda & A. McFarlane editors, *Psychobiology of Posttraumatic Stress Disorder*. Vol. 821, New York: New York Academy of Sciences. 225–37.

Spiegel, D., Lewis-Fernández, R., Lanius, R., Vermetten, E., Simeon, D., Friedman, M. (2013). Dissociative disorders in DSM-5. *Annual Review of Clinical Psychology*, **9**(1), 299–326.

Spreng, R.N., Mar, R.A., Kim, A.S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind and the default mode: a quantitative meta-analysis. *Journal of Cognitive Neuroscience*, **21**(3), 489–510.

Srull, T.K., Wyer, R.S. (1989). Person memory and judgment. *Psychological Review*, **96**(1), 58–83.

Stangl, M., Topalovic, U., Inman, C.S., *et al.* (2021). Boundary-anchored neural mechanisms of location-encoding for self and others. *Nature*, **589**(7842), 420–5.

Staniloiu, A., Markowitsch, H.J. (2014). Dissociative amnesia. *The Lancet Psychiatry*, **1**, 226–41.

Stella, F., Treves, A. (2015). The self-organization of grid cells in 3D. *Elife*, **4**, e05913.

Stemmler, M., Mathis, A., Herz, A.V. (2015). Connecting multiple spatial scales to decode the population activity of grid cells. *Science Advances*, **1**(11), e1500816.

Stiller, J., Dunbar, R.I.M. (2007). Perspective-taking and memory capacity predict social network size. *Social Networks*, **29**(1), 93–104.

Strombach, T., Weber, B., Hangebrauk, Z., *et al.* (2015). Social discounting involves modulation of neural value signals by temporoparietal junction. *Proceedings of the National Academy of Sciences*, **112**, 1619–24.

Sutcliffe, A., Dunbar, R., Binder, J., Arrow, H. (2012). Relationships and the social brain: integrating psychological and evolutionary perspectives. *British Journal of Psychology*, **103**, 149–68.

Tamir, D.I., Mitchell, J.P. (2010). Neural correlates of anchoring-and-adjustment during mentalizing. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(24), 10827–32.

Taube, J.S., Muller, R.U., Ranck, J.J.B. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of Neuroscience*, **10**, 420–35.

Tavares, R.M., Mendelsohn, A., Grossman, Y., *et al.* (2015). A map for social navigation in the human brain. *Neuron*, **87**(1), 231–43.

Theves, S., Fernandez, G., Doeller, C.F. (2019). The hippocampus encodes distances in multidimensional feature space. *Current Biology*, **29**, 1226–31.

Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychological Review*, **55**, 189–208.

Triandis, H.C., Leung, K., Villareal, M.J., Clack, F.I. (1985). Allocentric versus idiocentric tendencies: convergent and discriminant validation. *Journal of Research in Personality*, **19**(4), 395–415.

Trope, Y., Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, **117**(2), 440–63.

Tversky, B. (1993). Cognitive maps, cognitive collages, and spatial mental models. *European Conference on Spatial Information Theory*, 14–24.

Tversky, B. (2000). Levels and structure of spatial knowledge. In: Kitchin, R., Freundschuh, S., editors. *Cognitive Mapping Past Present Future*, Vol. 1, London: Routledge. 24–43.

Tversky, B. (2011). Spatial thought, social thought. In: *Spatial Dimensions of Social Thought*. Vol. 18, Berlin, Boston: De Gruyter Mouton. 17–38.

Tversky, B., Heiser, J., Morrison, J. (2013). Space, time, and story. *Psychology of Learning and Motivation*, **58**, 47–78.

Tversky, B. (2014). Visualizing thought. In: Huang, W. editor. *Handbook of Human Centric Visualization*, New York, NY: Springer. 3–40.

Tversky, B., Hard, B.M. (2009). Embodied and disembodied cognition: spatial perspective-taking. *Cognition*, **110**(1), 124–9.

Tzakis, N., Holahan, M.R. (2019). Social memory and the role of the hippocampal CA2 region. *Frontiers in Behavioral Neuroscience*, **13**, 233.

van Dijk, J.A., Fracasso, A., Petridou, N., Dumoulin, S.O. (2021). Laminar processing of numerosity supports a canonical cortical microcircuit in human parietal cortex. *Current Biology*, **31**, 4635–40.

Vann, S.D., Aggleton, J.P., Maguire, E.A. (2009). What does the retrosplenial cortex do? *Nature Reviews Neuroscience*, **10**(11), 792–802.

Viganò, S., Piazza, M. (2020). Distance and direction codes underlie navigation of a novel semantic space in the human brain. *The Journal of Neuroscience*, **40**(13), 2727–36.

Vogeley, K., Fink, G.R. (2003). Neural correlates of the first-person-perspective. *Trends in Cognitive Sciences*, **7**(1), 38–42.

Wang, H., Callaghan, E., Gooding-Williams, G., McAllister, C., Kessler, K. (2016). Rhythm makes the world go round: an MEG-TMS study on the role of right TPJ theta oscillations in embodied perspective taking. *Cortex*, **75**, 68–81.

Whittington, J.C.R., Muller, T.H., Mark, S., *et al.* (2020). The Tolman-eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. *Cell*, **183**(5), 1249–63.

Wilson, R.C., Takahasi, Y.K., Schoenbaum, G., Niv, Y. (2014). Orbitofrontal cortex as a cognitive map of task space. *Neuron*, **81**(2), 267–79.

Yeshurun, Y., Nguyen, M., Hasson, U. (2021). The default mode network: where the idiosyncratic self meets the shared social world. *Nature Reviews Neuroscience*, **22**(3), 181–92.