



Published in final edited form as:

J Cogn Neurosci. 2021 February ; 33(2): 159–166. doi:10.1162/jocn_a_01635.

Current Promises and Limitations of Combined Virtual Reality and Functional Magnetic Resonance Imaging Research in Humans: A Commentary on Huffman and Ekstrom (2019)

Adam Steel,

Caroline E. Robertson*,

Jeffrey S. Taube*

Dartmouth College

Abstract

Real-world navigation requires movement of the body through space, producing a continuous stream of visual and self-motion signals, including proprioceptive, vestibular, and motor efference cues. These multimodal cues are integrated to form a spatial cognitive map, an abstract, amodal representation of the environment. How the brain combines these disparate inputs and the relative importance of these inputs to cognitive map formation and recall are key unresolved questions in cognitive neuroscience. Recent advances in virtual reality technology allow participants to experience body-based cues when virtually navigating, and thus it is now possible to consider these issues in new detail. Here, we discuss a recent publication that addresses some of these issues (D. J. Huffman and A. D. Ekstrom. A modality-independent network underlies the retrieval of large-scale spatial environments in the human brain. *Neuron*, 104, 611–622, 2019). In doing so, we also review recent progress in the study of human spatial cognition and raise several questions that might be addressed in future studies.

INTRODUCTION

We learn about our spatial environment through active, self-generated movements: We move our eyes, turn our heads, and physically traverse our surroundings. In real-world conditions, these movements cause changes in the visual information sensed by the retina (optic flow), and these visual and body-based cues are integrated to form a *spatial cognitive map* (Figure 1), a modality-independent representation of a spatial environment that is fundamentally divorced from the manner in which it was encoded (Bellmund, Gärdenfors, Moser, & Doeller, 2018; Epstein, Patai, Julian, & Spiers, 2017; McNaughton, Battaglia, Jensen, Moser, & Moser, 2006; O'Keefe & Nadel, 1978; Tolman, 1948). This abstract representation is thought to underlie a variety of functions, including flexible wayfinding strategies that adapt to environmental changes (e.g., connecting novel routes; Schinazi, Nardi, Newcombe, Shipley, & Epstein, 2013; Ishikawa & Montello, 2006) and an ability

Reprint requests should be sent to Jeffrey S. Taube, Department of Psychological and Brain Sciences, Dartmouth College, 6207 Moore Hall, Hanover, NH 03755, or jeffrey.taube@dartmouth.edu.

*The last two authors contributed equally to this paper.

to express spatial representations obtained through one modality into other modalities or formats (e.g., locomotion signals to map-drawing; Huffman & Ekstrom, 2019a; Hegarty, Montello, Richardson, Ishikawa, & Lovelace, 2006). Yet, despite consensus regarding the existence of a spatial cognitive map (cf. Grieves & Dudchenko, 2013; Benhamou, 1996), key questions remain concerning the nature of spatial cognitive map representations in humans. First, during navigation, how are cues from multiple modalities (i.e., visual, motor, vestibular) integrated to form a modality-independent spatial cognitive map? Second, how is information from this modality-independent map accessed during recall, and what (if any) role do the original encoding modalities play at that point in time?

A large body of work has suggested that both visual and body-based (idiothetic) cues are crucial for spatial cognitive map formation in animals. Multiple studies have shown that vestibular cues are required for normal place (Russell, Horii, Smith, Darlington, & Bilkey, 2003; Stackman, Clark, & Taube, 2002), grid (Winter, Clark, & Taube, 2015), and head direction (HD) cell generation in rodents (Yoder et al., 2011; Muir et al., 2009; Stackman, Golob, Bassett, & Taube, 2003; Stackman et al., 2002; Stackman & Taube, 1997), as well as accurate updating of their spatial location and directional heading (Winter, Mehlman, Clark, & Taube, 2015; Yoder et al., 2011; Stackman et al., 2003). For example, lesions of the vestibular labyrinth disrupt the direction-specific firing of HD cells in the anterodorsal thalamus of rats (Stackman & Taube, 1997). Furthermore, when rats are wheeled passively on a cart into a novel environment, they do not maintain an accurate HD signal, as they do during active navigation of the same environment (Stackman et al., 2003), and grid cells in the medial entorhinal cortex lose their hexagonal firing patterns when the animals are passively moved around in a cart (Winter, Mehlman, Clark, et al., 2015). In addition, in humans, prior work has shown that the presence of idiothetic cues improves performance on a variety of spatial tasks, including virtual water maze tasks (Brandt et al., 2005), distance estimation (Witmer & Kline, 1998), and certain relative direction judgment tasks (Chance, Gaunet, Beall, & Loomis, 1998).

Yet, the importance of body-based cues for the formation of the cognitive map in humans is somewhat contentious, in large part because of the paucity of experimental paradigms that probe human navigation during naturalistic conditions (Taube, Valerio, & Yoder, 2013). For example, many studies employ *desktop-based virtual navigation*, in which participants view a spatial environment on a desktop computer screen and navigate through this environment using hand-held input devices, like keyboards or joysticks. These paradigms are limited, in the sense that participants are not provided body-based cues. Despite this limitation, many studies have shown that humans are capable of constructing a cognitive map under these conditions, presumably using visual information alone, without access to body-based self-movement cues (Nau, Schröder, Frey, & Doeller, 2020; Persichetti & Dilks, 2019; Bellmund, Deuker, Schröder, & Doeller, 2016; Doeller, Barry, & Burgess, 2010; Pine et al., 2002; Maguire et al., 1998). However, these studies only suggest that idiothetic cues are not *necessary* for constructing a cognitive map in humans; they do not preclude the possibility that idiothetic cues might play a central role in supporting the formation of the cognitive map during conditions of real-world navigation. In contrast with desktop-based virtual navigation, recent advances in head-mounted virtual reality (VR) and omnidirectional treadmill technologies allow participants to move their heads and bodies to explore a virtual

environment, thus providing somewhat realistic visual and body-based cues. In summary, *immersive VR-based navigation* has opened the possibility of probing the importance of body-based cues for forming and utilizing cognitive maps in humans, including a recent study published in *Neuron* by Huffman and Ekstrom (2019b).

HUFFMAN AND EKSTROM (2019)

Huffman and Ekstrom (2019b) used a novel combination of head-mounted VR and an omnidirectional treadmill to allow participants to learn novel virtual environments via the use of idiothetic cues associated with head movements, eye movements, and motor actions in addition to an immersive visual display (Bellmund et al., 2020; Huffman & Ekstrom, 2019b). Huffman and Ekstrom addressed three questions in their study. First, are cognitive map representations in the brain modality dependent or independent? Second, does access to idiothetic cues while learning a novel environment improve the fidelity of the spatial cognitive map or speed up the formation of such a map? Third, does the presence of idiothetic cues during encoding impact brain activity during recall and implementation of the cognitive map? To address these questions, Huffman and Ekstrom tested cognitive map formation using a judging relative direction (JRD) task after participants learned a spatial environment under three different conditions with varied access to idiothetic cues: 1) an enriched condition, where participants jointly used the head mounted display for heading direction and the omnidirectional treadmill for translation to navigate; 2) a limited condition, where participants used a head mounted display (for heading direction) and a joystick (for translation); and 3) an impoverished condition, where the joystick was used for both heading direction and translation. The authors reported that recall performance (i.e., pointing accuracy on the JRD task) and the rate of boundary alignment (a measure of global environment knowledge; Manning, Lew, Li, Sekuler, & Kahana, 2014; Kelly, Avraamides, & Loomis, 2007; Mou, Zhao, & McNamara, 2007; Shelton & McNamara, 2001) was equal across the different conditions, leading the authors to conclude that a spatial cognitive map of the environment was formed equally well regardless of whether or not idiothetic cues were available. These findings provide valuable behavioral evidence supporting the existence of a modality-independent spatial representation in humans.

Next, the authors assessed whether the presence of idiothetic cues during encoding (learning) impacted neural activity when the participants performed the JRD task. They utilized four separate analyses: 1) classifying task-state (JRD vs. rest) using functional connectivity; 2) comparing univariate activation of ROIs, including parahippocampal cortex, hippocampus, and retrosplenial cortex; 3) comparing univariate activation across the brain using a novel Bayesian analysis approach; and 4) classifying task condition based on multivariate activity (whole brain and ROIs). Across all methods, the results suggested that body-based cues did *not* impact neural representations during JRD performance: Neural representations of maps encoded during the enriched, limited, and impoverished conditions were statistically indistinguishable. Based on these behavioral and imaging findings, the authors concluded that body-based cues impacted neither the behavioral implementation of the modality-independent cognitive map, nor the neural substrates supporting recall of the spatial cognitive map. These findings provide further support for the existence of a modality-independent spatial cognitive map in humans.

AMODAL VERSUS MODALITY-DEPENDENT SPATIAL REPRESENTATIONS IN HUMANS: A FALSE DICHOTOMY?

What do these results tell us about the nature of cognitive map representations in humans? The authors aimed to distinguish between two possible accounts of the extent to which the encoding of spatial representations were modality dependent. The first view postulated that the neural representation of a learned space is not significantly influenced by the modality through which this information was encoded. Any combination of cues can, in theory, create the same cognitive map, which is ultimately supported by modality-independent neural systems. This hypothesis was termed the amodal spatial coding hypothesis. The alternate view postulated that the neural representation of a learned space is inherently linked to the original modality in which the information was encoded, so that neural systems associated with that modality must be called upon while performing subsequent spatial tasks. This distinction, however, is really an artificial formulation. Indeed, the bulk of research on spatial cognition supports the view that the brain will flexibly use whatever sensory and motor information is available to construct a representation of external space, as shown in Figure 1 (also see Figure 1 of Taube et al., 2013) (see also for discussions of how the cognitive map is constructed, see Gallistel, 1990; O'Keefe & Nadel, 1978; Tolman, 1948). As shown in Figure 1, inputs into the cognitive map representation may come from a number of different sources—some idiothetic and some based on visual input. In this view, vision alone is capable of creating an accurate spatial representation. Nevertheless, spatial systems function best when body-based movement cues are also available (see also Taube et al., 2013). Likewise, the expressed outputs derived from the cognitive map do not necessarily have to call upon neural systems associated with the encoding modality, although they can. Thus, when computing a navigational route, the output (recall) can be expressed in a number of different formats—actively taking a route (movement), drawing a graphical representation, verbally expressing the route, or performing visual imagery (Figure 1). In this sense, the inputs into the cognitive map come from many potential sources and, in turn, the outputs from the map can be expressed in a number of different ways. As such, it is clear that the output formats are independent of the way the map was originally constructed (input formats).

CURRENT LIMITATIONS TO INVESTIGATING THE NEURAL BASIS OF NAVIGATION USING fMRI

As Huffman and Ekstrom note, understanding navigation processes in humans is an area of great importance, and fMRI has been a critical tool for advancing this research. Although great progress has been made in recent years, several limitations are still noteworthy with regard to fMRI studies. First, it should be noted that *spatial cognition* encompasses many different processes such as perceived spatial orientation, spatial manipulation of objects in 3-D, distance estimation, and navigation (Figure 2). Second, *navigation*, itself, is a complex and multifaceted process that includes 1) a perception of one's spatial orientation relative to the surrounding environment, 2) computation of a route to a goal, and 3) the implementation of that route based on one's current location and directional heading. Understanding which of these processes are being monitored during fMRI studies needs to be considered in

relation to the spatial task the participant is performing. Furthermore, this broad definition of navigation includes two forms of navigation: 1) physical navigation, which involves moving the body through space in the real-world, and 2) mental or virtual navigation, where a person moves through a nonphysical space. Virtual and mental navigation do not require physical movement of the body in space and can therefore be studied with fMRI. However, mental and virtual navigation deprive participants of body-based self-motion cues. So, obtaining a complete and accurate picture of the mechanisms underlying physical navigation necessitates studying participants when they are using body-based self-motion cues. Whereas physical and mental/VR navigation certainly rely upon some shared neural mechanisms, some mechanisms will undoubtedly be proven to differ. For these reasons, fMRI studies, in which participants are immobile and in a supine position, will not lead to a full understanding of the neural mechanisms underlying navigation.

Multiple mental spatial cognitive tasks can be used to determine what mechanism might be impacted by the absence of idiothetic cues (and whether this mechanism might be shared with navigation). However, because of practical constraints, such as time and cost, fMRI studies typically use only one task, and the tasks most frequently used are ones that are better characterized as orientation tasks rather than navigational ones per se. For example, Huffman and Ekstrom only investigated performance on the JRD task, and behavioral performance across the conditions was equivalent. So, the finding that BOLD activation did not differ across conditions is perhaps unsurprising—the participants performed the same task, at the same level, and, most likely, using the same cognitive strategy across all conditions (for additional challenges in interpreting null results in fMRI, see Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017). As Huffman and Ekstrom correctly noted, it is possible that a task that encouraged the use of idiothetic cues (or idiothetic-cue recall) would dissociate the conditions more effectively at the behavioral level, and if that were the case, brain activation might differ across the conditions. Future studies might ask how idiothetic cues influence brain activity during tasks where such cues have been previously shown to benefit spatial memory recall.

Another crucial limitation of using fMRI to study spatial representations in humans is that the contribution of body-based cues to the neural instantiation of the cognitive map can only be evaluated through inference during recall and *not* during encoding because participants are immobile in the scanner and therefore cannot use self-motion cues for encoding when being scanned. Clearly, neural representations during active encoding, when participants are free to move their heads and bodies, will differ from a passive condition, when participants are seated and head-restricted. Understanding how these idiothetic and visual cues are integrated during encoding will be essential for building a complete model of the human navigation system, and fMRI may not be the correct tool for investigating this question (e.g., Aghajan et al., 2019). Thus, whether and how idiothetic cues impact brain activity during encoding in humans remains an open question. Encoding processes for navigation can be studied using fMRI techniques if a participant is engaged in learning a spatial task while in the scanner, but note that under these conditions, all learning is occurring in the absence of body-based self-motion cues.

Finally, aside from the limitations of the fMRI scanner, omnidirectional treadmills have unique considerations. Most notably, the extent to which the body-based cues afforded by movement on the omnidirectional treadmill match those that occur during real-world movement is also currently limited, which limits the generalization of findings using this technique to real-world navigation. Similarly, although some degree of body-based cues are afforded by omnidirectional treadmills, it is not a “given” that participants use these cues. Omnidirectional treadmills force participants into an awkward posture and require unnatural body movements. Participants are posed in a leaning position and drag their feet along the treadmill’s surface, and normal upper body and trunk movements are hampered by the ring enclosure. This positioning may interfere with the use of body-based cues. Furthermore, little data exist to demonstrate how well optic flow is calibrated relative to the stepping movements taken by the participant using these techniques and how optic flow timing was calibrated with the stepping motion (i.e., was it precisely 1:1?). These aspects of naturalism are critical for making the body-based movement cues appear realistic and consistently reliable, and are particularly important for interpreting studies such as Huffman and Ekstrom (2019b), where participants’ behavioral performance does not differ between enriched and impoverished encoding conditions. This absence of a difference between these conditions suggests that the participants could have disregarded body-based cues altogether and relied on visual cues alone across the three conditions, rather than using body-based cues when available to supplement visual cues during encoding (as suggested by the authors).

In summary, the absence of idiothetic self-motion cues when studying navigational mechanisms in the scanner would not provide a complete picture of how navigation works in the brain. Thus, to fully understand the neural mechanisms that underlie navigation will require methods that incorporate body-based cues when participants are navigating. This situation should serve as a reminder that researchers need to consider how the absence of self-motion-based systems impact their interpretations of fMRI experiments.

RECENT ADVANCES IN HUMAN SPATIAL COGNITION USING fMRI AND MORE NATURALISTIC APPROACHES

Despite the limitations detailed above, great strides have been made in our understanding of human spatial cognition using fMRI (Taube et al., 2013). This is due in no small part to improvement in the head-mounted VR systems. In the following section, we highlight several notable studies that combine head-mounted VR and fMRI, before raising several questions that remain to be addressed.

One advantage of head-mounted virtual reality is that participants have natural idiothetic cues during encoding, which might be reinstated during recall. Shine et al. leveraged this advantage to study HD coding when participants recalled an environment that was learned during VR-based virtual navigation (Shine, Valdés-Herrera, Hegarty, & Wolbers, 2016). Specifically, they had participants experience a virtual environment using a head-mounted display. After exposure in the head mounted display during fMRI scanning, participants viewed scenes and made judgments about whether their orientation was the same as in the preceding trial. Shine et al. observed reduced activation in the HD system when the

participant saw views of scenes from the same HD compared to different HDs (repetition suppression). The change in activation was present in the anterior thalamus, retrosplenial cortex, and precuneus (Shine et al., 2016). This finding was particularly noteworthy because the anterior thalamus was known to contain a high percentage of HD cells in rats (Taube, 1995), yet activation of the anterior thalamus had not been reported previously in participants performing spatial tasks in imaging studies.

Head-mounted virtual reality can also be used to study how active engagement influences neural representations of an environment. Robertson, Hermann, Mynick, Kravitz, and Kanwisher (2016) used this approach to investigate the neural structures that link discrete fields of view. Participants learned real-world panoramic environments using active movements with a head-mounted display. Behaviorally, they found that memory for the environments formed with body-based cues caused discrete views from within that environment to be linked into a broader representation of their shared environment: Linked views had a facilitatory priming effect on one another during subsequent memory recall. Moreover, multivoxel pattern analyses showed that the linked views were represented more similarly in retrosplenial cortex and the occipital place area (Dilks, Julian, Paunov, & Kanwisher, 2013), providing a mechanistic account of this effect.

In addition to providing an environment to be recalled in the fMRI scanner, head-mounted virtual reality can be used as a tool for detailed assessment of spatial-cognitive abilities that can be related to MRI measurements. For example, Stangl et al. had participants virtually navigate in the fMRI scanner to obtain a measurement of grid-coding in each participant's entorhinal cortex in a manner consistent with prior work (Stangl et al., 2018; Bellmund et al., 2016; Doeller et al., 2010). In a separate session, participants performed a path integration task that involved making distance and direction judgments while walking freely in an open arena wearing a head-mounted display. Stangl et al. reported that, in older adults, higher grid scores in the entorhinal cortex were correlated with better performance on the path integration task. In addition to showing a creative use of head mounted virtual reality, this study showed an interesting relationship between a measurement of the human navigation system derived from virtual navigation with an independent spatial-cognitive task outside the scanner.

Another novel and interesting approach, which shows good promise, combines the use of film (video) simulation or naturalistic activities in large-scale environments with fMRI. For example Javadi et al. (2017) had participants view films of a first person account traveling through a neighborhood in London, and then the participants had to devise a new route to a goal upon encountering a detour along the way. This task evoked activity bilaterally in inferior lateral prefrontal cortex that scaled with task difficulty. In an earlier study, Spiers et al. (Howard et al., 2014) had participants conduct a walking tour and view maps of routes through Soho, London. Later, during fMRI scanning, the participants viewed a video footage of the same routes and had to make navigational decisions. Howard et al. reported a dissociation between the information encoded by posterior hippocampus and entorhinal cortex: The posterior hippocampus activity appeared to encode the distance traveled, whereas the entorhinal cortex activation correlated with the Euclidean distance from the participant to the goal. What is interesting about this study is the use of real-world

navigation for training, which allowed the participants to experience naturalistic, body-based movements while gaining spatial knowledge about the environment. Although the decision events in the scanner were made based on visual views of the environment alone—the decision was made with knowledge that was, in part, acquired through naturalistic, body-based movement cues. Overall, video footage from real-world places, as employed by Howard et al. and Javadi et al., provides realistic visual input, such as optic flow and natural objects, which is an improvement over the artificial environments used in many VR tasks.

MOVING FORWARD: QUESTIONS FOR FUTURE RESEARCH INTO SPATIAL COGNITION IN HUMANS

In summary, Huffman and Ekstrom (2019b) used a novel combination of an omnidirectional treadmill, head-mounted VR, and fMRI to study the role of idiothetic cues when encoding a spatial cognitive map. Their study provides more evidence that vision, in the absence of body-based cues, is sufficient to form a spatial cognitive map. It also provides a unique insight into the mechanisms for spatial memory recall in humans.

Nonetheless, we are still left to puzzle about how body-based cues contribute to the neural representation of our environment. In particular, we highlight questions that still require addressing:

- How are body-based cues integrated with visual representations (e.g., landmark and optic flow systems), and how does the brain decide which cues to use during encoding when the spatial information from different systems conflict?
- What are the circumstances under which body-based, self-motion cues matter, and what are the circumstances in which they do not?
- How are different navigational processes (landmarks and body-based cues) updated to correct for errors, such as misorientation (Julian, Keinath, Marchette, & Epstein, 2018)? What brain areas signal that an error has occurred, and what enables an orientation reset to occur?
- How does the HD system operate in the fMRI setting? The animal literature suggests that, within the HD cell network, there will always be a subset of cells that are active. Furthermore, the active subset will continuously encode heading direction, even when the animal maintains the same HD for a period of time. In this sense, the HD network is always “on” whether or not a participant is using this information at a given moment. Given this operative, one would expect that brain areas that contain HD signals, like the anterior thalamus, would always be active. Yet, many imaging studies report activation of brain areas involved in directional heading when the participant is performing a spatial task, although that brain area should already be activated. Others have shown that just the opposite occurs, a decrease in activation (repetition suppression), in brain areas containing HD cells when participants repeatedly view scenes from the same heading direction within an environment (Shine et al., 2016; Baumann & Mattingley, 2010). However, repetition suppression contrasts with the way many

researchers believe the HD cell system operates, where there is little adaptation in cell firing over time periods greater than 5 sec (Shinder & Taube, 2014). These issues need to be resolved.

- Relatedly, given that the HD cell network is always encoding real-world heading direction, how are multiple reference frames maintained simultaneously? In the sections above, we mentioned that participants that are immersed in VR have an awareness of their directional heading with respect to the VR environment (i.e., cognitive heading direction), but at the same time also have an awareness of how they are oriented within the room they occupy in the real world. How does the HD system switch between these two reference frames? Does the same HD network or same population of cells capture both perceptions, or is the cognitive heading direction encoded by a different population of HD cells or a different neural system entirely? Finally, can fMRI distinguish these two different perceptual states?
- Using fMRI, Shine et al. have now reported HD coding in one subcortical brain area (anterior thalamus) where HD cells have been found in rodents (Shine et al., 2016). However, no studies have shown activation in other subcortical areas that are known to be important for generating the HD signal and directly drive the anterior thalamus, such as the lateral mammillary nuclei and dorsal tegmental nuclei (Bassett, Tullman, & Taube, 2007; Blair, Cho, & Sharp, 1999). Importantly, HD cells within these structures are thought to encode directional heading in its “purest” form and do not contain other conjunctive properties, which might interfere with interpreting what information was activating the cells. These subcortical areas should therefore be targets for future fMRI investigations.

In summary, to gain a complete understanding of mechanisms underlying navigation, researchers will need to continue to consider the role of body-based movement cues in spatial cognition. Head-mounted VR and other techniques for providing body-based cues (e.g., omnidirectional treadmills) as well as the use of naturalistic fMRI stimuli (e.g., walking tours and films), are significant improvements over previous approaches and will likely be important tools for answering these questions. Whether future advances in neural imaging techniques will allow one to image brains while participants engage in physical movement remains to be seen, but ultimately this approach is what is needed to gain a *complete* understanding of how the brain performs navigation. Future work should also consider distinguishing between real-world navigation signals (i.e., “I am in Washington, DC facing north”) from those generated by cognitive processes (i.e., “I am imagining that I am in Washington, DC facing north”). Similarly, it is important to remain mindful that, when a participant is engaged in a virtual task while in the fMRI scanner, they are simultaneously aware of the direction they are facing in the real world. The latter perception does not get “turned off” just because the participant is engaged in a virtual spatial task in the scanner. Abiding by this distinction will aid in interpreting discrepant experimental outcomes in animal and human models. As raised by Taube et al. in 2013: “As research moves forward in this field, particularly with developments enabling ever finer spatial and temporal resolution with fMRI techniques, it will be important that the dialogue among

researchers using real-world conditions and those using virtual reality systems refer to the same thing”.

Acknowledgments

The authors would like to thank Anna Mynick and Thomas Botch for helpful comments on this paper. This work was supported by National Institutes of Health grant NS104193 (J. S. T.).

Funding Information

National Institute of Neurological Disorders and Stroke (<http://dx.doi.org/10.13039/100000065>), grant number: NS104193.

REFERENCES

- Aghajian ZM, Villaroman D, Hiller S, Wishard TJ, Topalovic U, Christov-Moore L, et al. (2019). Modulation of human intracranial theta oscillations during freely moving spatial navigation and memory. *BioRxiv*, 738807. DOI: 10.1101/738807
- Bassett JP, Tullman ML, & Taube JS (2007). Lesions of the tegmentomammillary circuit in the head direction cell system disrupt the head direction signal in the anterior thalamus. *Journal of Neuroscience*, 27, 7564–7577. DOI: 10.1523/JNEUROSCI.0268-07.2007 [PubMed: 17626218]
- Baumann O, & Mattingley JB (2010). Medial parietal cortex encodes perceived heading direction in humans. *Journal of Neuroscience*, 30, 12897–12901. DOI: 10.1523/JNEUROSCI.3077-10.2010 [PubMed: 20881108]
- Bellmund JLS, de Cothi W, Ruiter TA, Nau M, Barry C, & Doeller CF (2020). Deforming the metric of cognitive maps distorts memory. *Nature Human Behaviour*, 4, 177–188. DOI: 10.1038/s41562-019-0767-3
- Bellmund JLS, Deuker L, Schröder TN, & Doeller CF (2016). Grid-cell representations in mental simulation. *eLife*, 5, e17089. DOI: 10.7554/eLife.17089 [PubMed: 27572056]
- Bellmund JLS, Gärdenfors P, Moser EI, & Doeller CF (2018). Navigating cognition: Spatial codes for human thinking. *Science*, 362, eaat6766. DOI: 10.1126/science.aat6766 [PubMed: 30409861]
- Benhamou S (1996). No evidence for cognitive mapping in rats. *Animal Behaviour*, 52, 201–212. DOI: 10.1006/anbe.1996.0165
- Blair HT, Cho J, & Sharp PE (1999). The anterior thalamic head-direction signal is abolished by bilateral but not unilateral lesions of the lateral mammillary nucleus. *Journal of Neuroscience*, 19, 6673–6683. DOI: 10.1523/JNEUROSCI.19-15-06673.1999 [PubMed: 10414996]
- Brandt T, Schautzer F, Hamilton DA, Brüning R, Markowitsch HJ, Kalla R, et al. (2005). Vestibular loss causes hippocampal atrophy and impaired spatial memory in humans. *Brain*, 128, 2732–2741. DOI: 10.1093/brain/awh617 [PubMed: 16141283]
- Chance SS, Gaunet F, Beall AC, & Loomis JM (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments*, 7, 168–178. DOI: 10.1162/105474698565659
- Dilks DD, Julian JB, Paunov AM, & Kanwisher N (2013). The occipital place area is causally and selectively involved in scene perception. *Journal of Neuroscience*, 33, 1331–1336. DOI: 10.1523/JNEUROSCI.4081-12.2013 [PubMed: 23345209]
- Doeller CF, Barry C, & Burgess N (2010). Evidence for grid cells in a human memory network. *Nature*, 463, 657–661. DOI: 10.1038/nature08704 [PubMed: 20090680]
- Epstein RA, Patai EZ, Julian JB, & Spiers HJ (2017). The cognitive map in humans: Spatial navigation and beyond. *Nature Neuroscience*, 20, 1504–1513. DOI: 10.1038/nn.4656 [PubMed: 29073650]
- Gallistel CR (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.
- Grieves RM, & Dudchenko PA (2013). Cognitive maps and spatial inference in animals: Rats fail to take a novel shortcut, but can take a previously experienced one. *Learning and Motivation*, 44, 81–92. DOI: 10.1016/j.lmot.2012.08.001

- Hegarty M, Montello DR, Richardson AE, Ishikawa T, & Lovelace K (2006). Spatial abilities at different scales: Individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, 34, 151–176. DOI: 10.1016/j.intell.2005.09.005
- Howard LR, Javadi A-H, Yu Y, Mill RD, Morrison LC, Knight R, et al. (2014). The hippocampus and entorhinal cortex encode the path and Euclidean distances to goals during navigation. *Current Biology*, 24, 1331–1340. DOI: 10.1016/j.cub.2014.05.001 [PubMed: 24909328]
- Huffman DJ, & Ekstrom AD (2019a). Which way is the bookstore? A closer look at the judgments of relative directions task. *Spatial Cognition & Computation*, 19, 93–129. DOI: 10.1080/13875868.2018.1531869 [PubMed: 31105466]
- Huffman DJ, & Ekstrom AD (2019b). A modality-independent network underlies the retrieval of large-scale spatial environments in the human brain. *Neuron*, 104, 611–622. DOI: 10.1016/j.neuron.2019.08.012 [PubMed: 31540825]
- Ishikawa T, & Montello DR (2006). Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. *Cognitive Psychology*, 52, 93–129. DOI: 10.1016/j.cogpsych.2005.08.003 [PubMed: 16375882]
- Javadi A-H, Emo B, Howard LR, Zisch FE, Yu Y, Knight R, et al. (2017). Hippocampal and prefrontal processing of network topology to simulate the future. *Nature Communications*, 8, 14652. DOI: 10.1038/ncomms14652
- Julian JB, Keinath AT, Marchette SA, & Epstein RA (2018). The neurocognitive basis of spatial reorientation. *Current Biology*, 28, R1059–R1073. DOI: 10.1016/j.cub.2018.04.057 [PubMed: 30205055]
- Kelly JW, Avraamides MN, & Loomis JM (2007). Sensorimotor alignment effects in the learning environment and in novel environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 1092–1107. DOI: 10.1037/0278-7393.33.6.1092 [PubMed: 17983315]
- Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, & Poeppel D (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron*, 93, 480–490. DOI: 10.1016/j.neuron.2016.12.041 [PubMed: 28182904]
- Maguire EA, Burgess N, Donnett JG, Frackowiak RSJ, Frith CD, & O’Keefe J (1998). Knowing where and getting there: A human navigation network. *Science*, 280, 921–924. DOI: 10.1126/science.280.5365.921 [PubMed: 9572740]
- Manning JR, Lew TF, Li N, Sekuler R, & Kahana MJ (2014). MAGELLAN: A cognitive map-based model of human wayfinding. *Journal of Experimental Psychology: General*, 143, 1314–1330. DOI: 10.1037/a0035542 [PubMed: 24490847]
- McNaughton BL, Battaglia FP, Jensen O, Moser EI, & Moser M-B (2006). Path integration and the neural basis of the “cognitive map.” *Nature Reviews Neuroscience*, 7, 663–678. DOI: 10.1038/nrn1932 [PubMed: 16858394]
- Mou W, Zhao M, & McNamara TP (2007). Layout geometry in the selection of intrinsic frames of reference from multiple viewpoints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 145–154. DOI: 10.1037/0278-7393.33.1.145 [PubMed: 17201558]
- Muir GM, Brown JE, Carey JP, Hirvonen TP, Della Santina CC, Minor LB, et al. (2009). Disruption of the head direction cell signal after occlusion of the semicircular canals in the freely moving chinchilla. *Journal of Neuroscience*, 29, 14521–14533. DOI: 10.1523/JNEUROSCI.3450-09.2009 [PubMed: 19923286]
- Nau M, Schröder TN, Frey M, & Doeller CF (2020). Behavior-dependent directional tuning in the human visual-navigation network. *Nature Communications*, 11, 3247. DOI: 10.1038/s41467-020-17000-2
- O’Keefe J, & Nadel L (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press.
- Persichetti AS, & Dilks DD (2019). Distinct representations of spatial and categorical relationships across human scene-selective cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 116, 21312–21317. DOI: 10.1073/pnas.1903057116
- Pine DS, Grun J, Maguire EA, Burgess N, Zarahn E, Koda V, et al. (2002). Neurodevelopmental aspects of spatial navigation: A virtual reality fMRI study. *Neuroimage*, 15, 396–406. DOI: 10.1006/nimg.2001.0988 [PubMed: 11798274]

- Robertson CE, Hermann KL, Mynick A, Kravitz DJ, & Kanwisher N (2016). Neural representations integrate the current field of view with the remembered 360° panorama in scene-selective cortex. *Current Biology*, 26, 2463–2468. DOI: 10.1016/j.cub.2016.07.002 [PubMed: 27618266]
- Russell NA, Horii A, Smith PF, Darlington CL, & Bilkey DK (2003). Long-term effects of permanent vestibular lesions on hippocampal spatial firing. *Journal of Neuroscience*, 23, 6490–6498. DOI: 10.1523/JNEUROSCI.23-16-06490.2003 [PubMed: 12878690]
- Schinazi VR, Nardi D, Newcombe NS, Shipley TF, & Epstein RA (2013). Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, 23, 515–528. DOI: 10.1002/hipo.22111 [PubMed: 23505031]
- Shelton AL, & McNamara TP (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, 43, 274–310. DOI: 10.1006/cogp.2001.0758 [PubMed: 11741344]
- Shinder ME, & Taube JS (2014). Self-motion improves head direction cell tuning. *Journal of Neurophysiology*, 111, 2479–2492. DOI: 10.1152/jn.00512.2013 [PubMed: 24671528]
- Shine JP, Valdés-Herrera JP, Hegarty M, & Wolbers T (2016). The human retrosplenial cortex and thalamus code head direction in a global reference frame. *Journal of Neuroscience*, 36, 6371–6381. DOI: 10.1523/JNEUROSCI.1268-15.2016 [PubMed: 27307227]
- Stackman RW, Clark AS, & Taube JS (2002). Hippocampal spatial representations require vestibular input. *Hippocampus*, 12, 291–303. DOI: 10.1002/hipo.1112 [PubMed: 12099481]
- Stackman RW, Golob EJ, Bassett JP, & Taube JS (2003). Passive transport disrupts directional path integration by rat head direction cells. *Journal of Neurophysiology*, 90, 2862–2874. DOI: 10.1152/jn.00346.2003 [PubMed: 12890795]
- Stackman RW, & Taube JS (1997). Firing properties of head direction cells in the rat anterior thalamic nucleus: Dependence on vestibular input. *Journal of Neuroscience*, 17, 4349–4358. DOI: 10.1523/JNEUROSCI.17-11-04349.1997 [PubMed: 9151751]
- Stangl M, Achtzehn J, Huber K, Dietrich C, Tempelmann C, & Wolbers T (2018). Compromised grid-cell-like representations in old age as a key mechanism to explain age-related navigational deficits. *Current Biology*, 28, 1108–1115. DOI: 10.1016/j.cub.2018.02.038 [PubMed: 29551413]
- Taube JS (1995). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *Journal of Neuroscience*, 15, 70–86. DOI: 10.1523/JNEUROSCI.15-01-00070.1995 [PubMed: 7823153]
- Taube JS, Valerio S, & Yoder RM (2013). Is navigation in virtual reality with fMRI really navigation? *Journal of Cognitive Neuroscience*, 25, 1008–1019. DOI: 10.1162/jocn_a_00386 [PubMed: 23489142]
- Tolman EC (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208. DOI: 10.1037/h0061626 [PubMed: 18870876]
- Winter SS, Clark BJ, & Taube JS (2015). Disruption of the head direction cell network impairs the parahippocampal grid cell signal. *Science*, 347, 870–874. DOI: 10.1126/science.1259591 [PubMed: 25700518]
- Winter SS, Mehlman ML, Clark BJ, & Taube JS (2015). Passive transport disrupts grid signals in the parahippocampal cortex. *Current Biology*, 25, 2493–2502. DOI: 10.1016/j.cub.2015.08.034 [PubMed: 26387719]
- Witmer BG, & Kline PB (1998). Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments*, 7, 144–167. DOI: 10.1162/105474698565640
- Yoder RM, Clark BJ, Brown JE, Lamia MV, Valerio S, Shinder ME, et al. (2011). Both visual and idiothetic cues contribute to head direction cell stability during navigation along complex routes. *Journal of Neurophysiology*, 105, 2989–3001. DOI: 10.1152/jn.01041.2010 [PubMed: 21451060]

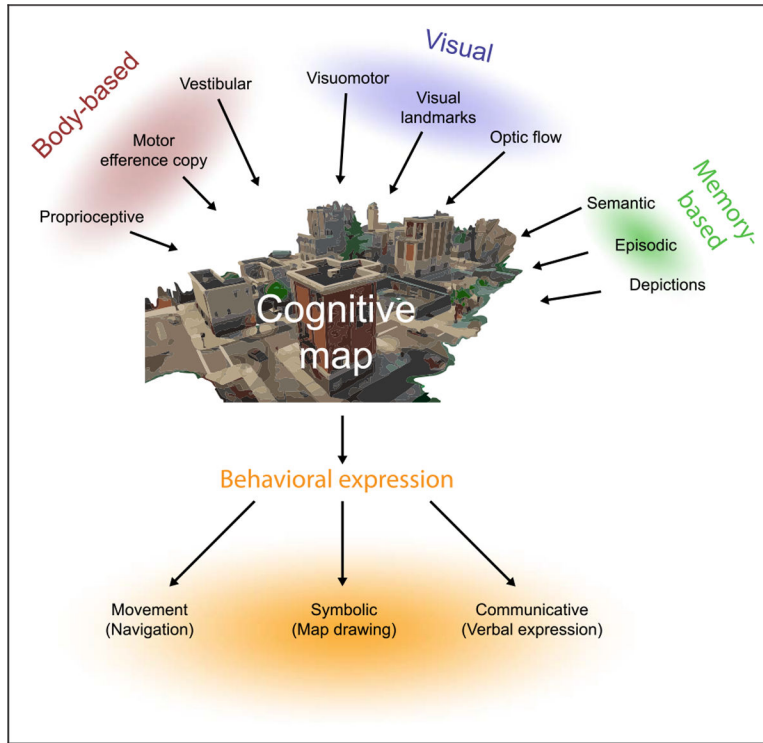


Figure 1.

The construction and implementation of the amodal cognitive map in humans. When exploring a new environment, modality-specific sensory input is integrated to form a cognitive map, an amodal representation of the environment. The sensory inputs include body-based cues as well as visual cues, and, in humans, can be informed by depictions (such as maps). Humans can also use prior knowledge from other sources to inform cognitive map formation, such as personal experience (episodic memory) and factual knowledge (semantic memory, e.g., rivers flow towards oceans). The amodal cognitive map can be implemented in a variety of modalities, including verbal descriptions, symbolic depictions, and through navigation. Importantly, the cognitive map can be constructed from any of these sources independently, such as vision alone, but is most reliable when all inputs are present. Desktop-based and fMRI studies of navigation predominantly deliver visual input and neglect body-based cues.

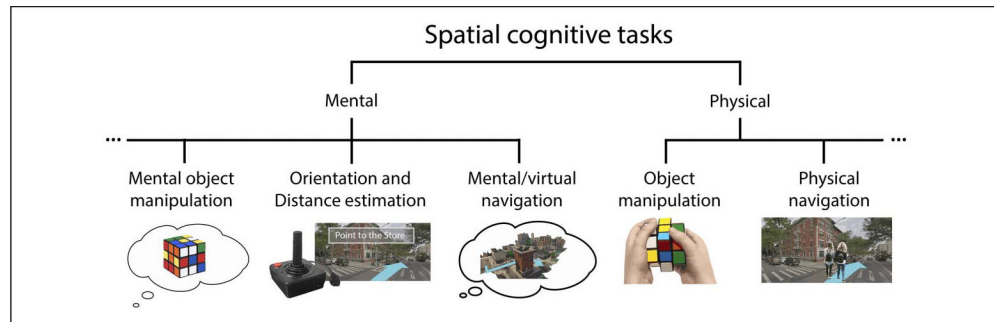


Figure 2.

Mental and physical spatial cognitive tasks. Spatial cognitive tasks can largely be considered mental or physical in nature. Examples of mental cognitive tasks include mental object manipulation/mean rotation, distance estimation, direction judgments, mental navigation, or desktop-based virtual navigation. Examples of physical tasks include physical manipulation of objects and navigation in physical space. Importantly, physical tasks produce body-based cues (such as vestibular and proprioceptive cues), whereas mental tasks do not. As such, physical tasks may have a mental component, but mental tasks do not have a physical component. Understanding the physical components of spatial cognition is critical to building a full model of the neural mechanisms underlying real-world behavior.