

COMMENTARY: Grid coding, spatial representation, and navigation: Should we assume an
isomorphism?

Arne D. Ekstrom¹ & Sevan K. Harootonian¹ & Derek J. Huffman²

¹Department of Psychology, University of Arizona, 1503 E. University Blvd.,
Tucson, AZ 85721

²Center for Neuroscience, University of California, Davis, 1544 Newton Court, Davis, CA 95616

PREPRINT: FINAL ACCEPTED MANUSCRIPT AT HIPPOCAMPUS

CORRESPONDING AUTHOR:

Dr. Arne Ekstrom, email: adekstrom@email.arizona.edu

Acknowledgements and funding: Research supported by grants from NSF (BCS-1630296) and NIH (R01NS076856) awarded to A.D.E. Research also supported by a Ruth L.~Kirschstein Postdoctoral Individual National Research Service Award from the National Institute of Mental Health of the National Institutes of Health (F32MH116577) awarded to D.J.H. We thank members of the Human Spatial Cognition Lab, especially Michael Starrett and Mingli Liang, for helpful conversations about this project.

Abstract:

Grid cells provide a compelling example of a link between cellular activity and an abstract and difficult to define concept like space. Accordingly, a *representational* perspective on grid coding argues that neural grid coding underlies a fundamentally spatial metric. Recently, some theoretical proposals have suggested extending such a framework to non-spatial cognition as well, such as category learning. Here, we provide a critique of the frequently employed assumption of an isomorphism between patterns of neural activity (e.g., grid cells), mental representation, and behavior (e.g., navigation). Specifically, we question the strict isomorphism between these three levels and suggest that human spatial navigation is perhaps best characterized by a wide variety of both metric and non-metric strategies. We offer an alternative perspective on how grid coding might relate to human spatial navigation, arguing that grid coding is part of a much larger conglomeration of neural activity patterns that dynamically tune to accomplish specific behavioral outputs.

Grid cells, neurons recorded in the medial entorhinal cortex, represent a striking example of a potential correlation between a neural response and a higher order cognitive coding scheme for how we might mentally represent space. Grid cells, typically recorded from freely navigating rats and humans (Hafting, Fyhn, Molden, Moser, and Moser, 2005; Sargolini, Fyhn, Hafting, McNaughton, Witter, Moser, and Moser, 2006; Jacobs, Weidemann, Miller, Solway, Burke, Wei, Suthana, Sperling, Sharan, Fried, and Kahana, 2013), display a regularly spaced pattern of neural firing (“grids”) that span the environment. In addition to their presence during spatial navigation, studies have also observed such grid coding during tasks involving representation of conceptual categories, imagination, and scene processing (Killian, Jutras, and Buffalo, 2012; Bellmund, Deuker, Schroder, and Doeller, 2016; Constantinescu, O'Reilly, and Behrens, 2016; Horner, Bisby, Zotow, Bush, and Burgess, 2016; Garvert, Dolan, and Behrens, 2017; Bao, Gjorgieva, Shanahan, Howard, Kahnt, and Gottfried, 2019). One theoretical interpretation of such grid coding is that it forms the basis of a universal spatial metric representation that underlies many, if not all, forms of spatial and non-spatial representations (Hasselmo, Giocomo, Brandon, and Yoshida, 2010; Bush, Barry, Manson, and Burgess, 2015; Behrens, Muller, Whittington, Mark, Baram, Stachenfeld, and Kurth-Nelson, 2018; Bellmund, Gardenfors, Moser, and Doeller, 2018; Hawkins, Lewis, Klukas, Purdy, and Ahmad, 2018).

Grid coding is certainly intriguing and potentially of importance in understanding a difficult and abstract concept like “space.” Such theoretical interpretations described above, however, assume a strict isomorphism between patterns of neural activity, mental representation, and behavior. Specifically, such an isomorphism assumes that we can infer between the metric properties of grid cells, a metric mental representation of space, and forms of spatial navigation that possess metric properties. In this commentary, we will explore both the assumptions of “strict” isomorphism as well the generalizability of metric assumptions of spatial processing, probing under what situations they may be valid. We also explore alternative perspectives that suggest such interpretations of strict isomorphism are both unnecessary and inflexible.

Grid coding as an underlying representational metric of space: The navigational and extended perspectives

As a rat navigates an environment, often by being offered rewards for exploring the environment as comprehensively as possible, grid cells fire in a relatively evenly spaced manner that covers much of the environment much like a piece of graph paper (Hafting et al., 2005; Sargolini et al., 2006). One of the key features of grid cells is that the grid often shows *6-fold symmetry*; in other words, the lattice surrounding each “node” in the grid forms a hexagon (Doeller, Barry, and Burgess, 2010; Dordek, Soudry, Meir, and Derdikman, 2016). Some studies have also found evidence that grid coding extends to other scales of neural recordings, such as functional magnetic resonance imaging (fMRI), intracranial EEG (iEEG), and MEG recordings (Doeller et al., 2010; Kunz, Schroeder, Hweeling, Montag, Lachman, Sariyska, Reuter, Stirnberg, Stocker, Messing-Floeter, Fell, Doeller, and Axmacher, 2015; Julian, Keinath, Frazzetta, and Epstein, 2018; Maidenbaum, Miller, Stein, and Jacobs, 2018; Stangl, Achtzehn, Huber, Dietrich, Tempelmann, and Wolbers, 2018; Staudigl, Leszczynski, Jacobs, Sheth, Schroeder, Jensen, and Doeller, 2018). Such studies typically involve humans and more complex measures of behavior, and thus provide potential links between grid coding at the single cell level and complex behavioral expression (Bellmund et al., 2018; Kunz, Maidenbaum, Chen, Wang, Jacobs, and Axmacher, 2019). We call the idea that grid cells form the basis for how we represent and navigate space the *navigational representational perspective* on grid coding (Moser and Moser, 2008).

Some have also argued that the representational perspective on grid coding may extend beyond navigation into the non-spatial domain as well. In one theoretical paper, Bellmund et al.

develop the idea that “grid cells can encode positions along dimensions of experience beyond Euclidean space for navigation, suggesting a more general role of hippocampal-entorhinal processing mechanisms in cognition” (p. 1). Accordingly, they define cognitive spaces as: “spanned by a set of quality dimensions, which can be closely related to sensory inputs but also comprise abstract features” (p. 1). For example, Constantinescu et al. (2016) tested participants’ knowledge about the “direction” of different concepts. In the paradigm, participants were trained to associate stimuli (a bird) with outcomes (Christmas symbols). The direction in which participants moved in concept space depended on the rates of change of the lengths of the neck and legs. They found a 6-fold modulation of activity based on the angles of such trajectories in entorhinal cortex, prefrontal cortex, and temporal parietal junction. Based on these and other studies (Bellmund et al., 2016; Horner et al., 2016; Garvert et al., 2017), Bellmund et al. (2018) argued that such grid coding might extend more broadly into abstract or “cognitive spaces” (see also: Behrens et al., 2018). Such grid-like coding also manifests during situations that involve coding temporal intervals (Kraus, Brandon, Robinson, Connerney, Hasselmo, and Eichenbaum, 2015; Tsao, Sugar, Lu, Wang, Knierim, Moser, and Moser, 2018), further suggesting a grid coding mechanism as a potential universal “metric” for processing information.

Based on what we term the *extended representational perspective*, grid coding then forms the basis of several different forms of representation of both spatial and non-spatial domains, including categories, imagination, processing of sensory input, and experiences more generally. As argued by Bellmund et al., such representations involve: “an underlying metric [which] follow [the] geometric constraints satisfying the mathematical notions of betweenness and equidistance” (p. 1). In this way, such cognitive spaces will typically involve the Euclidean rules of geometry (O’Keefe and Nadel, 1978; Gallistel, 1990). Moreover, the symmetrical nature of such grid coding would seem to imply that the representation of spatial and non-spatial relationships is metric and also six-fold symmetric. We attempt to summarize the representational perspective on grid coding in Figure 1A.

Grid coding and isomorphism with spatial and non-spatial representation

As we have summarized above, grid coding manifests across different scales of recordings and in both spatial and non-spatial tasks. The important question for our consideration here then is how we might interpret such grid coding. Specifically, what is the relationship between grid coding, mental representation, and behavior? The theoretical models and interpretations we have reviewed above, either implicitly or explicitly, assume an isomorphism between neural activity patterns, mental representation, and behavioral expression. In a similar vein, Gallistel (1989) argued for a weaker form of such an isomorphism between how animals navigated and the organization of their neural representations. Specifically, Gallistel (1989) argued that “the formal mathematical descriptions of many brain processes appear to mirror the formal descriptions of external realities to which those processes adapt the animal’s behavior” (p. 185). Gallistel based much of his reasoning on desert ants and other species who, in the absence of explicit cues, could return to their nest with remarkable accuracy. In this way, Gallistel argued, one might expect a strong similarity (although not necessarily an identical correspondence) between how an ant navigated and the representation needed to accomplish this task.

While isomorphism assumes that neural activity reflects the structure experienced by the navigator (i.e., a relationship between neural activity and the structure of the environment), the representational perspectives make the additional assumption that grid codes can be directly observed, and directly influence, behavioral expression. The idea that patterns of neural activity can be used to infer representations at the mental or behavioral level is a form of *reverse inference* (Poldrack, 2006, 2011). In this way, we could also think of the representational

perspective on grid coding as assuming a *strict isomorphism* whereby a property at one level must necessarily manifest at another.

We certainly do not challenge the idea that one of the principal jobs of the nervous system is to capture rules and structure that underlie how we navigate and behave more generally. Indeed, as pointed out, there may also be instances in which reverse inference can be valid (Hutzler, 2013). There are several challenges, however, to the assumption of strict isomorphism between a neural phenomenon and behavior, which we explore in detail here. This is because the assumption of strict isomorphism: 1) involves a *mereological fallacy* in which one incorrectly assumes a whole (behavior) can be derived from a limited number of parts (e.g., grid cells), 2) confounds levels of analysis (e.g., computational, algorithmic/representational, and implementational; Marr, 1982), 3) is at odds with behavioral studies of navigation, which suggest that navigation involves the dynamical use of a variety of different forms of representations, 4) is at odds with behavioral performance on spatial and non-spatial tasks, which are typically *non-metric*. We expand on these points in the following sections.

Potential problems with isomorphism: The mereological fallacy

How do we map from neural activity to behavior? While the assumption of isomorphism may be reasonable in some contexts, others have pointed out that observing a subset of neural activity does not provide sufficient information to extrapolate to the behavior that we wish to describe (Embick and Poeppel, 2015; Krakauer, Ghazanfar, Gomez-Marin, MacIver, and Poeppel, 2017). Embick and Poeppel approach this problem from a primarily linguistic perspective, exploring the issue of linking different levels of processing during language output and its putative neural underpinnings. For example, one may observe distinct neural activation patterns that correlate with the processing of morphemes, but this does not mean that such correlations constitute a direct link and/or explanation for how we understand words in sentences. More generally, Embick and Poeppel point out the flaws with assuming that we can draw direct links between theoretical considerations of linguists and properties studied in neuroscience because the two different levels provide fundamentally different forms of logic and explanation.

Krakauer et al. take this one step further and attempt to generalize the problem of linking behavior and neuroscience. They argue that the tendency to assume that the activity of individual neurons (or groups of neurons) maps meaningfully onto behavior involves what they term a “mereological fallacy.” Specifically, they point out that it is problematic to assume that one part (in our case, grid coding) of a larger whole (the behaving animal) can be meaningfully interpreted without all the other parts (the other neural coding schemes other than grid coding and their respective interactions). For example, they argue: “Modeling and studying the responses of the neural substrate on any scale—large or small—will not *by itself* lead to insights about how behavior is generated. One reason for this is that the properties of neural tissue may be more diverse than the subset actually exploited for natural behaviors” (p. 484). Similarly, the assumption of a strict isomorphic relationship between patterns of neural activity and behavior ignores the importance of task-specific configurations and dynamics (Eichenbaum and Cohen, 2014)—i.e., the idea that neural tuning itself may change to accomplish a set of behaviors.

It is useful to consider human category learning as an example of what we are attempting to point out above. We can consider classifying a novel set of stimuli according to different features or learning to associate stimuli A-B and B-C. As pointed out early by Wittgenstein, a hallmark of category learning often involves extracting a prototype, or running average, across the many different stimuli presented (Wittgenstein, 2001 [1953]; Ashby and O'Brien, 2005). Such running averages can readily be observed in neural signatures, which often manifest as a combination of other stimuli learned during the task (Mack, Love, and

Preston, 2017). For example, when learning to associate stimulus A-B and B-C, participants can learn the association between A-C, even if it is never explicitly presented. The neural representation for A-C can best be described as an integrated, or averaged version, of A,B,C, consistent with the idea that each has weak associations (Schlichting, Mumford, and Preston, 2015). Importantly, though, the representation for A-C evolves as part of the task demands. More generally, how participants categorize any set of dimensions can involve a variety of different strategies and brain systems, all of which are task-dependent and dynamic (Ashby and Maddox, 2005; Ashby and O'Brien, 2005; Ashby and Maddox, 2011).

Levels of analysis and description

As pointed out by Marr (1982), we can consider problems in cognitive neuroscience from three levels of description: computational, algorithmic/representational, and implementational (e.g., neural connectivity and activity). These three levels can briefly be described as: computation: *why* (problem), algorithm: *what* (rules), and implementation: *how* (physical) (Krakauer et al., 2017). Marr argued that “trying to understand perception by understanding neurons is like trying to understand a bird’s flight by studying only feathers. It just cannot be done” (quoted in Krakauer et al p. 482). Additionally, as pointed out previously in the context of navigation, these three levels cannot readily be bridged (Zhao, 2018).

In the case of the *extended representational perspective*, and, to some extent, the *navigational perspective*, the assumption is that grid coding at the single cell level corresponds to a mental form of representation involving a spatial metric (perhaps a piece of graph paper), which then maps, more or less directly, onto how we categorize, remember, and/or navigate. As Marr argued, however, each of these different levels regard qualitative and/or quantitative descriptions such that even a complete description of one level (implementational) does not necessarily provide insight into the other level (e.g., algorithmic). Warren has similarly pointed out that simply because a cartographic map, with its metric details, is useful for thinking about space, it does not follow that this is how we actually represent space (Warren, 2019). This is particularly important when one considers the dynamic nature of navigation, an issue we will expand upon shortly.

It is worth pausing, however, to note that in the context of neural codes related to navigation, it is clear that some computational models have effectively captured variables relevant to navigation such as position, velocity, distance, and direction (Samsonovich and McNaughton, 1997; Burgess, 2008; Burak and Fiete, 2009; Chaudhuri, Gerçek, Pandey, Peyrache, and Fiete, 2019). This does not mean, however, that such computational-level descriptions directly equate to knowledge that the participant has about the environment and specific implementations of *how* to navigate from A to B. While there are some intriguing demonstrations of situations in which participant patterns of path integration fit predictions of computational models of grid cells (Chen, He, Kelly, Fiete, and McNamara, 2015), there may be other potential explanations for such patterns of behavioral results (Hartley, Trinkler, and Burgess, 2004). In addition, while some studies have shown correlations between task performance and the degree of grid coding present in neural signals (Doeller et al., 2010; Constantinescu et al., 2016), as discussed above, a correlation between a neural pattern and a behavioral variable does not imply that the properties at either level automatically port to the other.

More generally, there are almost certainly many different solutions for how we walk to a hidden target, and it is unclear whether grid cells will inform our understanding of the strategies that the navigator is using to solve the task. While few would make the argument that grid coding accounts for *all* aspects of behavior, the argument we wish to pose here is that grid coding *in isolation* provides limited insight into the complex phenomenon that is human spatial navigation and cognition more generally. As just one example of this, the precision of

representations we employ to walk to hidden locations or point to unseen targets are poorer than what we might get from decoding of neural signals in such computational models (Ekstrom, Arnold, and Iaria, 2014). For example, even with extensive exposure to an environment, we typically show a “floor” level of about 15-20 degrees error in how accurately we can point to unseen targets (McNamara, Rump, and Werner, 2003; Frankenstein, Mohler, Bulthoff, and Meilinger, 2012; Huffman and Ekstrom, 2018; Starrett, Stokes, Huffman, Ferrer, and Ekstrom, 2018). While computational models often deal with this issue by assuming “noise,” the lack of direct correspondence between precision at the grid level and that of behavioral output suggests that strict isomorphism is incorrect at some level as well.

Navigation involves the dynamical use of a variety of different forms of representation

It might be reasonable to ask, though, why would 6-fold grid coding be present neurally if it were not in some form used as a representation? At present, it is difficult to answer this question without further experiments such as those we will discuss shortly related to the neural-behavioral interface. On the flip side, however, it seems that one could also observe many other forms of neural coding that would be unlikely to map in any obvious way onto behavior. The presence, for example, of multiple forms of conjunctive responses, such as combinations of grid X velocity or grid X head direction conjunctive responses (Sargolini et al., 2006) does not necessarily imply that such representations have an explicit manifestation in behavior or even underlie it. Instead, they suggest that such responses are present in sufficient number over sufficient amounts of time during the task to be observed during recordings, and related, collectively, to how the brain codes the task.

Additionally, we might consider more broadly the extent to which spatial metrics are necessary for situations involving navigation. Here, we use the term “metric representations” to refer predominantly to representations that obey the laws of Euclidean geometry, although we note that non-Euclidean spaces can be either metric, e.g., Hilbert spaces, or non-metric, e.g., topological graphs (for more on this topic, please see: Warren, 2019). While we think there are certainly instances that are consistent with employing a spatial metric to navigate (for example, in small-scale, regular, well-learned environments), we also think there are many situations in which participants do not readily employ spatial metrics when navigating.

As one example for navigating, participants often use beaconing strategies to find a landmark, which involves continuing to walk to a landmark as it grows in size on your retina – no spatial metric is required for this type of task (Morris, Hagan, and Rawlins, 1986; Wolbers and Wiener, 2014; Ekstrom and Isham, 2017). Beaconing, or responses-based procedural tasks, however, are widely recognized as non-spatial (McDonald and White, 1994). Even if we were to restrict our considerations to only situations involving “spatial” navigation, we still think that is difficult to assume that this can be distilled to a single strategy involving spatial metrics (Ekstrom et al., 2014). For example, rats navigating in situations that might appear to involve either a “spatial” or “non-spatial strategy” to solve the 8-arm maze typically show a mixture of both strategies and lesions that should impair one system often have more complex effects (Ferbinteanu, 2016; Goodroe, Starnes, and Brown, 2018).

As another example, remembering the locations of cities on a map would appear to be a strong candidate for using a form of a spatial metric. Interestingly, though, when participants are asked which is further west, San Diego or Reno, they often answer incorrectly due the category heuristic that California is further west than Nevada (Stevens and Coupe, 1978; Okabayashi and Glynn, 1984). If one had “pure” metric knowledge of the map of the United States, which might be expected if one grows up in the U.S. and studies such maps during school, one might not expect these errors. It could certainly be the case, however, that these are exceptions and that in the majority of situations, we employ an explicitly “metric” representation. As reviewed and studied in situations both involving remembering locations in the world and locations of dots

on a screen, however, participants often substitute heuristic information (e.g., the angle was about 90 degrees, the dots were somewhere near the center of lines) for any explicit metric knowledge (Moar and Bower, 1983; Sadalla and Montello, 1989; Huttenlocher, Hedges, Corrigan, and Crawford, 2004).

Similarly, other studies suggest situations involving what would appear to be distinctly spatial representations, such as distance judgments, nonetheless violate Euclidean propositions. For example, participants typically exhibit asymmetric distance estimations between landmarks and non-landmarks (McNamara, 1991; McNamara and Diwadkar, 1997), which violates the postulate of Euclidean symmetry ($AB = BA$). In addition, distance estimates often violate Euclidean propositions when comparing within and between “cluster” distance estimations (clusters were defined by an algorithm that found landmarks that were consistently recalled together: Hirtle and Jonides, 1985; McNamara, Hardy, and Hirtle, 1989). Moreover, the distance between locations within a sub-environment was systematically underestimated relative to the distance between locations in different sub-environments (i.e., across an environmental boundary) even when the Euclidean distance between these locations was matched (McNamara, 1986).

Based on these and other findings, previous theories have suggested that participants could use non-metric spatial strategies in many situations. Such mental operations, however, could appear metric due the mapping of similarity between simple relations such as “contained in, next to, and perhaps, above/below, [and] left/right” (McNamara, 1991). It is important to note that such situations could, with enough learning, or explicit task demands, eventually involve a Euclidean metric (e.g., is the chair 1 meter or 1.1 meters from the desk? (Montello, 1998)). It seems, however, that navigators might use whatever combination of heuristics and strategies are needed to produce the desired behavioral output and not necessarily ones that require spatial metrics (Ekstrom, Spiers, Bohbot, and Rosenbaum, 2018).

Non-metric spatial representations

Interestingly, numerous situations involving navigation in humans often involve non-metric representations. As recently argued by Warren (2019), much of what we refer to as navigation – path integration, using short cuts, and wayfinding, can best be considered non-metric and non-Euclidean in terms of how we perform the task. As one example, Warren et al. tested participants navigating an environment in a large room by having them wear a head-mounted display. In this way, the path integration cues would be comparable to what we would experience with our body senses in a “normal” environment yet the virtual rendering was altered to test ideas about metric postulates by embedding, unbeknownst to the participant, wormholes into the environment. This allowed Warren et al. to directly contrast two different hypotheses: when participants went through wormholes, did they maintain a metric representation or did the wormholes alter their representations (Warren, Rothman, Schnapp, and Ericson, 2017)? Warren et al. found that, consistent with a non-metric topological representation, participants readily (and without any direct knowledge) incorporated the wormholes into their spatial representations in a way that warped space in a non-metric fashion.

Even in situations with a regular geometry in a small-sized room and with sufficient exposure, participants may still favor non-metric task-specific heuristics. For example, when learning a novel spatial environment, participants employ the intrinsic structure of objects or the boundaries of the room to bind their spatial representation (Mou and McNamara, 2002; Starrett et al., 2018). Specifically, pointing accuracy is superior when participants imagine headings aligned with a 4-fold axes defined by the objects, room, or other features the participant selects compared to misaligned. This might appear to be consistent with a metric perspective because memory is based on the boundaries of the room (although note such coding is not 6-fold but rather 4-fold). Notably, however, such a heuristic (“I can remember things relative to the

boundaries of the room”) involves distortions that favor more accurate memory for locations aligned with the boundaries compared to those that are offset (McNamara et al., 2003; Mou, McNamara, Valiquette, and Rump, 2004; Chan, Baumann, Bellgrove, and Mattingley, 2013; Peacock and Ekstrom, 2018; Starrett et al., 2018). This type of distortion instead could best be considered non-metric because location memory is warped based on the imagined heading and the room boundaries.

What about situations involving neural recordings of grid cells?

Besides the central presence of metric coding, another important prediction of the *representational grid coding perspective* is that when the neural activity underlying grid coding is perturbed, this should impair the ability to navigate. While there is still debate in the field of navigation regarding what type of spatial metric we employ during navigation and when (e.g., allocentric vs. egocentric, Ekstrom et al., 2014), it is clear that there are situations in which we code space using some form of relative spatial information (i.e., the bookstore is closer to the gym than the library), such as the Morris Water Maze (Morris et al., 1986 although see: Wolbers and Wiener 2014 and Wang 2017). In this task, participants are asked to remember the location of a hidden platform from a novel start location based on its position relative to the distal landmarks. Focal lesions to the medial entorhinal cortex impair retention of the hidden platform in the Morris Water Maze. These animals, however, can learn new hidden platform locations although at a slower rate than control animals (Steffenach, Witter, Moser, and Moser, 2005; Hales, Schlesiger, Leutgeb, Squire, Leutgeb, and Clark, 2014). Additionally, in one study, reducing grid cell activity using optogenetics had no observable effect on Morris Water Maze performance (Kanter, Lykken, Avesar, Weible, Dickinson, Dunn, Borgesius, Roudi, and Kentros, 2017). Overall, the effects of perturbing medial entorhinal cortex are somewhat task specific and do not consistently impair allocentric navigation or situations that might appear to necessitate a spatial metric (for a review, see: Save and Sargolini 2017).

Additional evidence suggests that neural coding in the entorhinal cortex, rather than read out in a static fashion, is dynamically influenced by the environment and behavioral demands. For example, deformations (i.e., changes from the typical 6-fold hexagonal symmetry) of grid cells have been observed in trapezoidal and irregular environments (Krupic, Bauza, Burton, Barry, and O'Keefe, 2015; Krupic, Bauza, Burton, and O'Keefe, 2018), suggesting that the medial entorhinal cortex is sensitive to the local shape of the environment (Krupic, Bauza, Burton, and O'Keefe, 2016). Additionally, while many studies have investigated grid coding during free foraging tasks, a recent study found that introducing reward locations altered the structure of grid cells in medial entorhinal cortex to preferentially represent the reward location (Butler, Hardcastle, and Giocomo, 2019). More generally, other findings have suggested that the involvement of entorhinal cortex in navigation can vary based on environmental features (Barry, Ginzberg, O'Keefe, and Burgess, 2012; Stensola and Moser, 2016) and behavioral demands (Rodo, Sargolini, and Save, 2017; Save and Sargolini, 2017; Yoo and Lee, 2017). Together, these findings reveal that grid cell coding in the entorhinal cortex is inconsistently hexagonally symmetric, rarely static, and is flexible.

Such considerations above also raise interesting questions for the *extended representational view* on grid coding. For example, the boundaries of abstract cognitive spaces would typically be unknown, and thus the metrics of such spaces could likely be highly non-linear. Moreover, one might ask why non-spatial tasks might be amenable to a spatial metric if not somehow metric in the first place. It is intriguing to consider that the link between spatial and non-spatial tasks is in the heuristics and strategies used to solve the tasks. For example, knowing that the grocery store is located generally east of your house provides useful categorical information about where things are located. Similarly, thinking of (non-spatial) categories as existing along a qualitative continuum (lions are more like cats than trucks) is

certainly useful to remembering facts about mammals. It does not follow, however, that such categories necessitate a Euclidean spatial metric to underlie such thinking.

Dynamical systems perspectives on neural coding

As we have suggested so far, many aspects of navigation, and behavior more generally, involve a high degree of dynamic tuning. Notably, past efforts have suggested the value of dynamical systems broadly to understanding behavior (Kelso, Saltzman, and Tuller, 1986; Sternad, 2000). While our efforts within this commentary are intended for a much more limited scope than past work on this topic, we think it is useful to illustrate the potential value of such models to understanding how neural activity and navigation might interact. Our illustration here builds off motor systems research, where this perspective has been effective in better linking motor output with underlying neural representations (Shenoy, Sahani, and Churchland, 2013; Elsayed, Lara, Kaufman, Churchland, and Cunningham, 2016).

As Shenoy et al. state: “the dynamical systems perspective stresses the view that the nervous system... must generate a pattern of activity appropriate to derive the desired movement” (p. 341). According to this framework, cortical activity is a time-varying vector $[r(t)]$ that must be mapped onto muscle activity $[m(t)]$ by downstream circuitry by some transformation function G (a non-invertible function, see Figure 1B). Importantly, the population activity $[r(t)]$ evolves based on the local circuitry activity on the current input $[h(t)]$ and does not represent or “map” onto a specific behavior. Instead, the neural patterns themselves may undergo dynamics as part of tuning the (intended) correct output rather than the “representation” of the behavior itself (e.g., which muscle fiber to activate or the direction to move).

How does this perspective relate to the patterns of neural activity observed during navigation? The dynamical systems perspective argues that there are many neural patterns that one might observe during a task involving, for example, remembering a straight vs. curvy path from point A to point B. None of these, however, necessitate a direct mapping between neural activity and the behavioral output. As one example of this in the context of grid coding, Dordek et al. (2016) found *four-fold* periodic behavior emerged when the input to their model was not constrained to be excitatory.

Put another way, and consistent with the dynamical systems perspective, there are several different physiologically realistic, viable solutions for how a complex neural network might solve something like navigating even a simple trajectory. Instead of being a metric representation for space, grid coding could instead be considered a multidimensional piece of a highly dynamic system “attempting” to map the correct pattern of motor outputs via feedback (Figure 1B). Note that the patterns of input and output during navigation, in particular, would be expected to be highly dynamic. This is because it involves multimodal sensation, continuously encoding and retrieving novel memories, and tracking one’s position (coarsely) using the path integration system (Ekstrom, Huffman, and Starrett, 2017; Zhao, 2018).

Successor representation models

One new exciting frontier in the interpretation of some of the cells in the medial temporal lobe is the *successor representation* (SR), which derives from the field of reinforcement learning (Stachenfeld, Botvinick, and Gershman, 2014; Russek, Momennejad, Botvinick, Gershman, and Daw, 2017; Stachenfeld, Botvinick, and Gershman, 2017; Behrens et al., 2018; Gershman, 2018). Successor representations stem from the idea that the more one experiences an environment, the better each state represents a future state (for review see Gershman 2018). If the environment is sampled equally, then the SR of a single state or place closely resembles a place field, and in cases of reward, SR nicely predicts an uneven distributions of place cells near the rewarded locations (Stachenfeld et al., 2014). For our current considerations, one

important property of SR for a circular or square environment is that some of the eigenvectors (vectors that explain the most variance) have a similar spatial patterns to that of grid cells (Stachenfeld et al., 2017). These grid-like eigenvectors have also been shown to be warped in triangular environments, similar to grid cells recorded in the rodent (Krupic et al., 2016). In this way, the eigenvector (basis function) of a grid cell representation is a form of information reduction and state prediction rather than a structure isomorphic to behavior.

We believe that ideas at the core of SR also tie in with the properties listed above about dynamical systems and, more generally, the notion that space can be metric or non-metric depending on the task demands. For example, SR can account for non-metric spaces because this is an inherent property of graphs (Stachenfeld et al., 2017). In addition, a study by Momennejad, Russek, Cheong, Botvinick, Daw, and Gershman, 2017 showed that while participants behaviors are more sensitive to reward changes, in favor of SR models, they are also sensitive to transition changes, in favor of model-based reinforcement learning (not predicted by SR alone). Importantly, such findings suggest that SR need not always be isomorphic to behavior; rather, as others have suggested, there is a cooperative and competitive dynamic between multiple systems of model-based, model-free, SR, episodic memory, working memory, and other coding schemes (Collins, Ciullo, Frank, and Badre, 2017; Gershman and Daw, 2017; Kool, Gershman, and Cushman, 2017; Kool, Cushman, and Gershman, 2018). Thus, taken together, the results from SR suggest that the brain exhibits dynamical coding, which limits the likeliness of a strict isomorphism between a group of neurons and complex behavior.

Where do we go from here?

While there is insufficient data currently to falsify either the navigation or extended representational perspectives on grid coding, we do think, however, that there are situations in which a dynamical systems perspective can be falsified and some of its inherent circularity can be overcome. As we have elucidated for SR models, one key prediction of dynamical models, in contrast to the representational perspective, is that any given navigational task (or other cognitive task) will have a range of potentially different “optimal” neural configurations that could be used to solve the task, some of which might involve grid coding and others that do not. These optimal solutions, in principle, could be derived using network modeling for any given navigation task and compared in terms of the patterns of neural activity.

Indeed, the presence of 4-fold symmetry in at least one empirical and one theoretical study might appear to falsify at least the universality of 6-fold grid coding and instead favor a form of dynamical configuration (Dordek et al., 2016; He and Brown, 2019). Another key prediction is that even small changes in the behavioral demands of the task, particularly if a condition involves non-stationarities, will evoke fairly dramatic changes in the neural configurations underlying the task (Shenoy et al., 2013). In this way, grid coding would be one example of a set of neural configurations that could be useful to solve the task, but which is not necessary for the normal expression of it. A key experiment that would more directly pit the two models against each other would involve some form of systematic stimulation of putative grid codes to determine whether their structure itself is necessary for solving non-spatial tasks. A dynamical systems perspective would argue that a range of different solutions might be possible, and thus, by perturbing grid coding (e.g., microstimulation of the LFP), one could determine the “boundary conditions” for grid coding and behavioral expression.

A recent study by Gil et al. manipulated grid coding optogenetically in rodents and determined how this affected path integration (Gil, Ancau, Schlesiger, Neitz, Allen, De Marco, and Monyer, 2018). Gil et al. knocked out NMDA receptors within the entorhinal cortex and parts of the hippocampus and observed specific impairments in the “gridness” score of grid cells (a measure of 6-fold symmetry). These same mice showed impairments in a path integration

task involving swimming within a corridor and then finding the same location in an open pool, which was measured by the accuracy of the vector representing their swim patterns. Notably, mutant mice still had grid cells that showed comparable “gridness” to the best grid cells in the control mice. This suggests that even a small number of grid cells are not sufficient for normal performance of the task, consistent with earlier arguments regarding mapping between cellular responses and behavior (Krakauer et al., 2017).

More importantly, however, for our current considerations, another class of cells, the irregularly spatially selective (ISS) cell, was also disrupted and correlated with reductions in path integration performance. The ISS cell is an excellent example of response that has no obvious manifestation in navigational behavior yet their disruption also related, in some form, to normal path integration performance. This type of response would be consistent with earlier arguments regarding dynamic configurations of neural ensembles and the difficulty of mapping them to any explicit task-based representation. Finally, at least in some cases, head-direction cells were also impaired, suggesting that dynamic interactions between ensembles of neurons were important to normal expression of path integration in their task. While these findings do not directly falsify the grid coding perspective, they make clear that even relatively simple experimental tasks like path integration involve a complex mapping between single cell responses and navigational behavior. While this study did not test the interactions between these different cell types, one line of evidence that would more strongly support the dynamical systems vs. representational perspectives would be a demonstration that interactions between these multiple cell types (grid cells, ISS cells, and head directions), rather than their activity in isolation, are necessary for normal expression of path integration.

Conclusions

In this commentary, we have explored the extent to which assuming an isomorphism between grid coding, mental representation, and behavioral expression can help us to understand human spatial navigation and non-spatial tasks like categorization. In particular, we explored both the logical foundations of isomorphism and how well this assumption jibes with what we know about how humans navigate. Overall, we question the strict assumption of such an isomorphism, particularly because humans are capable of a wide range of different behaviors that dynamically tune to accomplish task demands. We suggest some alternative considerations for how we might instead conceive of the link between grid cells and behavior and some experiments that might better test the nature of this interaction.

References

- Ashby, E. G. & Maddox, W. T. 2005. Human category learning. *Annual Review of Psychology*, 56, 149-178.
- Ashby, F. G. & Maddox, W. T. 2011. Human category learning 2.0. *Year in Cognitive Neuroscience*, 1224, 147-161.
- Ashby, F. G. & O'brien, J. B. 2005. Category learning and multiple memory systems. *Trends in Cognitive Sciences*, 9, 83-89.
- 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*.
- Barry, C., Ginzberg, L. L., O'keefe, J. & Burgess, N. 2012. Grid cell firing patterns signal environmental novelty by expansion. *Proc Natl Acad Sci U S A*, 109, 17687-92.
- Behrens, T. E., Muller, T. H., Whittington, J. C., Mark, S., Baram, A. B., Stachenfeld, K. L. & Kurth-Nelson, Z. 2018. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron*, 100, 490-509.
- Bellmund, J. L. S., Deuker, L., Schroder, T. N. & Doeller, C. F. 2016. Grid-cell representations in mental simulation. *Elife*, 5.
- Bellmund, J. L. S., Gardenfors, P., Moser, E. I. & Doeller, C. F. 2018. Navigating cognition: Spatial codes for human thinking. *Science*, 362.
- Burak, Y. & Fiete, I. R. 2009. Accurate Path Integration in Continuous Attractor Network Models of Grid Cells. *Plos Computational Biology*, 5.
- Burgess, N. 2008. Grid cells and theta as oscillatory interference: theory and predictions. *Hippocampus*, 18, 1157-1174.
- Bush, D., Barry, C., Manson, D. & Burgess, N. 2015. Using grid cells for navigation. *Neuron*, 87, 507-520.
- Butler, W. N., Hardcastle, K. & Giocomo, L. M. 2019. Remembered reward locations restructure entorhinal spatial maps. *Science*, 363, 1447-1452.
- Chan, E., Baumann, O., Bellgrove, M. A. & Mattingley, J. B. 2013. Reference frames in allocentric representations are invariant across static and active encoding. *Front Psychol*, 4, 565.
- Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A. & Fiete, I. 2019. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nature Neuroscience*, 1-9.
- Chen, X. L., He, Q. L., Kelly, J. W., Fiete, I. R. & Mcnamara, T. P. 2015. Bias in Human Path Integration Is Predicted by Properties of Grid Cells. *Current Biology*, 25, 1771-1776.

- Collins, A. G. E., Ciullo, B., Frank, M. J. & Badre, D. 2017. Working Memory Load Strengthens Reward Prediction Errors. *The Journal of Neuroscience*, 37, 4332-4342.
- Constantinescu, A. O., O'reilly, J. X. & Behrens, T. E. J. 2016. Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352, 1464-1468.
- Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M.-B. & Moser, E. I. 2009. Fragmentation of grid cell maps in a multicompartment environment. *Nature neuroscience*, 12, 1325.
- Doeller, C. F., Barry, C. & Burgess, N. 2010. Evidence for grid cells in a human memory network. *Nature*, 463, 657-61.
- Dordek, Y., Soudry, D., Meir, R. & Derdikman, D. 2016. Extracting grid cell characteristics from place cell inputs using non-negative principal component analysis. *Elife*, 5, e10094.
- Eichenbaum, H. & Cohen, N. J. 2014. Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function? *Neuron*, 83, 764-770.
- Ekstrom, A. D., Arnold, A. E. & Iaria, G. 2014. A critical review of the allocentric spatial representation and its neural underpinnings: toward a network-based perspective. *Front Hum Neurosci*, 8, 803.
- Ekstrom, A. D., Huffman, D. J. & Starrett, M. 2017. Interacting networks of brain regions underlie human spatial navigation: A review and novel synthesis of the literature. *Journal of Neurophysiology*, jn. 00531.2017.
- Ekstrom, A. D. & Isham, E. A. 2017. Human spatial navigation: representations across dimensions and scales. *Current Opinion in Behavioral Sciences*, 17, 84-89.
- Ekstrom, A. D., Spiers, H. J., Bohbot, V. D. & Rosenbaum, R. S. 2018. *Human Spatial Navigation*, Princeton University Press.
- Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M. & Cunningham, J. P. 2016. Reorganization between preparatory and movement population responses in motor cortex. *Nature Communications*, 7, 13239.
- Embick, D. & Poeppel, D. 2015. Towards a computational(ist) neurobiology of language: Correlational, integrated, and explanatory neurolinguistics. *Lang Cogn Neurosci*, 30, 357-366.
- Ferbinteanu, J. 2016. Contributions of Hippocampus and Striatum to Memory-Guided Behavior Depend on Past Experience. *J Neurosci*, 36, 6459-70.
- Frankenstein, J., Mohler, B. J., Bulthoff, H. H. & Meilinger, T. 2012. Is the Map in Our Head Oriented North? *Psychological Science*, 23, 120-125.
- Gallistel, C. R. 1989. Animal Cognition - the Representation of Space, Time and Number. *Annual Review of Psychology*, 40, 155-189.
- Gallistel, C. R. 1990. *The Organization of Learning*, Cambridge, MA, MIT Press.

- Garvert, M. M., Dolan, R. J. & Behrens, T. E. 2017. A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *Elife*, 6, e17086.
- Gershman, S. J. 2018. The Successor Representation: Its Computational Logic and Neural Substrates. *The Journal of Neuroscience*, 38, 7193-7200.
- Gershman, S. J. & Daw, N. D. 2017. Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, 68, 101-128.
- Gil, M., Ancau, M., Schlesiger, M. I., Neitz, A., Allen, K., De Marco, R. J. & Monyer, H. 2018. Impaired path integration in mice with disrupted grid cell firing. *Nature neuroscience*, 21, 81.
- Goodroe, S. C., Starnes, J. & Brown, T. I. 2018. The Complex Nature of Hippocampal-Striatal Interactions in Spatial Navigation. *Front Hum Neurosci*, 12, 250.
- Hafting, T., Fyhn, M., Molden, S., Moser, M. B. & Moser, E. I. 2005. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436, 801-6.
- Hales, J. B., Schlesiger, M. I., Leutgeb, J. K., Squire, L. R., Leutgeb, S. & Clark, R. E. 2014. Medial entorhinal cortex lesions only partially disrupt hippocampal place cells and hippocampus-dependent place memory. *Cell Rep*, 9, 893-901.
- Hartley, T., Trinkler, I. & Burgess, N. 2004. Geometric determinants of human spatial memory. *Cognition*, 94, 39-75.
- Hasselmo, M. E., Giocomo, L. M., Brandon, M. P. & Yoshida, M. 2010. Cellular dynamical mechanisms for encoding the time and place of events along spatiotemporal trajectories in episodic memory. *Behavioural brain research*, 215, 261-274.
- Hawkins, J., Lewis, M., Klukas, M., Purdy, S. & Ahmad, S. 2018. A framework for intelligence and cortical function based on grid cells in the neocortex. *Frontiers in neural circuits*, 12.
- He, Q. & Brown, T. I. 2019. Environmental Barriers Disrupt Grid-like Representations in Humans during Navigation. *Curr Biol*.
- Hirtle, S. C. & Jonides, J. 1985. Evidence of hierarchies in cognitive maps. *Memory & cognition*, 13, 208-217.
- Horner, A. J., Bisby, J. A., Zotow, E., Bush, D. & Burgess, N. 2016. Grid-like processing of imagined navigation. *Current Biology*, 26, 842-847.
- Huffman, D. J. & Ekstrom, A. D. 2018. Which way is the bookstore? A closer look at the judgments of relative directions task. *Spatial Cognition & Computation*, 1-37.
- Huttenlocher, J., Hedges, L. V., Corrigan, B. & Crawford, L. E. 2004. Spatial categories and the estimation of location. *Cognition*, 93, 75-97.
- Hutzler, F. 2013. Reverse inference is not a fallacy per se: Cognitive processes can be inferred from functional imaging data. *Neuroimage*.

- Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X. X., Suthana, N., Sperling, M. R., Sharan, A. D., Fried, I. & Kahana, M. J. 2013. Direct recordings of grid-like neuronal activity in human spatial navigation. *Nat Neurosci*, 16, 1188-90.
- Julian, J. B., Keinath, A. T., Frazzetta, G. & Epstein, R. A. 2018. Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature neuroscience*, 21, 191.
- Kanter, B. R., Lykken, C. M., Avesar, D., Weible, A., Dickinson, J., Dunn, B., Borgesius, N. Z., Roudi, Y. & Kentros, C. G. 2017. A Novel Mechanism for the Grid-to-Place Cell Transformation Revealed by Transgenic Depolarization of Medial Entorhinal Cortex Layer II. *Neuron*, 93, 1480-+.
- Kelso, J. S., Saltzman, E. L. & Tuller, B. 1986. The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, 14, 29-59.
- Killian, N. J., Jutras, M. J. & Buffalo, E. A. 2012. A map of visual space in the primate entorhinal cortex. *Nature*, 491, 761-4.
- Kool, W., Cushman, F. A. & Gershman, S. J. 2018. Chapter 7 - Competition and Cooperation Between Multiple Reinforcement Learning Systems. In: MORRIS, R., BORNSTEIN, A. & SHENHAV, A. (eds.) *Goal-Directed Decision Making*. Academic Press.
- Kool, W., Gershman, S. J. & Cushman, F. A. 2017. Cost-Benefit Arbitration Between Multiple Reinforcement-Learning Systems. *Psychol Sci*, 28, 1321-1333.
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., Maciver, M. A. & Poeppel, D. 2017. Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93, 480-490.
- Kraus, B. J., Brandon, M. P., Robinson li, R. J., Connerney, M. A., Hasselmo, M. E. & Eichenbaum, H. 2015. During running in place, grid cells integrate elapsed time and distance run. *Neuron*, 88, 578-589.
- Krupic, J., Bauza, M., Burton, S., Barry, C. & O'keefe, J. 2015. Grid cell symmetry is shaped by environmental geometry. *Nature*, 518, 232-5 LID - 10.1038/nature14153 [doi].
- Krupic, J., Bauza, M., Burton, S. & O'keefe, J. 2016. Framing the grid: effect of boundaries on grid cells and navigation. *J Physiol*, 594, 6489-6499.
- Krupic, J., Bauza, M., Burton, S. & O'keefe, J. 2018. Local transformations of the hippocampal cognitive map. *Science*, 359, 1143-1146.
- Kunz, L., Maidenbaum, S., Chen, D., Wang, L., Jacobs, J. & Axmacher, N. 2019. Mesoscopic Neural Representations in Spatial Navigation. *Trends Cogn Sci*.
- Kunz, L., Schroeder, T. N., Hweeling, L., Montag, C., Lachman, B., Sariyska, R., Reuter, M., Stirnberg, R., Stocker, T., Messing-Floeter, P. C., Fell, J., Doeller, C. & Axmacher, N. 2015. Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease. *Science*, 350, 430-433.
- Mack, M. L., Love, B. C. & Preston, A. R. 2017. Building concepts one episode at a time: The hippocampus and concept formation. *Neuroscience Letters*.

- Maidenbaum, S., Miller, J., Stein, J. M. & Jacobs, J. 2018. Grid-like hexadirectional modulation of human entorhinal theta oscillations. *Proceedings of the National Academy of Sciences*, 115, 10798-10803.
- Marr, D. C. 1982. *Vision: A computational investigation into the human representation and processing of visual information*, Cambridge, MA, MIT Press.
- Mcdonald, R. J. & White, N. M. 1994. Parallel Information-Processing in the Water Maze - Evidence for Independent Memory-Systems Involving Dorsal Striatum and Hippocampus. *Behavioral and Neural Biology*, 61, 260-270.
- Mcnamara, T. P. 1986. Mental Representations of Spatial Relations. *Cognitive Psychology*, 18, 87-121.
- Mcnamara, T. P. 1991. Memory's view of space. *Psychology of learning and motivation*. Elsevier.
- Mcnamara, T. P. & Diwadkar, V. A. 1997. Symmetry and asymmetry of human spatial memory. *Cognitive Psychology*, 34, 160-190.
- Mcnamara, T. P., Hardy, J. K. & Hirtle, S. C. 1989. Subjective Hierarchies in Spatial Memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, 15, 211-227.
- Mcnamara, T. P., Rump, B. & Werner, S. 2003. Egocentric and geocentric frames of reference in memory of large-scale space. *Psychon Bull Rev*, 10, 589-95.
- Moar, I. & Bower, G. H. 1983. Inconsistency in Spatial Knowledge. *Memory & Cognition*, 11, 107-113.
- Momennejad, I., Russek, E. M., Cheong, J. H., Botvinick, M. M., Daw, N. D. & Gershman, S. J. 2017. The successor representation in human reinforcement learning. *Nature Human Behaviour*, 1, 680-692.
- Montello, D. R. 1998. A new framework for understanding the acquisition of spatial knowledge in large-scale environments. In: ENGENHOFER, M. J. & GOLLEDGE, R. G. (eds.) *Spatial and temporal reasoning in geographic information systems*. New York: Oxford University Press.
- Morris, R., Hagan, J. & Rawlins, J. 1986. Allocentric spatial learning by hippocampectomised rats: a further test of the "spatial mapping" and "working memory" theories of hippocampal function. *The Quarterly journal of experimental psychology*, 38, 365-395.
- Moser, E. I. & Moser, M. B. 2008. A Metric for Space. *Hippocampus*, 18, 1142-1156.
- Mou, W. & Mcnamara, T. P. 2002. Intrinsic frames of reference in spatial memory. *J Exp Psychol Learn Mem Cogn*, 28, 162-70.
- Mou, W., Mcnamara, T. P., Valiquette, C. M. & Rump, B. 2004. Allocentric and egocentric updating of spatial memories. *J Exp Psychol Learn Mem Cogn*, 30, 142-57.
- O'keefe, J. & Nadel, L. 1978. *The Hippocampus as a Cognitive Map*, Oxford, Clarendon Press.

- Okabayashi, H. & Glynn, S. M. 1984. Spatial cognition: systematic distortions in cognitive maps. *J Gen Psychol*, 111, 271-9.
- Peacock, C. E. & Ekstrom, A. D. 2018. Verbal cues flexibly transform spatial representations in human memory. *Memory*, 1-15.
- Poldrack, R. A. 2006. Can cognitive processes be inferred from neuroimaging data? *Trends Cogn Sci*, 10, 59-63.
- Poldrack, R. A. 2011. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, 72, 692-7.
- Rodo, C., Sargolini, F. & Save, E. 2017. Processing of spatial and non-spatial information in rats with lesions of the medial and lateral entorhinal cortex: Environmental complexity matters. *Behav Brain Res*, 320, 200-209.
- Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J. & Daw, N. D. 2017. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLOS Computational Biology*, 13, e1005768.
- Sadalla, E. K. & Montello, D. R. 1989. Remembering Changes in Direction. *Environment and Behavior*, 21, 346-363.
- Samsonovich, A. & McNaughton, B. L. 1997. Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci*, 17, 5900-20.
- Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B. & Moser, E. I. 2006. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, 312, 758-62.
- Save, E. & Sargolini, F. 2017. Disentangling the role of the MEC and LEC in the processing of spatial and non-spatial information: contribution of lesion studies. *Frontiers in systems neuroscience*, 11, 81.
- Schlichting, M. L., Mumford, J. A. & Preston, A. R. 2015. Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nat Commun*, 6, 8151.
- Shenoy, K. V., Sahani, M. & Churchland, M. M. 2013. Cortical control of arm movements: a dynamical systems perspective. *Annu Rev Neurosci*, 36, 337-59.
- Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. 2014. Design principles of the hippocampal cognitive map. *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*. Montreal, Canada: MIT Press.
- Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. 2017. The hippocampus as a predictive map. *bioRxiv*, 097170.
- 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. e6.

- Starrett, M. J., Stokes, J. D., Huffman, D. J., Ferrer, E. & Ekstrom, A. D. 2018. Learning-dependent evolution of spatial representations in large-scale virtual environments. *J Exp Psychol Learn Mem Cogn.*
- Staudigl, T., Leszczynski, M., Jacobs, J., Sheth, S. A., Schroeder, C. E., Jensen, O. & Doeller, C. F. 2018. Hexadirectional Modulation of High-Frequency Electrophysiological Activity in the Human Anterior Medial Temporal Lobe Maps Visual Space. *Current Biology*, 28, 3325-3329. e4.
- Steffenach, H.-A., Witter, M., Moser, M.-B. & Moser, E. I. 2005. Spatial memory in the rat requires the dorsolateral band of the entorhinal cortex. *Neuron*, 45, 301-313.
- Stensola, T. & Moser, E. I. 2016. Grid cells and spatial maps in entorhinal cortex and hippocampus. *Micro-, meso-and macro-dynamics of the brain*. Springer, Cham.
- Sternad, D. 2000. Debates in dynamics: a dynamical systems perspective on action and perception. Elsevier.
- Stevens, A. & Coupe, P. 1978. Distortions in Judged Spatial Relations. *Cognitive Psychology*, 10, 422-437.
- Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M.-B. & Moser, E. I. 2018. Integrating time from experience in the lateral entorhinal cortex. *Nature*, 561, 57.
- Wang, R. F. 2017. Spatial updating and common misinterpretations of spatial reference frames. *Spatial Cognition & Computation*, 1-28.
- Warren, W. H. 2019. Non-Euclidean navigation. *J Exp Biol*, 222.
- Warren, W. H., Rothman, D. B., Schnapp, B. H. & Ericson, J. D. 2017. Wormholes in virtual space: From cognitive maps to cognitive graphs. *Cognition*, 166, 152-163.
- Wittgenstein, L. 2001 [1953]. *Philosophical Investigations*, Blackwell Publishing.
- Wolbers, T. & Wiener, J. M. 2014. Challenges in identifying the neural mechanisms that support spatial navigation: the impact of spatial scale. *Frontiers in Human Neuroscience*, 8, 1-12.
- Yoo, S.-W. & Lee, I. 2017. Functional double dissociation within the entorhinal cortex for visual scene-dependent choice behavior. *Elife*, 6, e21543.
- Zhao, M. 2018. Human Spatial Representation: What We Cannot Learn from the Studies of Rodent Navigation. *J Neurophysiol*.

Figure 1: Navigation in everyday situations: Getting from the Psychology Building to the Student Union at University of Arizona based on the representation and dynamical systems perspectives

(a) Representational perspective: Grid cells, which arise as a result of input from other neural systems, represent the Euclidean coordinates of the environment and provide a direct relationship with behavior. Note, for convenience, only one grid cell is shown, but we assume multiple grid cells of different scales, orientations, and phases are present. M is a transform function that maps this grid activity to behavior and M^{-1} is the inverse transform function. To navigate to the student union, this grid code could serve as the basis for navigating from one location (Psychology Building) to the next (Student Union) based on aligning the grid to a landmark and calculating the vector addition of linear paths

(b) Dynamical systems perspective: Here, $r(t+1)$ is the cortical activity at time $t+1$, which is modulated from neuronal inputs $\{h_n(t+1) \dots h_{n+N}(t+1)\}$, which themselves are modulated by behavior $b(t)$. G is a non-invertible transform function that maps neuronal activity to behavior. The gray box signifies the internal representation of the environment and whether it is Euclidean or not depends on task demands $T()$. $T()$, in turn, a function of environmental complexity $c()$, prior experience $e()$ and other latent factors. Navigating, in this scenario will involve a dynamic interaction between neural activity from multiple senses (visual, vestibular, somatosensory, and auditory), memory systems, and other knowledge (west is to my left) to optimize the strategy employed. For example, when navigating from the Psychology Building, one may initially take a sub-optimal path around a building, but then seeing another building reminds one that they are close to the Student Union and need only head north-west to get there. Once the Student Union is in sight, a beaconing strategy is sufficient to arrive there.

