A map of spatial navigation for neuroscience

Eloy Parra-Barrero a, b, Sandhiya Vijayabaskaran a, Eddie Seabrook a, Laurenz Wiskott a, b, Sen Cheng a, b, *

a Institute for Neural Computation, Faculty of Computer Science, Ruhr University Bochum, Bochum, Germany
b International Graduate School of Neuroscience, Ruhr University Bochum, Bochum, Germany

ABSTRACT

Spatial navigation has received much attention from neuroscientists, leading to the identification of key brain areas and the discovery of numerous spatially selective cells. Despite this progress, our understanding of how the pieces fit together to drive behavior is generally lacking. We argue that this is partly caused by insufficient communication between behavioral and neuroscientific researchers. This has led the latter to under-estimate the relevance and complexity of spatial behavior, and to focus too narrowly on characterizing neural representations of space—disconnected from the computations these representations are meant to enable. We therefore propose a taxonomy of navigation processes in mammals that can serve as a common framework for structuring and facilitating interdisciplinary research in the field. Using the taxonomy as a guide, we review behavioral and neural studies of spatial navigation. In doing so, we validate the taxonomy and showcase its usefulness in identifying potential issues with common experimental approaches, designing experiments that adequately target particular behaviors, correctly interpreting neural activity, and pointing to new avenues of research.

1. Introduction

We can define spatial navigation as the process of determining and maintaining a course from one spatial location to another, irrespective of the intervening distance or medium of travel (Gallistel, 1990). The importance of spatial navigation for neuroscience is neatly illustrated by the sea squirt. In its larval stage, the sea squirt is a tadpole-like creature with a rudimentary eye and a brain-like ganglion. After swimming around in search of a place to settle in, it attaches itself to a solid surface where it will remain for the rest of its life, and proceeds to digest most of its own brain. The sea squirt thus confirms a common suspicion: that brains are only needed by creatures that move around (Llinás, 2001). Understanding how brains enable spatial navigation thus amounts to understanding one of the brain’s most important functions. Navigation is also a paradigmatic example of a complex task engaging a wide array of cognitive processes. To navigate, we need to perceive our environment, learn and deploy abstract representations, plan courses of action, and coordinate distinct modes of behavior, all of which applies to much of cognition. For these and other reasons, spatial navigation has received a tremendous amount of attention across different scientific fields that study cognition and the brain.

Research on the cognitive and neural bases of mammalian navigation dates back at least to Tolman’s experiments in rats in the 1940s (Tolman, 1948). It most famously continues with the ongoing discovery, again mostly in rodents, of a whole “zoo” of cell types responding to different aspects of spatial navigation tasks (Moser et al., 2017; Bicanski and Burgess, 2020). However, research on navigation also involves fields as disparate as primate ecology, human behavioral and functional imaging studies, computational modelling, or robotics. The resulting body of literature is vast and somewhat fragmented across systems and scientific disciplines, further compounding the difficulty of understanding of how the myriad pieces fit together.

In particular, the disconnect between neural and behavioral studies of spatial navigation has led certain areas of neuroscience to focus narrowly on characterizing neural representations of space, often neglecting the computations that those representations are meant to enable and that ultimately drive behavior. For example, we know much about how hippocampal place cells tile the environment with place fields of different sizes, shapes and densities (Lee et al., 2020; Eliav et al., 2021; Tanni et al., 2022), how their fields remap across different environments (Colgin et al., 2008), or how cells fire at different phases of the theta oscillation as animals cross the cells’ fields (O’Keefe and Recce, 2001).

* Corresponding author at: Institute for Neural Computation, Faculty of Computer Science, Ruhr University Bochum, Bochum, Germany.
E-mail address: sen.cheng@rub.de (S. Cheng).

https://doi.org/10.1016/j.neubiorev.2023.105200
Received 25 January 2023; Received in revised form 13 April 2023; Accepted 24 April 2023
Available online 11 May 2023
0149-7634/© 2023 Published by Elsevier Ltd.
However, we have only limited understanding of the computations that make use of place cells to drive spatial behavior. This problem is aggravated by spatial navigation involving more distinct cognitive processes than commonly appreciated. When we navigate to a goal we might, for example, follow a street, aim for a prominent landmark, or triangulate our way to the goal based on some configuration of distal cues. We might follow a familiar route, or attempt to discover a shortcut. Failing to appreciate this complexity leads to the use of oversimplified experimental paradigms that do not adequately probe for specific navigation processes. As a result, often it is not only unclear how the spatial representations observed in experiments contribute to spatial computations, it is not even known which computations they are contributing to.

For these reasons, we argue that the neuroscientific study of spatial navigation requires an increased focus on the various kinds of navigation processes animals engage in, the conditions under which they are expressed, and the ways in which they interact with one another. To this end, we advance a taxonomy of spatial navigation processes building and expanding upon proposals by O’Keefe and Nadel (1978), Trullier et al. (1997) and Franz and Mallot (2000) among others. These earlier suggestions include concepts from both neuroscience and robotics. These fields have worked towards the similar goal of understanding (biological or artificial) agents’ ability to navigate complex environments, often drawing inspiration from one another, e.g., in the development of biologically inspired solutions to the simultaneous localization and mapping problem (Milford et al., 2004; Barrera and Weitzenfeld, 2008; Fox and Prescott, 2010). In acknowledging these interdisciplinary exchanges and potential for further collaboration, we use the term “agent” while describing general features of the taxonomy that may apply to both biological and robotic agents.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Aiming</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Beacon navigation</td>
</tr>
<tr>
<td><strong>Direction field navigation</strong></td>
<td>Guidance</td>
<td>Body alignment &amp; target approaching</td>
<td>Taxon navigation</td>
<td>Aiming</td>
<td>Guidance</td>
<td>Beacon navigation</td>
<td></td>
</tr>
<tr>
<td><strong>Path following</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trail following (Allen, 1999)</td>
</tr>
<tr>
<td><strong>Vector movement</strong></td>
<td>Orientation</td>
<td>Guidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dead reckoning (Gallistel, 1990)</td>
</tr>
<tr>
<td><strong>Guidance</strong></td>
<td></td>
<td>Guidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Landmark navigation</td>
</tr>
<tr>
<td><strong>Route navigation</strong></td>
<td>Route, Taxon system</td>
<td>Place-recognition triggered response</td>
<td>Route navigation</td>
<td>Recognition-triggered response</td>
<td>Route-based navigation</td>
<td>Sequential-egocentric navigation</td>
<td></td>
</tr>
<tr>
<td><strong>Topological navigation</strong></td>
<td>Map, Locale system</td>
<td>Topological navigation</td>
<td>Topological navigation</td>
<td>Map-based navigation</td>
<td>Place navigation</td>
<td>Piloting (Allen, 1999; Biegler, 2000)</td>
<td></td>
</tr>
<tr>
<td><strong>Metric navigation</strong></td>
<td></td>
<td>Metric navigation</td>
<td>Locale navigation</td>
<td>Survey navigation</td>
<td></td>
<td></td>
<td>Piloting (Gallistel, 1990)</td>
</tr>
</tbody>
</table>

Our taxonomy divides navigation processes into navigation behaviors (gray cells) and navigation strategies (white cells). Note that some terms have previously been used to refer to different processes, e.g., guidance was introduced as the process of approaching or maintaining a fixed distance to a single landmark (O’Keefe and Nadel, 1978), but was then used to refer to navigating to a goal based on an array or configuration of landmarks (Trullier et al., 1997; Franz and Mallot, 2000). Piloting is also invoked when navigating with the help of landmarks, but authors differ in whether a configuration of landmarks defines the goal position (e.g., Biegler, 2000), helps to determine the agent’s position within a known metric map of the environment (e.g., Gallistel, 1990), or whether landmarks are approached sequentially defining a route (e.g., Allen, 1999; Biegler, 2000). In some cases, it is also the placement of the boundaries separating navigation processes that differs across taxonomies. For instance, many taxonomies define a process similar to what we call aiming, which also includes things like following an odor gradient, which we place in direction-field navigation. Similarly, praxic (Redish, 1999) or response navigation (Nyberg et al., 2022) involves egocentric responses, e.g., take the corridor to the left, go two meters to the right, which we classify as either path following or vector movement. However, in our characterization, these two behaviors also allow for allocentric responses, and so they encompass more than praxic navigation.
Using the taxonomy as a scaffold, we then review and comment on behavioral and neural studies of spatial navigation in several mammalian species including rodents and humans. In doing so, we validate the taxonomy and showcase its usefulness along the following lines. First, consider that agents can often solve a given navigation task in various ways. Paying attention to the different options available to agents and their defining characteristics can inform the design of experiments that better probe specific behaviors and avoid potential confounds. Second, different navigation processes suggest the need for different kinds of spatial representations. Taking this into account can help us make better sense of neural data and generate hypotheses about what to look for in it. Third, by mapping out the repertoire of navigation processes, the taxonomy can highlight areas in need of more study and point to new avenues of research. And fourth, the terminology used in spatial navigation is inconsistent (Table 1). Different researchers place the boundaries separating navigation processes in different ways, and even when the boundaries are aligned, they sometimes refer to the same behaviors by different names, and apply the same names to different behaviors. This variation is partly idiosyncratic, and partly an expression of different underlying assumptions and characterizations, e.g., as with the different flavors of route navigation discussed in Section 2.4.1. Furthermore, while these differences are already present within given disciplines, they are even more pronounced when comparing across them. The elements of the taxonomy thus provide a common conceptual framework that can enhance communication and collaboration both within and across disciplines.

2. A taxonomy of spatial navigation

In categorizing navigation processes, we draw the lines separating the different elements of our taxonomy at the computational level (Marr, 1982). That is, we focus here on what computational problems are being solved and not on how they are solved at the level of algorithms or neural implementation (that will be discussed in Section 4). We also focus on “core” navigation processes that take as inputs goals (e.g., go home), spatial knowledge (e.g., the layout of the neighborhood) and highly processed sensory information (e.g., there is a pharmacy store on the left), and produce as outputs abstract movement instructions (e.g., get to the point 2 m ahead). Hence, we do not cover supporting behaviors and computations involved in the generation of those inputs (e.g., detection of landmarks, estimation of distances, etc.) or in the execution of those outputs (e.g., orienting, locomotion, avoidance of small obstacles, etc.). We also focus on processes involved in navigating towards specific goal locations, as opposed to other spatial behaviors such as exploration, cruising, or escape (Wiener et al., 2009).

2.1. The architecture of navigation

Our taxonomy distinguishes between two levels of processes: navigation strategies and navigation behaviors (Fig. 1A, blue and green). At the top level, navigation strategies receive the final navigational goal and output a moment-by-moment goal. For a simple navigation task, such as approaching a visible target, this moment-by-moment goal coincides with the final goal. However, more complex tasks demand that they be decomposed into a series of sub-tasks, each of which has its own subgoal. For example, to go abroad on holidays one might have to go to the bus station near home (subgoal 1), then to the city’s main train station (subgoal 2), then to the airport (subgoal 3), and so on. Within this top level of organization there can be some degree of recursion. For example, to reach subgoal 2, the train station, the same type of navigation strategy that introduced it, or a different one, could further subdivide the task and introduce additional subgoals, e.g., the exit of the bus station (SG2.1), and a traffic light along the way (SG2.2). Navigation strategies may also output pseudo-subgoals that guide the navigation for a period of time but are not meant to be reached. For instance, upon exiting the bus station, one might walk towards a clock tower in the distance, but with the intention of turning at a traffic light that lies along the way. Since generally more than one navigation strategy will be available to an agent, a process we refer to as organization of strategies must select or integrate the most adequate strategies to solve the task at hand.

The goals defined by the top level of our taxonomy correspond to sets of points in space. These sets of points could be specified in some kind of reference frame, or more likely, in terms of the object(s) located within
or around that set of points (e.g., the train station situated at the intersection between Main Street and Oak Street). Thus, in general, an agent will have different sources of information available regarding how to reach a goal (e.g., what the train station looks like, its position along a street, the trains that stop there, etc.). At the bottom level of the taxonomy, more elementary navigation behaviors take these goals and transform them into operational objectives (e.g., follow Main Street downhill, aim for a yellow building, etc.), finally producing movement instructions. Since multiple navigation behaviors might be available at any given time, a process that we refer to as organization of behaviors must orchestrate them.

The taxonomy most closely follows that of Franz and Mallot (2000), but with a more fine-grained categorization of navigation behaviors (Table 1) and the addition of the organization layers.

2.2. Navigation behaviors

As described above, navigation behaviors are basic navigation processes that produce movement instructions which guide an agent towards a given goal. The movement instructions are not considered to be motor commands, but rather higher-level representations of, for example, desired movement speed and direction, or distance and direction to a via-point that needs to be reached next. We distinguish five navigation behaviors (Fig. 2) that differ in how they convert the goal into concrete operational objectives that guide the agent towards the goal.

2.2.1. Aiming

In aiming (sometimes referred to as beaconing), the operational objective is simply to move towards the perceived position of a landmark at the goal position until the goal is reached. This navigation behavior is available when cues emanate from the goal in rather straight lines, such as is the case of light and sound, enabling an agent to directly compute the position of the goal—or at least the direction to it. Targets in aiming could also be moving, such as when one agent follows another (O’Keefe and Nadel, 1978).

2.2.2. Direction field navigation

In direction field navigation, the direction to the goal is specified relative to a local directional signal defined over some extended area. The operational objective is then to move at some angle with respect to this signal until the goal is reached. An example of this could be to move in some compass direction guided by the earth’s magnetic field, or by the position of the sun. Another common case involves moving relative to a gradient defined based on some extended property of the environment (e.g., altitude) or based on a signal emanating directly from the goal (e.g., an odor). A potential drawback of this method of navigating is that since the objective is defined in terms of a local direction, lateral deviations from the path to the goal cannot be detected and corrected.

2.2.3. Path following

For path following, the agent has to have knowledge that the goal location lies along a particular path such as an odor trail, a street, a river, the walls of a hallway, etc. The agent also needs to know in which direction to take the path to arrive at the goal. For example, we might know that after exiting our home, we need to follow the street to the right to get to the bus stop. The operational objective is then to follow the path in that particular direction until the goal is reached. Unlike in direction field navigation, lateral deviation from the path can be detected and corrected.

2.2.4. Vector movement

In vector movement, the goal location is defined in terms of a vector that indicates the direction and distance to the goal. Such a vector could have been obtained through path integration (also known as dead reckoning (Darwin, 1873) in a previous traversal to or from the goal.

---

Fig. 2. Navigation behaviors. A: In aiming, movement is directed towards a perceptible beacon. B: In direction field navigation the agent follows a local directional signal (an odor gradient in this example) towards the goal. C: For path following, the agent follows some form of linear marking leading to the goal. D: In vector movement, the agent uses a vector pointing to the goal direction for navigation. The vector could have been obtained, for example, through path integration during the outbound travel. E: When using guidance, the agent navigates to an occluded goal defined by a configuration of landmarks.
location. Alternatively, the vector could be computed from previously acquired metric knowledge about the environment or from perceptible cues (e.g., estimating the distance and direction to a target on the fly when a direct line of sight becomes momentarily available). The operational objective is then to reach the position specified by the vector. This requires path integration to monitor the progress towards the goal and indicate when the desired displacement has been achieved. Because path integration is subject to accumulating noise in practice, this strategy works accurately only for small distances.

2.2.5. Guidance

The most complex navigation behavior involves navigating to a goal location defined in terms of its relationship to a configuration of perceivable landmarks, or to at least one landmark combined with distal orienting cues. For example, a rodent might localize its burrow as being in the midpoint between three prominent trees, or being a certain distance to the left of a tree when facing a mountain. The operational objective is to reach this point. More specifically, guidance could rely on matching the perceived angles between landmarks to those known to obtain at the goal location. In such a case, the computation performed would be that of triangulation. Alternatively, in trilateration, only distances to landmarks would be matched, whereas in triangulation, both distances and angles would be employed.

2.3. Organization of navigation behaviors

As mentioned above, multiple navigation behaviors may be available to an agent at any given time. For example, when trying to get to the peak of a hill, one can either follow an available footpath (i.e., path following), or take the steepest route up (i.e., direction field navigation). In realistic settings, the navigation behaviors available to an agent will be in constant flux as the agent moves through space. For example, a path being followed might end abruptly or a new shortcut to the goal might become visible. Moreover, both the internal state of an agent as well as the environment itself are often subject to change over time, e.g., reaching a certain goal might become more or less urgent, or an odor gradient being followed might suddenly disappear in the wind. The different navigation behaviors might also have different reliabilities, which might vary over time. All of this points to the need for flexibly organizing behaviors to reach a goal. We outline two general possibilities for how this might be accomplished: selection and integration.

In the case of selection, at each point in time a single behavior is active, and agents switch between behaviors during navigation. In the simplest case, a switch only occurs when a current behavior terminates naturally, meaning either (i) a subgoal is reached, or (ii) the current behavior is unable to continue because, for example, an obstacle is encountered or sensory cues are lost (Fig. 3A, dashed lines). We refer to this case as “relay selection”, since different behaviors pass control on to one another. Alternatively, a new behavior might be selected as soon as a better option becomes available, even when the current behavior has not terminated yet (Fig. 3A, red solid line). Such a “dynamic selection” appears preferable, since it is more adaptive in dynamically changing environments. This would require multiple behaviors to run in parallel and continuously compete for control. Therefore, even if less powerful, relay selection might be the only choice when an agent does not have the computational resources needed to engage multiple parallel processes, e.g., when the agent is not paying sufficient attention to the navigational task.

In the case of integration, multiple behaviors running in parallel have shared control over an agent’s movements (Fig. 3B). Each behavior can itself be partitioned into an extraction phase, whereby the information needed for a given operational objective is computed (e.g., visually identifying the position of a target to aim for, judging the direction with which to follow a path), and an execution phase, in which such an objective is implemented. Integration of behaviors could take place at the level of either of these phases. As an example of integration within the extraction phase, consider an agent navigating to a known tree that is surrounded by other similar trees that the agent cannot distinguish visually from a distance. Reaching the goal using aiming alone is not feasible, since an initial step of visual search would fail to identify the correct target. However, provided that the agent has some additional knowledge about the configuration of landmarks surrounding the correct tree, it could use guidance to narrow down the visual search space in which to look for the landmark to aim for, hence combining the extraction phases of guidance and aiming. Integration within the execution phase could happen if independently operating behaviors

---

**Fig. 3.** Illustration of the organization of navigation behaviors. **A:** An agent uses path following through a forest in order to reach a castle on the other side. Exiting the forest, the agent sees that a shortcut through open terrain is available. Dynamic selection (solid red line) immediately makes use of this information and selects aiming. By contrast, relay selection (dotted lines) does not allow premature interruptions and so the agent continues path following until the path is blocked by an obstacle, at which point it selects aiming. **B:** Integration: A rat navigating back to its pups could use a combination of auditory and odor cues, thus integrating aiming (green) and direction field navigation (blue).
each made proposals for the next movement instruction, and these proposals were then integrated. For example, each behavior could propose a direction and speed, which are then combined into a weighted average to determine the direction and speed of the actual movement. Similarly, each proposal could consist of a distribution over movement variables and then a movement would be chosen that maximizes the product of these distributions, which would in some sense provide an optimal form of integration. Finally, note that integration could encompass selection as a limit case that uses a winner-take-all mixing function.

2.4. Navigation strategies

After having discussed how navigation behaviors can be used to arrive at one given goal, we address more complex forms of spatial navigation (Fig. 4A), in which a series of subgoals need to be determined according to some heuristics. Navigation to each subgoal is then accomplished by an appropriate navigation behavior introduced above.

We will discuss three navigation strategies. The simplest one, route navigation, consists of following a sequence of memorized subgoals to reach a (final) goal (Fig. 4B). The other two are built upon the notion of cognitive maps, and we therefore refer to them collectively as map navigation. Here, we interpret the concept of map broadly, in line with Tolman’s proposal of cognitive maps that indicate “routes and paths and environmental relationships” (Tolman, 1948). Within map strategies, we distinguish between those that use topological maps (e.g., a subway map, Fig. 4C), and those that use metric maps (Fig. 4D).

2.4.1. Route navigation

The concept of route navigation is rooted in the idea of actions being guided by stimulus-response associations (Thorndike, 1898; Guthrie, 1935; Hull, 1943). In spatial navigation, stimulus-response associations have previously been termed “response learning” (Tolman et al., 1946, 1947) or “habitual locomotion” (Allen, 1999). They involve simple associations between sensory stimuli and motor responses such as “turn left”. Although it might be possible to navigate using these simple associations, arguably it would be wasteful with respect to the behavioral repertoire available to sophisticated agents such as mammals, i.e., the navigation behaviors just discussed, and would render navigation unrealistically rigid and brittle. In addressing this concern, some authors have reinterpreted the stimulus-triggered responses to involve navigation behaviors instead of motor responses per se (O’Keefe and Nadel,
We take this one step further and propose that what is triggered by the recognition of the arrival at a place is neither a particular motor response nor a navigation behavior, but the navigation toward the next goal in a sequence of goals defining a route. The organization of navigation behaviors outlined above would then integrate or select among potentially several navigation behaviors available for reaching the goal. While the agent could still develop a strong bias towards a preferred sequence of navigation behaviors due to repeated experience of the route, there is nevertheless room for flexibility in choosing alternative behaviors. For example, one of the steps to reach a final goal might normally involve aiming. Yet, on a foggy day one might switch to vector movement for that part of the trajectory. Despite this flexibility, the strategy is still fragile in that it cannot recover from missed cues. If a trigger location is not recognized, the whole sequence of subsequent goals becomes useless, as the agent might lack the knowledge to reach the next goal from the current location. On the positive side, this strategy is relatively fast and has low cognitive load, since it only requires the recall of a sequence of goals.

2.4.2. Topological navigation

Topological navigation uses a graph representation composed of nodes (representing locations in the environment) and connecting edges (representing the possibility of navigating from one node to another based on one or more navigation behaviors) (Fig. 4C). This type of representation has previously been referred to as topological knowledge (Kuipers, 1978; Poucet, 1993), (cognitive) graph knowledge (Chastel and Warren, 2014; Peer et al., 2021), and network knowledge (Byrne, 1979). The graph captures topological properties of the environment such as closure, proximity, separation, and continuity (Poucet, 1993). These properties are preserved under continuous deformations. For example, a square can be continuously deformed to give rise to a circle, and therefore share the same topology. To build such a graph, the agent must detect whether different routes pass through the same location and then merge them into a graph. Planning on the graph requires the definition of the starting node (the current location) and the goal node, and finding a path between them. When more than one path is available, the choice of the path might depend on additional information such as the distance, time of travel or metabolic cost associated with the edges of the graph. This type of navigation is more flexible than route navigation, since the agent can plan a path from any starting node in the graph to any other node that it is connected to. Despite this flexibility, topological navigation does not allow for navigating through unexplored terrain or discovering short-cuts. A topological graph lends itself particularly well to hierarchical organization, encoding spatial relationships at different levels of detail (Wiener and Mallot, 2003; Stevens and Coupe, 1978; McNamara, 1986; Hirtle and Jonides, 1985).

2.4.3. Metric navigation

A prerequisite of metric navigation is the ability to build a (usually two-dimensional) Euclidean map that encodes information about the positions of relevant spatial elements within an embedded coordinate system (O’Keefe and Nadel, 1978; Gallistel, 1990; McNaughton et al., 2006; Bellmund et al., 2018). Portions of this metric map could be built based on direct sensory access to the region of the environment that can be apprehended roughly from a single viewpoint, known as the vista space (e.g., rooms, town squares, valleys, etc.) (Montello, 1993). Multiple such maps could then be stitched together to form a representation of environmental space (e.g., buildings, cities, forest) (Montello, 1993). Generally speaking, the spatial extent and accuracy of the resulting map would depend on the range of sensory access to the environment (e.g., how far out you can see), as well as the degree of familiarity. Metric navigation then involves self-localizing within the metric map, and computing a trajectory expressed as a series of subgoals that take the agent from its current position to the final goal. The optimization procedure used to compute this trajectory can be approximate and based on various criteria or combinations thereof – such as distance, metabolic cost, time, or even loosely defined measures like the most scenic or familiar route.

Metric navigation allows for the most navigational flexibility, because it also represents locations that have never been visited before. It also offers a more robust alternative to topological navigation, since it can deal with some degree of environmental change or map degradation owing to its metric nature, which in turn supports finding short-cuts, and new detours around blocked paths. However, constructing, storing, and operating on a metric map is computationally expensive, i.e., metric navigation is often less efficient than the other two navigation strategies.

2.5. Organization of navigation strategies

Agents often have more than one navigation strategy available and must decide which one to follow or how to combine them. We suggest that agents can organize navigation strategies also using either selection or integration. Selection of a navigation strategy means that only one is active at a given time. The selection might be based on a number of factors, e.g. availability of spatial information, cognitive capacity, and navigational demands. The preferred strategy might shift on longer time scales based on the amount of experience with the environment. For instance, someone navigating to a work place in a new city might initially use metric navigation based on a physical city map, and then switch to route navigation based on extensive familiarity with one particular route. Selection may also occur quickly and flexibly in areas or points of overlap in the spatial representations used by the strategies. Following the previous example, if the familiar route to work is suddenly blocked by construction, the person could switch back to metric navigation to find an alternative route.

Integration could occur at three different levels. First, at the level of the information used by the strategies, with a single strategy deciding the navigational subgoals. For instance, a topological graph might be enhanced by incorporating information from a metric map regarding the relative distances between the nodes of the graph, making topological navigation more efficient. Second, at the level of subgoals. Different navigation strategies could run in parallel and vote for different subgoals, which are then integrated. However, this form of integration appears unlikely since it could result in unviable or incoherent series of subgoals. Lastly, the planning processes employed by the strategies may be integrated at the process level, making use of information and heuristics available to both strategies to work together and find good overall subgoals. For example, processes operating on a metric map could guide and constrain the search for a path from the starting position to the goal in a topological graph, akin to the search in the A* algorithm (Hart et al., 1968).

2.6. Navigation with incomplete information

In this review, we focus on navigation to known goals in static and known environments, which excludes cases in which agents search for resources whose location is variable (e.g. prey or mates) in a familiar territory or for resources in an unfamiliar territory. Wiener et al. (2009) refer to these two possibilities as informed and uninformed search, respectively. We believe that in both cases, search builds upon and makes use of the navigation processes described in our taxonomy, but adds another layer of complexity on top of the navigation processes we are considering in this review. This view contrasts starkly with Redish (1999) or Franz and Mallot (2000) who place search at the base of their taxonomy.

In the simplest cases of search, the behavior of the agent could be described as a random walk characterized by a series of apparently random displacements. Examples of this include Brownian motion, where the length of displacements follow a normal distribution, or the Lévy flight, where infrequent longer displacements enable agents to explore a larger area (Auger-Méthé et al., 2016). However, even in such
cases, the agent is likely engaging in topological or metric navigation, relying on knowledge of the environment that the agent acquired previously or during the search itself, or that is directly available in vista space. This would allow the agent to plan paths around obstacles and set subgoals it then reaches with behaviors such as aiming, path following, or vector movement.

Spatial knowledge would also allow the agent to add structure to its search patterns, for example by implementing some inhibition of return mechanism to avoid visiting the same location multiple times. Exploiting knowledge of statistical regularities could also increase the efficiency of the search. This is clear for known environments, e.g., a monkey might engage a route navigation strategy to visit a series of known trees in search for ripe fruit, or navigate to common meeting points in search for mates (DiFiore and Suarez, 2007). Exploiting statistical knowledge can also be advantageous in unknown environments, e.g., a squirrel might aim towards oak trees to search for acorns, or use direction field navigation to go downhill in search for water. The higher-level processes that guide search behavior fall outside the scope of this review.

3. An introduction to spatial representations in the brain

Much of the neuroscientific research on spatial navigation has focused on studying the spatial representations displayed by individual neurons (Moser et al., 2017). The presence of a given neural representation does not necessarily define which navigation process is being used by the animal, but it does provide some hints (Vijayabaskaran and Cheng, 2022). Here we introduce some neural representations of space that may play a role in multiple navigation processes. In Section 4, we will explicitly relate them to navigation processes, and introduce other more specialized neural representations.

3.1. Place cells

Interest in spatial representations began with the discovery of place cells in the hippocampus proper and dentate gyrus (O’Keefe and Dostrovsky, 1971; Hartley et al., 2014). According to the classical view, each place cell codes for one region of space by increasing its firing rate when the animal is located there. These regions of space are known as the cells’ place fields, and they can be defined in two- or three-dimensional space (Yartsev and Ulanovsky, 2013; Grieves et al., 2020). More recent recordings in larger environments show, however, that each cell can have multiple place fields (Rich et al., 2014; Lee et al., 2020; Eliav et al., 2021). It therefore seems to be the combined activation of the population as a whole that can provide a code for spatial location. This code is composed of representations at multiple spatial scales. Recording from dorsal CA1 of bats flying along a 200-m-long tunnel, Eliav and colleagues found that the multiple fields of cells varied widely in size. The mean ratio between each cell’s largest and smallest field was 4.4, with the largest recorded fields being ca. 30 m long. In addition, place field size increases along the longitudinal dorso-ventral axis of the hippocampus. For instance, in rats running on a 18-m-long linear track, sizes increased from less than 1 m at the dorsal pole to ca. 10 m at the ventral pole (Kjelstrup et al., 2008).

The place cell population encodes each environment differently. This is due to cells shifting their place fields relative to one another, or popping in our out of the neuronal ensemble representing each environment in a phenomenon known as global remapping (Leutgeb et al., 2005b, see Colgin et al., 2008 for a review).

Owing to these properties, place cells are generally thought to underlie a cognitive map of the environment. For this to be true, however, the activation of place cells should be relatively stable in time, and the evidence for this is mixed. Recordings in mice have showed that the place cell code is dynamic, with only about 15% of cells retaining the same place fields for 30 days (Ziv et al., 2013). Still, in that experiment, mice were simply running back and forth along a linear track, which arguably does not require detailed spatial knowledge. Place fields appear to be more stable under conditions that produce or require increased attention to spatial cues (Kentros et al., 2004). Additionally, in rats, some place fields have been shown to be stable for as long as 6 months (Thompson and Best, 1990).

In monkeys, hippocampal cells appear to respond to the animal’s position less frequently. Instead, they are more selective for the position in space the monkeys are facing (Mao et al., 2021) or looking at (Rolls et al., 1997). Similar results have been found in humans (Titsiklis et al., 2020).

3.2. Head direction cells and speed cells

Head direction cells, as their name suggests, fire when an animal is heading in a particular allocentric (i.e., world-centered) direction. In rats, they have been observed in dorsal presubiculum and para-subiculum, medial entorhinal cortex, and a number of regions outside of the hippocampal formation (Taube et al., 1990a; Taube, 2007). The directional preference of head direction cells is anchored by visual cues, and rotating these visual cues leads to an equivalent rotation in the directional preference of all cells (Taube et al., 1990b; Yoganarasimha et al., 2006). In addition to the influence of landmarks, the activity of head direction cells can also be updated by idiothetic (self-motion) cues (Knierim et al., 1998; Yoder et al., 2011). Classical head direction cells respond to heading direction along the horizontal plane (azimuth), but more recent recordings in rats (Angelaki et al., 2020), bats (Finkelstein et al., 2015), and monkeys (Laurens et al., 2016) moving in 3D reveal cells coding not only for azimuth but also pitch and roll or combinations thereof.

A complete representation of the animal’s movement through space also requires representations of linear and angular speed. Indeed, cells coding for angular head velocity have been observed in multiple brainstem and diencephalic structures (Taube, 2007), and across the cortex (Hennestad et al., 2021; Spalla et al., 2022; Keshavarzi et al., 2022; Long et al., 2022); and cells with positive linear responses to linear speed have also been found throughout the hippocampal formation (Kropff et al., 2015; Spalla et al., 2022).

3.3. Grid cells

Grid cells are similar to place cells, except that each cell displays multiple firing fields arranged in a regular lattice of equilateral triangles. Grid cells were first discovered in the medial entorhinal cortex of rats (Hafting et al., 2005). Subsequently they have been reported in the pre- and parasubiculum of rats (Boccara et al., 2010), in mice (Fyhn et al., 2008), bats (Yartsev et al., 2011), and humans (Doeller et al., 2010; Jacobs et al., 2013). Grid cells are arranged in discrete but anatomically overlapping modules. Within each module, grid cells share the same scale and orientation, but different cells have their fields offset from each other. The ratio between grid spacings of successive modules is about 1.4. Four or five modules have been found when recording from up to 50% of the dorsoventral axis of MEC (Stensola et al., 2012). Extrapolating these observations, there could be up to 10 grid modules reaching a grid spacing of ~ 9 m (see also Brun et al., 2008).

Grid cells have generally been thought to provide a metric representation of space based on path integration (McNaughton et al., 2006). This is supported by their regular firing pattern, but also by their reliance on self motion cues (Chen et al., 2019) and by the fact that firing can be maintained in darkness (Hafting et al., 2005). Furthermore, the fields of different grid cells move and rotate in concert across different environments (Fyhn et al., 2007; Yoon et al., 2013), or when fields drift in the absence of sensory cues (Waaga et al., 2022), suggesting that grid cells underlie a coherent and reusable metric representation of space. Conjunctive grid × heading direction cells in deeper layers of MEC may enable the update of the position represented by grid cells during self-motion (Sargolini et al., 2006). Path integration errors are then
corrected by encounters with boundaries (Hardcastle et al., 2015) and landmarks (Pérez-Escobar et al., 2016), possibly mediated by place cell inputs. Indeed, reliable grid cell firing emerge after that of place cells during development (Langston et al., 2010; Wills et al., 2010) and depends on continued input from place cells (Bonnevie et al., 2013). The metric nature of grid cells, however, has come into question due to the observation of numerous instances of irregular firing patterns (Krupic et al., 2015; Derdikman et al., 2009; Hägglund et al., 2019; Stensola et al., 2015). Grid fields are also irregularly arranged in 3D (Ginosa et al., 2021; Grieses et al., 2021).

3.4. Representation of boundaries and object locations

In addition to representing aspects of their own state such as position and heading direction, animals also represent various features of the outside world in order to navigate successfully. These representations are often vectorial, that is, they encode the distance and direction to some environmental feature in allocentric or egocentric coordinates (see Bicanski and Burgess, 2020), for a comprehensive review. One prominent example are boundary vector cells in the subiculum, which respond to the presence of boundaries at specific allocentric distances and directions from the animal (Lever et al., 2009). Similar cells have also been identified in pre- and parasubiculum and MEC, although firing in closer proximity to the borders. This latter group has been referred to as border or boundary cells (Solstad et al., 2008; Boccara et al., 2010; Savelli et al., 2008), although it could perhaps be functionally described as a subclass of boundary vector cells (Bicanski and Burgess, 2020). Cells responding to the presence of objects at specific allocentric positions have also been observed. These cells are modulated by object identity in CA1 (Deshmukh and Knierim, 2013), but not in MEC, where they are active for a wide spectrum of object dimensions and shapes (Haydal et al., 2019).

Egocentric counterparts of these neural codes have also been reported. Egocentric boundary cells have been observed in dorsomedial striatum, and retrosplenial, posthirinal and lateral entorhinal (LEC) cortices (Himnan et al., 2019; Alexander et al., 2020; van Wijngaarden et al., 2020; Gofman et al., 2019; Wang et al., 2018). Other cells have been shown to respond to the egocentric position of certain points in space, like LEC cells coding for the egocentric position of objects (Wang et al., 2018), bat CA1 cells (Sarel et al., 2017) and rat posterior parietal cortex cells (Alexander et al., 2022) coding for the egocentric position of the target, or human parahippocampal cells coding for the egocentric position of anchor or reference points (Kunz et al., 2021).

3.5. Integration and heterogeneity of spatial representations

Most of the cell types encoding spatial information form a coherent spatial representation. For instance, when rotating visual cues within a given environment, place cells rotate coherently with simultaneously recorded head direction cells (Knierim et al., 1995; Yoganarasimha and Knierim, 2005). Also, head direction cells, grid cells, border cells and object vector cells all rotate in concert both within and across environments (Sargolini et al., 2006; Solstad et al., 2008; Haydal et al., 2019).

The evidence reviewed so far suggest a neat division of labour between different well-defined cell types. However, reality appears to be more complex, with cells displaying a high degree of mixed selectivity to various spatial variables such as position, speed or head direction (Hardcastle et al., 2017; Ledergerber et al., 2021; Spalla et al., 2022; Mao et al., 2021; Finkelstein et al., 2015). For instance, in tasks where movement is very stereotypical, such as in linear tracks, the place cell code for a given location depends on running direction (McNaughton et al., 1983; Markus et al., 1995). Furthermore, some spatial cells also display mixed selectivity with non-spatial variables. For example, the activity of place cells can depend on the route the animal is taking through a maze (Wood et al., 2000; Grieses et al., 2016), the position of a reward or a conspecific in the environment (Smith and Mizumori, 2006; Sarel et al., 2022), the number of laps run on a circular maze (Sun et al., 2020), or the evidence accumulated for a left-right choice (Nieh et al., 2021). Cells also present highly heterogeneous response profile shapes. For instance, while canonical speed cells increase their firing rate with speed (Kropff et al., 2015), other cells decrease their firing rate or display U- or inverted U-shaped response profiles (Hardcastle et al., 2017).

4. The taxonomy in action

In this section, we use our taxonomy as a guide for providing a broad overview of behavioral and neurobiological studies of spatial navigation.

4.1. Aiming

Navigation towards directly perceivable goals is pretty straightforward and often used in laboratory settings as a control condition to compare against tasks requiring spatial knowledge (Vorhees and Williams, 2006; Bolding et al., 2020). For instance, in cued controls of the Morris water maze task, a platform placed in a circular pool of water is made visible by either rising it above water, or marking its location with a flag, while curtains around the pool reduce the availability of distal cues. Other studies have focused on aiming itself. For instance, Collett (1987) trained gerbils to approach a shining light bulb. When a different light bulb was switched mid-trajectory, gerbils quickly switched to approaching the newly illuminated bulb in most trials. Aiming has also been observed in the wild. Drickamer and Stuart (1984) analyzed tracks left by mice on the snow immediately following a snowfall. They found that most tracks went from tree to tree, as if the mice were aiming for the tree trunks. Furthermore, it appeared that bigger trees could be aimed for from further away, as there was a positive correlation between the length of the tracks and the diameters of the tree trunks being aimed for.

However, despite the simplicity of aiming, mammals—at least rodents—appear to avoid relying on it exclusively. Learning to aim for a goal did not overshadow learning the goal’s position with respect to the shape of the test environment (Pearce et al., 2001) and indeed guidance with respect to the boundaries of the environment or configurations of extra-maze cues seemed to help define which beacon to aim for or even dominate over the aiming itself (Cheng, 1986; Harrison et al., 2006). In a natural setting, Devenport and Devenport (1994) also showed that squirrels kept visiting the learned previous location of a feeder after it was dislocated, despite the feeder being a prominent beacon in an open area. These results suggest that rodents, and perhaps other mammals as well, are biased toward using guidance, or at least integrating aiming with guidance. This bias could reflect a tendency to perceive geometrical and configurational information as more stable or predictive of the presence of a certain reward (Cheng, 1986). Therefore, experimenters focusing on aiming should make sure that the animal is motivated to aim for the beacon itself, for example because they can directly recognize it as valuable (e.g., a piece of food), and focus on first trials in novel environments or change the position of the beacon frequently.

Aiming experiments generally use relatively proximal visual landmarks as beacons. However, we note that aiming towards distant landmarks can also be a useful strategy. Seals can be trained to swim in the direction of the star Sirius (Mauk et al., 2008), and human Polynesian and Micronesian navigators are also known to steer based on a succession of guide stars (Lewis, 1970). Furthermore, many mammals can accurately locate sound sources. Elephants and humans, for instance, can do so with an acuity of about 1° (Hefner, 1997). Therefore, audition can also provide the basis for aiming or work together with vision by guiding visual search to the location of the object to aim for.

The neural basis of aiming has received relatively little attention. A series of experiments suggest that the dorsolateral striatum is involved in aiming-like tasks whereas the hippocampus is not (for a review, see White et al., 2013). For instance, Packard et al. (1989) trained rats in a radial maze to enter randomly selected arms signaled by a light. Subjects
with dorsal striatum lesions were impaired relative to controls and to those with fimbria-fornix lesions. Other experiments using the Morris water maze found that lesions to the dorsolateral striatum, but not to the fornix, shifted goal preference to the previous position of a visible platform as opposed to its new location after displacement (McDonald and White, 1994; Devan and White, 1999; Devan et al., 1999). This suggests that striatal lesions impair aiming and facilitate guidance. This apparent role of the dorsolateral striatum in aiming has been hypothesized to depend on the structure’s ability to support stimulus-response associations, e.g., when the goal is on the left of the visual field, turn left (White and McDonald, 2002).

However, aiming based on these simple responses, similar also to what a Braitenberg vehicle would do (Braitenberg, 1986), is not very useful. Navigators often need to estimate the position of the target so that they can adjust their speed, and also so that they can switch to vector movement when they look away from the target or the target becomes temporarily occluded. The neural processes underlying these computations remain largely unknown. Nevertheless, electrophysiological studies have identified certain neural representations that could play a role in them. Most notably, these include posterior parietal neurons coding for the egocentric position of visual targets (Alexander et al., 2022), sometimes in conjunction with allocentric heading direction (Wilber et al., 2014). Object-vector cells in the medial entorhinal cortex (Høydal et al., 2019), perhaps working together with landmark vector cells in the hippocampus (Deshmukh and Knierim, 2013), could also serve to situate the target location within the metric representation provided by grid cells.

### 4.2. Direction field navigation

The vertebrate forebrain evolved in underwater habitats where vision and hearing contribute little to no sensory information, and distributed chemical and magnetic gradients dominate. Therefore, navigation using chemical and magnetic direction fields has been argued to reflect an evolutionarily ancient form of navigation behavior (Jacobs and Schenk, 2003; Phillips et al., 2006).

How mammals follow odour gradients, in particular, has received the most attention. For example, Catania (2013) has shown that eastern American moles climb an odour gradient to localize a food source. Moles employ both serial sampling (comparing odour intensity at different positions) when further away from the odour source and inter-nostril comparisons when closer to the goal to compute the direction of the gradient. Similarly, mice can follow odour gradients making use of bilateral cues (Parthasarathy and Bhalla, 2013) as well as of serial sampling by moving the body when far away from the target and moving the head at closer distances (Liu et al., 2020), and by synchronizing sniffs with nose movements (Findley et al., 2021). The navigation of mice to an odour source in the presence of moderate wind also seems consistent with gradient climbing (Gire et al., 2016). Humans are also able to navigate to a position defined by two perpendicular odour gradients (Jacobs et al., 2015), possibly also using bilateral cues (Wu et al., 2020). This demonstrates in a mammal the capacity to combine information from more than one gradient, which has been proposed to explain mid- to long-range navigation in birds and other vertebrates (Wallraff, 1990; Phillips et al., 2006).

Magnetic field navigation in mammals has been explored relatively little, although it has received more attention in other taxa such as migratory birds, insects or sea turtles (Wiltschko and Wiltschko, 2005; Mouretsen, 2018). Nonetheless, there is mounting evidence that mammals can sense magnetic cues and use them for navigation (Begall et al., 2014). For example, some bats have been shown to use the position of the sun at dusk to calibrate a magnetic compass that guides their movements at night (Holland et al., 2006; Lindecke et al., 2019). Analysis of live strandings of cetaceans also suggest that they tend to travel parallel to contour lines of the earth’s magnetic flux density (Klinowska, 1990).

Beyond odour gradients and magnetic fields, there may be many other direction fields that mammals employ for navigation but that have received little attention. Some examples could include water or air currents (Yu et al., 2016), hydrostatic pressure gradients in the ocean, polarized light (studied in birds), celestial compass information, up- or down-hill gradients, patterns of shadow or moss, varying concentrations of fauna and flora, climatic variables, etc. Direction field navigation is also likely to be more complex than generally appreciated. The main challenge is that direction fields are often noisy, e.g., odour gradients are distorted by turbulent air flows. In these situations, simple bilateral and serial strategies break down (Gumaste et al., 2020), and agents must use more complex filtering and compensatory mechanisms that remain to be characterized.

The neural bases of direction field navigation also remain poorly understood. However, there has been some research aimed at uncovering the neurons responsive to certain types of direction fields such as those based on magnetic fields or odour gradients. For instance, early work showed the sensitivity of pineal cells to magnetic fields in rats and guinea pigs (Semm et al., 1980; Reuss et al., 1983; Rudolph et al., 1988). More recently, Némec et al. (2001) found that neurons in the superior colliculus of Zambian mole rats are sensitive to magnetic fields and show direction-specific responses. The most direct evidence for the involvement of magnetic field information in navigation comes from experiments with Ansell’s mole rats. Burger et al. (2010) showed that shielding or periodically switching the direction of the magnetic field resulted in reduced c-Fos expression in multiple regions of the navigation circuit including hippocampal CA1 and CA3, postsubiculum, and retrosplenial and entorhinal cortices in animals that explored a novel circular arena. Reduced c-Fos expression was also observed in hippocampal CA1 and CA3 and dorsal subiculum of animals that rested in a familiar home cage when exposed to magnetic field manipulations. While these findings suggest that magnetosensory information might affect brain regions in the navigation circuit of mole rats, the lack of an explicit goal-directed task makes it harder to pinpoint the nature of this contribution.

Neural evidence for coding of odor concentration gradient in both the serial sampling and inter-nostril comparison strategies has been found in mice. A subset of mitral/tufted cells in the mouse olfactory bulb have been found to code for changes in odor concentration in subsequent inhalations, corresponding to the odor gradient in the serial sampling case. These neural representations of concentration are also direction sensitive—some cells code for an increase in concentration, while others for a decrease (Parabucki et al., 2019). Cells coding for inter-nostril odour comparisons by neurons have also been reported in the anterior olfactory nucleus Kikuta et al. (2010). These neurons exhibit excitatory responses to odor stimulation of the ipsilateral nostril and an inhibitory response to stimulation of the contralateral nostril. Furthermore, the spike responses are phase-locked to the respiratory cycle, suggesting that these neurons might compare inputs from the two nostrils within each sniff.

Future work is required to determine the neural structures and processes that translate these representations of direction fields into navigational decisions.

### 4.3. Path following

Following paths and trails is a common navigation behavior for many mammals. For instance, odour trails can be followed by mice (Jones and Urban, 2018), rats (Khan et al., 2012), and even humans (Porter et al., 2007) (for a review, see Marin et al., 2021). Paths are also often created by the navigators themselves. Humans do this with roads and streets, but many other species also form “game trails” that allow for easier travel. For instance, in a semi-natural 30 × 30 m enclosure, rats developed trails connecting basic goals such as food sources, harborage boxes and burrow entrances, and they usually confined their movements to the same routes following these trails (Calboun, 1963). The trails were kept cut and cleared of all vegetation, making them easy to detect and travel
through. Rats also develop complex burrow systems with numerous interconnected tunnels. Navigation within tunnels would also constitute an important form of path following and would presumably be based on thigmotaxis, a kind of movement where contact with a solid object is the directive factor. Thigmotaxis is also observed in common laboratory settings, such as when rats swim around the borders of a pool during initial learning in the Morris water maze (Vorhees and Williams, 2006).

Despite how widespread path following behavior seems to be among animals and humans, its underlying computations and neural processes have received very little attention among researchers of spatial navigation. The reason is that path following appears to be trivial. For instance, in an fMRI study investigating the neural correlates of spatial navigation while driving, path following (referred to as “coasting”) was used as the baseline condition against which to contrast the neural responses of other spatial behaviors (Spiers and Maguire, 2007). However, as illustrated by the difficulties facing self-driving cars, path following is far from trivial. Some challenges involve the identification of the path on the basis of sparse and noisy data (think of odour trails, or occasional footprints and broken vegetation left by a passing animal), or the control of movements when navigating along a winding path at high speed.

Unsurprisingly, driving is perhaps one of the only areas in which the mechanisms of path following have been studied explicitly. Such studies highlight the need for anticipatory mechanisms based on perception of the upcoming path and the use of internal models (Lappi and Mole, 2018; Land and Lee, 1994). An fMRI study has also identified some of the neural systems engaged by the visual control of steering (Field et al., 2007). The study showed that areas in the superior parietal lobule were highlighted in the need for anticipatory mechanisms based on perception of path following. Electrophysiological studies often analyze the spatial firing properties of cells while animals run on linear or circular tracks. These tracks are tiled with place cells (Rich et al., 2014) and grid cells (Yoon et al., 2016), but it is unclear what functional role, if any, these cell types play in following the track. Perhaps boundary cells could be more directly connected to path following. They are known to respond to both walls and drops (Lever et al., 2009) and could potentially respond to other types of boundaries that define a path. This also raises the question to what extent border cells can represent upcoming boundaries (and therefore upcoming paths), which, as just mentioned, is beneficial for path-following behavior (Lappi and Mole, 2018).

### 4.4. Vector movement

Vector movement has been studied extensively in the context of homing by path integration. In the absence of perceptible cues, subjects leave a starting position (home), follow a complex trajectory and then attempt to return to the starting point along a straight path. To do this, subjects must first integrate their velocity along the outbound path (path integration), so that they can estimate their final position relative to the initial one. Then, subjects return to the starting position using vector movement. This vector movement includes again a path integration component, this time to monitor the progress back to the home and determine when the return journey has been completed.

Mittelstaedt and Mittelstaedt (1980) famously demonstrated this kind of behavior in gerbils. In a circular arena with a diameter of 130 cm, female gerbils went out of their nest in search of a displaced suckling and, upon finding them, brought them back to the starting point following a rather straight path. They were able to do this in complete darkness and with the nest taken to different places before each excursion. If the nest was shifted after the animal had left, they returned to where the nest had been when they had left, demonstrating that they were not following cues emanating from the nest. Furthermore, if the animal was rotated smoothly when retrieving a suckling from the center of the arena, the nest was missed by precisely the rotation angle. These observations provide strong evidence for the animals performing vector movements guided by path integration.

Similarly, hamsters can estimate the distance and direction back to a starting point (Seguinot et al., 1993), even in novel environments (Siegrist et al., 2003). However, subjects tend to commit systematic errors. In triangle completion tasks, hamsters (Seguinot et al., 1993), dogs (Seguinot et al., 1998) and humans (Harootonian et al., 2020) tend to overestimate the angle they need to turn and underestimate the distance they need to walk to get back to the starting point, a pattern of error which is also observed in ants (Müller and Wehner, 1988). This pattern of error might act as a safety device, as it brings the animal closer to its outbound path, where it is more likely to recognize familiar positions (Seguinot et al., 1993). These experiments are generally conducted in small laboratory settings. Testing human participants in an omnidirectional treadmill, Harootonian et al. (2020) showed that the underestimation of distance in the triangle completion task increased logarithmically within the tested range, up to an error of about 40% for a homing distance of 200 m.

In addition to homing behavior, vector movement has been demonstrated in navigation towards known goals. For instance, Etienne et al. (1998) showed that, in the dark, hamsters could reach a familiar feeding place after having been led from their nest to various arbitrary positions. However, this task, like the homing task, includes an initial path integration phase prior to the vector movement. Path integration and vector movement are partly overlapping and have often been considered synonymous (Etienne and Jeffery, 2004; Foo et al., 2005; Gil et al., 2018). They are, however, distinct: path integration can happen outside of vector movement; and vector movement incorporates an additional fundamental component beyond path integration, namely, the computation of the movement instructions required to reach the goal. Because of this, vector movement could be studied more specifically using tasks that do not involve an initial path integration phase. An example of this would be navigating between a pair of points of a well-known environment in the absence of sensory cues. In this case, the vector to be executed could already be known, or could be computed from a metric map of the environment. An example of this can be seen in the experiment by Foo et al. (2005), where they trained participants to navigate from a home position to two known invisible goals.

Other tasks that could rely on vector movement more than it might seem at first are aiming and guidance tasks. This is because vector movement could serve as a fall-back mechanism that is engaged whenever the sensory cues supporting aiming or guidance become unavailable, e.g., the animal looks away from the landmarks, or the landmarks become temporarily occluded. This is nicely illustrated by the finding that gerbils using guidance could still arrive at the goal location when the lights were switched off mid-trajectory (Collett et al., 1986). Similarly, human participants could navigate blindfolded to the location of a previously seen target (Mittelstaedt and Mittelstaedt, 2001). This last experiment also revealed a curious effect of walking speed: when asked to walk slower, participants seemed to underestimate their travelled distance and overshoot the target, and vice-versa for walking faster. This suggests that participants were also relying on measuring walking time, or that they performed path integration using a speed signal biased towards the average walking speed.

Overall, these behavioral studies highlight the somewhat severe errors incurred during path integration, which limits the applicability of vector movement to navigation over a relatively small range of distances.

Of all the navigation behaviors, vector movement is perhaps the one whose neural implementation we are closest to understanding. That is because we now know of several cell types that seem very relevant for vector movement. Of these, grid cells take center stage. Because of their metric firing properties, it has been widely hypothesized that grid cells underlie the two main components of vector movement: path integration (McNaughton et al., 2006), and the computation of movement
instructions required to reach a goal, either directly (Edwardsen, 2016) or via the computation of a vector pointing to the goal (Stemmler et al., 2015; Bush et al., 2015). This view is supported by evidence that lesions, or inactivations, of the MEC affect performance in vector movement tasks (Van Cauter et al., 2012; Tennant et al., 2018) possibly due to the MEC’s role in estimating distances (Jacob et al., 2017b; Winter et al., 2013). A couple of studies provide more direct evidence for the involvement of grid cells in particular in vector movement. Genetically modified mice lacking NMDA-glutamate receptors in the retro-hippocampal region (i.e., entorhinal cortex, subiculum, pre- subiculum, postsubiculum and parasubiculum) showed disrupted grid cell activity, which affected performance in a vector movement task (Gil et al., 2018; Allen et al., 2014).

Head direction cells have also been implicated in vector movement, with their activity correlating with the directional errors subjects incur in during homing tasks (Butler et al., 2017; Valerio and Taube, 2012; van der Peet et al., 2010). Arguably, this is because head direction cells perform angular path integration (Knierim et al., 1998; Yoder et al., 2011; Clark and Taube, 2011; Valerio and Taube, 2012) and are generally believed to provide the directional signal required for grid cells to perform path integration. Indeed, head direction cells are necessary for sustaining grid cell activity (Winter et al., 2015). However, an inconvenient truth about head direction cells is that, as their name suggests, they seem to be more indicative of the heading direction of the animal than of its movement direction, and the two significantly deviate from each other (Raudies et al., 2015). As a result, it seems that head direction cells cannot be used directly for path integration. Perhaps we are yet to find cells coding for allocentric, i.e., world-reference, movement in mammals similar to those recently observed in flies (Lyu et al., 2022).

The hippocampus has also been implicated in vector movement, although the link is less clear. Rats with lesions to the hippocampus were severely impaired in their ability to perform homing (Whishaw and Maaswinkel, 1998; Kim et al., 2013; Winter et al., 2013), but see (Alyan and McNaughton, 1999). Vector signals have also been observed in neuroimaging studies of the human hippocampus while subjects were involved in navigating toward spatial goals (Howard et al., 2014; Viard et al., 2011; Patai et al., 2019). Cells encoding vector representations of goals (even when occluded) have been observed in bats Sarel et al. (2017). However, these are likely to be specific to particular goals and particular tasks, and therefore might not support vector movement more generally.

Taken together, these studies suggest that grid cells provide a metric representation of space based on path integration, which is supported by speed and head direction cells. The starting and goal locations are represented in this metric, perhaps with the help of place cells, and then translated into movement instructions.

However, there is a striking counterpart to this general picture. Humans with lesions to the medial temporal lobe are still able to perform homing (Shrager et al., 2008; Kim et al., 2013). This has been attributed to the ability of humans to use spatial working memory to keep track of where they are within the environment (Kim et al., 2013; Sapiurka et al., 2016), but that in turn suggests that grid cells might not be crucial for computing and representing movement vectors.

4.5. Guidance

Much of the experimental research on spatial navigation has been conducted using relatively small and open arenas where subjects learn to navigate to a position specified relative to a configuration of surrounding boundaries or landmarks. The most notable example is the Morris water maze, where animals swim to a hidden platform in a pool of water guided by some configuration of distal landmarks (Morris, 1981, 1984; Sutherland and Dyck, 1984; Vorhees and Williams, 2006). These experimental paradigms have often been construed as testing some form of map navigation. However, map navigation requires pinpointing both the starting and the goal positions in a map and computing an efficient route between them. While it is conceivable that animals solve these kinds of tasks in this way, we believe it is more likely that they use guidance instead. This is because, in these experiments, landmarks are visible from every point in the environment and animals can simply move towards the target location by projecting allocentric vectors from each landmark to the goal (Wolbers and Wiener, 2014; Benhamou, 1996; Eichenbaum, 2017).

Some studies have attempted to determine precisely how the position of the goal is calculated from configurations of landmarks and/or boundaries. Collett et al. (1986) trained gerbils to find sunflower seeds at locations indicated by constellations of landmarks placed on the floor. Gerbils learned to complete this task, even when the array of landmarks and the starting positions were randomly displaced from trial to trial, both with respect to each other, and with respect to the room, demonstrating that gerbils indeed relied on the configuration of landmarks. In some experiments, gerbils were trained to search in some position within triangles defined by three landmarks. When one or two landmarks were then removed, such that the position of the target could no longer be unambiguously determined, gerbils searched in the positions consistent with the remaining landmarks. When one landmark was moved further apart, the animals kept searching at the original distance of the target from the landmarks that did not move, or from the landmark that was closest to the target. Their experiments point at independent and differently weighted contributions of vectors pointing from each landmark to the goal. This is also supported by the finding that performance in tasks like the Morris water maze is not severely impaired by removal of redundant cues (Fenton et al., 1994; Nakazawa et al., 2002; Gold and Kesner, 2005).

Some authors have studied whether guidance operates by matching the distances from the goal to the landmarks (trilateration) or by matching the relative angles between landmarks as seen from the goal position (triangulation). Manipulation studies on rats (Maurer and Derivaz, 2000) and humans (Waller et al., 2000) lend stronger support to the former. Waller et al. (2000) conclude that humans rely primarily on information about relative distances except when (i) it violated an enclosure relationship (e.g., the target was originally enclosed within the area defined by three landmarks, but maintaining the relative distances after manipulation brings the target outside of it) or (ii) angular information was very salient during learning (e.g., containing all right angles).

Other studies have focused on guidance with respect to environmental boundaries. Hartley et al. (2004) asked humans to learn the positions of objects in empty rectangular arenas that were then stretched or compressed. The responses of the participants were best described as trying to minimize the deviation between the learned and observed proximities (1 / (d + c)) of the target to the four walls of the arena, where d is the distance to the wall and c is a constant. This tends to maintain fixed distances to nearby walls, which are weighed more strongly than far away ones. For targets towards the center of the arena, it predicts instead that the ratio between distances to the walls should be preserved, since the walls are more equally weighed. A similar experiment, however, suggests a different mechanism by showing that the remembered positions of objects in deformed environments shift from trial to trial depending on the boundary of origin of the participants’ movement trajectories (Keinath et al., 2021).

In addition to studying how subjects match distances or directions to landmarks, a key question related to how the mere presence of the landmarks themselves might affect guidance. In particular, it has been debated whether guidance operates purely based on geometric (e.g., the shape of the arena) as opposed to featural (e.g., colors, textures, etc.) information. The debate originated when Cheng (1986) found that in rectangular arenas, rats often confuse diagonally opposite corners, which are geometrically equivalent, even when distinct visual features clearly differentiate between them (see also Keinath et al., 2017). This led to the postulation of a geometric module that operates independently of the
processing of featural information. Although a strict independence is now contested (Cheng, 2008), there is still evidence for the differential treatment of geometric and non-geometric information.

Hermer and Spelke (1994) and Gouteux and Spelke (2001) tested human participants’ ability to return, after being disoriented, to one of three or four identical boxes where an object had been hidden. Adults were able to use the geometrical configuration of boxes to localize the goal. When the boxes were arranged in a rectangle, adults could also use a non-geometric polarizing cue (a colored curtain along a wall) to disambiguate the correct corner. However, children of up to 4 years old failed to orient in these circumstances. When the rectangular configuration of boxes was then framed by four walls, children searched in the two geometrically correct corners, like rats did (Cheng, 1986). This suggests a double dissociation between geometric and non-geometric cues on the one hand, and within geometrical cues, between those based on arrangements of discrete objects and those based on the shape of extended surfaces like environmental boundaries. Indeed, participants prefer to learn the position of a target relative to boundaries rather than relative to an intra-maze landmark (Vouillas et al., 2010). Similarly, rats fail to learn the location of a feeder when specified in relation to a moving intra-maze landmark (Biegler and Morris, 1993). Furthermore, the dissociation between boundaries and intra-maze landmarks can also be observed based on the different learning rules that make use of them. Doeller and Burgess (2008) show that in humans, learning the position of an object with respect to an intra-maze landmark obeys associative reinforcement and can thus be overshadowed and blocked by learning its position with respect to another landmark or with respect to the boundaries. However, learning a position with respect to the boundaries is incidental and does not show overshadowing or blocking.

Guidance has also been observed outside the laboratory. For instance, Tsoar et al. (2002) captured wild bats and released them up to 84 km away from their cave and foraging areas. The bats could return easily to their caves when distal visual landmarks, such as hills or town lights, were not occluded, suggesting that the bats were relying on guidance with respect to these landmarks.

Regarding the neural structures underlying guidance, rodent lesion studies using the Morris water maze or dry versions of it point to a key role for the hippocampus. In particular, the hippocampus appears to be involved in both learning to solve the task (Morris et al., 1982, 1990; McDonald and White, 1994; de Brun et al., 2001; Jeffery et al., 2003; Miyoshi et al., 2012), and in continuing to solve it for at least up to 3 months after acquisition (Sutherland et al., 2001; Clark et al., 2005a,b; Ocampo et al., 2017). More fine-grained lesioning or genetic manipulation studies have found a role specifically for CA3 in solving these kinds of tasks when some of the visual cues surrounding the maze that were used during training are removed (Nakazawa et al., 2002; Gold and Kesner, 2005). Rat lesion studies (Pearce et al., 1998, 2004; Kosaki et al., 2015) and human fMRI studies (Doeller and Burgess, 2008; Vikbladh et al., 2019) also indicate that the hippocampus is selectively involved in guidance with respect to environmental boundaries and distal (extra-maze) landmarks, whereas the striatum is involved in guidance with respect to a discrete proximal (intra-maze) landmark.

The importance of the hippocampus in guidance can also be intuited from the characteristics of place cell firing. For instance, Keinath et al. (2017) trained rats to find rewards in a corner of a rectangular box. As mentioned above, animals often made errors, searching in the other geometrically equivalent corner. Simultaneous place cell recordings showed that after disorientation, the virtual place map aligned within the rectangle predicted where the rats would navigate to. Furthermore, place cell activity is degraded in conditions in which performance in water maze tasks is reduced (Nakazawa et al., 2002; Hales et al., 2014). Additional circumstantial evidence for the involvement of place cells in guidance comes from the fact that the pattern of responses of human participants in the task by Hartley et al. (2004) can be explained by a model of place cells (Barry et al., 2006). The activity of some hippocampal cells has also been shown to shift relative to movable landmarks (Gothard et al., 1996), which suggests that the hippocampus could also play a role in navigating to a point specified relative to them. An intriguing exception to the association between place cells and guidance is a study that showed that performance in a hippocampus-dependent guidance task was not affected by an intervention that induced place cell remapping (Jeffery et al., 2003).

Theta phase coding in the hippocampus also seems to play a role in guidance. Disrupting theta oscillations severely impaired navigation to a hidden goal, even though the spatial selectivity of place cells was largely preserved (Bolding et al., 2020). Intriguingly, theta phase coding is also disrupted in multiple diseases associated with cognitive and spatial impairments (Speers et al., 2021; Munn et al., 2022; Talbot et al., 2018).

The hippocampus is connected to the neocortex mainly through the entorhinal cortex (Witter et al., 2017). Hence, numerous lesion studies have looked at the role of the entorhinal cortex in solving the water maze task. Some studies did not find a significant contribution of the entorhinal cortex (e.g. Galani et al., 1998; Pouzet et al., 1999; Bannerman et al., 2001; Oswald et al., 2003), but others find a contribution of the entorhinal cortex in general (Eijkelboom et al., 2000; Hales et al., 2014), specifically for the dorsolateral entorhinal cortex (Steffenach et al., 2005), or specifically when using distal as opposed to proximal landmarks (Parron et al., 2004; Poitreau et al., 2021). The entorhinal cortex has also been implicated in the expression of remote (1 month old) spatial memories (Hales et al., 2018). These roles of entorhinal cortex in guidance could possibly be mediated by combinations of grid cells and object vector cells (Hoydal et al., 2019).

Other structures connected to the hippocampus and often involved in spatial navigation are the medial prefrontal cortex and the retrosplenial cortex. The medial prefrontal cortex does not appear to play a major role in solving guidance tasks (de Brun et al., 2001; Lacroix et al., 2002; Sloan et al., 2006), but the retrosplenial cortex does. Lesion studies in rats have shown that retrosplenial cortex contributes to solving the water maze task (Vann and Aggleton, 2004; Whishaw et al., 2001; Sutherland et al., 1988; Harker and Whishaw, 2002). Nelson et al. (2015) showed that rats with retrosplenial cortex lesions could not learn to navigate to a goal by being passively placed at the goal location, but could learn to actively navigate to it. This finding is consistent with the proposal that retrosplenial cortex is involved in switching between egocentric and allocentric viewpoints (Vann et al., 2009; Mitchell et al., 2018). Human fMRI studies have also implicated the retrosplenial cortex in the processing of stable landmarks (e.g. Auger et al., 2012). This is consistent with single cell activity in rodent studies. Czajkowski et al. (2014) showed increased reporter activity in the water maze when the position of the platform was fixed and determined by extra-maze cues, as opposed to when it was movable and marked by a metal rod. Another study revealed that landmarks served as dominant reference points anchoring the receptive fields of many spatially selective neurons in retrosplenial cortex (Fischer et al., 2020). Landmarks also exert strong control over a kind of head direction cell in retrosplenial cortex (Jacob et al., 2017a). Egocentric boundary vector cells found in retrosplenial cortex (Alexander et al., 2020) could also arguably facilitate guidance with respect to boundaries.

The fact that the hippocampus and related structures appear to be required for tasks like the Morris water maze, together with the fact that the hippocampus contains place cells, has been interpreted as evidence that these tasks involve the use of a cognitive map. However, using a cognitive map is not synonymous with metric or topological navigation. Even if place cells constituted some kind of cognitive map, it could still be used for guidance, e.g., the map could be used to retrieve the relationship between the goal and landmarks, which is then used in triangulation. Future research is required to fully dissociate the use of guidance and map navigation in these kinds of tasks, and to determine the exact role that place cells play in solving them. Experiments focusing on guidance should also change as many features of the environment as possible, except for the goal and surrounding landmarks, to discourage the use of map navigation.
4.6. Organization of navigation behaviors

To date, relatively few behavioral studies have examined how navigation behaviors are organized. A key reason for this is that it is already difficult to ascertain that subjects engage in a particular behavior, and studying the switching between or integration of different behaviors further complicates the issue. However, studies examining the degree to which subjects rely on certain types of information for navigation indirectly shed light on this question. By identifying each cue set with the navigation behavior that most likely makes use of it, we can infer which navigation behavior is active and whether subjects select or integrate behaviors.

One line of work has indicated that rodents can shift from using one set of cues to another within a single navigation task. For instance, rats in a Morris water maze initially navigated towards the platform using configurations of distal visual cues (guidance) and subsequently navigated directly to the visible platform (aiming) (Hamilton et al., 2004). Furthermore, the authors were able to identify the point at which the rats switched between behaviors, suggesting that animals select one behavior at a time. Another example of switching comes from gerbils. When they approached a goal using configurations of visible landmarks (guidance) and the lights were turned off, the gerbils usually continued their trajectory using self-motion cues (vector movement) and located the goal (Collett et al., 1986). Since this switch occurred only when the lights were turned off, this result agrees with other observations that rodents preferentially follow distal visual cues in the absence of cue conflicts, and resort to self-motion cues only when visual information is weak or unavailable (Teroni et al., 1987; Maaswinkel and Whishaw, 1999), and have an intermediate preference for olfactory cues (direction field navigation) (Maaswinkel and Whishaw, 1999). Together, these studies support the idea that rodents select navigation behaviors in a way that (i) reflects any biases that they may have and (ii) is flexible enough to deal with a variety of situations.

Other work has found evidence that navigators combine information from distinct cue systems to estimate the position of the goal, which in our framework would be consistent with integration at the level of extraction. For example, information about the geometry of the environment (related to guidance) seems to guide and constrain aiming towards a perceptible landmark at the goal location (Cheng, 1986). Other studies have focused on cue conflict paradigms, which is motivated by the observation that when agents navigate complex environments they are subject to different types of spatial information that have different reliabilities, i.e., they elicit responses with varying degrees of precision. Bayesian theory dictates that cue sets are weighted inversely to their relative reliability to obtain an optimal combination (Cheng et al., 2007; Chen et al., 2017). The benefit of this Bayesian integration is twofold: (i) in the case of disparity between cue sets, the mean response direction tends to be more accurate, and (ii) responses will be maximally precise (i.e., minimally variable). For small disparities, young adult humans have been observed to integrate self-motion cues (vector movement) and landmark (aiming) or geometric cues (guidance) in a near-optimal fashion (Nardin et al., 2008; Chen et al., 2017; Sjolund et al., 2018). By contrast, older adults displayed sub-optimal integration on a similar task (Bates and Wolbers, 2014) and children do not seem to integrate these cue sets at all, instead alternating between them (i.e., selection) (Nardin et al., 2008). However, conflicting results suggest the need for more research on these age related effects (Petri et al., 2016). A limitation of Bayesian cue integration is that for large disparities, condition (i) runs the risk of producing responses with low accuracy, and in such circumstances it may be more beneficial for an animal to rely on a single set of cues (Cheng et al., 2007; Zhao and Warren, 2015). Indeed, under such circumstances both rats and humans have been reported to follow self-motion cues (vector movement) and ignore other available cue sets (Etienne et al., 1999; Shettleworth and Sutton, 2005; Zhao and Warren, 2015; Sjolund et al., 2018). These results support the idea that navigators can flexibly engage in both selection and integration depending on internal, task and environmental parameters.

The neural correlates underlying the organization of navigation behaviors has not received much attention. However, partial insight on this topic may be obtained from cue conflict experiments investigating the relative dominance of visual and self-motion cues on the firing properties of place cells (Rotenberg and Muller, 1997; Knierim et al., 1998) and/or head direction cells (Taub and Burton, 1995; Knierim et al., 1998; Zugaro et al., 2000) while rats forage for randomly scattered food rewards (see Etienne, 2003, for a review). While there are some disagreements between studies, they all observed that both cell types were predominantly controlled by visual cues in the case of little to no conflict. Conversely, for more severe conflicts (i.e., those that are introduced suddenly rather than gradually, or those that have a large disparity in the directive information), the firing of both place and head direction cells are typically less predictable. The most common response in this case is a shift in control to self-motion cues (Rotenberg and Muller, 1997; Knierim et al., 1998), although in at least one study head direction cells were observed to fire maximally in a direction intermediate between those corresponding to each cue type (see Fig. 2D from Knierim et al., 1998).

The changes in neural responses observed in these studies likely reflect the process of reorienting or self-localizing within the environment based on different sets of cues, rather than the organization of navigation behaviors per se. Nonetheless, these studies might offer some hints about what kinds of brain activity changes could underlie the selection or integration of navigation behaviors that make use of different sets of cues. In particular, we could hypothesize that during selection, the set of cues relevant for the selected behavior takes control of the firing of spatially tuned cells. For instance, when switching from guidance to vector movement, the firing of place cells, grid cells, etc., could shift from reflecting a spatial code that is more strongly anchored to landmarks to one that is more dependent on path integration. Similarly, during integration, the activity of spatially tuned cells could reflect the combined influence of the cues guiding each of the integrated navigation behaviors. Alternatively, selection and integration could happen at the level of more specialized representations, for example, of the position of the current subgoal or of the required movement instructions. Further research could try to elucidate this question as well as the neural mechanisms that guide the organization process and determine when selection or integration occur.

4.7. Route navigation

Route navigation has often been characterized as consisting of a chain of habitual associations between stimuli and egocentric motor responses (Goodman, 2021; Nyberg et al., 2022; Packard and McGaugh, 1996; Hartley et al., 2003; Igloi et al., 2009; Allen, 1999). Research on this kind of behavior dates back at least to the study of “response learning” by Tolman and colleagues (Tolman et al., 1946, 1947), where they tested whether rats could learn to always make the same egocentric turn at the intersection point of a plus-shaped maze irrespective of the rats’ starting position. Rats learned to perform this task, although they did so with greater difficulty than rats trained to always approach the same point in the maze from different starting positions (“place learning”). This kind of navigation has also been evaluated in more complex mazes composed of multiple decision points. For instance, Rosenberg et al. (2021) trained mice to navigate in darkness in a maze designed as a symmetrical binary tree with 6 levels of branches. All nodes at a given level of the tree had the same local geometry, so the maze could only be solved by learning the correct sequence of left-right choices. While it is conceivable that subjects retrieve motor commands (e.g., “at the intersection, turn left”), their behavior is also compatible with the interpretation that they retrieve subgoals (e.g., “at the intersection, go to the end of the corridor on the left”) and navigate there using navigation behaviors (e.g., path following), in line with our characterization of route navigation.
Some experimental observations lend support to our hypothesis that route navigation involves the retrieval of subgoals (Schinazi and Epstein, 2010). Participants were taught to navigate along a circuitous route around a university campus. They were then shown pictures of buildings and asked to indicate whether the buildings could be found along the learned route. For test buildings at decision points along the route, participants reacted faster when primed with pictures of the buildings immediately preceding the test buildings, as opposed to pictures of buildings immediately succeeding them. This suggests that experiencing a subgoal triggered the retrieval of the next one. Further research is required to confirm this hypothesis and assess whether there really is the kind of behavioral flexibility in reaching each of the subgoals along the route that we advocate for.

Experiments in route following have also focused on determining the impact of cues at or around decision points. Not surprisingly, humans are better at navigating mazes if landmarks identify each decision point uniquely (Lingwood et al., 2015). Typically, visual landmarks are used, but auditory (Hamburger and Röser, 2014) and odour (Hamburger and Krauff, 2019) landmarks have also been shown to enable this kind of navigation.

Once a decision point has been identified, a key question is in which reference frame the next subgoal is specified. While the traditional view of response learning would suggest that the next subgoal is specified in egocentric coordinates, our taxonomy also allows for the possibility that it is specified in terms of allocentric information, e.g., at an intersection, “take the street between the bakery and the pharmacy” or “take the street leading north”. The experimental evidence supports the view that subjects can use both of these reference frames. In a virtual reality experiment, human participants navigating a city had to choose whether to go right or left based on configurations of three buildings surrounding the intersections (Mallot and Gillner, 2000). Each of the buildings appeared to vote for an egocentric direction to the next subgoal independently of the others, because it did not matter that they were rearranged within the same intersection, or swapped across different intersections as long as they were all associated to the same egocentric direction. In the experiments by Wiener et al. (2013) and de Condappa and Wiener (2016), however, young adult participants became increasingly better at choosing the correct corridor defined relative to the spatial configuration of landmarks at the intersection. Waller and Lippa (2007) also showed that routes were learned more easily when distinct landmarks placed next to each of the available paths could be used to choose the correct one, as opposed to having to associate the correct path defined in egocentric coordinates to a single landmark placed between the paths. Finally, Steck and Mallot (2000) showed that participants could also use distant landmarks (e.g., a faraway hilltop) to determine the allocentric direction to the next subgoal.

Beyond learning routes composed of a sequence of path following, Loomis et al. (1993) showed that humans can learn routes composed of a series of vector movements, although not very accurately.

Our architecture of navigation processes allows for different navigation strategies to be nested within each other (Fig. 1). An instance of this might have been observed by Shamas et al. (2021). Mice learned to navigate from a threat zone back to shelter in the presence of an obstacle. The mice memorized the edges of the obstacle as allocentric subgoals during practice runs. Then, they often went through these subgoals, even in darkness or long after the obstacle had been removed. This suggests that the mice were engaging in route navigation, but possibly using metric navigation in a nested fashion to arrive at the obstacle’s edge.

The neural bases of route navigation have often been studied in the context of response learning. Numerous studies in rodents point to a role of the dorsal striatum, and not of the hippocampus, in this kind of route navigation (e.g., Packard and McGaugh, 1996, and see Goodman, 2021 for a review). This finding is also consistent with several human fMRI experiments. Participants showed increased activation in the caudate nucleus when following previously learned routes (Hartley et al., 2003; Voermans et al., 2004), and individuals who tended to stick to learned routes (as opposed to taking shortcuts) showed more activation in the caudate nucleus and less in the hippocampus (Marchette et al., 2011).

Some electrophysiological studies in rodents have looked at the responses of neurons in the striatum, with mixed findings. Some report the presence of cells with spatial tuning similar to that in the hippocampus (Schmitzer-Torbert and Redish, 2004; Yeshenko et al., 2004). Others find cells tuned to certain sections of a route regardless of spatial position, e.g., a cell that always fires as the rat leaves the initial arm of the maze, regardless of which is the initial arm (Mulder et al., 2004; Berke et al., 2009). Yet others observe a “bracketing” pattern where cells are most active at the beginning and end of the route (Smith and Graybiel, 2013; Regier et al., 2015).

However, the classical association of route-following behavior with the striatum is now being contested (Goodroe et al., 2018). For instance, a study has shown that when rats learn a response task at the same time as a place task, both tasks depend on both the hippocampus and the striatum (Ferbinteanu, 2016). The hippocampus also seems to be involved when routes are more complex than the Y- or plus-mazes. For instance, mice with hippocampal lesions or a lack of CA1 NMDA receptors were impaired at a route navigation task involving three turns (Fouquet et al., 2013; Rondi-Reig, 2006). In a similar experimental paradigm with human subjects, the left hippocampus showed increased activation when using a route navigation strategy (Iglói et al., 2010).

Patients with temporal lobe epilepsy have also been shown to be impaired at route following, where the number of errors they made is inversely correlated with their left hippocampal volume (Bell, 2012). Other human imaging studies have also identified a role for the hippocampus in route following (Ghaem et al., 1997), including an increase in functional connectivity between the hippocampus, caudate nucleus and orbitofrontal cortex when different routes contained overlapping segments (Brown et al., 2012). Finally, a meta study of the effect of brain lesions has shown impaired route learning in patients with damage to the right medial temporal lobe (including the hippocampus) and the left parahippocampal gyrus, as well as other regions such as the medial occipitotemporal and right ventral inferotemporal cortices (Barrash et al., 2000). These findings fit well with our broader notion of route navigation which encompasses all navigation behaviors and therefore should naturally involve multiple brain regions. The involvement of the hippocampus, in particular, could be due to its role in learning complex sequences of subgoals, or due to its involvement in navigation behaviors such as vector movement or guidance.

The behavioral and neural evidence reviewed is thus more consistent with our richer notion of route navigation than with one based on habitual egocentric responses. However, this richer characterization of route navigation makes it harder to dissociate from map navigation, especially in well known environments where both strategies may be available. In some cases, route navigation might be identified on the basis of fast and stereotypical behavior that is the result of extensive training. However, this does not need to be the case because, first, route navigation could make flexible use of multiple navigation behaviors and, second, it does not need to be habitual, as when following directions someone has given you, or navigating a new route for the second or third time. A clear signature that an animal is engaging in route navigation would be getting stuck if a subgoal is missed or one of the paths connecting two subgoals is blocked. Similarly, subjects using route navigation should not take shortcuts that suddenly become available. However, being able to find shortcuts and alternative routes would not rule out that the animal had been using route navigation before switching to another strategy.

The flexibility granted by the possibility of dynamically switching to a more powerful strategy thus makes it challenging to determine whether subjects were using route navigation before switching to the more powerful strategy, or whether they were using the more powerful strategy all along. Still, one could imagine ways to dissociate between these two cases using behavioral data. For example, route navigation
4.8. Map navigation: topological and metric

Mammals often need to plan new routes in response to previous ones being blocked, or to travel between new pairs of starting and goal positions. Thus, there is a general consensus that route navigation is not sufficient to account for the navigational abilities of mammals, and that some forms of navigation rely on cognitive maps to provide more flexibility. A great deal of research has gone into probing the nature of such maps, with researchers often wondering whether the maps and the navigation strategies that make use of them are best described as topological or metric. The main diagnostic feature that can dissociate these two options is that metric, but not topological, navigation allows for the use of efficient shortcuts over unexplored terrain and in the absence of direct perception of cues at or around the goal that could support aiming or guidance. Here we will review behavioral and neural studies examining this question and supporting the view that both topological and metric navigation are likely to coexist in parallel, each operating at a different spatial scale.

4.8.1. Behavioral studies on map navigation

Some behavioral evidence for topological maps comes from Alvernhe et al. (2012). They compared rats’ exploratory behavior in a multi-compartment environment after opening or closing doors that did or did not alter the ability of the animals to run between the different compartments. The animals reacted differently to these two types of manipulation, suggesting that they had built a topological map of the connectivity of the space. To probe whether rats could also make use of metric maps, Grieves and Dudchenko (2013) trained rats on a stereotypical route and then tested their ability to take novel shortcuts in mazes composed of 3 or 4 boxes connected by corridors. Rats failed to do this, unless they had been familiarized with the shortcut prior to being trained on the route. This pattern is precisely what one would expect if the animals could use topological, but not metric, navigation to take the shortcut.

Whether cognitive maps are best described as topological or metric has also received a great deal of attention in the primate ecology literature. Findings have been mixed but are more consistent with topological than metric maps (see Trapanese et al., 2019, for a review). For instance, animals spend most of the time traveling through a complex network of well-demarcated routes, specifically woolly and spider monkeys spend 78% and 95% of their time in locations falling within 50 m of their route networks, which were stable over years (Di Fiore and Suarez, 2007). Bearded capuchin monkeys (Presotto et al., 2018) and chacma baboons (de Raad and Hill, 2019) also navigate route networks and frequently change directions at route intersections, thus flexibly combining different route segments. Furthermore, the initial travel direction of the baboons was significantly different from the direction to the next goal, suggesting that the routes were not the result of metric navigation. Accidental encounters with other groups can also be used as a natural experiment (Nosé and Byrne, 2007). When a group of baboons is traveling to some resource and detects the presence of another group blocking their way, they try to avoid a direct encounter. Baboons perform large detours around the other group only when prominent visual landmarks enabling aiming or guidance are visible, otherwise they wait for extended periods of time or change plans and follow familiar routes to alternative destinations. It thus appears that in the absence of distal landmarks, the animals would be lost if they tried to take a large-scale detour, which suggests that they do not possess a metric map of the environment.

By contrast, the behavior of several other species in the wild is suggestive of the use of shortcuts in metric navigation. Chimpanzees travel in straight lines between food sources, and return to them from many different directions (Normand and Boesch, 2009). Similarly, Egyptian fruit bats forage in goal-directed, long and straight flights that included potential shortcuts, and arrive at goals from multiple initial positions (Toledo et al., 2020). Humpback whales’ seasonal migrations span more than 6500 km of open ocean, during which they maintained their course with extreme precision over segments of more than 200 km spanning several days, despite changing ocean floor depths, earth magnetic profiles, positions of the sun or sea surface currents (Horton et al., 2011). Although much more work is needed to clarify the strategies employed by these animals, the findings are suggestive of large-scale cognitive maps with metric qualities.

In humans, the knowledge of metric information can be probed directly. For instance, human hunter-gatherers in rainforests are very accurate at pointing to out-of-sight targets (median error of 6°) (Jang et al., 2019). By contrast, the metric knowledge of city dwellers seems rather poor. For instance, roads with several major bends were estimated to be longer than equivalent linear roads; and road junctions at angles between 60° and 120° were all remembered as 90° (Byrne, 1979). This bias towards 90° junctions consistently leads to the remembered inner angles between three roads intersecting in a triangle to add up to more than 180° (Moar and Bower, 1983).

Metric and topological inconsistencies in the spatial representations of humans have also been revealed through clever experiments in virtual environments that violate principles of Euclidean geometry (Zetzsche et al., 2009; Kluss et al., 2015; Warren et al., 2017; Murry and Glennerster, 2018). These violations include wormholes or jumps between different parts of the environment. A common result of many of these experiments is that participants could navigate the impossible environments as efficiently as physically possible ones, and in general could not tell whether an environment was physically possible, even when forced to decide (Zetzsche et al., 2009). Furthermore, participants seem to acquire locally accurate but globally inconsistent cognitive maps of the impossible environments. This is demonstrated by their ability to navigate the environments without visual feedback (Kluss et al., 2015). It is also demonstrated by the way in which participants attempt to take shortcuts in the impossible environments. Warren et al. (2017) trained participants to navigate a hedge maze with wormholes (Warren et al., 2017). The maze was then removed and participants were asked to walk directly to a goal location. If the shortest route would have involved a wormhole, participants went to the position where the goal would have been as perceived through the wormhole, as opposed to its actual position in the maze. The wormholes also introduced rips and folds in the spatial representation of the maze that could not be accommodated by a continuous metric map. Interestingly, it seems that even topological properties can be violated. In a virtual reality experiment, participants could navigate a 2D environment composed of a circular path and a meandering path that left the circular path towards the inner region and reunited with it from the outer region, without ever crossing the circular path (Zetzsche et al., 2009). Participants did not notice this kind of inconsistency either.

Another question that has been studied in the human literature is how subjects represent and navigate vista as compared to multi-compartment environments. Subjects could quickly create an internalized map of the environment directly in front of them and then walk with their eyes closed avoiding obstacles for up to 9 m (Thomson, 1980). However, participants did not integrate spatial knowledge of a room and the campus in which it was located (Wang and Brockmole, 2003). Instead, participants dynamically switched representations as they moved in and out of the room, reorienting to the environment they were approaching. The fragmentation in the representation of multi-compartment environments is also revealed by the way in which
subjects appear to construct survey estimates of these environments on the fly (Meilinger et al., 2018). After participants learned a multi-corridor route, the latency of pointing towards certain targets increased with the distance to them along the route, and pointing error distributions were sometimes bimodal, consistent with participants mixing up turns or forgetting segments of the route. These kinds of effects appear to be specifically caused by the presence of visual boundaries (Meilinger et al., 2016). Similarly, humans have been shown to keep independent representations at multiple spatial scales (Marchette et al., 2017). Participants learned a virtual environment composed of four rooms containing objects. Later, when navigating to the position where one of the objects had been, participants often went to the correct position within the wrong room. Again, walls were necessary for this effect.

Finally, caution is in order when generalizing specific findings across individuals, as some acquire metric knowledge of new complex environments quickly while others never do (Ishikawa and Montello, 2006; Newcombe, 2018). Furthermore, in a recent study of humans navigating in virtual environments, it was observed that the abilities of test subjects depended on whether they had grown up in cities and, for those who had, how regular the cities’ street topology was (Coutrot et al., 2022). This emphasizes how difficult it is to generalize across test subjects of one species and supports the need for more comparative analyses of the navigation abilities of laboratory-grown and wild animals.

Beyond assessing the topological or metric nature of map navigation, some research has been directed at determining how mammals decide which route to take based on a map — be it topological or metric — of their environment. For instance, when the target was hidden from view and cats had the choice between minimizing route length or initial angular deviation to the target, cats took the most appropriate route based first on length, and then on deviation. By contrast, humans walking in cities preferentially chose paths that pointed more directly in the direction of the goal, even if that led to longer trajectories (Bongiorno et al., 2021). Wiener and Mallot (2003); Wiener et al. (2008) analyzed how human participants select routes in environments composed of separate identifiable regions. They found that subjects minimized the number of regions traversed, and that when alternative routes were equally long, subjects preferred to enter the regions containing the targets sooner rather than later. The authors interpret these findings as consistent with a hierarchical planning heuristic, according to which subjects first plan a route that takes them to the target region, and only then update their route taking into account more detailed spatial information about that region. More work of this kind would be needed to elucidate the mechanisms and planning heuristics used in topological or metric navigation.

We can extract some general conclusions from the behavioral experiments reviewed above. The evidence for metric navigation is mixed, however, most of it is consistent with the view that mammals can engage in metric navigation over vista space (e.g., rooms, valleys) but not over environmental space (e.g., cities, forests). Metric navigation in vista space has rarely been studied explicitly, perhaps because it might appear trivial. However, calculating efficient routes around obstacles in vista space (e.g., navigating around cars in a car park) is a non-trivial task and could amount to a large fraction of the navigation decisions animals and humans face. Topological navigation would then account for a large fraction of mammalian navigation over environmental space. Although perhaps the graph supporting it is best characterized as a labeled graph incorporating approximate information about edge lengths and angles between edges (Chraitil and Warren, 2014; Warren et al., 2017; Warren, 2019).

4.8.2. Neural studies on map navigation

The neural basis of map navigation has received a great deal of attention. Various lines of evidence point to a prominent role for the hippocampus and related structures in the storage and use of spatial representations that enable agents to flexibly plan complex routes from a starting position to a goal via a series of subgoals. In a human fMRI experiment, the hippocampus was shown to encode the positions of goals and subgoals during spatial planning (Brown et al., 2016). Spatial planning in the hippocampus and related structures could involve cells representing navigational goals, which have been observed in bats (Sarel et al., 2017) and humans (Ekstrom et al., 2003; Tissiökis et al., 2020). Place cells are also known to represent positions other than the animal’s current location, which could help sample and evaluate different possibilities during the planning process (Zielinski et al., 2020). One way in which place cells do this is based on theta phase coding. Within each cycle of the theta oscillation, place cells represent trajectories reaching ahead of the animal (Johnson and Redish, 2007; Gupta et al., 2012; Parra-Barrero et al., 2021). These so-called theta sequences are modulated by the distance to the goal (Wikenheiser and Redish, 2015) and can cycle constantly between the representation of different possibilities (Kay et al., 2020; Johnson and Redish, 2007). Place cells also rehearse possible trajectories through the environment in a phenomenon known as replay, which occurs during sharp wave ripple events (Foster, 2017; Bulry et al., 2011). The occurrence of replay is predictive of correct future navigation (Singer et al., 2013). However, there is no straightforward association between the positions being replayed and subsequent behavior, with some studies finding a positive correlation (Pfeiffer and Foster, 2013) and others finding negative correlations (Wu et al., 2017; Carey et al., 2019) (as if replay could remind the animal where not to go) or no correlations at all (Gillespie et al., 2021). Trajectories during theta and sharp wave ripples can both occur in coordination between the hippocampus and medial prefrontal cortex (Zielinski et al., 2019; Shin et al., 2019; Jones and Wilson, 2005). This is in line with the proposed roles of prefrontal cortex in goal coding planning and decision making (see Patai and Spiers, 2021, for a review).

While the studies above do not clearly point to either topological or metric navigation, there are some properties of the two that may help shed light on the extent of the role that the hippocampus and place cells play in each of them. For instance, the computation of metric distances and directions is a distinctive feature of metric navigation. Thus, if the hippocampus is the seat of a cognitive map used for metric navigation, it should contain metric representations. Human fMRI studies suggest that this is the case as hippocampal activation correlates with the distance along a route to a goal (Howard et al., 2014; Viard et al., 2011; Patai et al., 2019) or with the distance between successively presented landmarks (Morgan et al., 2011). This could be the basis for taking shortcuts, another signature of metric navigation. Indeed, increased hippocampal activation has been shown to correlate with finding novel shortcuts through complex mazes both within (Iglói et al., 2010) and across (Hartley et al., 2003; Marchette et al., 2011) subjects.

These metric properties of hippocampal representations are, however, difficult to relate to known properties of hippocampal place cells, whose firing does not reveal any clearly spatial metric properties. For instance, the density with which place fields tile a certain region of space, as well as the size of the place fields, vary across the environment as a function of rate of visual change, typical running speed, or the presence of rewards in some conditions (Tanni et al., 2022; Parra-Barrero et al., 2021; Hollup et al., 2001; Dupret et al., 2010; Zaremba et al., 2017).

In contrast to place cells, the regular firing patterns of medial entorhinal grid cells observed in initial recordings made them an obvious candidate for the neural substrate of a metric map (Hutting et al., 2005; Moser et al., 2008). Subsequent studies have revealed that grid patterns are not so regular after all, with distortions occurring in trapezoid-shaped environments (Krupic et al., 2015) or near borders (Hägglund et al., 2019), or with grid fields being attracted to goals (Boccara et al., 2019). These distortions, however, are consistent with the idea introduced above that metric navigation operates mostly within a limited range. Broadly speaking, there are two views for how grid cells could provide a spatial metric. According to the first, the overlay of grid patterns of different spatial scales and phase offsets leads to unique
combinations of grid cell activity for each position within a very large area (Gorchetchnikov and Grossberg, 2007; Fiete et al., 2008; Sreenivasan and Fiete, 2011). Models based on this idea would allow grid cells to support metric navigation over long distances (Bush et al., 2015). However, these models require learning very large numbers of combinations of grid cell activity, especially if grid patterns are distorted. That would go against the purported benefit of grid cells as providing a universal metric code that can be instantaneously deployed wherever the animal goes. The second view proposes that grid cell modules of different scales operate in a nested fashion, like the hours, minutes and seconds hands of a clock (Stemmler et al., 2015; Wei et al., 2015). The scale of the largest grid module defines the range of possible navigation and provides a rough estimate of the animal’s position, and smaller modules add increasingly more precision. These models are thus consistent with grid cells being used for navigating over moderate distances, for which grid distortions are rendered relatively harmless (Edwardsen, 2018).

The distortions in the grid patterns also offer a rare opportunity to establish a connection between grid cells and metric navigation. The distortion of grid patterns in trapezoids seems to underlie a distortion of spatial memory of trapezoid environments in humans, as evidenced by the pattern of errors made when subjects navigate to remembered positions or attempt distances between pairs of the same (Bellmund et al., 2020). Similarly, expanding or compressing environments leads to an apparent expansion or compression of grid patterns (Barry et al., 2007) and corresponding errors in distance estimation when human subjects navigate in such environments (Chen et al., 2015). The apparent expansion or compression of the grid pattern in deformed environments could result from time averaging a grid pattern that shifts dynamically, anchored to the last boundary encountered (Keinath et al., 2018). These history-dependent shifts in grid phase are, again, accompanied by corresponding shifts in the remembered position of targets in the environment (Keinath et al., 2021).

If grid cells were to provide a coherent metric map of complex environments, grid patterns should span across the differentiated regions that constitute many environments. In an experiment by Carpenter et al. (2015), rats foraged in two identical compartments, and grid patterns replicated between them. When rats were then allowed to run between the compartments through a connecting corridor, a single coherent grid pattern emerged spanning both compartments. In a similar experiment in a larger environment, however, only the grid pattern in the transition zone reorganized to maintain approximate local periodicity, whereas the original grid patterns further away from the transition zone were largely preserved (Wernle et al., 2018). On the other end of the spectrum, grid patterns were completely fragmented in a hairpin maze composed of many repeating corridors (Derdikman et al., 2009). More experiments would be required to understand under which conditions grid patterns become fragmented.

Like place cells, grid cells are also known to phase precess with respect to theta oscillations (Reifenstein et al., 2012; Hafting et al., 2008) and to generate replay sequences (O’Neill et al., 2017). Thus, ideas about the potential involvement of theta sequences and replay in the planning process could extend to grid cells, with the added benefit that grid cells could provide a metric component.

It is also worth noting that distortions in the grid pattern and differences in firing rates across different grid fields (Ismakov et al., 2017; Diehl et al., 2017) could serve as a substrate for richer cognitive maps that incorporate environmental information beyond what would be required for a simple metric of space.

In addition to grid cells, head direction cells could also play an important role in guiding metric navigation. Like grid cells, if they are to form part of a coherent metric map, their directional selectivity should be preserved across different regions of the environment. This is indeed the case (Yoder et al., 2011; Carpenter et al., 2015), even in environments where grid patterns become fragmented (Whitlock and Derdikman, 2012).

Much like the computation of metric quantities is a distinguishing feature of metric navigation, topological navigation may be distinguished by the use of computations that leverage the graph structure of the topological map. While participants navigated through a simulation of the city of London, BOLD activity in the hippocampus and prefrontal cortex were correlated with graph theoretical measures (Javadi et al., 2017). More specifically, activity in the right posterior hippocampus was correlated with a measure of local topological connectivity, and activity in the prefrontal cortex scaled with the demands of planning using a breadth-first search (an algorithm for planning on graphs). In another fMRI study where participants were tasked with navigating in a virtual subway network, hierarchical planning was associated with increased activity in the dorsomedial prefrontal cortex and premotor cortex (Balaguer et al., 2016). The hippocampal activity was correlated with the distance to the goal, but it did not contain any evidence of a hierarchical signal. This again suggests that the hippocampus is involved in monitoring local topological connections, while the prefrontal cortex is involved in planning trajectories.

It has been argued that a topological representation of space relies on place cells (Dabaghian et al., 2014; Dabaghian, 2019). Indeed, the ensemble spiking activity of place cells has been shown to code for a topological graph that defines the connectivity of locations within the environment (Chen et al., 2014). Another indirect line of evidence in support of this theory comes from experiments showing that place cell deformations brought about by gradual environmental changes preserve topological properties. For instance, when a linear track was shortened compared to the training runs, the place cell population vector translated smoothly through intermediate locations to adjust to the new track length (Gothard et al., 1996). Similar evidence in 2D comes from gradually morphing a square environment into a circle (Leutgeb et al., 2005a; Wills et al., 2005), and by rotating cue cards together or apart in a cylindrical environment (Fenton et al., 2000a,b; Tourretzky et al., 2005). There is also evidence that CA1 and CA3 place cell firing changes in response to changes in the environment’s topology — place cells remap near newly introduced shortcuts and barriers (Alvernhe et al., 2008, 2011). However, a more recent study did not find this effect (Duvelle et al., 2021). Another line of evidence is based on properties of hippocampal replay. Replay in rats appears to capture the topological structure of a Y-maze (Wu and Foster, 2014), flexibly go around new barriers in an environment and predict future behavior (Widloski and Foster, 2022), much like topological path planning.

The evidence reviewed so far points to the hippocampus, working together with related structures such as the medial entorhinal and medial prefrontal cortex as playing a key role in topological and metric navigation. There are, however, some studies that cast doubt on the direct involvement of the hippocampus and related structures in navigation as such, suggesting instead that the hippocampus is only involved in the creation and temporary storage of a cognitive map. Human patients (Teng and Squire, 1999; Rosenbaum et al., 2000; Maguire et al., 2006) and rats (Winocur et al., 2005, 2010) with hippocampal lesions cannot navigate in novel environments but are generally able to navigate environments they learned long before their lesions occurred. Their navigation shows signatures of metric navigation, since they can generally execute efficient routes between arbitrary pairs of points, and humans can successfully estimate distances and directions between pairs of landmarks. Although subjects are impaired at navigating particularly complex routes (Maguire et al., 2006) or dealing with obstacles (Winocur et al., 2010), overall, these findings suggest that flexible navigation can largely be accomplished on the basis of spatial representations external to the hippocampal system. This fits well with the recent discovery of spatially tuned cells throughout the brain in areas such as the claustrum (Jankowski and O’Mara, 2015) and in prefrontal (Zielinska et al., 2019; Wikenheiser et al., 2021), anterior cingulate (Bota et al., 2021), retrosplenial (Mao et al., 2017), piriform (Poo et al., 2022), primary somatosensory (Long and Zhang, 2021), or primary visual cortex (Saleem et al., 2018). In fact, one study reported spatially selective cells in all neocortical areas examined (Esteves et al., 2021). These...
cells partly depended on an intact hippocampus to develop their spatial selectivity, but could sustain it in familiar environments in the absence of a hippocampus (Esteves et al., 2023). Furthermore, fMRI studies in humans have also found signatures of grid cells in a range of areas outside the hippocampal formation including posterior and medial parietal, cingulate and medial prefrontal areas (Doeller et al., 2010; Jacobs et al., 2013). These findings suggest that cognitive maps are complex and distributed, with different cortical areas contributing content of different modalities, e.g., visual, somatosensory, olfactory, etc. This also implies, however, that the neural underpinnings of topological and metric navigation could extend well beyond those classically considered.

4.8.3. An integrative view of topological and metric navigation

The evidence reviewed in these sections points to the existence of both topological graphs and local metric maps that are distorted, fragmented and globally incoherent, each fulfilling different roles. We summarize this view in Fig. 5. These ideas are not entirely new. Several previous proposals for the structure of cognitive maps have already combined topological graphs with local metric maps, although they have not considered that the metric maps could be deformed (Mellinger, 2008; Pouzet, 1993; Kuijpers, 2000). In these previous models, local reference frames corresponding to different vista spaces become embedded into a global topological graph by learning the transitions between pairs of them. This idea has even been applied in robotics to solve the SLAM (simultaneous localization and mapping) problem (Blanco et al., 2008; Konolige et al., 2011). One of these approaches also uses the hybrid map for navigation, first planning an overall trajectory based on the topological graph, and then navigating locally using overlapping local metric maps (Konolige et al., 2011). We believe that this is similar to what mammals actually do, as captured in the model developed by Edvardsen et al. (2019), which combines topological navigation using place cells to find key subgoals, with local metric navigation using grid cells and border cells.

4.9. Organization of navigation strategies

It might be argued that response strategies, as discussed above, are examples of what we call route navigation, whereas place strategies could correspond to map navigation. The typical paradigm for studying these two strategies are so-called dual-solution tasks, in which animals are trained to perform a spatial task that can be completed using either type of strategy. In the most common type of task, test subjects start from a fixed arm in a maze (mostly a T-maze or plus maze) and must navigate past a single intersection to the end of another arm in order to receive a reward. In probe trials, subjects are placed in a different arm of the maze to observe whether they (i) approach the previously rewarded position in space (indicative of a place strategy), or (ii) perform the same egocentric turn as before, arriving at a different position (indicative of a response strategy). Several studies using T- or plus mazes have shown that rodents predominantly use a place strategy in early stages of training, while probe trials carried out at later stages indicate a preference for a response strategy (Ritchie et al., 1950; Hicks, 1964; Packard and McGaugh, 1996; Packard, 1999; Chang and Gold, 2003). Moreover, the same phenomenon has been observed in humans using a multiple T-maze (Schmitzer-Torbert, 2007). These results are often interpreted as a type of organization by selection: as an environment or task becomes more familiar, subjects transition from a place to a response strategy.

However, linking place/response strategies in T or plus mazes to what we call navigation strategies is challenging since these environments provide a poor dissociation of route and map navigation. On the one hand, place and response strategies could both be consistent with route navigation where animals navigate first to a subgoal at the intersection and then to the final goal. If the final goal is defined in allocentric terms (e.g., the end of the corridor to the west, or the corridor under the lamp) then the animal would be considered to use a place strategy. If the final goal is defined in egocentric terms (e.g., the end of the corridor to the left), then the animal would be considered to use a response strategy. On the other hand, place and response strategies could both be consistent with map navigation, too. This is obvious for the place strategy; for the response strategy an explanation is required. If the map of the environment were always aligned with the starting position, then in probe trials animals would make the same turn not because they have encoded a particular egocentric response, but because they are following a rotated map. Alternatively, subjects could have a correctly aligned map and use it for navigation, but define the goal location relative to the starting position.

The link between place and response strategies, and map and route navigation, respectively, may be more clear in more complex environments that involve a larger number of sensory cues and decision points, such as in the starmaze (Rondi-Reig, 2006; Iglö et al., 2009). Interestingly, results in these environments conflict with those in simpler mazes in two ways: response and place strategies were both acquired at similar speeds, and bidirectional transitions between them occurred. This deviation might arise because the larger number of decision points makes it more difficult for subjects to learn and consistently execute the same navigation strategy.

Besides the amount of training, studies indicate that other experimental factors impact the interaction of place and response strategies. For instance, the use of response strategies is enhanced by stress, by longer inter-trial intervals during training, and by more homogeneous visual surroundings (see Packard and Goodman, 2013, for a review). Furthermore, older human subjects show a deficit relative to younger subjects in their ability to transition from route to map navigation (Harris et al., 2012; Harris and Wolbers, 2014). In both studies, this deficit was measured by the inability to make use of new shortcuts in a known environment that became available in test trials. The authors hypothesize that this deficit could underlie various age-related navigation impairments observed in humans. Lastly, some aspects of the interaction between route and map navigation seem to depend on whether tasks are based on positive or negative reinforcement. While all the studies cited thus far used rewards, other experiments have utilized escape tasks involving water submerged mazes and have either observed no training-related transition (Kealy et al., 2008) or a transition in the opposite direction (i.e., from response to place strategy) (Asem and Holland, 2013; Gasser et al., 2020). This might be linked to a shift towards more habitual, i.e. egocentric, responses when animals are exposed to stress (Schwabe and Wolf, 2013).

The neural correlates underlying the interaction of place and
response strategies have also received much attention, particularly the role of the hippocampus and dorsal striatum (Goodman, 2021). One idea emerging from this research — sometimes called the multiple memory systems theory (Ferbinteanu, 2019) — is that, together with the amygdala, they constitute three independent memory systems (McDonald and White, 1993; White and McDonald, 2002; Poldrack and Packard, 2003; Goodman, 2021). This theory posits that each of these neural structures is sensitive to different forms of information and has distinctive capabilities for processing and storing that information, with the hippocampus and dorsal striatum being sensitive to relationships between stimuli (S-S) and stimulus-response associations (S-R), respectively. An additional claim is that these systems can interact both competitively (leading to distinct behaviors) and cooperatively (leading to similar behaviors). In the case of spatial tasks involving the hippocampus and dorsal striatum, competitive interactions have been observed in a number of experiments (White and McDonald, 2002; Poldrack and Packard, 2003; McDonald et al., 2004; Hartley and Burgess, 2005), and at least one study observed a type of cooperation (McDonald and White, 1995). Given the strong relationships between these structures and place/response strategies, these results could be interpreted as examples of organization of navigation strategies. While this is plausible, we take issue with this interpretation since many of the experiments used to dissociate the functions of these brain regions are typically performed in the eight-arm radial maze, mostly with the win-shift, win-stay, and conditioned cue preference tasks (Packard et al., 1989; McDonald and White, 1993; White and McDonald, 2002). Relating such tasks to the navigation strategies of our taxonomy is problematic since they either require working memory, are too simple, or do not pick out specific strategies due to the absence of probe trials.

However, neural recordings of these brain regions have also been made while animals and humans perform in dual-solution tasks, and results from these studies support the idea that transitions between place and response strategies are accompanied by corresponding changes in activity in the hippocampal-striatal system. In the cross-maze task, selection of a place (response) strategy can be induced by pharmacological inactivation of the dorsal striatum (hippocampus), respectively (Packard and McEachern, 1996). In the same maze, when animals were trained and then required to alternate between both strategies, animals made errors primarily to arms associated with the other strategy, revealing a form of integration. The interpretation of the results from these experiments remains unclear, however, as some studies have suggested patterned activity in the hippocampus that may differ between strategies (Hartley and Burgess, 2005). This posits that, together with the amygdala and stress hormone systems, it is the hippocampus that stores memories in their entirety, only indices or indices of indices or other systems. Regardless, test subjects more often chose the third arm in the cross-maze. As explained above, spontaneous bidirectional shifts in strategy use have been observed behaviorally in starmaze tasks. An fMRI study using the same paradigm then revealed that increased activity in the right (left) hippocampus at the beginning of each trial was predictive of whether the participants would use a place (response) strategy during the trial (Iglói et al., 2010). Since this trend persisted throughout the experiment, despite the strategies alternating, it indicates that a behavioral shift in either direction is accompanied by a shift of activity between the two hippocampi. These results indicate that the organization of place and response strategies depends on shifts of activity within the hippocampal-striatal system, however it remains unclear what this implies about the organization of map and route navigation for the reasons pointed to above. Because the hippocampal-striatal system, and other neural systems have also been observed to modulate the interaction of place and response strategies in dual-solution tasks. Firstly, several studies support the idea that activity in the amygdala and stress hormone systems enhances (impairs) the ability of animals to navigate using response (place) strategies (Goodman et al., 2017; Packard et al., 2018), consistent with the effect of emotional stress on these two strategies (Packard and Goodman, 2013). Secondly, a number of neurotransmitters have been shown to affect the relative use of the two strategies (see Goodman, 2021, for a review).

## 5. Two cross-cutting issues

### 5.1. The role of hippocampus in spatial navigation remains unclear

As reflected in Sections 3 and 4, research on the neural basis of spatial navigation has focused predominantly on the hippocampus. However, the hippocampus has also received a lot of attention from researchers of memory, especially episodic memory (memory of particular experienced events) (Tulving, 1972, 2002; Dicker and Eichenbaum, 2010), and efforts to reconcile these apparently disparate functions have led to a wide array of proposals for the overall role of the hippocampus.

The various proposals generally agree in considering that the hippocampus does not store memories in their entirety, only indices or pointers to content residing throughout the cortex or elsewhere in the brain (Teyler and DiScenna, 1986). However, they give different answers to these two questions: What type of information can the hippocampus encode and store? And for how long?

A prominent hypothesis within the spatial navigation community is that the hippocampus holds a cognitive map composed of places and their spatial relationships as well as what can be found where (“feature-in-place”) (O’Keefe and Nadel, 1978; O’Keefe and Krupic, 2021). Others claim that the hippocampus plays a more general role in relational memory and that it is important for navigation only to the extent that navigation processes require memory (Eichenbaum, 2017). This is supported by findings that the hippocampus is required for non-spatial tasks such as transitive inference (Dusek and Eichenbaum, 1997; Barron et al., 2020) or remembering the order of a sequence of odours (Fortin et al., 2002) or images (Konkel et al., 2008). Hippocampal cells have also been shown to map other dimensions such as time (Kraus et al., 2013) or auditory frequency (Aronov et al., 2017) and to reflect the structure of the task as opposed to simply space (Sun et al., 2020; Nieh et al., 2021; Smith and Mizumori, 2006; Markus et al., 1995; Wood et al., 2000; Grieves et al., 2016; Sarel et al., 2022). (Nevertheless,
spatial memories) are eventually copied over to the cortex and become structures common to multiple episodes. In a more radical take, systems required to settle the question of what the exact function of the hippocampus in spatial navigation is. More parsimonious account of the role of the hippocampus in spatial navigation is by the involvement of numerous brain areas in navigation tasks revealed by human imaging studies (Hartley et al., 2003; Doeller et al., 2010; Iglói et al., 2010; Marchette et al., 2011; Maguire et al., 1998). Some of the areas that have been identified as particularly engaged during navigation and spatial memory in a recent imaging metastudy include the parahippocampal gyrus, the posterior cingulate cortex (particularly the retrosplenial cortex), the posterior parietal cortex, and the middle occipital gyrus (Cona and Scarpazza, 2019).

5.2. Scene representations: a new key player?

Regardless of whether they account for the role of the hippocampus in navigation, scene representations could still play a crucial role in many navigation processes. A scene representation of the surrounding environment held in working memory could combine information that is required for spatial navigation but not all available at the same time, such as spatial features that are currently out of view and information stored in long-term memory. For example, the scene representation could incorporate the fact that there is a door behind us, which we know is there because we looked at it before, or because we have a cognitive map of the room. Navigation processes could then operate based on this richer scene representation. We illustrate this with the yellow boxes and arrows in Fig. 6.

An additional benefit of scene representations is that they could constitute a shared interface that facilitates the integration of navigation behaviors. Recall, for instance, the example of an agent using guidance to narrow down the area in which to look for the tree to aim for. This constraint of one behavior by the other would occur most easily if they were both operating on the same representation. Similarly, the shared scene representation could facilitate the selection of navigation behaviors. For instance, during aiming we might not be constantly looking at the goal. Every time we look away, we are presumably switching to vector movement. To do this efficiently, aiming and vector movement could share the same goal representation, perhaps a location pin-pointed in the scene representation. This also opens the possibility for the execution phase of navigation behaviors to be shared by all or many of the behaviors, simply translating the situated goal within the scene representation into movement instructions. This translation could employ a mechanism similar to that proposed by Edwardse et al. (2019), which combines grid cells and border cells to enable an agent to reach a nearby subgoal while avoiding small obstacles along the way. The situated goal could even be fed back to the navigation strategies for cases in which it is beneficial to use a navigation strategy to plan how to reach a goal identified by a navigation behavior, e.g., you pin-point the goal location using guidance but then need to use metric navigation to plan a path around some big obstacles blocking your way. These ideas are illustrated in Fig. 6 with the green arrows going from the extraction phase of the navigation behaviors to the situated goal, and from there back to the execution phase of the navigation behaviors and to the navigation strategies.

As discussed in the previous subsection, scene representations could depend on the hippocampus (Barry and Maguire, 2019), perhaps working together with a metric code in entorhinal cortex. A highly relevant observation in this regard is that entorhinal grid-like codes tile visual space in primates and humans (Killian et al., 2012; Julian et al.,
This could constitute an important component of the scene representation in strongly visual animals. Other areas strongly associated with the processing of scenes and their spatial layout include the parahippocampal, medial, and occipital place areas (Epstein and Baker, 2019; Epstein, 2008).

6. Conclusions

We have proposed a comprehensive taxonomy of navigation processes building on previous efforts (O’Keefe and Nadel, 1978; Gallistel, 1990; Trullier et al., 1997; Redish, 1999; Franz and Mallot, 2000), among others) and extending them to incorporate novel experimental findings. This taxonomy includes more fine-grained distinctions among navigation behaviors, a reconceptualization of route navigation, and the addition of new levels corresponding to the organization of navigation behaviors and strategies.

Using this taxonomy as a guide to parse the literature highlights two issues affecting much of the experimental literature. The first is that most common experimental paradigms do not properly dissociate or identify which navigation processes subjects are engaging in. In part, this is because laboratory settings are too simple compared to real-world navigation. Consider, for instance, that wild rats navigate foraging ranges several hundreds of meters wide with many complex spatial features (Taylor, 1978; Russell et al., 2010). However, most of what we know about the neuronal basis of spatial navigation comes from experiments where rodents shuttle back and forth on a linear track or forage randomly in open arenas, which are only one or two meters in size. Because these tasks are too simple, animals can often solve them in multiple ways. And since there are typically no probe trials, it is generally unclear what behavior or strategy animals are using to solve them. For example, tasks like the Morris water maze are ambiguous between guidance and metric navigation. Spatial strategies in a plus maze could correspond either to topological or metric navigation, or to a navigation behavior like guidance within the context of route navigation. Running back and forth on linear tracks is compatible with almost all navigation processes we describe. Random foraging in open environments is also compatible with most processes, perhaps in combination with mechanisms that help the animal avoid locations it has visited recently. Fortunately, the use of environments where animals need to make more complex navigational decisions is becoming more popular (e.g., Vallianatou et al., 2021) and starting to reveal some of the complexity hidden by simpler experimental paradigms. As an example, consider findings on how route segments, decision points, and goals affect the hippocampal theta phase code (Gupta et al., 2012; Wikenheiser and Redish, 2015; Kay et al., 2020). The second common issue in the literature is that studies of the neural basis of navigation have tended to focus on neural representations (e.g., place cells, grid cells, etc.) and their properties as opposed to focusing on the computations they enable or are involved in. The second point is aggravated by the first. Since different navigation processes make use of different neural representations (Vijayabaskaran and Cheng, 2022), in order to understand the representations observed in experiments it is vital to know what computational problems they are contributing to solve.

Another benefit of a comprehensive taxonomy of spatial navigation processes is that it helps us recognize uneven, and perhaps unjustified, allocations of resources and attention by the research community to the various processes. A majority of the research seems to have focused on vector movement, guidance, and route and metric navigation. Aiming, direction field navigation and path following seem to have been taken for granted but are, in our opinion, far more complex than they appear at

Fig. 6. The architecture of navigation processes including a scene representation. Like in Fig. 1A but adding the scene representation and showing the extraction and execution phases of the navigation behaviors. See text for details.
first glance and deserve more scrutiny. Topological navigation is often considered how researchers of grid cells (thought to underlie a metric representation of space) would benefit from learning about the limitations of metric navigation identified in human behavioral studies.

The topics discussed in this review could potentially be relevant for the study of other cognitive functions as well, since spatial navigation involves generic cognitive processes such as planning based on internal representations and orchestrating multiple available behaviors. Perhaps following a cooking recipe is in some sense analogous to route navigation. Involves generic cognitive processes such as planning based on internal representations and orchestrating multiple available behaviors. Perhaps following a cooking recipe is in some sense analogous to route navigation.

Acknowledgements

We thank Thomas Walther and Mohammadreza Mohagheghi Nejad for participating in many early discussions, Kamran Diba for feedback on an initial draft of the taxonomy, and the two anonymous reviewers for insightful criticisms and constructive comments.

Funding Support

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) — project number 316803389 – from SFB 1280, projects A14 (S.C.).

References

Brown, T.L., Carr, V.A., Loffler, K.F., Favilla, S.E., Gordon, A.M., Bowles, B.


Ganali, R., Weiss, I., Cassel, J.C., Kelche, C., 1998. Spatial memory, habituation, and reactions to spatial and nonspatial changes in rats with selective lesions of the hippocampus, the entorhinal cortex or the subiculum. Behav. Brain Res. 96 (1), 70–80.


Yonelinas, A.P., 2013. The hippocampus supports high-resolution binding in the service of perception, working memory and long-term memory. Behav. Brain Res. 254, 34–44.


