5

Animal Navigation

A Synthesis

Jan Wiener, Sara Shettleworth, Verner P. Bingman, Ken Cheng, Susan Healy, Lucia F. Jacobs, Kathryn J. Jeffery, Hanspeter A. Mallot, Randolf Menzel, and Nora S. Newcombe

Abstract

Navigation, the ability to organize behavior adaptively to move from one place to another, appeared early in the evolution of animals and occurs in all mobile species. At the simplest level, navigation may require only movement toward or away from a stimulus, but at a more sophisticated level, it involves the formation of complex internal representations of the environment, the subject's position within it, the location of goals, the various routes from current position to goal and possible obstacles along the way. The vast array of navigational capabilities in various species has made it challenging for students of comparative cognition to formulate unifying frameworks to describe and understand these capabilities, although the variety also confers an exciting opportunity for asking comparative questions that are hypothesis driven.

A unifying framework, the *navigation toolbox*, is proposed to provide a way of formulating common underlying principles that operate across many different taxa. The toolbox contains a hierarchy of representations and processes, ranging in complexity from simple and phylogenetically old sensorimotor processes, through the formation of navigational "primitives" such as orientation or landmark recognition, up to complex cognitive constructs such as cognitive maps, and finally culminating in the human capacity for symbolic representation and language. Each element in the hierarchy is positioned at a given level by virtue of being constructed from elements in the lower levels and having newly synthesized spatial semantic contents in the representations that were not present in the lower levels. In studying individual species, the challenge is to determine how given elements are implemented in that species, in view of its particular behavioral and anatomical constraints. The challenge for the field as a whole is to understand the semantic structure of spatial representations in general, which ultimately entails understanding the behavioral and neural mechanisms by which semantic content is synthesized from sensory inputs, stored, and used to generate behavior.

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Introduction

Navigation is one of the most fundamental problems that animals and humans confront. It is based on a complex interplay of a large number of different processes and components, and requires the integration of spatially relevant information across sensory modalities, the formation and retrieval of memories, and the selective activation of task-specific representations. Thus, navigation comprises a paradigmatic case of cognitive functions operating across several levels of complexity ranging from sensorimotor loops to higher forms of cognitive processing.

Because navigational tasks are performed by most animal species, in a wide variety of environmental conditions, over very different spatial scales, and using a variety of sensorimotor systems, the analysis of navigation poses an exciting but difficult challenge for comparative cognition. That such a variety of sensory inputs and motor outputs could all be subserved by the same general mechanisms seems a priori unlikely. Nevertheless, some fundamental processes may have been conserved by evolution, whereas others may have evolved convergently in different taxa. Comparison of animal species and their associated environmental adaptations provides us with vital information about the potential representations and processes that are involved in navigation and is a key strategy in identifying both the general as well as task- and species-specific components.

Comparative studies of navigation in a wide range of species have revealed certain processes which appear to be fundamental, inasmuch as they appeared early in evolution and have persisted, and other processes that seem to be more recent and more complex. Furthermore, evidence, some of it reviewed in this chapter, suggests that the more complex and recent processes are, in many (if not all) cases, synthesized from the older and simpler processes. This can be concluded not only from behavioral studies which look at the different abilities displayed by different animals, but also by studies of the underlying neurobiology and of how these processes are organized in the brain. These considerations have led us to propose a taxonomy of navigational processes, organized hierarchically, to try and capture the elements of this synthetic process. This taxonomy, which we call the "navigation toolbox," forms the basis for organizing the subsequent discussion of what we know, and what we have yet to discover, about how animals and humans navigate. Having outlined the processes in the toolbox, we then use it as the basis to explore the synthetic processes that gave rise to the higher-level elements before concluding with an examination of how these tools may be used in decision making and planning.

The Navigation Toolbox

The toolbox that we propose (Figure 5.1) was compiled by considering findings in the literature on animal behavior and neuroscience concerning navigation

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

and its building blocks. It comprises a collection of processes and representations ("tools") that have been identified by various research enterprises as contributing to navigation in species of different kinds. Each species has access to a subset of the tools in the toolbox, and it uses these tools to construct navigational behaviors. We acknowledge that alternative typologies and vocabularies exist in other disciplines, such as the spatial ontologies proposed by computer scientists (e.g., Kuipers 1978, 2000), the space syntax developed by architects (Hillier and Hanson 1984), the concepts employed by geographers and used in applied sciences that rely critically on the analysis of spatial relationships such as geographic information system (GIS) and cartography, the formulations of mathematicians working in metric geometry or topology, and even the long-standing philosophical debate about absolute and relative space. Ultimately, intensive cross-disciplinary discussion is needed to delineate the points of correspondence and difference among these approaches. Many of the differences may derive from the different data types considered, as well as differing assumptions and goals in the various fields.

The toolbox is organized hierarchically, distinguishing between lower-level sensory processes and higher-level processes thought to operate on the lower-level components. This is, inevitably, an artificial and hence imperfect hierarchy; for convenience we have made it discrete, whereas in reality there is no clear dividing line between simpler and more complex processing capabilities. However, we hope the organization of the toolbox may be useful for students of navigation in thinking about how complex behaviors are synthesized from more elemental processes in the nervous system.

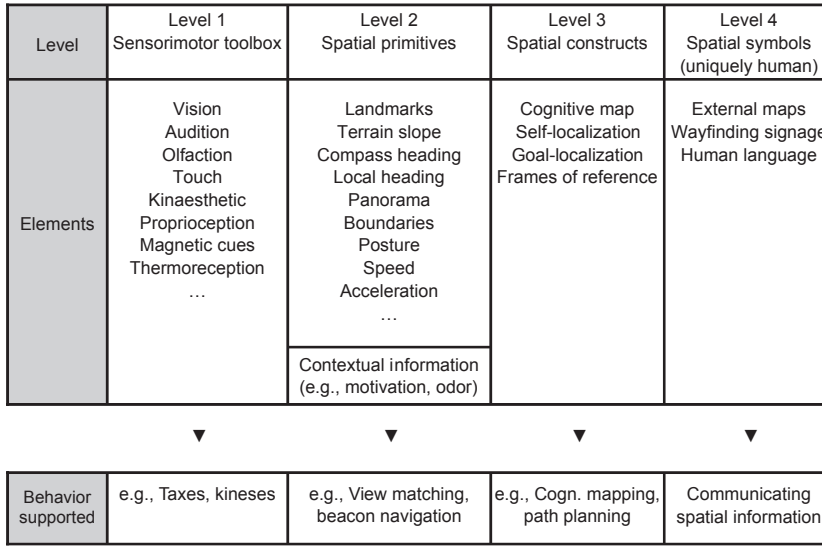

Figure 5.1 The navigation toolbox.

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Navigational elements (representation or processes) are assigned to different levels of the hierarchy according to two principles:

1. Elements at a given level are synthesized by integration of elements at lower levels.
2. Elements at a given level possess new spatial semantic content not present in the lower levels.

Although the first principle is intuitively obvious, the second requires some explanation. By "spatial semantic content" is meant the "meaning," in spatial terms, of the given element; in other words, how the content of that element relates to some real-world parameter. Examples could be heading, distance, or "place." An example of "new spatial semantic content" might be construction of a position vector by integrating information about distance with information about direction. Not only is the new representation synthesized by integration of elements at lower levels, it also contains additional spatial information. With this new, semantically richer representation, it is possible to perform more sophisticated navigational feats, such as homing or shortcutting. Thus, the hallmark of navigational complexity is the synthesis of internal representations. We will return to this point in the discussion of cognitive mapping.

The navigation toolbox has four levels:

1. *The sensorimotor toolbox* comprises different sensory modalities (e.g., vision, olfaction) and simple motor processes (e.g., approach and avoidance), and provides information relevant for locomotion and navigation. In mammals, these processes take place in, or near, primary sensory or motor cortices. The elements of the sensorimotor toolbox are also involved in other, nonspatial behaviors (e.g., mating, predator avoidance, tool use, or social interaction).
2. *Spatial primitives* are a set of representations that animals build, using the sensorimotor tools in Level 1. For example, an animal may combine the pattern of optic flow, computed in the visual system, together with linear acceleration signals extracted from the otolith organs in the inner ear, to help construct a representation of velocity. Velocity is a spatial primitive in the sense that while it is metric, being a measure of distance traveled per unit time, it does not in itself contain positional information. It can, however, be used to *compute* position, a Level 3 representation. Another example is landmark identification, another higher-order process that may make use of several sensory modalities. Again, landmarks in themselves do not contain positional information, but can be used to compute position if the animal can determine their relative distances and directions with respect to itself. These navigational building blocks are formed using integrative processes (discussed in detail later in the chapter).

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

3. *Spatial constructs* are even higher-level, more complex, spatial representations constructed from the spatial primitives (Level 2). An example could be position, calculated by identifying landmarks and determining their distances and directions—all Level 2 processes in our hierarchy—and using these to create a representation of current location by a process known as piloting. In a complementary manner, velocity signals computed using Level 2 processes can be used to update a positional calculation, by interaction with the piloting process. Moreover, by relating such positional calculations with landmarks, an integrated representation of space can be formed at Level 3. As with Level 2, the underlying integrative processes in Level 3 are explored later.
4. *Spatial symbols*, including human language and graphic representations, allow for external storage and interindividual communication of Level 3 spatial constructs.

Generating Behavior with Navigational Tools

A given organism does not need all of the tools, or even all of the levels in the toolbox, to generate navigational behavior. Spatial behavior can be supported by any or all four levels of the navigation toolbox, as follows:

- The sensorimotor toolbox can support spatial behavior that is based purely on sensory processes, or sensorimotor loops. Examples of such behavior are taxes and kineses (e.g., positive phototaxis in flies), which do not require forming representations of space or spatial concepts, as may occur in the higher levels of the navigation toolbox. All that is required is simple detection of a stimulus, such as light or odor, and then organization of movement either proportional in its intensity (kinesis) or toward or away from that stimulus (taxis).
- At the level of spatial primitives, information from the sensory toolbox has been integrated to form simple representations that may lack positional information, but which are useful in some kinds of navigation, and are used to build complex spatial representations. At this level, simple locomotion and navigational behavior (such as centering by keeping optic flow equal on the two visual hemispheres, beacon navigation, or view matching) is supported (see Shettleworth 2010b, Chapter 8), but the assembly of more spatially complex internal representations is not yet implied.
- The third level of the navigation toolbox, spatial constructs, consists of representations that are built from the spatial primitives. At its more sophisticated level, integration of primitive nonspatial processes, such as landmark identification, together with primitive relational information, such as the relative distances and directions of landmarks, can result in a spatial memory structure (an internal representation of space) that is

From “Animal Thinking: Contemporary Issues in Comparative Cognition,“ edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

in many ways map-like, and is often referred to as a "cognitive map." The extent to which animals of various kinds do or do not possess a memory organized like a cognitive map has been the source of lively debate over many years (see below).

- The fourth level of the toolbox, the spatial symbols, allows for unique behavioral feats in humans, such as navigating to an arbitrary place given a street address anywhere around the world. However, more importantly, humans exhibit the ability to communicate their navigational computations, to a high level of sophistication, to other humans, via maps, language, and mathematics. Map drawing exists, to a very restricted degree, in other species (e.g., bees) as does vocal communication, but the level attained by humans far surpasses anything seen in the rest of the animal kingdom in its flexibility and productive power.

Spatial behavior supported by Level 2 (spatial primitives) and Level 3 (spatial constructs) of the navigation toolbox is of particular interest for comparative cognition: at these levels, we are able to move away from species-specific mechanisms toward general principles of navigation. For example, migrating birds may have access to magnetoreception whereas insects are capable of perceiving polarization patterns in the sky. Despite these differences on the sensory level (Level 1), both sources of information are utilized or integrated to derive semantically equivalent information about space—a compass heading (Level 2). Similarly, honeybees and desert ants rely primarily on different sensory cues (optic flow and proprioception) to estimate the distance that the animal has traveled (odometry). Bees flying over a large range of local cues and following long-ranging landmarks, however, are equipped with the capacity to integrate these spatial structures in such a way that they localize themselves and navigate to the intended goal by novel shortcuts (see Menzel, this volume). Differences occur not just on the sensory but also the motor side. An example is the navigation behavior of the desert ant *Cataglyphis* which, when running over a featureless landscape, seems to primarily rely on an elementary form of path integration, whereas ants navigating in a landscape rich in local cues, such as *Melophorus bagoti*, appear to learn sequences of turns in relating measures of path integration with these cues in their rather stereotypical foraging routes. These examples illustrate that the processes operating during navigation are heavily constrained by the animal's action space.

One might think of these higher-level representations (compass heading or distance) as being *supramodal* (i.e., independent of or "lying above" specific sensory modality) and of having meaning (semantic content) independent of the neural implementation. Such abstractions allow not only comparison across species, but also for interactions between students of natural and artificial navigational agents.

The navigation toolbox thus provides an organizational framework that allows for more systematic cross-species comparisons of higher-level principles

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

supporting spatial navigation, independent of species-specific implementations. We suggest that by attempting to fit the diversity of navigational inputs and outputs seen among species into this proposed framework, we will make significant progress in understanding the similarities and differences in their navigational (and cognitive) abilities. However, having extracted general principles, it is then important to consider whether and how species-specific constraints influence information processing. For example, does an animal's action space and complexity of neural integration determine the set of spatial primitives that are used by the animal? In addition, what is the influence of the animal's navigation range, and how do the properties of its habitat influence how it uses the spatial primitives and spatial constructs available (Figure 5.1, Levels 2 and 3)?

Before examining the integrative processes that lie behind the construction of these navigational elements, let us consider the four levels in the toolbox in more detail.

Level 1: Sensorimotor Processes

The simplest level of the navigation toolbox contains a set of sensory and motor processes that can support simple navigational processes, such as taxes. These sensory processes are used to build higher-order tools in the toolbox and will not be considered further here. Increasing evidence supports the notion of top-down modulation of sensory processing, and it may be that some kinds of simple navigational process, such as template matching for view-based navigation, might occur at very early stages of neural processing, such as primary sensory cortical areas.

Level 2: Spatial Primitives

Many navigational phenomena, particularly those seen in laboratory settings using small-scale environments, can be explained by one or more mechanisms which are more sophisticated than simple sensory-driven Level 1 processes such as taxes, but are nevertheless spatially still relatively unsophisticated (i.e., belonging to Level 2 of the navigation toolbox). Spatial primitives are considered to exist at a level of complexity above sensorimotor elements for two reasons: (a) they are synthesized *from* Level 1 elements, and (b) they have spatial (or proto-spatial) semantic content not present in those elements.

An example of a spatial primitive is landmark recognition. The majority of species that have been studied to date rely on landmarks, at least some of the time, to help organize their spatial behavior; thus this competence seems to be phylogenetically old, although landmark recognition may require different levels of processing sophistication: an ant may simply recognize a landmark

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

as being a particular retinal pattern as part of a panorama (Collet and Collet 2002). In this case, the "snapshot" is a code of the place or viewpoint from which it has been taken. In addition to this snapshot memory, primates might have a multimodal, viewpoint-independent representation built using object-processing capabilities in the perirhinal cortex. Here, one landmark memory may be used in the recognition of more than one place, and the places recognized may be remote from the actual landmark position. Landmark identity, being (usually) multimodal, is a semantically different category from a simple stimulus pattern impinging on a sense organ. Note, however, that the simplest forms of recognition (e.g., odor detection or the ant landmark recognition example above) could be considered as bordering on Level 1 processes, reinforcing the notion that this is not a hierarchy to be taken too rigidly. However, while a landmark may be necessary for spatial navigation, it is not sufficient inasmuch as it contains no spatial information; unless the goal is *at* the landmark, the animal needs other information in order to navigate. In this sense, the landmark, being space-free, is a spatial primitive. If the goal is at the landmark, then a simple kind of navigation (e.g., beacon homing) is supported, and evidence suggests that many species can do this.

Spatial primitives can support behavior that looks remarkably complex to an observer. Navigation by view matching provides an example: it requires a representation of the goal (a standard, e.g., a snapshot or a panorama recorded at the home location) and a record of the current location. According to the snapshot model (Cartwright and Collet 1982) the difference between the standard and the current record is computed by a comparator, resulting in an error. This error is thought to drive the movement of the navigator and in this way the error is reduced. View matching can thus be conceptualized as a servomechanism (Cheng 2006). Beacon navigation requires landmark identification, which may be a higher-order object recognition process, but the behavior that follows is simple approach or avoidance. At a more sophisticated level, a turning movement could be organized to the left or right on the basis of a perceived stimulus; this introduces the beginnings of true spatial (i.e., relational) processing and such behavior lies on the boundary between Levels 2 and 3.

Navigating along a well-known route is another example of behavior supported at Level 2, allowing animals to reach distant destinations even if these are beyond the current sensory horizon. It can be explained assuming spatial primitives, such as a string of beacons or landmarks, to identify the place at which a particular turn has to be conducted. These turns can be informed by motor responses associated to the place. A series of such recognition-triggered responses, therefore, is sufficient for explaining route navigation. Importantly, such processes do not require the operation of a cognitive map, although this does not in itself rule out that some animal species use a map, even in route following, if one is available.

Heading calculation is another well-known example of what we call here a spatial primitive. Many species are able to extract earth-relative directional

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

information from either magnetic or celestial compass cues. Behavioral and neurobiological studies in the laboratory have additionally revealed the existence of local orientational reference frames which are conferred by local cues such as distal landmarks. Heading calculation is an interesting process from a theoretical point of view because it may require the bringing together of very different sensory information streams (directional cues, e.g., with respect to a compass, landmarks and multimodal self-motion cues) to synthesize a semantically new, supramodal representation (such integration processes are discussed in more detail below). As with landmarks, heading alone is of limited use in navigation but is of very great use when combined with other spatial primitives, such as distance leading to an integration of traversed path (i.e., path integration), for example, in the formation of goal or homing vectors. In its basic form, path integration integrates rotational and distance information (both spatial primitives), resulting in a home vector (i.e., a working memory representation of the distance and direction to the home or to other places encountered during travel). When heading home, the animal moves so as to reduce this home vector to zero. It may be debated whether such a home vector represents a spatial primitive or a spatial construct (see next section). If, however, the status of the path integration measure is associated with other spatial primitives, such as landmarks, this would allow the formation of long-term representations of metric relationships between different places in the environment. In this case, path integration undoubtedly becomes part of a spatial construct.

Level 3: Spatial Constructs

Having established a collection of relatively primitive tools which animals may have access to for navigational purposes, we focus now on Level 3: the spatial constructs. As with the previous levels, the elements in this level are characterized by having been synthesized from lower-level elements and by having new semantic content. An important new semantic category is that of *position*, a relational term that implies specification of a subject or an object by means of adjacencies or neighborhoods, distances, or bearings to other memory items. We will call such relations a frame of reference without implying that it has to take the form of a metrical coordinate system. As such, one of the most interesting and contentious issues has been the extent to which it is appropriate to place, in this level, a representation that could be considered map-like.

A good example of a spatial construct is the local vector (Collet and Collet 2002). In executing a local vector operation, contextual cues, such as a view, are associated with path integration to enable travel in a particular compass direction. A view of the surrounding scene, one spatial primitive, serves as the trigger to execute a behavior that relies on another spatial primitive, a compass upon which a direction of travel is derived. Local vectors based on the

From “Animal Thinking: Contemporary Issues in Comparative Cognition,“ edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

current state of path integration have been well demonstrated in desert ants (*Cataglyphis*: Collet et al. 1998; *Melophorus*: Legge et al. 2010). It is possible that route-following behavior consists of a string of such local vectors, with the end of one local vector providing the initial conditions (e.g., the appropriate view) to trigger the next (Trullier et al. 1997; Cheng 2006). If such associations between places and local vectors are spread in the vicinity of a goal, they can support homing from an area exceeding the sensory horizon, or catchment area, of a single snapshot marking the home (Cartwright and Collet 1987).

A closely related example is that of path integration. As discussed in the last section, path integration can be viewed as a spatial primitive combining velocity and heading over the course of a journey so as to maintain a constant, ongoing representation. Alternatively, when combined with landmark information, it can be seen as an example of a spatial construct including relational information of landmarks. The exact nature of this representation has not been fully elucidated and may vary according to species or settings: it could be a vector (e.g., the distance and direction back to a home base at a particular location or to another known position) or it could be a location in a cognitive map (see below). If the path integrator transiently fails, or if the animal has just arrived in the environment, it can undertake a position fix using a different set of spatial primitives: namely, available landmarks together with their computed distances and directions. The process of taking a position fix from familiar landmarks is, as mentioned earlier, known as piloting.

The sophistication of Level 3 constructs also allows for navigation toward unseen goals. Insofar as a given species uses such relational spatial constructs, goals can be defined as attractive locations in the spatial reference system, and the task for the animal is to reach the goal from its current location. For spatial behavior at Level 3, it is assumed that the goal cannot be approached simply by steering toward a beacon at the goal or executing a set of landmark-stimulated responses or a learned sequence of body turns. Instead, Level 3 constructs are relational, whereby more than one spatial primitive is required to be integrated to solve the spatial problem.

The Cognitive Map

In the section on spatial primitives, we explored navigational behaviors that can be achieved with relatively simple processes, such as view matching and beacon homing. Other navigation behaviors, in contrast, cannot be explained without reference to a spatial memory in which places are related to one another in a common reference frame. Such a reference frame-based spatial memory system is usually called a *cognitive map*. This is an interesting class of spatial semantic information that is likely possessed by humans (Gillner and Mallot 1998; Hartley et al. 2003), and for which the existence in other animals has been disputed (Bennett 1996; Shettleworth 2010b).

From “Animal Thinking: Contemporary Issues in Comparative Cognition,“ edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

The strongest evidence for the operation of a cognitive map, in the sense of a connected allocentric representation of familiar space, has been seen as the ability to take a shortcut or a novel route under conditions in which path integration, view matching, and beacon orientation can be excluded. While the exclusion of these processes is difficult to achieve experimentally, this can be demonstrated, for example, when an animal is displaced from a goal-oriented route and needs first to self-localize, on the basis of local cues, and then to correct for its displacement in order to reorient toward a goal. What kind of memory structure would allow an animal to relocate its position relative to the goal and subsequently compute a corrective direction to the goal? It can be argued that only a spatial representation with features that resemble a map could enable such corrective reorientations. The critical feature of this representation which makes it map-like is its ability to support spatially relational processes: self-localization and vector calculation.

Thus, the term "cognitive map" has frequently been used to refer to internal spatial representations that organize spatial knowledge about different locations in the world by relating them to one another either by adjacencies or in a common reference frame. The term originated in experimental psychology with Tolman (1948) and was memorably utilized in a widely cited book by O'Keefe and Nadel (1978). While the concept of cognitive maps has been criticized for being used too often as a metaphor without reference to mechanisms, researchers working on navigation in the field where the animal is allowed to navigate in an open space find *map* a useful concept and see evidence for *cognitive maps* in several species and settings (see Menzel and Bingman, both this volume).

The organization of spatial information in the mammalian (rat) brain can also be seen in terms of mapping. In fact, the place and grid cell systems in rats constitute a neural substrate of metric information/coordinate system (see Jeffery this volume; Hafting et al. 2005). Similarly, functional brain imaging studies in humans navigating large-scale virtual environments demonstrate that novel shortcutting or route-planning behavior recruits a different cortical network than route-following behavior (e.g., Hartley et al. 2003); the "map task" used in this and similar studies can be solved with a memory based on place adjacencies, leaving the possibility that the underlying "map" is not using metric coordinates but only a simpler graph structure (Gillner and Mallot 1998).

As discussed above, cognitive maps need to encode spatial relations between locations in the environment, either as a graph of adjacent places or via an allocentric reference frame. Spatial behavior specified in Level 3 of the navigation toolbox—such as self-localization, goal-localization, and the formation of a plan designed to get to the goal (cf. definition of spatial planning below)—require such relational knowledge. Although spatial relations between places can be encoded in a topological (i.e., nonmetric) graph-like format, in which places are represented as nodes and transitions between places are represented as edges (graph-theoretic representation), the inclusion of metric information

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

appears to be crucial to explain a number of behavioral findings (Menzel, this volume).

In humans, evidence of metric information in spatial long-term memory is debatable (see, however, Schinazi et al. 2009). For example, Foo et al. (2005) trained subjects to walk along two straight paths from a starting point, thereby acquiring clear knowledge of distances and directions of each path. When released at the end of one path, shortcutting to the end of the other path was poor (for an analogous study in dogs, see Chapuis and Varlet 1987). Note that this shortcutting is not a type of path integration but requires metric recombination of local distance and angle information from long-term memory. The question about the amount of metric information encoded in the cognitive map remains controversial. A challenge for the future will be to answer a number of questions that are related to this point:

1. Is metric knowledge a prerequisite for cognitive maps?
2. Is metric knowledge equivalent to cognitive maps?
3. How does local and global metric information relate to the concept of the cognitive map?

A different notion of the map has been suggested for birds. Often discussed in the context of bird navigation is a map, referred to as the "navigational map," formed by intersecting stimulus gradients that form a bi-coordinate system by which any point in space is characterized by a unique combination of coordinates (Wallraff 2005). These coordinate values could provide a global allocentric reference frame with respect to which items of spatial long-term memory (places, landmarks, "home") could be represented (for further details, see Bingman, this volume).

Some of the controversy on the nature of the cognitive map may arise from a confounding of separable functions which must be integrated for novel shortcutting but which may be used independently. For example, it may be necessary to analyze separately the cognitive map component derived from distributed cues, such as compass bearings and gradients, from those derived from discrete cues, as done by the parallel map model for the integration of such cues (Jacobs and Schenk 2003; reviewed in Menzel, this volume). Because of the necessity of an underlying distributed cue representation (i.e., a bearing map) for shortcutting, it may be difficult to demonstrate true cognitive map shortcutting in small laboratory settings. In the notable exceptions (e.g., Roberts et al. 2007), the experimental space is characterized by an extreme paucity of discrete cues, which may force the animal to rely on bearing-map function, even in the relatively small space of a laboratory maze. In contrast, free-flying birds and insects may naturally recruit a strategy based on such distributed cues, which would enable them to shortcut flexibly. Therefore, one of the challenges for future navigation research is to develop well-controlled laboratory experimental paradigms that will allow us to distinguish spatial behaviors supported by Level 2 processes and Level 3 processes.

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Level 4: Spatial Symbols

The final class of navigational elements in our hypothetical navigation toolbox is an almost uniquely human capacity: the ability to use external symbols to represent space. This ability allows semantic information to be developed beyond that which can be achieved by the neural navigation systems embedded in the brain. Two classes of symbolic manipulation are particularly important in this regard: language and mathematics. Here we concentrate primarily on language, because its role in ordinary day-to-day spatial navigation processes has been the focus of intensive recent study.

Are There Uniquely Human Navigational Strategies? Is There a Role for Language?

Humans have a distinctive, and arguably unique, capacity for symbolic representation and communication (see Figure 5.1, Level 4). Symbolic systems include language, of course, but also encompass more spatial systems such as gesture, sketching, drawing, and external maps. In the context of comparative cognition, two questions need to be considered. First, to what extent does the availability of Level 4 have a top-down influence on the prior levels? This classic issue concerns the extent to which language (as well as other symbolic systems) influences thought. Second, to what extent does the availability of Level 4 extend or augment the capabilities available in species that only have access to the first three levels? That is, are symbolic systems powerful tools for augmenting spatial functioning?

To What Extent Does the Availability of Level 4 Have a Top-Down Influence on Prior Levels?

Thinking about this question has concentrated almost exclusively on spatial language. It is interesting that there are some notable design mismatches between space and language: space is continuous whereas language is categorical; multiple spatial relations are available simultaneously in the world, but must be discussed sequentially in language. Given these considerations, one might suspect that spatial language can be helpful in spatial functioning, but perhaps to a limited extent.

Strong claims have been made that spatial language *shapes* spatial thought (e.g., Levinson 2003). A prominent and often-cited example of such shaping concerns languages that require the use of compass-referenced coordinates to describe spatial relations (e.g., the cup is to the *north* of the plate). By contrast, many other languages describe this spatial relation in a relative fashion (e.g., the cup is to the *right* of the plate). Experiments have been done in which speakers of these contrasting types of languages have been asked to view objects arranged in a line on one table, and then to place them in the same way

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

on a second table after rotating the objects 180 degrees. In many studies, as reviewed in Levinson (2003), speakers of absolute languages maintain the correct compass directions in their reproductions on the second table, while speakers of relative languages use a relative coding internal to the array. However, Li and Gleitman (2002) have shown that these results can be altered through small environmental manipulations so that, for example, speakers of relative languages will produce absolute reproductions when a landmark is available (e.g., a pond at one end of the table). They suggest that there is a great deal of flexibility in spatial representations and thought, so that language picks out certain systems of spatial coding but does not preclude the existence of others (see also Gallistel 2002; Newcombe 2005).

The issue of the uses of spatial language has special resonance in terms of thinking about human development, since, of course, children begin life without having a language and acquire language(s) over the first few years of life. Two lines of research have been important in recent thinking about how spatial language affects spatial development. First, it has been argued that the spatial language which infants hear structures their acquisition of spatial concepts (e.g., Bowerman 1996). For example, Korean children, who hear a language in which the distinction between "tight fit" and "loose fit" is an important semantic contrast, acquire the perceptual and conceptual basis for the distinction, whereas English children do not. A contrasting position is that infants begin with an array of spatial concepts (either innately specified or acquired in the first year) and subsequently map certain of those concepts to the language they hear around them (e.g., Mandler 1996). Data gathered on this issue seem to favor the position that spatial concepts exist before spatial language, but that spatial language draws attention to some of these concepts and dampens others, making the concepts used frequently more available and the ones not used more difficult to activate (see review by Göksun et al. 2010).

Second, Spelke and her coworkers claim (Hermer and Spelke 1996) that language (specifically, productive control of the terms "left" and "right") is essential to being able to use features for reorientation in the geometric module paradigm introduced by Cheng (1986; reviewed by Cheng and Newcombe 2005). Others claim that children younger than 6 years behave like nonlinguistic, nonhuman animals by using geometric information alone to reorient, ignoring useful featural information. They also report a transition to a uniquely human use of features as well as geometry at the age of 6 years (with the advent of the productive use of the words "left" and "right"). However, the failure of toddlers to use features turns out to be specific to the use of extremely small spaces of limited ecological validity (Learmonth et al. 2002). Recently, Twyman and Newcombe (2010) reviewed the extensive literature that has accumulated on this issue, arguing that a variety of evidence indicates that feature use is far more flexible than has been argued by the Spelke group, and that language is not necessary to the development that does occur (although it could be helpful, as could other kinds of experience).

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

What about symbolic systems other than language? Few, if any, strong Whorfian positions have been taken on the topic of how maps (or other visuo-spatial systems) might mold spatial thinking. Uttal (2000) reviewed how maps affect spatial development, arguing that they are helpful in acquiring abstract spatial concepts and in systematizing spatial thinking. This view, however, merely augments the position related to our second question; it does not take a deterministic stance. Dehaene, Izard, Pica, and Spelke (2006b) claim that the basic geometry of maps is innately available without experience and suggest that maps merely capture knowledge rather than mold it. However, their data from Western cultures show development in the use of some map-relevant concepts that does not occur in their Amazonian population, suggesting a role of cultural input (see Newcombe and Uttal 2006).

Does the Availability of Level 4 Extend or Augment the Capabilities Available in Species that Only Have Access to the First Three Levels?

Spatial language and other spatial symbolic systems seem to aid spatial thought in a powerful way, even if they do not have a strictly determinative effect. For example, babies learn spatial concepts better when the concepts are labeled (Casasola et al. 2009), and, similarly, teaching children a particular spatial word (such as "middle") seems to support their learning of the corresponding spatial concept (Simms and Gentner 2007). Preschool children whose parents use more spatial words (such as "outside" or "corner") show stronger spatial growth than children who do not hear as much language of this kind (Pruden et al. 2010). Spatial thinking in adults is also clearly affected by spatial language (Boroditsky 2001; Casasanto 2008), although some studies of this kind take a turn back in the Whorfian direction.

Maps may be as important as or more important than language in giving human navigation a distinctive quality (Uttal 2000). Maps have been used by humans for millennia (possibly up to 8,000 years), though there have been notable technical advances over the past centuries (e.g., the discovery of a means of measuring longitude and navigation assistant systems that guide navigators to the destination). External maps offload memory and cognitive processing demands, allow for communication and social interaction toward a spatial goal, and highlight areas of ignorance (i.e., territories not explored or separate territories that are known internally but whose relation needs to be determined).

While maps are very elaborate external navigation aids, humans also design and utilize less complex tools to communicate spatial information (e.g., signposts). Such simpler means of communicating spatial information are not unique to humans but are also found in animals. Many animal species, for example, mark their territories with "scent posts" which could be seen as aids to conspecifics in detecting territory boundaries. The most famous example of spatial communication in animals is the waggle dance of honeybees (von

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Frisch 1965). In fact, bees use symbolic communication to convey distance and direction information to food locations. However, this is a closed system which lacks the productivity and flexibility found in a true language.

Synthesis and the Navigation Toolbox

The navigation toolbox identifies several sets of processes and representational structures, some taxon-specific and some more general, that can be integrated in the service of navigation. This integration is useful for two reasons. First, for a given type of information, two sensors carrying information of the same semantic content are better than one: the accuracy of detection can always be improved by increasing the number of detectors. Second, as discussed earlier in this chapter, integration of information streams having different semantic content allows for the formation of higher-order representations that allow computations not possible at the lower levels. These two types of integration are discussed in turn.

Integration of Information Streams that Have Similar Semantic Content

In many cases it is advantageous to combine semantically equivalent information sources to derive maximum advantage from all the information available. Homing pigeons, for example, may use both sun compass information and magnetic compass information (both Level 2 primitives, and both belonging to the semantic category of "heading") to compute the direction toward the home. The information sources may even come from the *same* sensory modality (e.g., from two visual landmarks both intermittently visible in foggy conditions). Cue integration presents an interesting problem, however. What should the brain do with these two sources of information? Should it average them, or choose one or the other? If the former, should all sources be given equal weight or should some count more toward the average than others? If the latter, how is the decision made to prefer one over the other?

Bayesian approaches provide a normative framework for modeling how and whether animals should combine versus select information from multiple cues to make inferences or judgments. The importance of the Bayesian approach is its reliance on the existence of "priors," which comprise preconfigured knowledge about the nature of the incoming information. An example of a Bayesian prior would be past experience that a landmark is positionally unstable. Prior knowledge allows a decision maker to discount information known to be unreliable, or at least to rely more heavily on sources known to be trustworthy. Specifically, the Bayesian framework suggests that the available sources of information should be weighted in inverse proportion to their reliability, as measured in terms of variance in the estimate. The principle applies when the animal has multiple sources of information (such as Level 2 primitives) which

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

may be combined and the animal has a past history of relevant experiences (the Bayesian prior) which may be used to weight the current information sources. While Bayesian probability theory has been extensively used in human spatial perception, it has also been proposed for the integration of spatial information in other species (Cheng et al. 2007). It is important to note that controversy remains concerning the value of the Bayesian approach for the explanation of animal navigation behavior. This is because (a) at present there is little evidence for Bayesian integration outside of humans and (b) the Bayesian approach provides a formalism that addresses the statistics of information sources rather than the animals' responses to these sources. While Bayesian approaches are often accepted as useful for hypothesis generation, the direct relation between the formalism and actual behavior is arguable (Jones and Love 2011).

While it is advantageous to combine semantically equivalent information sources in many cases, it sometimes makes more sense to choose one or the other in a winner-take-all (or "first past the post") manner. The question then arises as to how to determine which approach to take. Cheng and colleagues (2007) and Körding (2007) argue that the Bayesian framework predicts the integration of different sources of spatial information only if these sources indicate directions near to each other. If, however, the two sources indicate directions very different from each other, integrating and thus averaging will result in a direction that is indicated by neither of the two sources, which makes integration unreasonable. In cue-conflict experiments, this lack of integration is often exploited to compare the behavioral relevance of different information sources, such as different Level 2 primitives. For example, phase- or clock-shift experiments in homing pigeons result in a large difference (conflict) in the indicated compass direction to the home between the sun compass and magnetic compass—a situation where integration, in the Bayesian view described above, should be unlikely. Indeed, under phase-shift conditions, homing pigeons rely almost exclusively on the sun compass with very little integration of magnetic information in the final behavioral output (Wallraff 2005).

Similarly, studies of place neurons in the rat hippocampus have found that the ensemble location-specific activity will rotate to follow small shifts (10–20°) of a landmark, but will fail to follow large shifts (180°) which introduce a large discrepancy between the landmark and the animal's internal sense of direction (Rotenberg and Muller 1997). Prior history of the landmark also influences how the neurons respond: they will tolerate reasonably large angular rotations of the landmark if the animal did not see these occur, but will not follow the landmark if the animal saw it moved (Jeffery and O'Keefe 1999). An interesting and unresolved question concerns whether the place system is able to reweight the reliability of various cues based on the animal's own internal state. For example, if an animal has been deprived of vision for some time, allowing time for its path integrator to drift, it may be more inclined to tolerate large discrepancies between external and internal cues than if its path integrator was recently reset and hence more reliable.

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Integration of Information Streams that Have Different Semantic Content

As well as simply improving detection accuracy for information of a single semantic category, integration can occur when streams of semantically different information converge to form semantically new kinds of information. A prototypical example of integration to form a semantically new representation is the head direction system in rats: vestibular, proprioceptive, motor efferent, and visual signals are combined to form a representation of the orientation of the animal's head. The incoming sensory information is semantically different in nature. Vestibular signals, for example, code for accelerations of the animal's head, whereas visual information conveys the presence and/or location of static landmarks. The Level 2 primitive that results from integrating these Level 1 sensory inputs, the head direction signal, is a supramodal representation because it is not anchored to a particular sensory modality and is not a particular information type; instead, it arises through the convergence of several different kinds. Supramodal representations allow an organism to make use of different kinds of information having different properties. In the case of the head direction signal, the self-motion cues are imprecise but always present, whereas landmarks are very precise but only available sporadically (e.g., when the animal is looking in their direction). By using both, the organism derives an adaptive advantage. Indeed, it can be speculated that the great enlargement of the cerebral cortex during vertebrate evolution derives from the development of multiple supramodal representations in many different domains.

Even more importantly than simply exploiting the advantages of different cue types, the resulting representation is of a semantically different class: "heading," which was not present in the original inputs. This allows the brain to construct signals with new properties and enables an organism greater flexibility in its behavior. In the head direction cell example, the construction of a heading signal permits an animal to navigate even in directions in which there is no specific cue present. Furthermore, it is one of the building blocks to formation of yet another semantically new class of information, the cognitive map, which in turn allows an animal to perform such feats as navigational planning.

Bayesian rules can be used to model integration of semantically different information, just as they can be used to integrate cues of the same "sort." For example, in using head-mounted virtual displays, Warren et al. (2001), found that humans integrated two kinds of information, both providing a heading direction toward an object (beacon). Subjects used both the focus of expansion in optic flow, in which the optic flow pattern expands from the point toward which we face and the identified beacon object. This was shown by dissociating the two cues in virtual reality. In human path integration, both visual cues (optic flow) and cues from the body senses (kinaesthetic and proprioceptive senses) are used (Kearns et al. 2002; Nico et al. 2002).

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Bayesian priors might not just affect whether cue combination or choice occurs, but also the hierarchy of cue preference in cases when choice is necessary. Cue-conflict experiments have been a classic method for examining the hierarchy of cue use in spatial problems (e.g., Brodbeck 1994). The hierarchy appears to be dependent on the species, their ontogenetic experience, and on reliability of the cues. For example, when solving spatial problems in the field, free-living hummingbirds use information in a context-dependent manner: They will use the color of a flower, but only if spatial information is not relevant (Hurly and Healy 2002). They will learn the location of a flower relative to other flowers if the other flowers are relatively close (up to 40 cm) but will ignore the information once they are further away (> 40cm; Healy and Hurly 1998). In squirrels, the use of a hierarchy or a majority strategy appears to change by season under natural conditions, suggesting that physiological changes induced by photoperiod and/or experience may constrain or mold the attentional resources available for spatial mapping and hence the final strategy used for orientation (Waisman and Jacobs 2008). In adult humans, the choice of a geometric or featural cue depends on the size of the enclosure (Ratliff and Newcombe 2008). Geometric cues seem to be preferred in smaller spaces, whereas featural cues are followed in larger ones.

In summary, cue integration has the advantage of allowing an animal to maximize its use of information: it can choose the most reliable cues and discount the less reliable; it can use different cues in different ways, depending on their characteristics; and it can combine cues to form supramodal, semantically new representations that allow more sophisticated calculations. For some navigation behavior, Bayesian approaches appear to describe how information can be combined. Determining how general Bayes's rule is in navigation, and how it is implemented at the neural level, remains a challenge for the future.

Cue Integration and Cognitive Mapping

Integrating spatial primitives (Level 2) into spatial constructs (Level 3) is a crucial feature of cognitive mapping. Essentially, different Level 2 primitives, such as landmarks that may define particular locations in space, are related to one another either by adjacency or in a global reference frame (see above section on "The Cognitive Map"). One proposal is a common coding system that integrates location on a coordinate map constructed from distributed cues (the bearing map) with the topological relations of discrete objects (the sketch map) encoded in relation to their location on the coordinate map (the integrated, i.e., cognitive map; Jacobs and Schenk 2003).

Localization of self, other, goal, etc. is a central process of Level 3 and requires either a place representation, which may be recognized from landmarks or some other context in the neighborhood, or a specification of the locality (i.e., frame of reference). It is usual in the field of spatial cognition

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

to distinguish between egocentric and allocentric frames of reference (both Level 3 constructs), where egocentric is self-referred and allocentric (literally: "other-centered") is referenced to the external world. Egocentric frames of reference follow the animal around, as it were, and neurobiological studies in mammals indicate that there are many such reference frames, at least in this taxon, encoded in parietal cortex. The mammalian frames are anchored variously to the eye, head, arm, etc., and need to be integrated for actions to be coordinated. Likewise, there are probably many different allocentric reference frames. Three have thus far been identified: (a) object-centered, in which encoding is specified relative to an object; (b) earth-centered, in which encoding is specified with respect to global latitude/longitude and north/south; and (c) encoding referenced to the local environment (as delineated by boundaries, landmark arrays, etc.). These allocentric frames have been identified in a variety of taxa, including insects, birds, and mammals (Burgess 2006; Lohmann et al. 2008; Nitz 2009).

An open question in navigation concerns whether and/or how integration might take place across allocentric reference frames. This can be examined at the level of the neural substrate and/or at the behavioral level. In the literature on egocentric integration, Anderson et al. (1985) have proposed that interactions between reference frames occur whereby neurons encoding one frame modulate the activity of those in another through a process known as gain field modulation. For example, in monkeys, neurons in parietal cortex that respond to the position of a visual stimulus on the retina are modulated by how the head is oriented with respect to the torso, and their activity thus reflects both parameters. An analogy in the allocentric domain might be found in rodent hippocampal place cells whose response to a boundary is modulated by how that boundary is oriented with respect to direction (Jeffery, this volume).

At the behavioral level, we often see examples of choice rather than integration. Cue dissociation experiments suggest that animals can plan navigational trajectories in more than one reference frame. There are several strategies to utilize the memory of a location encoded by multiple reference frames. An animal could simply pick one frame from the list of remembered frames, or it could orient to frames in a fixed hierarchy of preferences, as seen in juncos and black-capped chickadees, respectively (Brodbeck 1994). An animal could also construct a majority strategy, where the animal chooses the location indicated by the majority (2 of 3) of the remembered frames of reference, with no preference for one frame over another, as has been shown in squirrels (Gibbs et al. 2007; Waisman and Jacobs 2008).

Animals that undertake long-range migration often use an earth-centered allocentric reference frame, but animals with local ranges may have reference frames that are defined by local environmental features. In rodents, data from single neuron recordings suggest that encoding of local environments is metric; that is, it contains information about distances and directions (Jeffery, this volume). An unknown question concerns how independent local reference frames

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

(e.g., patches in a heterogeneous foraging environment or, as in humans, different rooms in an office block) can be related to each other, and whether these between-room relationships are metric or whether they are simply topological (reflecting adjacency relations only). Furthermore, we do not know whether animals are able to represent environments in which they are not currently placed and whether (if so) they do this in a strictly metric way. Recently, research has begun on how spatial maps at local levels can be interconnected and coordinated to allow for navigation in large and complex (and changing) natural environments (Derdikman and Moser 2010), and pursuing this issue is a clear challenge for the future.

Efficient navigation in three-dimensional space requires computing and integrating both horizontal and vertical position. The computation of position in two dimensions requires integration of distance and direction, or (equivalently) distance in two orthogonal directions, in a trigonometric-like fashion. How this is done remained mysterious until recently, when studies of the neural representation of space in mammals revealed the existence of grid cells and so-called conjunctive cells, whose activity seems to combine distance and direction in the horizontal plane (Fyhn et al. 2004; Sargolini et al. 2006). The next question concerns whether the third dimension (i.e., the vertical dimension) is integrated in a similar fashion, using neurons sensitive to distances in all three dimensions. Until now, there are very few experimental data on this. Ecological considerations suggest that many species—those that swim, glide, fly, or climb—would benefit from an integrated three-dimensional representation, because they move freely in volumetric space. On the other hand, the theoretical considerations outlined by Jeffery (this volume) suggest that this would be computationally expensive, so that a modified two-dimensional representation might have to suffice. Resolution of this remains a matter for future research, and comparative studies will be particularly important here.

Using the Navigation Toolbox: Decision, Planning, and Memory Processes in Navigation

As well as self-localization (a sensory process) and orchestration of movement (a motor process), navigation requires decision making, planning, and memory, because there may be more than one way to navigate to a goal, or more than one goal, or the usual route to a goal may be blocked. In cognitive science, planning is usually conceptualized as searching for a path (i.e., a solution) in a problem space. A problem space is a mental representation of the problem containing knowledge of the initial state and goal state as well as possible intermediate states. Our discussion below centers on the questions of how to define decision making and planning in the context of animal navigation. We consider these processes in light of the contributions made by elements in the navigation toolbox.

From “Animal Thinking: Contemporary Issues in Comparative Cognition,“ edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Decision Making versus Planning

It is useful to distinguish between decision making and planning, although these are often conflated in discussion. Decision making can be defined as choosing between alternative options and, depending on the exact definition, may not require any form of internal representation at its simplest level (for more detailed discussions on definitions of decision making, see Seed et al., this volume). Decision making could thus, in principle, be supported by simple processes in Levels 1 and 2 of the navigation toolbox. By contrast, planning occurs if this choice is informed by simulated future states of the system. In the context of navigation, this means that an animal has to form expectations about future payoff if it makes a particular spatial choice. At its most sophisticated level, navigational planning could potentially involve expectation or simulation of future *location*. The operation of such simulations could be said to involve recruitment of spatial constructs (Level 3) of a map-like representation (the cognitive map), and the question of whether nonhuman animals can do this is so far unresolved (see, however, Menzel, this volume). Next, we discuss what kinds of navigational decision making could take place without the need for a map, and what kinds would need true map-dependent planning.

Mapless Navigational Choices

Many quite sophisticated navigational behaviors can be orchestrated without the need for activation of a cognitive map. Navigating along a well-known route, for example, requires choosing between different options at intersection or choice points. These choices, however, can be informed solely by recognition-triggered responses or servomechanisms (i.e., mechanisms that do not require reference to a planning process as discussed here) nor to any kind of map-based representation. Navigation along a route can therefore be explained by only assuming spatial primitives (i.e., Level 2 of the navigation toolbox) such as landmarks or panoramas to identify a specific choice point, together with an associated local heading to inform about the required movement direction. In the context of this discussion it is important to stress that such choices can therefore be explained without internal simulations of future states and as such, they constitute decision making but not necessarily planning. When these decisions are automatic and not influenced by any representation of a goal, they are said to be *habits*.

Planning, in contrast, requires goal-directed actions. Much of the traditional support for the distinction between habits and goal-directed actions discussed by Dickinson (this volume) comes from studies of rats in mazes and runways. Habits continue automatically even when their outcome is devalued, whereas goal-directed behavior is sensitive to the value of its outcome. For example, if the animal is satiated on the food that it will find at the end of its trip, goal-directedness would mean not setting out, or setting out in search of a different

From “Animal Thinking: Contemporary Issues in Comparative Cognition,“ edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

goal. Good evidence that some small-scale spatial behavior has this property has been obtained in laboratory studies with a few species (Dickinson, this volume). To what extent does behavior of other species in more naturalistic spatial tasks provide evidence that it is goal directed in the same sense? The behavior of honeybees, which routinely travel among multiple goals in a flexible way, indicates that bees choose to fly either along novel shortcuts to one of the goals or to apply the memory for a flight direction that would have taken them to the intended goal if they had not been transported to an unexpected location. Honeybees also choose between the information they receive from dance communication and their own experience from former foraging flights. After they have made such a decision they can correct themselves if the outcome did not meet the expectations, and fly along a novel shortcut to the other location, a behavior that meets our definition of planning (see also Menzel, this volume, who interprets these findings as support for planning in relation to a cognitive map).

Navigational Planning Requiring a Cognitive Map

Which experimental paradigms in animal navigation research address planning processes that require a cognitive map? It is generally agreed that to confirm the operation of a cognitive map, it is necessary to demonstrate behavioral planning that has a spatial component (e.g., showing that an animal can compute a novel shortest-path route to a goal).

An example of the minimal test of map-based planning is a route-planning experiment in vervet monkeys (Gallistel and Cramer 1996), in which baited locations are arranged in a diamond shape. When starting from the lower corner, the shortest possible path depends on whether the animals are required to return to the start location. The most efficient path to the upper corner is a zigzag route. If, however, a monkey intends to return to the starting position, because it was baited after the monkey left it, a different route is optimal, resembling a diamond in this traveling salesman task. Choice at one of the baited locations is reportedly influenced in a flexible way by options that are present only later in the navigation task (i.e., the absence or presence of a food reward at the start location). In other words, animals use memory of the options and information about their current state to take different paths according to the circumstances (for a related paradigm in humans, see Wiener et al. 2008).

Experiments like these certainly seem suggestive of the operation of map-like planning processes. However, in situations that require training about each possible path, as in many experiments on small spatial scales (such as in the experiments described above), it seems difficult to provide cognitive map-like knowledge without conditional discrimination training that, for example, in the presence of cue *A*, path *A* is the most profitable, in the presence of cue *B*, path *B* is, and so on. If the animal then chooses appropriately, depending on whether *A* or *B* is present at the start, planning cannot automatically be assumed. Because

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

Gallistel and Cramer did not describe how their monkeys were trained, their experiment is potentially subject to this objection.

Experiments like these, in which animals can be shown to make decisions informed by the activation of internal spatial representations, are needed to demonstrate the operation of cognitive maps convincingly. The difficulty is that such demonstrations must take place in a single probe trial, because from the second trial onward, rapidly acquired associative reinforcement processes could, in principle, explain successful navigation. However, one-trial processing is difficult to show in animals, because novelty responses often override their goal-directed inclinations on the first probe trial. Thus the question of whether nonhuman animals use map-based planning in navigation is still not fully resolved.

An alternative approach would be to observe internal cognitive representations directly and, indeed, some neurobiological studies are beginning to provide evidence suggestive of spatial simulation processes. Recently, van der Meer and Redish (2009, 2010) provided evidence of this from electrophysiological experiments investigating decision making and planning in navigating rats. The animals were trained to run loops on an elevated track to receive food reward. Between sessions, the rewarded side is varied such that at the beginning of each session, rats were uncertain about the correct choice. During this period of relative uncertainty, rats paused longer at the critical choice points than at other choice points. Moreover, while pausing at the choice points, sequential activation of place cells with place fields around the choice point may be observed (Johnson and Redish 2007). Van der Meer and Redish interpret these findings as rats representing future locations (i.e., simulating the outcome of a spatial decision), which is a crucial aspect of spatial planning. The rats' behavior here seems to be an example of the classic phenomenon of "vicarious trial and error" (VTE), in which animals spend time sampling the cues associated with the options in a difficult discrimination. As discussed by Seed et al. (this volume), one criterion for true decision making is that the latency to choose between options is greater than to accept either one alone. Evidence for behavioral and/or neurological VTE suggests that observations of spatial behavior may provide evidence for decision making according to this criterion.

Because of the difficulties in demonstrating convincingly cognitive mapping in small-scale laboratory settings, it has been suggested that navigation experiments in large-scale spaces, such as the animal's natural habitat (or a city-like environment in the human case), may be a more promising arena for collecting planning relevant data. For example, Wiener and colleagues used large-scale virtual environments to investigate route-planning behavior in humans (Wiener and Mallot 2003; Wiener et al. 2004). They demonstrated that the hierarchical organization of spatial memory influenced participants' route choice behavior and proposed a planning mechanism that uses spatial information at different levels of detail simultaneously. Analogous data might, in principle, be obtained by observing how the paths animals take through

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

their familiar home ranges vary across days and seasons. Short-term changes in routes may appear to reflect planning, for instance, to visit certain kinds of trees when their fruits are available. However, it may be difficult to infer anything about spatial knowledge and plans from such data, which are typically gathered without experimentally manipulating the animals' knowledge or goals (see Janson and Byrne 2007).

In summary, one of the challenges for future navigation research is to develop controlled experiments to investigate spatial planning. According to the definition introduced above, planning requires the animal to form internal representations of simulated future states, given a particular movement decision scenario. Furthermore, spatial (i.e., map-based) planning requires these simulated states to incorporate spatial information, as shown, for example, in the ability to calculate shortest or least-effort paths, or to find novel detours that reveal a knowledge about the spatial relations of connected spaces. One possible way to demonstrate that animals/humans do, in fact, form such representations might be to demonstrate behavioral or physiological responses to the violation of their expectations (which may be conceptualized as "surprise"). An alternative approach, which is just beginning, is to observe the underlying neural processes directly. It may be easier to see future simulations (sometimes called "preplay") in neural activity, although interventional studies would be needed to show that these processes are causally related to navigational behavior. Emerging technologies, such as optogenetic manipulation of neural circuits, will be very exciting in this regard.

Concluding Remarks and Future Directions

Our review of the current status of animal navigation research has focused on overarching principles that arise independently of a given animal substrate. Comparing across a range of species whose navigational competence varies from simple to sophisticated, we note that complex and more phylogenetically recent abilities appear to be synthesized from simple, phylogenetically older ones. Using this observation as a starting point, we organized the panoply of navigational behaviors loosely into a hierarchical framework—the navigation toolbox—which is a collection of processes that can support, either alone or collectively, navigational behaviors of varying complexity. This collection is organized such that elements in higher levels are synthesized from elements in lower levels, acquiring new semantic content in the process.

We argue that the ability to integrate across sensory modalities and semantic classes, so as to generate semantically new information (such as "position"), expanded during brain evolution. It reaches a peak in the human ability to represent spatial information symbolically, using language and mathematics, deriving entirely new semantic content in the process (e.g., multidimensional space, non-Euclidean space, complex space).

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.

The task now, in comparative cognition, is twofold. First, at the level of individual taxa, we need to know how a given element in the navigation toolbox is implemented: in route following or beacon homing, for example, which sensory modalities, and spatial primitives, are recruited to enable decision making? Second, more broadly, we need to understand the underlying neural principles behind certain types of spatial computation. For example, how does the brain do "trigonometry," and is this process the same across all taxa or have multiple solutions evolved independently?

Finally, there remains the open question of whether any animal other than humans make use of the complex internal representation of spatial relations, which is sometimes called a "cognitive map," and even to what extent humans do so. To address this question requires a combination of more sophisticated behavioral experiments, controlling for the possibly occult operation of more primitive processes, and neurobiological studies capable of probing the existence of putative simulation phenomena, such as neural sequence "preplay." Unequivocally, we must be able to undertake sophisticated interventional experiments to disable the processes in question and test hypotheses about spatial representation.

From "Animal Thinking: Contemporary Issues in Comparative Cognition," edited by R. Menzel and J. Fischer. Strüngmann Forum Report, vol. 8, J. Lupp, series ed. Cambridge, MA: MIT Press. ISBN 978-0-262-01663-6.