Hippocampal spatial view cells, place cells, and concept cells: View representations

Edmund T. Rolls

Abstract

A commentary is provided on issues raised in the Special Issue of Hippocampus (2023) on hippocampal system view representations. First, the evidence for hippocampal and parahippocampal spatial view cells in primates including humans shows that the allocentric representations provided by at least some of these cells are very useful for human memory in that where objects and rewards are seen in the world “out there” is a key component of episodic memory and navigation. Spatial view cell representations provide for memory and navigation to be independent of the place where the individual is currently located and of the egocentric coordinates of the viewed location and the facing direction of the individual. Second, memory and navigation in humans are normally related to the visual cues encoded by spatial view cells that define a location “out there” such as a building, hill, and so forth, not to an unmarked place without local cues and identified only by distant environmental/room cues. Third, “mixed” representations, for example of particular combinations of spatial view and place, can arise if the training has been for only some combinations of place and view, for that is what can then be learned by the hippocampus. Fourth, rodents, with their much less good visual acuity (<1 cycle/° in rats, compared with ~60 cycles/° for the human fovea), and rodents’ very wide viewing angle for the world (~270°) might be expected, when using the same computational mechanisms as in primates, to use widely spaced environmental cues to define a place where the rodent is located, supported by inputs about place using local olfactory and tactile cues. Fifth, it is shown how view-point dependent allocentric representations could form a view-point independent allocentric representation for memory and navigation. Sixth, concept cells in humans and primates with connectivity to the hippocampus are compared.

KEYWORDS
concept cells in humans and macaques, episodic memory, hippocampus, navigation, parahippocampal scene area, place cells, spatial view cells

1 | INTRODUCTION

This is a commentary on issues important in understanding hippocampal function that are raised by papers in the Special Issue of Hippocampus (2023) entitled “Hippocampal system neurons encoding views in different species” (Alexander et al., 2023; Corrigan et al., 2023; Donoghue et al., 2023; LaChance & Taube, 2023; Lee et al., 2023; Quiroga, 2023; Rolls, 2023b; Ryom et al., 2023; Wang et al., 2023; Wirth, 2023; Yang et al., 2023; Zhu et al., 2023).
2 | VIEW CELL SPATIAL COORDINATES: ALLOCENTRIC VERSUS FACING DIRECTION VERSUS EGOCENTRIC

Several contributions to the Special Issue of Hippocampus (2023) provide evidence that many primate hippocampal neurons respond to the location being viewed “out there,” not the place where the individual is located (Corrigan et al., 2023; Rolls, 2023b; Wirth, 2023; Zhu et al., 2023). Complementary evidence is also available for humans, in that in another paper in this Special Issue of Hippocampus, it was found that significant numbers of spatial view cells but not a significant number of place cells were found in the medial temporal lobe in a VR navigation task to the remembered location in a scene of a Treasure Chest (Donoghue et al., 2023). The spatial coordinates of these view cell representations are considered next. A key test for allocentric encoding is whether a neuron responds to a location in a scene independently of the head-based (craniotopic) location of the scene. If a neuron responds in head-based coordinates, then it is egocentric.

In Section 2.1 the background is set by reviewing the evidence that hippocampal spatial view cells have allocentric representations. In Section 2.2, the role of facing location in the encoding of these allocentric neurons is considered (Mao et al., 2021; Zhu et al., 2023). In Section 2.3, evidence for some egocentric encoding in the hippocampal system of rodents (Alexander et al., 2023; Wang et al., 2023) is discussed. In Section 3, the case is made that the issue of the coordinate framework for spatial representations, allocentric versus egocentric, is separate from the interesting issue of view-dependent spatial representations, which relates to which side a scene is viewed from by an observer (Wirth, 2023).

Allocentric spatial coordinate frameworks are with respect to the head or body, for example to the left of the head in craniotopic coordinates. An egocentric coordinate framework is important in primates for visually guided reaching and grasping objects in nearby space (Andersen & Cui, 2009; Gambarini et al., 2020; Passarelli et al., 2021; Rolls, Deco, et al., 2023a; Rolls, Deco, et al., 2023b), and in rodents might be used for obstacle avoidance such as a boundary on the right. An allocentric spatial coordinate framework is in terms of locations in the world, and is independent of egocentric coordinates in that an allocentric representation is independent of whether the location in space being viewed is to the left or right of the head. Allocentric spatial representations are important for remembering where objects or people are in the world (e.g., an episodic memory), and for navigation in the world (O’Keefe, 1990; O’Keefe & Nadel, 1978; Rolls, 1999, 2023a, 2023b).

2.1 | Allocentric spatial view

The hippocampal and parahippocampal gyrus spatial view cells described by Rolls and colleagues (Feigenbaum & Rolls, 1991; Georges-François et al., 1999; Robertson et al., 1998; Rolls, 1999; Rolls, 2023b; Rolls et al., 1989, 1998, 2005; Rolls & O’Mara, 1995; Rolls, Robertson, & Georges-François, 1997; Rolls & Xiang, 2005, 2006) have allocentric representations as shown by three types of evidence.

First, for neurons that respond to the location on a screen in which a stimulus is shown, the majority of neurons respond to the allocentric location on the screen when the screen is moved left or right or up or down to different egocentric locations with respect to the macaque (Feigenbaum & Rolls, 1991). Further experiments showed that the neurons did not respond in retinotopic coordinates (Feigenbaum & Rolls, 1991).

Second, when a location in a scene is viewed with different head directions, spatial view neurons respond to the same location in a scene independently of the head direction (provided of course that the eyes can fixate the location in the scene; Georges-François et al., 1999; Rolls, Robertson, & Georges-François, 1997). This type of evidence is illustrated in figure 2 of Rolls, Robertson, and Georges-François (1997) and figures 3, 5, and 6 of Georges-François et al. (1999) (see also figure 2 of Rolls, Deco, et al., 2023b), with versions of these figures with the firing rates shown in color available at https://www.ocn.org.

Third, when a macaque is freely locomoting in a large open laboratory that provides a rich spatial scene on all its walls, the firing rate of hippocampal spatial view neurons and the information encoded by hippocampal spatial view neurons is about the allocentric location in space being looked at, and is relatively independent of place, head direction, and eye position, or the egocentric location in space of what is being viewed (Georges-François et al., 1999; Rolls et al., 1998), with movies to illustrate this available at https://www.ocn.org (Rolls, 2023b; Rolls & Wirth, 2018).

In this analysis, it is important to be clear about different coordinate frameworks. If the locations of the stimuli are in coordinates relative to the head (craniotopic) or body, these representations are in an egocentric framework. If the representations of the locations of things such as parts of a scene are fixed relative to other parts of a scene, and are encoded provided that location in the world is viewed, independently of the location relative to the head, body, or retinal position, then this is an allocentric representation. That type of allocentric representation is provided by hippocampal spatial view cells, given the evidence described here and by Rolls, Deco, et al. (2023b). As noted in Section 3, the allocentric representation provided by spatial view cells may be viewpoint dependent, with the view of the scene seen from one side of the scene what is stored and remembered. In addition, as shown in Section 3, an allocentric representation may be formed that is viewpoint independent.

In a virtual reality (VR) navigation task, further evidence was found that many macaque hippocampal neurons responded to the location being viewed, and not the place where the individual was located (Corrigan et al., 2023). They were not able to determine what coordinate framework (allocentric or egocentric) was being used. In humans, the ‘target neurons’ found in the medial temporal lobe in the Treasure Hunt task (Donoghue et al., 2023; Tsitsiklis et al., 2020) can be considered as being responsive to spatial views. The responses of these spatial-target cells are consistent with allocentric encoding, since they responded to a specific viewed spatial location in the
environment, from multiple different subject places, such that the egocentric orientation of the subject with respect to the target location varied.

2.2 Facing direction

Angelaki and colleagues have also presented evidence that many hippocampal neurons in primates respond to view, and not to the place where the individual is located (Mao et al., 2021; Zhu et al., 2023). These neurons are described as mainly responding to the allocentric location in space towards which the macaque is facing (Mao et al., 2021). However, typically one does face in the direction towards which one is navigating, and that can affect and limit the data available for analysis. So specific tests are needed to determine whether neurons respond to the part of the scene that is being looked at independently of facing direction. When specific tests of this type are performed for hippocampal spatial view cells, it is found as noted above that they respond to the location in the scene being looked at, relatively independently of head direction. This is illustrated in figure 2 of Rolls, Robertson, and Georges-François (1997) and figures 3, 5, and 6 of Georges-François et al. (1999) (see also figure 2 of Rolls, Deco, et al., 2023).

Facing direction is computationally not very useful for scene/episodic memory, for if facing direction is used in the memory of where an object or a reward location in a typical spatial environment is located, then that memory in not invariant with respect to the head direction. That is, if facing direction was the information that formed part of a memory of where an object or reward was in a spatial environment, then that memory would only be useful when the individual had exactly the facing direction present when the memory was stored, whereas if we are searching for a location in a scene, we may well find the searched for item using eye position changes and saccades round the scene to one side or the other of where we are facing. If all we had were facing direction neurons, we could only find objects in a scene by moving our heads to search the scene with head movements, whereas much visual search is made using saccades to different locations in a scene (Rolls et al., 2003). Moreover, hippocampal spatial view cells respond well when the spatial field in the scene requires different eye positions (Georges-François et al., 1999; Robertson et al., 1998; Rolls, Robertson, & Georges-François, 1997). Thus for memory and navigation, facing direction is not computationally very useful, whereas the viewed location in an environment as encoded by spatial view cells that is invariant with respect to the head direction is very useful (Rolls, 2023a).

Of course marmosets are a special case, as they use fast head movements to look at viewed locations. So in marmosets, allocentric-facing direction towards a location in a scene would be useful in memory and navigation, if what is encoded is invariant with respect to the place where the individual is located.

The point that eye movements are important in understanding hippocampal function has been made (Zhu et al., 2023), but it is useful to consider why eye movements are closely related to hippocampal function. The reason that I suggest is that saccadic eye movements are made in primates to search scenes for objects or rewards or other desired targets, and this is well illustrated by the eye movements evident in the movies of freely locomoting macaques made available with the contribution to the Special Issue of Hippocampus on View Representations by Rolls, (2023b). So eye movements per se are not what is remembered by the hippocampus, or used in navigation, but information about the allocentric locations in space is important for understanding hippocampal function in primates including humans, and viewing these locations of course involves eye movements.

2.3 Egocentric location

In the investigation in which the screen was moved relative to the macaque, a small proportion of neurons did provide evidence for egocentric encoding, that is location relative to the head in craniotropic coordinates (Feigenbaum & Rolls, 1991).

As emphasized elsewhere, in order to separate spatial view from place representations, and allocentric from egocentric encoding, tests are essential for the same set of views seen from different places in a factorial design, and of the same set of views seen with different head directions in a factorial design, in order to test these hypotheses (Rolls, 2023a, 2023b). Those designs were utilized in the investigations of primate spatial view cells (Feigenbaum & Rolls, 1991; Georges-François et al., 1999; Robertson et al., 1998; Rolls, 2023b; Rolls et al., 1998, 2005; Rolls & O’Mara, 1995; Rolls, Robertson, & Georges-François, 1997).

It is worth noting that any hypothesis that hippocampal view cells encode information in an egocentric spatial framework would render the hippocampus useless for episodic memory of where an object or reward had been seen. Imagine storing some food or other reward or object in egocentric coordinates at 45 degrees left of head center in egocentric head-based (craniotropic) coordinates. When one wanted to find the food or other reward or object later, then wherever one was, and whatever one’s head direction, one would look 45 degrees to the left, and almost never find the food or other reward or object. One would look for almost ever.

Egocentric view representations are shown by some rodent hippocampus-related neurons. For example, neurons in the retrosplenial cortex show egocentric coding of the position of boundaries in relation to the rodent (Alexander et al., 2023), and neurons in the rodent lateral entorhinal cortex represent the angular bearing of objects and boundaries in an egocentric frame of reference (Wang et al., 2023). Neurons with egocentric encoding of this type might be suitable for obstacle avoidance (rather than remembering where objects are in the world, or for allocentric navigation). An example might be detection of a boundary located on one side of the head.

Thus, it is argued that allocentric encoding of locations in scenes that are relatively independent of the place where the individual is located and of head direction is what is needed for any useful episodic memory of what happened where; and that is not only a key function of the hippocampal system (Rolls, 2023a), but is exactly what is encoded by hippocampal allocentric spatial view neurons (Feigenbaum & Rolls, 1991; Georges-François et al., 1999; Robertson...
Spatial View Cells Provide an Allocentric Representation That May Be View-Point Dependent or View-Point Independent, and That Is Useful for Memory and Navigation

3.1 Spatial view cells may provide an allocentric representation even on the other side of a scene: View-point dependence versus view-point independence of spatial representations

If in humans and primates, the representation is to the other side of the scene, then of course when the viewing point changes in this way the left–right relations of the component locations in the scene will appear as reversed (Wirth, 2023). However, the representations can be allocentric in that they do not depend on the craniotopic coordinates with which the scene is viewed, but nevertheless view-points on either side of the scene will have the left–right relations reversed.

However, it is shown next that what has been learned on one side of the scene by the overlap of the representations of nearby locations, in for example a continuous attractor network (Rolls, 2023a), can still be useful for navigation if it contains evidence about the allocentric, world-based, relations between the different parts of the scene. In this sense, even when the viewing point changes by going behind the scene, the representation is still really about the relative locations of spatial positions in the environment, and so is in fact allocentric in that sense. The following describes this. But, it is important not to confuse viewpoint-dependent encoding with egocentric encoding, because the scenes when viewed from either viewpoint are in allocentric coordinates, in that they are represented by spatial view neurons that encode in allocentric and not in craniotopic egocentric spatial coordinates (the definitions of allocentric vs. egocentric coordinate frameworks for spatial representations are defined at the start of Section 2).

Let us consider further how the parts of a scene are “stitched together” in primates. When a primate including humans with spatial view cells is on one side of a scene, then the nearby parts of the scene become associated together by co-firing of neurons that produce synaptic modification based on how much co-firing there is, and therefore the distance between the spatial view fields (some sort of Gaussian-shaped receptive field spatial sensitivity is assumed, as shown for primate hippocampal spatial view cells; Rolls, 2021b, 2023b; Rolls et al., 1998). That forms a continuous attractor network for spatial view in primates (Rolls et al., 2008; Rolls & Stringer, 2005; Stringer et al., 2005). Exactly the same computational mechanism in rodents potentially builds a continuous attractor network for place fields (Samsonovich & McNaughton, 1997; Stringer et al., 2002).

Now a continuous attractor network (Rolls, 2023a) can be regarded as a type of map or chart, for nearby locations in the map are represented by co-firing neurons, and one can imagine moving through a series of nearby items in a trajectory, which may be thought of as a type of navigation (Rolls & Stringer, 2005; Stringer et al., 2002, 2005). Indeed, if one continued the navigation to get behind the scene, the order in which the landmarks in the scene are linked would be the same (e.g., house then lake then hill), so the representation of the scene would be allocentric, world based, in that the parts of the map would be linked together in the correct order independently of which side of the scene it was viewed from. The view-dependent relations, for example, which landmark was on the left, would change, but not the allocentric world-based representation in which the landmarks are linked in the same order regardless of the exact viewpoint. Thus, spatial view cells could be linked in a continuous attractor network to provide an allocentric representation of the structure of the scene, which encodes the relations between the parts in a viewpoint-independent way. That allocentric representation could be useful for navigation from landmark to landmark (Rolls, 2021b).

Spatial view neurons are thus allocentric in that they represent features in a scene stitched together in the correct spatial relationship with respect to each other, in a continuous attractor network as described above. Some of the clear evidence that the structure of the scene is encoded is that view neurons start to respond in VR to a part of the scene towards which an eye movement is being performed even before the view of the scene has appeared on the VR screen (Wirth et al., 2017). Remarkably similar evidence for predictive encoding is that hippocampal spatial view neurons respond to a location in a scene when a macaque moves the eyes in the dark to look towards that location in a scene, even though the scene is hidden by curtains and it is dark (Robertson et al., 1998). That is evidence that there is a hippocampal representation that has structural information about the spatial relations of parts of a scene, and that predictive encoding is a feature of primate hippocampal function.

All of this is consistent with the theory that spatial view neurons are in a sense viewpoint dependent, in that they have fired at a scene that is being viewed, for example, when the individual is on one side of the scene. In such a representation Part 1 of the scene may be to the left of Part 2 of a scene, Part 2 may be to the left of Part 3, and so forth, and that relationship will remain as long as the individual is on that side of a scene. In this sense, hippocampal spatial view neurons may have an allocentric world-based representation, based on the evidence above, which may nevertheless be stored with the scene viewed from one side, in what can be described as a viewpoint-dependent spatial framework. However, if one moves to the other side of the scene, a continuous attractor network representation already built could still reflect the closeness of the landmarks in a scene, in that Part 1 would be close to Part 2 but not Part 3, and so forth, and that could still potentially be used in terms of the closeness of parts of a scene, even though viewed from the other side the
left/right relations will be reversed consistent with a viewpoint-dependent framework. The closeness of the parts of a scene will still be useful for truly allocentric navigation independently of which side the scene is viewed from.

My point can be further illustrated by a different example. We all know what clockwise rotation is. But now imagine going behind the clock, and from that viewpoint the observer will see the hands of the clock rotating anticlockwise. The world-based representation of the clock hands in terms of itself is unchanged, but the viewpoint-based direction reverses when the observer moves behind the clock face. And that occurs independently of our egocentric coordinate framework, that is whether the clock is to the left or right of us in craniotopic egocentric coordinates. So, we must be careful to distinguish egocentric coordinate frameworks such as craniotopic spatial frameworks, from viewpoint-based reference frameworks. In that sense, hippocampal spatial view cells code in an allocentric framework, even though the viewpoint will alter the left–right ordering, the direction of rotation, and so forth.

In fact, this issue has arisen before, when we discovered neurons in the cortex of the macaque superior temporal sulcus (STS) that code for the rotation of objects including of the body on the head in a world-based framework, in that some neurons there code for the rotation of the head on the body in the world based framework of the body. For example, neurons coded for clockwise rotation of the head on the body independently of inversion which reverses the optic flow, but does not reverse the rotation of the head on the body in its own framework (Hasselmo, Rolls, Baylis, & Nalwa, 1989). This is termed an object-based framework (Rolls & Stringer, 2006). In the same way, scenes can be represented in a world-based framework, which may appear different when viewed from different viewpoints.

Thus it is proposed that spatial view cells form allocentric representations of scenes, even though the left–right viewpoint-dependent relationships are different on different sides of the scene when the viewpoint changes. The spatial representation provided by spatial view cells is allocentric in that it is not craniotopic, even when it may look “egocentrically, first-person” different to a viewer who is on different sides of a scene, in what should most clearly be described as a view-dependent allocentric representation. Moreover, that allocentric representation even though it may have viewpoint dependence may however also have important viewpoint-independent topological properties, for example, about the closeness of the different parts of the scene which is important for navigation (Rolls, 2021b, 2023a).

With these allocentric representations, primate including human spatial view cells are ideal for episodic memories of where rewards and objects have been seen in spatial scenes (Rolls et al., 2005; Rolls & Xiang, 2005, 2006), and for navigation (Rolls, 2021b, 2023a).

3.2 Human and primate navigation is normally towards allocentric locations that are being viewed

Humans normally navigate to a viewed object or reward in a scene such as a building, and nonhuman primates might locomote to particular locations in the viewed environment with local cues where there are sources of food, water, shelter and a place to sleep, and so forth. In these cases, identifiable environmental features identify each location, and spatial view cells encode the features in these locations, and their spatial overlap with features of nearby locations. The mechanism proposed is the learning of spatial associations in a continuous attractor network (Rolls et al., 2008; Rolls, Deco, et al., 2023a; Rolls & Stringer, 2005; Stringer et al., 2005) using spatial view cells (Donoghue et al., 2023; Georges-François et al., 1999; Robertson et al., 1998; Rolls, 2022; Rolls et al., 1998; Rolls, Robertson, & Georges-François, 1997; Rolls & Wirth, 2018; Tsitsiklis et al., 2020; Wirth et al., 2017) in the parahippocampal scene (or place) area, which in turn connects to the hippocampus to provide the “where” component of episodic memory (Rolls, 2023b; Rolls et al., 2022b; Rolls, Deco, et al., 2023a). A model of navigation towards viewed locations in scenes with intermediate viewed locations that use spatial view cells, and that can include whole body motion/body turn (“speed”) cells (O’Mara et al., 1994), has been described (Rolls, 2021b).

In this context, the type of navigation to an unmarked place without local visual cues and relying on relatively distant environmental cues (as in the Morris water maze; Morris et al., 1982) is unusual in humans and other primates. It is like looking for a needle in a haystack using distant hills as navigational cues. What is often studied in rodents as a model of navigation without any local cues as in the Morris water maze may not be at all commonly how humans and other primates navigate. Spatial view cells provide the mechanism often used for navigation in humans and other primates which is typically towards visible landmarks such as a building, the top of a hill, and so forth (Rolls, 2021b). Navigation may thus be performed in primates including humans (Rolls, 2021b, 2023a) in very different ways to those utilized in rodents (McNaughton et al., 1996; Moser et al., 2015, 2017).

3.3 The invariance of spatial view cells is useful for episodic memory and navigation

The primate hippocampal and parahippocampal spatial view cells respond to a location in a viewed scene relatively invariantly with respect to the place where the primate is located, to head direction, and to eye position, as shown by experiments in which a location in a scene is viewed from different places with different head directions and different eye positions (Georges-François et al., 1999; Rolls et al., 1998; Rolls & O’Mara, 1995; Rolls, Robertson, & Georges-François, 1997).

In this situation, the key property of invariance with respect to the exact view of the location is a great asset, for the view representation is relatively invariant even with the small differences in the view that occur from different places on the same side of the scene. Nearby locations in a scene are typically viewed close together in time, and this temporal information can also help to build representations of scenes (Rolls, 2021a).

Moreover, it is rather difficult to establish one’s exact place from the slightly different views of distant scenes that are present...
from different places on the same side of the scene, that is with viewpoints that are on the same side of the scene, as described in Section 3.1.

The allocentric view representation provided by spatial view cells in the hippocampus can then be associated with objects or rewards at those world-based locations to implement episodic memory (Rolls, 2022; Rolls et al., 2005; Rolls & Xiang, 2005, 2006).

4 | MIXED ENCODING BY HIPPOCAMPAL NEURONS

Some primate hippocampal neurons have been described as having mixed encoding, in that they respond to, for example, a combination of the spatial view and the place where the macaque is located (Tan et al., 2021; Wirth et al., 2017). One way in which such mixed selectivity can arise is if the training or experience of the individual has been unbalanced, with for example, certain views particularly likely to be seen from only certain places. In the model of spatial view cells and place cells shown in Figure 1 (De Araujo et al., 2001), great care was taken to provide equal training across all views, places, and head directions; that is, every spatial view was seen equally from all places; and for every place, all views were seen with all head directions. However, if particular views were seen from only certain places, then the neurons would learn to respond to those views only from those places. Indeed, that is a property that is likely to be key to understanding hippocampal function, that hippocampal neurons learn to respond to particular combinations of the inputs that are present at the same time, and this is ideally suited to episodic memory (Rolls, 2023a).

Consistently, in rodents, grid cells have place fields that respond best for head directions in which the rat can run, in for example a hairpin maze, as that reflects the statistics of the inputs received (Derdikman et al., 2009). A similar argument applies to whether place cells are invariant with respect to head direction.

This emphasizes the importance of a factorial design when evaluating whether neurons encode for spatial view, for place, or for a mixture of the two. The data analyzed must be carefully balanced to ensure that, for example, every spatial view is seen equally from every place in the data analyzed, otherwise a misinterpretation is a risk (Tan et al., 2021). Exactly that factorial type of design was utilized in the data analyzed for hippocampal spatial view neurons where it was shown that the majority of spatial view neurons encode information about spatial view, but relatively little information about place (or head direction, or eye position; Georges-François et al., 1999; Rolls et al., 1998). There may of course be other types of neuron in the primate hippocampus, including some place cells (Rolls & O'Mara, 1995; Wirth et al., 2017).

5 | SCENES REPRESENTED IN THE PARAHIPPOCAMPAL SCENE AREA AND HIPPOCAMPUS IN HUMANS AND OTHER PRIMATES ARE ENCODED BY A VENTRAL STREAM VISUAL PATHWAY USING VISUAL FEATURE COMBINATION NEURONS

Recent evidence has revealed a ventral stream visual path in humans from V1 > V2 > Ventromedial visual areas 1–3 (VMV1–3) > the medial parahippocampal cortex PHA1–3 (Rolls, Deco, et al., 2023b;...
which corresponds to TH in macaques where spatial view cells are found (Feigenbaum & Rolls, 1991; Georges-François et al., 1999; Robertson et al., 1998; Rolls et al., 1989, 1998, 2005; Rolls & O’Mara, 1995; Rolls, Robertson, & Georges-François, 1997; Rolls & Xiang, 2005, 2006). VMV1-3 and PHA1-3 are where in humans the parahippocampal place area (better termed the parahippocampal scene area because it responds to scenes: Epstein, 2005, 2008; Epstein & Baker, 2019; Epstein & Julian, 2013; Epstein & Kanwisher, 1998; Kamps et al., 2016; Natu et al., 2021; Rolls, 2023b; Rolls, Deco, et al., 2023b) is found (Sulpio et al., 2020). This pathway is illustrated in figure 6 of Rolls (2023b) which also shows the retrosplenial scene area (Cherstil et al., 2015; Sherrill et al., 2015) and the occipital scene area.

The implication is that the spatial view representations provided by spatial view cells are built by combinations of visual features nearby (i.e., close to the fovea) in a spatial scene. Thus, a representation of “where” is built using ventral stream processing, probably in the same way as representations of objects are built (Rolls, 2021a, 2023a) in the ventrolateral visual stream projecting to the inferior temporal visual cortex (Rolls, 2023b; figure 8 of Rolls, Deco, et al., 2023b).

In contrast, in rodents, it has usually been assumed that the visual “where” representations, used to define where place cells respond, come from the dorsal visual stream (Bicsanski & Burgess, 2018), although no detailed evidence is available on the rodent pathways.

In humans and other primates, it is proposed that the dorsal visual stream leading to the parietal cortex (Rolls, Deco, et al., 2023b) (figure 7 of Rolls, 2023b) is used primarily for the idiothetic update of locations in a scene (Rolls, 2020; Rolls, Deco, et al., 2023b), not for building the visual scene representations from ventral stream visual features (Rolls, 2023a, 2023b). The parietal areas are of course involved when actions such as reaching are made in nearby space using egocentric coordinates (Rolls, 2023a; Rolls, Deco, et al., 2023a).

6 | RODENT PLACE CELLS VERSUS PRIMATE SPATIAL VIEW CELLS

What is encoded by primate hippocampal and parahippocampal cortex spatial view cells is the location being looked at in an environment (Rolls, 1999, 2023a, 2023b; Rolls & Wirth, 2018). There is a clear influence of features in the spatial scene on the firing of primate spatial view cells.

How does this relate to place cells in rodents? There are clear similarities. Rodent place cells are influenced by the room cues, and indeed rotating the room cues alters the place encoded by a place cell so that it remains consistent with the place based on the environmental room cues (Muller & Kubie, 1987). But, what is encoded in rodents is different to what is encoded in primates. In rodents, what is encoded in contrast may be the place where the individual is located (Burgess & O’Keefe, 1996; Hartley et al., 2000; O’Keefe, 1979), as contrasted with the location in space being viewed that is encoded by spatial view cells in primates (Rolls, 1999, 2023a, 2023b).

A simple computational theory and model that accounts for how view-related information is important in both primates and rodents, but results in encoding of place in rodents and of spatial location out there in the environment in primates including humans, is shown in Figure 1 (De Araujo et al., 2001). The theory is that because primates including humans have a fovea, learning a combination of features over a small visual angle in the environment to activate a hippocampal neuron results in a spatial view cell (Figure 1). In contrast, because rodents do not have a fovea and have a very large visual field of view subtending ~270 degrees, learning a combination of widely spread out features in the environment to activate a hippocampal neuron results in a place cell (Figure 1; De Araujo et al., 2001).

To investigate the difference in what is analyzed by the primate and rodent visual systems, a scene was filtered with the contrast sensitivity function of humans and mice (Haun, 2021), with the results illustrated in Figure 2. The following points can be made. First, the field of view of the rodent is large, ~270° (135° each way from the center, with a calibration circle shown at 128° in Figure 2), whereas in humans and other primates with forward-facing eyes little is visible in fact beyond 90° from the fovea, with calibration circles provided at 64° and 128° in Figure 2, and the invisible region for primates shown as dark in Figure 2. Second, the spatial resolution for the whole of the field of view is poor in rodents, whereas humans have a very high-resolution fovea (with ~60 cycles/° for the acuity), which, as illustrated in Figure 2b when expanded, is sufficient for face recognition and reading text. The implication for the rodent is that what is analyzed by the rodent visual system might be sufficient to process a few, well spread out, visual landmarks to help form rodent place fields (or at least reset them) based on what is visible in the environment, as proposed by De Araujo et al. (2001; see Figure 1). The implication for humans and other primates is that very high resolution is available close to the fovea that can enable great detail in a scene available over a relatively small visual angle to be used to associate particular objects or faces or rewards (which as illustrated in Figure 2b might subtend 1° or less) with their location in a scene. This is completely in line with the theory (Rolls, 1989, 1999, 2023a, 2023b) and evidence (Rolls et al., 2005; Rolls & Xiang, 2005, 2006) that an important function of primate including human hippocampal spatial view cells is to participate in forming associations between objects including faces, rewards, and so forth and their location in a scene. The visual system of primates including humans thus supports this key function in episodic memory, by providing sufficient resolution at the fovea for objects, and so forth together with sufficient resolution more peripherally to represent scenes and landmarks. Further, the high visual spatial resolution of the fovea in primates that provides the ability to use fine detail in a distant scene is also very valuable for navigation that is guided by distant viewed landmarks, for it enables even somewhat similar distant landmarks to be distinguished and used as guides for navigation (Rolls, 2021b, 2023b). The high resolution close to the fovea in primates including humans can be appreciated better by expanding the image used for Figure 2b that is available in the Supplementary Material S1.
The spatial view cells in macaques are set up to form a representation of space, in that when a macaque moves the eyes in the dark to look at where the spatial view field is, spatial view neurons respond to that location in space towards which the macaque is looking (Robertson et al., 1998). Further evidence that the spatial view cells are present and are organized to form a spatial structural representation of a spatial scene as we have hypothesized and modeled (Rolls et al., 2008; Rolls & Stringer, 2005; Stringer et al., 2005) is that primate hippocampal neurons can fire in a VR task to a location towards which a saccade is made even before the scene has been shown in the VR (Wirth et al., 2017), with similar findings when the view details are hidden by curtains and in the dark (Robertson et al., 1998).

The above provides evidence for path integration for a spatial scene. Similarly, rodent hippocampal cells can perform path integration between places (McNaughton et al., 1996). However, whereas in rodents the path integration is hypothesized to be performed in the hippocampus (McNaughton et al., 1996), or medial entorhinal cortex by grid cells (Moser et al., 2017), at least a large part of the path integration in primates involves the dorsal visual stream leading to the parietal cortex, in part because the path integration in primates must include eye position, as well as head direction and place (Rolls, 2020, 2023b; Rolls, Deco, et al., 2023a).

Although spatial view cells are a feature of the cell types found in the primate hippocampus (Feigenbaum & Rolls, 1991; Georges-FIGURE 2 Legend on next page.
François et al., 1999; Robertson et al., 1998; Rolls et al., 1989; Rolls et al., 1998, 2005; Rolls & O’Mara, 1995; Rolls, Robertson, & Georges-François, 1997; Rolls & Xiang, 2005, 2006), place cells have also been described in the primate hippocampus (Rolls & O’Mara, 1995; Wirth et al., 2017).

7 TASK-DEPENDENT ACTIVITY BY PRIMATE INCLUDING HUMAN HIPPOCAMPAL NEURONS

Interesting evidence that in humans medial temporal lobe neurons can respond differently in different tasks (navigation vs. n-back memory) was presented in this Special Issue (Donoghue et al., 2023). There is some consistent evidence in macaques, as follows.

First, in a running recognition memory task in which recently presented items need to be remembered, macaque hippocampal neurons are not very responsive to objects or faces when they are presented, even though they need to be remembered (Rolls et al., 1993). (Recognition memory of this type is implemented in the perirhinal cortex [Baxter & Murray, 2001a, 2001b; Buckley, 2005; Buckley & Gaffan, 2000; Murray & Mishkin, 1998]). In contrast, when the locations of objects in a scene need to be learned and remembered, which is a hippocampal system-dependent task (Gaffan, 1994; Murray et al., 1998), neurons are found that respond selectively to objects, or to combinations of a particular object with a particular location in a viewed scene (Rolls et al., 2005; Rolls & Xiang, 2006).

Second, in a visual object to reward association task which is not hippocampus-dependent (and is implemented by the primate orbitofrontal cortex; Rolls, 2019a, 2019b; Rolls et al., 1996), macaque hippocampal neurons are not very responsive to the object-reward associations (Rolls & Xiang, 2005). In contrast, in a memory task in which the locations of rewards in a scene need to be remembered, which is a hippocampal system-dependent task, some hippocampal neurons respond to rewards and especially to spatial locations in a scene (Rolls & Xiang, 2005).

Thus in macaques, there is evidence that hippocampal neurons respond more to objects and to rewards when these need to be associated with spatial locations than when objects need to be remembered, or when reward associations of objects need to be remembered. These neuronal responses will facilitate the computations performed by the hippocampus (Rolls, 2023a).

The task dependence of hippocampal neurons in macaques is echoed by the findings in humans (Donoghue et al., 2023). However, the finding that individual human medial temporal lobe neurons can respond to different task variables in different tasks (Donoghue et al., 2023) needs further comment. One possibility is that in an associative memory system individual neurons might not reflect what can be encoded by the whole population, and sparse distributed representations do have some overlap with each other (Rolls, 2023a). Another possibility is that hippocampal neurons learn combinations of inputs that co-occur in particular situations, and thus provide associative maps suitable for different learned tasks. In this case, the associative maps might be different for the different tasks. The hippocampus could still recall task-relevant information to the neocortex, as described below in Section 8. Consistent with this latter possibility, in bats that learn the same route using visual or sonar cues, the locations in the two maps do not correspond (Geva-Sagiv et al., 2016), so that each map appears to consist of a set of useful conjunctive features for the two tasks, without a single map with corresponding features, such as places or locations “out there,” in the two maps.

8 THE HIPPOCAMPAL RECALL MECHANISM DOES ACT AS A “POINTER” TO WHERE THE INFORMATION SHOULD BE RECALLED IN THE NEOCORTEX

Teyler and DiScenna (1986) had the idea that it would be nice if the hippocampus had a pointer to where in the neocortex a hippocampal memory should be recalled to, but they had no theory or proposal for how any pointer could be implemented. Similarly, David Marr...
promised a theory of recall from the hippocampus to the neocortex (Marr, 1971), but did not produce a theory or model. Rolls (1989) produced a theory of the recall of information from the hippocampus, which specified that this was achieved by associative synaptic modification of synapses from the backprojection pathways from the hippocampus to the neocortical pyramidal cells that are active during the formation of the memory. Thus, for example, if a part of a spatial scene was being represented in the parahippocampal scene area during the formation of a hippocampal episodic memory, then the active backprojection synapses from the hippocampal system would strengthen onto just the active spatial view neurons in the parahippocampal scene area. Then during the later recall of that episodic memory from the hippocampus, just the correct spatial scene location will be recalled to the parahippocampal scene area through the pattern association effect of the backprojections onto the parahippocampal neurons that were active during the formation of the episodic memory (Rolls, 1989; see e.g., Rolls, 2018, 2023a, 2023b). The theory was made quantitative by Rolls and Treves (1994), and Treves and Rolls (1994). McClelland et al. (1995) agreed with the theory of recall, but suggested that some stages in the backprojection pathway for the pattern association learning might be especially important.

This is an important step in understanding how the hippocampus can help the neocortex to build semantic information, by recalling episodic memories to the neocortex for further processing (Rolls, 2022). Exactly how the neocortex builds its semantic memories with the different components in different cortical regions that are heavily interconnected as shown above and as illustrated in Figures 3 and 4 (Rolls et al., 2022a) is not clear, but models are being developed (Boboeva et al., 2018; Rolls, 2023a; Ryom et al., 2023; see also Sections 9 and 10).

9 | THE HIPPOCAMPUS IS FOR EPISODIC MEMORY, NOT FOR SEMANTIC MEMORY, BUT CAN CONTRIBUTE TO THE FORMATION OF SEMANTIC MEMORY

The organization of the hippocampus with a single network in CA3 is ideal for episodic memory, for any input to the hippocampus (e.g., a spatial location, an object, or a reward) can be associated with any other input, all occurring on a single occasion, to form an episodic memory. There is a well-developed and fully quantitative theory for how this episodic memory system operates and recalls information back to the neocortex (Kesner & Rolls, 2015; Rolls, 1989, 2018, 2023a; Rolls & Treves, 1994; Treves & Rolls, 1994). I now consider what the relation is of this hippocampal episodic memory system to semantic memory.

First, semantic information can reach the hippocampus, and can become part of an episodic memory. For example, a territory is a semantic representation, about for example the territory that is France, that incorporates much knowledge from many learning experiences, and the hippocampus can be used to encode, for example, that I saw Alain in France on a particular occasion.

Second, recall of episodic events from the hippocampus can be used to help build a semantic representation (sometimes termed a schema). For example, my knowledge about the territory France is derived in part from many different particular journeys to different parts of France. That is one reason why it is proposed that the learning of new semantic representations is impaired by hippocampal system damage, though semantic information acquired before the hippocampal damage is spared (Rolls, 2022, 2023a).

Third, a chart, or trajectory through a set of spatial locations, can be built in the hippocampus in a continuous attractor network (McNaughton et al., 1996; Battaglia & Treves, 1998; Rolls & Stringer, 2005; Stringer et al., 2005; Rolls et al., 2008). Moreover, if the spatial arrangement of the chart remains constant, but the visual room cues are changed, the new room cues can be rapidly associated onto the existing chart (Baraduc et al., 2019). In this case, the chart is the semantic information, the “cognitive map.”

Fourth, the semantic information (or schemas) is represented in the neocortex, with specialization of different parts of the human neocortex for different components of the semantic representation. The organization of these semantic representations in humans has been clarified by understanding cortical effective (directed, causal) connectivity (Rolls et al., 2022a). One group of interconnected cortical regions includes regions in the ventral bank of the STS (Figure 3) and includes the inferior temporal visual cortex with visual object and face representations (Rolls, 2021a, 2023a), and the inferior parietal cortex (region PGI) which probably allows incorporation of how actions can relate to objects, for example in tool use (Rolls, Deco, et al., 2023a). A second group of interconnected regions includes regions in the dorsal bank of the STS (Figure 4) in which neurons in macaques have been discovered that respond to face expression and socially relevant moving stimuli such as the head and eyes turning to make or break social contact (Hasselmo, Rolls, & Baylis, 1989; Hasselmo, Rolls, Baylis, & Nalwa, 1989), with consistent evidence from functional magnetic resonance imaging in humans (Pitcher et al., 2019; Pitcher & Ungerleider, 2021). This group of regions includes effective connectivity with key language regions including the temporoparieto-occipital region TPOJ1, the peri-Sylvian Language area PSL, the superior temporal visual region STV, Broca’s area 45, and inferior frontal gyrus regions closely connected with and extending Broca’s area, IFJa and IFSp (Rolls et al., 2022a). Importantly, both of the neocortical semantic systems have effective connectivity with the hippocampal system (Figures 3 and 4), providing for semantic information to become a component of an episodic memory; and allowing information about an episodic memory to be recalled back to the neocortex to contribute to the information in semantic memory (Rolls, 2022, 2023a, 2023b).

The paper by Treves and colleagues in this Special Issue (Ryom et al., 2023) makes the point that the storage capacity of information in coupled attractor networks such as those in the multiple cortical brain regions involved in semantic memory just described will be adversely affected by elements that are common to different semantic memories, such as the fact that many animals have four legs (see Section 11). It is suggested that a hierarchical structure for semantic memory, and the use of word symbols, may help with this issue.
10 | CONCEPT CELLS AND SEMANTIC REPRESENTATIONS IN HUMANS AND NONHUMAN PRIMATES

10.1 | Are there concept cells in nonhuman primates?

Quian Quiroga (2023) has argued that concept cells, which can be activated by any one of a range of attributes such as the sight of a person, the sight of a spatial view or location associated with that person, the sound of the person’s voice, and the name of the person (De Falco et al., 2016; Fried et al., 2014; Gastaldi et al., 2021; Ison et al., 2015; Quian Quiroga, 2012; Quian Quiroga et al., 2005; Rey et al., 2015; Rutishauser, 2019), may be unique to humans. This deserves further consideration.

When we discovered neurons in the macaque cortex in the anterior part of the STS that respond to moving objects such as the sight of a head making or breaking social contact (Hasselmo, Rolls, & Baylis, 1989; Hasselmo, Rolls, Baylis, & Nalwa, 1989; Rolls et al., 1987), we also discovered nearby neurons that responded to auditory stimuli including vocalization (Rolls et al., 1987). The concept was that these two representations might be brought together in these cortical regions to form representations that could be activated by the sight of, for example, the lips moving to vocalize, and by the sound of the vocalization being made (Rolls et al., 1987). That has in fact now been shown to be the case, in that 76% of neurons in face patch AF in the macaque cortex in the STS were significantly influenced by the auditory component of a movie with faces that elicited neuronal responses, most often through enhancement of visual responses but sometimes in response to the auditory stimulus alone.
This type of neuron in macaques is thus multimodal, and can be activated by corresponding stimuli in the visual and auditory sensory modalities. That would appear to satisfy at least part of what a concept cell is. The concepts represented by these neurons in the third visual stream are often about the social significance of stimuli (Pitcher et al., 2019; Pitcher & Ungerleider, 2021; Rolls, 2023a; Rolls, Deco, et al., 2023b).

In another example from macaques, hippocampal spatial view cells in macaques are involved in memory, in that they can become associated with an object in an object-location-in-scene memory task (Rolls et al., 2005; Rolls & Xiang, 2006), and with a reward in a reward-in-scene memory task (Rolls & Xiang, 2005). Thus a macaque single hippocampal neuron might respond to the sight of an object, or the sight of the place where it is located. In this sense, the responses of macaque hippocampal cells can be somewhat like those of concept cells (Rolls, 2023b). However, the macaque cells may be essentially specialized for particular associations that occur in an episodic memory, whereas neurons in the human cortical regions outlined in black in Figures 3 and 4 (Rolls et al., 2022a) are more likely to be long-term semantic representation cells.

One property that is different of course for human concept cells is that they can be activated by the word that represents the concept, for example, by the person’s name, where the name is an arbitrary symbol to stand for the object or concept. Moreover, these semantic representations can then be used in syntactic operations, in ways that are at least starting to be investigated in neuroscience using computational modeling of how syntactic operations might be performed by biologically plausible attractor networks, and using investigations of effective connectivity (Rolls, 2023a; Rolls et al., 2022a; Rolls & Deco, 2015). However, in humans, the situation may be further

FIGURE 4 Effective connectivity of region TPOJ1 chosen as an example of a region in a group in a semantic system in the human brain that includes cortex in the dorsal bank of the superior temporal sulcus. This group of cortical regions has connectivity with a superior temporal sulcus auditory–visual stream extending from STGa through STSda, A5, STSdp to TPOJ1. The regions in this group are indicated in bold italic font and are outlined in black, and are A5 STGa STSda STSdp PSL STV TPOJ1. TPOJ1 also has effective connectivity with PSL, STV, TPOJ2, the inferior frontal gyrus IFJA and IFSp and 45, and with premotor area 55b and the midcingulate cortex p24pr. There is also effective connectivity with parahippocampal TF. The regions are those defined in the Human Connectome Project Multimodal Parcellation atlas (Glasser et al., 2016; Huang et al., 2022), with further details and a list of abbreviations in Rolls et al. (2022a). The widths of the lines and the size of the arrowheads indicate the magnitude and direction of the effective connectivity.
subdivided, in that the left hippocampal system may be specialized for these language and word-related types of episodic memory (Bonelli et al., 2010; Sidhu et al., 2013), whereas the right hippocampal system may be more specialized for spatial representations and perhaps more like that in macaques (Barkas et al., 2010; Burgess et al., 2002; Crane & Milner, 2005).

In summary, it appears that some neurons with properties typical of concept cells are found in macaques. However, the concept cells in humans may be specialized for language-related processing, and in that respect are more developed than the concept cells present in macaques.

10.2 | Semantic representations and concept cells

Given the presence of concept cells in the human brain, including in the hippocampus, I next consider where semantic representations are formed in the human neocortex, and how they reach the hippocampus.

Semantic representations are likely to be constructed in the human brain in the regions outlined in black in Figures 3 and 4 (Rolls, 2023a; Rolls et al., 2022a). Each semantic representation is likely to be multimodal, to describe the properties of an object, person, scene, territory, and so forth, and instantiated by the connectivity between the different regions shown in Figures 3 and 4 each with their specializations. Each semantic representation is likely to be implemented by a distributed population of neurons, each of which will have properties like those of the concept cells described by Quian Quiroga (2023) in the Special Issue of Hippocampus 2023 and also as described elsewhere (De Falco et al., 2016; Fried et al., 2014; Gastaldi et al., 2021; Ison et al., 2015; Quian Quiroga, 2012; Quian Quiroga et al., 2005; Rey et al., 2015; Rutishauser, 2019). A concept cell might respond for example to the sight of an object or person, and also to the sight of where they were located.

Most semantic representations will be about viewed things at locations “out there,” not of the place where one oneself is located, and in that sense are like spatial view cells, which are about locations “out there.”

The findings that there are some concept cells in the medial temporal lobe including the hippocampus are consistent with the point made above that some episodic memories can include semantic information, for example, that a particular person is in a particular territory (which is a semantic-level concept) at a particular time.

The human cortical regions outlined in black in Figures 3 and 4 (Rolls et al., 2022a) are involved in language, and do have effective connectivity directed to Broca’s area regions 45 and/or 44 (Rolls et al., 2022a). In some parts of these regions, words are represented, and these word-level representations may be important in helping semantic representations to be kept somewhat discrete, and not so overlapping in their components that there is a big impact on the neocortical memory capacity for semantic representations. That may be part of a solution to the issues about the representation of information in neocortical semantic memory systems when the memory patterns to be stored are somewhat correlated with each other that are raised by Treves and colleagues (Ryom et al., 2023; see Section 11). The point here is that words are arbitrary, relatively orthogonal, symbols, and may thus help to decorrelate neocortical semantic representations.

11 | MODELS OF HIPPOCAMPAL FUNCTION WITH CORRELATIONS BETWEEN THE EPISODIC MEMORIES

If the same several people are all involved in a number of different episodic memories to be stored in the hippocampus, then the different memory patterns of neuronal firing (the set of active neurons for any one memory) will be correlated, as some of the components of each memory will be the same. It has been known since the start of research on autoassociation attractor models of memory that the optimal storage capacity, the number of memory patterns than can be stored, is reduced if the different memory patterns are correlated (Kohonen, 1984; Kohonen et al., 1981).

The standard approach to analyzing the memory capacity of single attractor networks uses random binary patterns of neuronal activity for each memory, which of course have some but well-defined correlations between them (Hopfield, 1982). This approach was developed in the theory of the hippocampus which includes a single attractor network in CA3 (Kesner & Rolls, 2015; McNaughton & Morris, 1987; Rolls, 1987, 1989, 2018, 2023a; Rolls & Treves, 1994; Treves & Rolls, 1994) by analytic approaches (Treves, 1990, 1991a, 1991b; Treves & Rolls, 1991a) that calculate the capacity of attractor networks with sparse distributed representations (in which only a small proportion of neurons is active for any one pattern; as are found in the brain [Franco et al., 2007; Rolls & Treves, 2011]), with graded firing rates (as are found in the brain [Franco et al., 2007; Rolls & Treves, 2011]), and with diluted connectivity (which is found in the brain [Rolls, 2023a]).

These approaches show that the maximum number of graded firing rate patterns, \( p_{\text{max}} \), that can be stored and correctly retrieved is approximately:

\[
 p_{\text{max}} = \frac{C^{ RC}}{a \ln(1/a)} k
\]

where \( C^{ RC } \) is the number of recurrent collateral connections onto each neuron, and \( k \) is a scaling factor that depends weakly on the detailed structure of the rate distribution, on the connectivity pattern, and so forth, but is roughly in the order of 0.2–0.3 (Treves & Rolls, 1991a). For example, for \( C^{ RC } = 12,000 \) associatively modifiable recurrent collateral synapses onto each neuron, and \( a = 0.02 \), \( p_{\text{max}} \) is calculated to be \( \sim 36,000 \). In these analyses, the neuronal population sparseness \( a \) of the representation can be measured by extending the binary notion of the proportion of neurons that are firing to any one stimulus or event as:
where $r_i$ is the firing rate (e.g., spikes/s, typically in the range 0–100 spikes/s) of the $i$th neuron in the set of $N$ neurons. The sparseness ranges from $1/N$, when only one of the neurons responds to a particular stimulus (a local or grandmother cell representation [Rolls & Treves, 2011], to a value of 1.0, attained when all the neurons are responding at the same rate to a given stimulus [Franco et al., 2007; Rolls & Treves, 2011; Treves & Rolls, 1991a]).

These analytic approaches are complemented by simulations of these single attractor networks, which also draw out the implications for the noise in the system, and the impact of graded firing rates (Rolls, Treves, Foster, & Perez-Vicente, 1997; Rolls & Webb, 2012; Simmen et al., 1996; Webb et al., 2011). These analyses emphasize the utility of having a sparse representation in the hippocampus, for sparse distributed representations increase the number of different memories that can be stored (Treves & Rolls, 1991a), essential for an episodic memory (Rolls, 2010, 2016b, 2018, 2023a).

Quian Quiroga (2023) considers the effect of correlations of the type referred to above in decreasing the memory capacity of an autoassociation attractor network. In doing so, he refers to the interesting paper of Gastaldi et al. (2021). That paper however considers a rather limited approach with binary neurons (either firing at a high rate, or not at all), which cannot be easily extended to the biologically plausible case of neurons with graded firing rates, that is different firing rates to each of the set of stimuli (Franco et al., 2007; Rolls, 2023a; Rolls & Treves, 2011). Another limitation is that Gastaldi et al. (2021) consider the case where there are two memories, $m_1$ and $m_2$, which share a fraction $c$ of their active units, but are encoded in a recurrent network in which all other memories are uncorrelated. They use this lack of correlation (see the sentence after equation 25 on p. 20) to derive analytically the storage capacity, which is essentially like that for networks of uncorrelated memories (and the rest of the details as in their model), except for the specific correlation between $m_1$ and $m_2$, the only pair with correlations in an ocean of otherwise uncorrelated memory patterns. Then in the final section of the main Results text (Gastaldi et al., 2021), which starts on their page 11 “How does a network embed?,” they state that so far they have “mainly” focused on neurons that are shared between a single pair of memories (i.e., $m_1$ and $m_2$), but now to compare with human recordings they assume that similar correlations are widespread among all memories. Gastaldi et al. (2021) then considered a number of statistical models (including that of Boboeva et al. (2018) which had already dealt with correlations between the patterns stored in an associative memory), but they did not redo the capacity analysis nor the simulations with the correlations among all memories. If Gastaldi et al. (2021) had done that (with simulations, because the analytical approach is difficult without the random correlation assumption of their equation 25), they would have found, subject to differences in the details, the same reduction in capacity that was reported by Ryom et al. (2023) in their paper in this Special Issue of Hippocampus.

To summarize the implications of the point made by Quian Quiroga (2023) about memory capacity, it has been known for a long time that correlations between the patterns to be stored in an autoassociation attractor network reduce the number of memories that can be stored below the maximal storage capacity shown in Equation (1). The best analyses to date are those of Boboeva et al. (2018) and Ryom et al. (2023), with the approach Quian Quiroga (2023) refers to by Gastaldi et al. (2021) having limitations as shown above. In this context, the pattern separation performed by the dentate granule cells and mossy fiber to CA3 synapses helps to reduce the effects of correlations between the memory patterns stored in CA3, and this pattern separation is considered in Section 12.

The issue of the sparseness of the representations in the hippocampus and in neocortical regions in primates including humans deserves further analysis. Quantitative studies of well-isolated single neurons and of populations of neurons using the measure of sparseness defined in Equation (2) have been performed for the hippocampal system and for neocortical regions in macaques, and have provided evidence for sparse distributed representations, which appear to be more sparse in the hippocampus than in neocortical regions (Baddeley et al., 1997; Franco et al., 2007; Rolls, 2023a; Rolls et al., 1998; Rolls & Tovee, 1995; Rolls & Treves, 2011). That is consistent with pattern separation for hippocampal representations. The measure of sparseness shown in Equation (2) is particularly useful, because it is a simple statistical measure that can be directly related to the storage capacity of neural networks using the tools of statistical mechanics (Boboeva et al., 2018; Rolls, 2023a; Rolls & Treves, 1990; Ryom et al., 2023; Treves, 1991a; Treves & Rolls, 1991a).

In practice, one can measure sparseness in two ways (Franco et al., 2007; Rolls, 2023a):

1. One can take a single cell, and use a large number of test stimuli (at least 20, preferably 32 or 64 or more) and measure the sparseness with Equation (2) where $r_i$ is now the responses of the cell to each of the $i = 1, N$ stimuli. That is the single-cell sparseness, $a^s$.
2. One can take a whole population of cells (at least 20, preferably 32 or 64 or more), and measure the sparseness using any one typical stimulus from the set with Equation (2). That is conceptually a way to measure the population sparseness, $a^p$.

In practice, in at least some brain regions, these two measures are close to each other, consistent with low correlations between the response profiles of single neurons, and this is a form of weak ergodicity (Franco et al., 2007; Rolls, 2023a; Rolls & Treves, 2011).

When measuring the sparseness of a neuronal representation, or the information in a neuronal representation, the values obtained depend on the stimulus set (Rolls, 2023a; Rolls & Treves, 2011; Rolls, Treves, & Tovee, 1997; Rolls, Tovee, & Panzeri, 1997). For example, if the sparseness of the representation of inferior temporal cortex face cells is measured with faces and a whole set of natural nonface visual stimuli, then the representation will be more sparse than when measured with faces only, because the face neuron will hardly respond to the nonface stimuli (Rolls, 2023a). Thus the stimulus
set must be clearly specified when describing measures of sparseness and information content (Rolls, 2023a; Rolls & Treves, 2011).

One particularly interesting approach to this is to estimate the probability distribution of the number of spikes from a neuron to natural environmentally representative visual stimuli using a fixed time window of, for example, 250 ms. This can be performed when the individual is viewing natural stimuli such as scenes with people, objects, and so forth in for example a movie, and this has shown an exponential distribution of firing rates, with a large number of low rates, and fewer, and fewer high rates, for the inferior temporal visual cortex, which provides evidence for a sparse representation (Baddeley et al., 1997; Franco et al., 2007; Rolls & Treves, 2011).

Given the interesting points made about the sparseness of representations in the hippocampal/medial temporal lobe regions in humans (Donoghue et al., 2023; Quian Quiroga, 2023), it would be of interest in future investigations in humans as well as other species to present sparseness and information measures using quantitative approaches of the type described here.

12 | PATTERN SEPARATION IN THE HIPPOCAMPAL SYSTEM

The several ways in which pattern separation is implemented in the dentate to CA3 system have been described (GoodSmith et al., 2017; Johnston et al., 2016; Knierim & Neunuebel, 2016; Leutgeb et al., 2007; Leutgeb & Leutgeb, 2007; Rolls, 1989, 2013, 2016a; Toda et al., 2019). Quian Quiroga (2020, 2023) has proposed that pattern separation does not occur in the human hippocampus. However, it is not clear that the representations in the inputs to the human hippocampus, for example, in the neocortex, have been compared with respect to their sparseness and interpattern correlations with those in the hippocampus. That essential type of comparison has been made for the macaque hippocampus, for spatial view cells (Rolls, 2023a; Rolls & Treves, 2011). For the representation of 64 locations around the walls of the room by hippocampal spatial view cells, the mean single-cell sparseness \( a^v \) was 0.34, and the mean population sparseness \( a^p \) was 0.33 (Rolls et al., 1998; Rolls & Treves, 2011). For comparison, the corresponding values for macaque inferior temporal cortex neurons tuned to objects and faces were 0.77 (Franco et al., 2007); for taste and oral texture neurons in the insular cortex the population sparseness was 0.71; for taste and oral texture neurons in the orbitofrontal cortex was 0.61; and for taste and oral texture neurons in the amygdala was 0.81 (Rolls & Treves, 2011). Thus the evidence is that the hippocampal CA3/pyramidal cell representation is more sparse in macaques than in neocortical areas and the amygdala, and this is consistent with the importance in hippocampal CA3 of using a sparse representation to produce a large memory capacity. Although the value of \( a \) for the sparseness of the representation does not seem especially low, it must be remembered that these are graded firing rate representations, and the graded nature appears to increase the value of \( a \) (Rolls & Treves, 2011). Moreover, these values were obtained in just one spatial environment, and if measured over many spatial environments, might be more sparse.

In summary, there is considerable evidence from research in many species that pattern separation in the dentate to CA3 system does occur and is important for function, and the proposal that there is no pattern separation in the human hippocampus (Quian Quiroga, 2020, 2023; and note that it is needed in CA3) requires more evidence, including a comparison of encoding differences between hippocampal CA3 and neocortical regions (Rolls, 2021c; Suthana et al., 2021).

13 | THE HUMAN ORBITOFRONTAL CORTEX CAN CONTROL MEMORY CONSOLIDATION IN THE HIPPOCAMPUS AND NEOCORTEX

When episodic memories involving, for example, spatial view cells and objects (Rolls et al., 2005; Rolls & Xiang, 2006) or rewards (Rolls & Xiang, 2005) are formed in the hippocampus, recent connectivity studies in humans provide evidence for connectivity in humans from the orbitofrontal cortex in part via the pregenual anterior cingulate cortex to the septal nuclei and basal forebrain nucleus of Meynert (Rolls et al., 2022b; Rolls, Deco, et al., 2023c), which contain cholinergic neurons that project to the hippocampus and neocortex respectively (Mesulam, 1990; Zaborszky et al., 2008, 2018). In that the orbitofrontal cortex represents reward value, aversive stimuli, and nonreward (Rolls, 2019a, 2019b; Rolls et al., 2020), this provides a way in which the storage of episodic memories can occur especially well when environmental reinforcement contingencies are present or change. In a corresponding way, the orbitofrontal cortex via the vmPFC has connectivity to the basal forebrain cholinergic neurons, which provide for the consolidation of semantic memories to be influenced particularly when there is a reward, punishment, or nonreward component (Rolls, 2022). These cholinergic systems are implicated in the synaptic mechanisms involved in episodic memory storage in the hippocampus, and memory consolidation involving the formation of semantic memories in the neocortex (Giacomo & Hasselmo, 2007; Hasselmo, 1999; Hasselmo & Bower, 1993; Hasselmo & Giacomo, 2006; Hasselmo & McGaughy, 2004; Hasselmo & Sarter, 2011; Newman et al., 2012).

The implication for spatial view representations in the primate including human hippocampus is that their use in memory storage and also thereby in navigation is influenced by reward / punishment contingencies that reflect the importance of the events and so modulate the storage of information that includes spatial view information. Mechanisms such as this are thus proposed to be key in understanding memory consolidation in the hippocampus and neocortex (Rolls, 2022, 2023a).

14 | SUMMARY AND CONCLUSIONS

1. Primate hippocampal and parahippocampal gyrus spatial view cells implement allocentric encoding of locations in scenes (Rolls, 2023b).
2. These neurons are idiothetically updated by eye position, head direction, and probably place (Rolls, 2023b).

3. The idiothetic update is implemented in the primate dorsal visual system, as it involves eye position (Rolls, 2023b), and not in the hippocampus as sometimes suggested for rodents (McNaughton et al., 1996). The idiothetic update in for example, the dark when

4. These spatial view cells can rapidly form associations in the hippocampus with the object or reward at a location in a spatial scene, which is prototypical of episodic memory (Rolls, 2023b).

5. The presence of the primate fovea with its high spatial resolution (60 cycles/°) is key to object-in-scene and reward-in-scene memory, because this resolution enables small objects or rewards to be seen when fixated, and the fixated location used to help define the location of the object or reward in the scene. This is a key function of the primate hippocampus in memory.

6. Similarly, the high visual acuity of the primate fovea enables even similar distant landmarks to be discriminated, and used for visually guided and implemented navigation from viewed landmark to viewed landmark.

7. In contrast, the poor visual acuity of rodents (~1 cycle/°) renders it poor for object or reward in scene memory, and for visually guided location to distant landmarks.

Instead rodent hippocampal cells respond to the place where the rodent is located as defined for example by vibrissa-based or olfactory local sensing. Correspondingly, in rodents navigation may be from place to place using different navigational rules to those used in primates including humans (except when the primate is navigating in the dark).

8. In rodents, the view inputs from scenes may be used instead to reset the idiothetic navigation being performed (place cells are locked to the room cues).

9. Rodent navigation may thus, because of the differences in the visual systems, not provide a model for navigation in primates. The same applies to the episodic memory for the locations in scenes where objects or rewards have been seen.

10. Several papers (including several in this Special Issue) have highlighted the presence of hippocampal and related neurons in primates including humans that respond to views of scenes, and not to the place where the individual is located (Corrigan et al., 2023; Donoghue et al., 2023; Mao et al., 2021; Tan et al., 2021; Tsitsiklis et al., 2020; Yang et al., 2023; Zhu et al., 2023). This helps to establish the encoding of views as key to understanding hippocampal function in primates including humans (Rolls, 2023b).

11. Although hippocampal spatial view neurons respond to where in a scene the macaque is looking and not to where in a scene the macaque is facing (Rolls, 2023b), some investigators have described some primate hippocampal neurons that respond to where the monkey is facing (and by inference not to where the macaque is looking, in that their eyes typically use saccades to inspect scenes; Mao et al., 2021; Zhu et al., 2023). If some neurons do respond to facing direction, that would be very inefficient for object-in-scene memory or reward-in-scene memory, for the only way that the object or reward could be found again would be by moving the head constantly until the head happened to be facing in the direction of the object or reward, whereas primates including humans search scenes with saccadic eye movements, which enable locations in scenes to be found without relying on slow and clumsy head movements to look for objects or rewards.

12. The present commentary makes the point that it is fundamental to separate the spatial coordinate system being used (egocentric or allocentric), from the viewer-dependent properties that arise when an observer views a scene. An egocentric coordinate system would be for example relative to the head, in what is described as craniotopic coordinates. So, a location in space might be to the left of the head in egocentric coordinates. An allocentric reference frame is world-based, and independent of whether the scene is on the left or the right of the head: it is a location in the world. In contrast, when an observer looks at a scene from one side of the scene, a mountain might appear to the left of a lake; and when the scene is viewed from behind (the other side), the mountain might appear to the right of the lake. However, those relations of an observer viewing the world from one side of a scene compared with the other side, are independent of whether for the individual the representation is in an egocentric coordinate frame related to for example the head. In fact, it is a property of allocentric encoding of scenes that items in the scene will be left-right reversed on the two different sides of the scene.

Second, it seems there is so far not yet a great deal of evidence for egocentric encoding of locations in scenes by primate hippocampal neurons. A few neurons were described by Feigenbaum and Rolls as having egocentric encoding, but almost all of their view cells used allocentric encoding (Feigenbaum & Rolls, 1991).

Third, it is shown in this commentary that egocentric encoding of spatial locations is not a strong candidate for use in navigation, and also in “object or reward” location in a “scene episodic memory.” When egocentric encoding is found, in rodents, for example, by “ego-centric boundary cells” (Alexander et al., 2020, 2023; Wang et al., 2018, 2020, 2023), the most parsimonious account of its utility is in obstacle avoidance (a running rat with poor vision using vibrissae and olfactory cues does not want to run into a boundary), for maintaining distances from other individuals as in flying flocks of birds or bats, and so forth.

13. Neurons with some properties like those of human concept cells (Quian Quiroga, 2023) are found in nonhuman primates, but clearly with language in humans concept cells can be much more elaborate. Pathways by which semantic information (which is
what concept cells represent) reach the hippocampus for use in episodic memories, and the return pathways for recall of episodic information back to semantic neocortical areas in humans, are described.

In conclusion, hippocampal and parahippocampal cortical spatial view cells found in primates including humans may be key to understanding the functions of the hippocampal system in memory and navigation in humans (Georges-François et al., 1999; Robertson et al., 1998; Rolls, 1999, 2021a, 2021b, 2021c, 2023a, 2023b; Rolls et al., 1998; Rolls, Robertson, & Georges-François, 1997); these cells go beyond rodent place cells, and indeed they compute something different and are used in different ways.

ACKNOWLEDGMENTS
Edmund Rolls has worked on some of the experiments and concepts described here with N. C. Anggelopoulos, I. Araujo, A. Berthoz, P. M. B. Cahusac, W. Cheng, G. Deco, J. D. Feigenbaum, J. Feng, L. Franco, P. Georges-François, C.-C. Huang, R. P. Kesner, S. M. O’Mara, W. P. C. Mills, Y. Miyashita, H. Niki, S. Panzeri, R. G. Robertson, S. M. Stringer, A. Treves, J. Tromans, S. Wirth, and J.-Z. Xiang, and their collaboration is sincerely acknowledged. Stimulating discussions with Alessandro Treves (SISSA, Trieste) and Sylvia Wirth (CNRS, Lyon) are greatly appreciated. Collaboration with Professor Andrew M. Haun (Department of Psychiatry, University of Wisconsin-Madison, Wisconsin, USA) who modified his software made available in connection with Haun (2021) at https://osf.io/8xf9w/ is warmly acknowledged. This research was supported by the Medical Research Council (PG8513790 and the Oxford Interdisciplinary Research Centre in Cognitive Neuroscience), and by a grant from the Human Frontier Science Program that enabled a very fruitful and appreciated collaboration with Carol Barnes, Jay McClelland, Bruce McNaughton, and Alessandro Treves.

CONFLICT OF INTEREST STATEMENT
The author does not have a conflict of interest to report.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID
Edmund T. Rolls https://orcid.org/0000-0003-3025-1292

REFERENCES


SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.