

(Abstract) neural representations of spaces and concepts

Neil Burgess
Institute of Cognitive Neuroscience
University College London



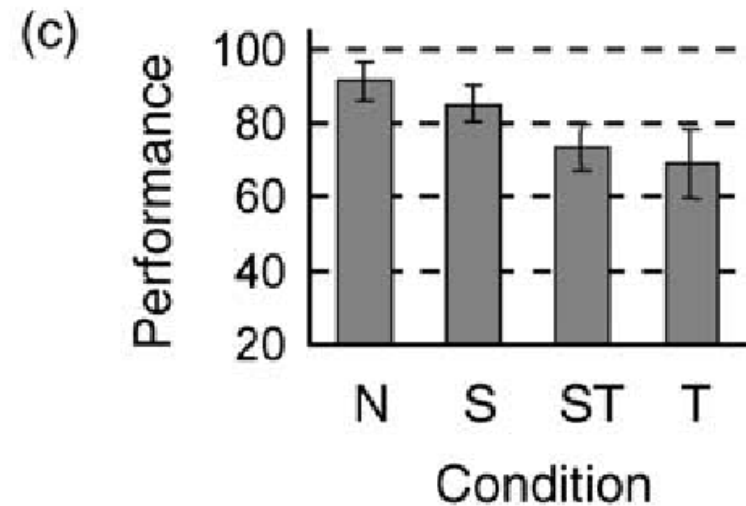
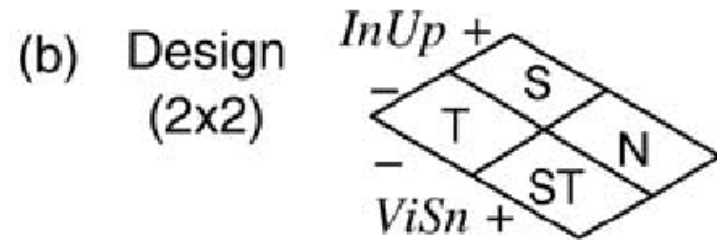
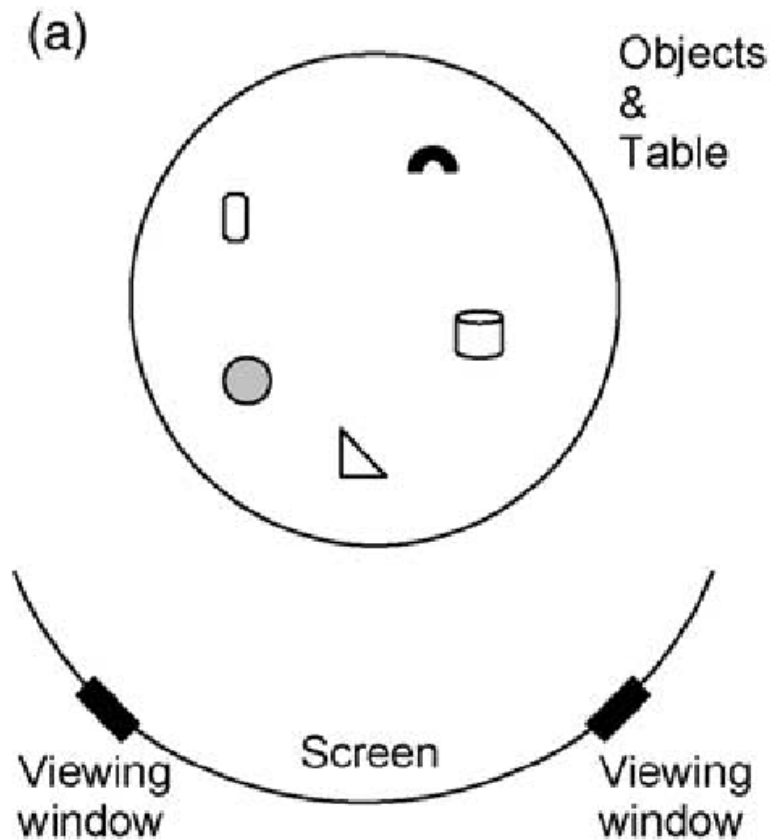
Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) *Place cells & boundary vector cells*
- 3) *Neural level model of Spatial Memory and Imagery*
- 4) *Place and grid cells, environmental and self-motion inputs?*
- 5) *Grid cells as dynamic imagery?*
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

Multiple parallel representations in spatial memory.

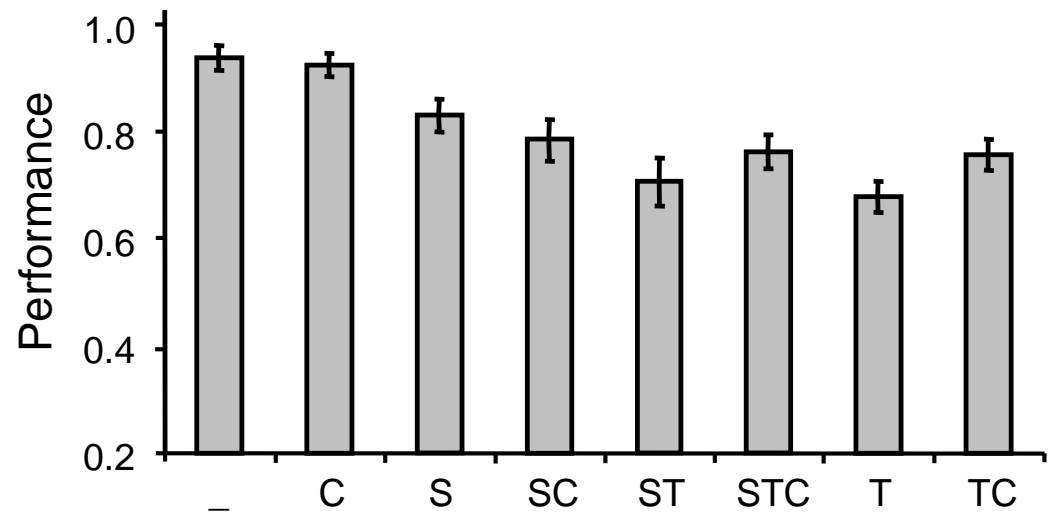
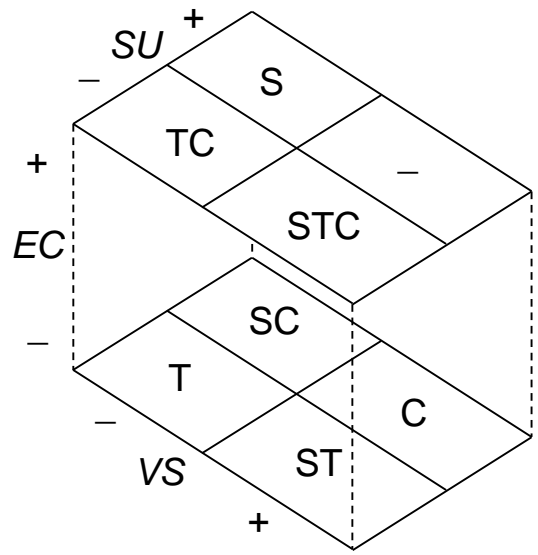
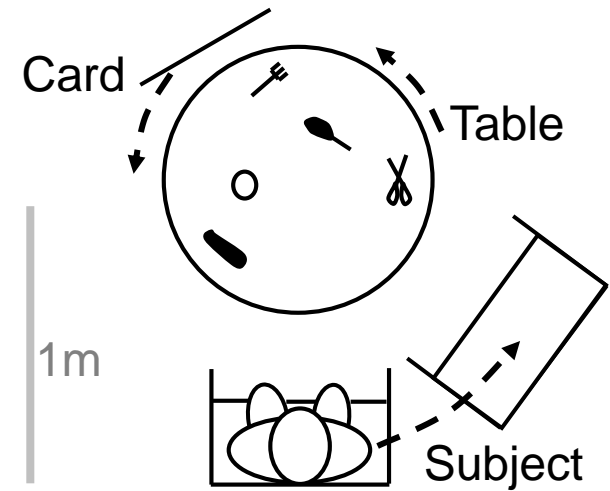
Effects of consistency with 'Visual Snapshots' & Internal 'Spatial Updating'

Wang & Simons 1999



Multiple parallel representations in spatial memory.

Visual Snapshots (*egocentric*),
 Spatial Updating (*egocentric*) and
 External Cues (*allocentric*).



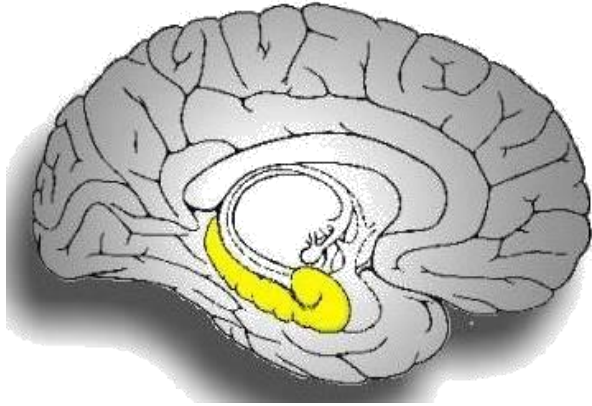
| Consistency | - | C | S | SC | ST | STC | T | TC |
|-------------|---|---|---|----|----|-----|---|----|
| InUp: | + | + | + | + | - | - | - | - |
| ViSn: | + | + | - | - | + | + | - | - |
| ExCu: | + | - | + | - | - | + | - | + |



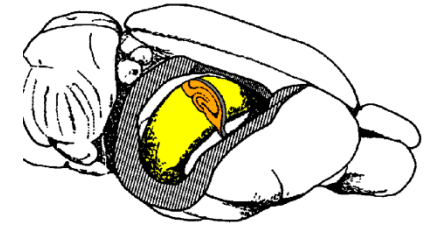
Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) ***Place cells & boundary vector cells***
- 3) *Neural level model of Spatial Memory and Imagery*
- 4) *Place and grid cells, environmental and self-motion inputs?*
- 5) *Grid cells as dynamic imagery?*
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

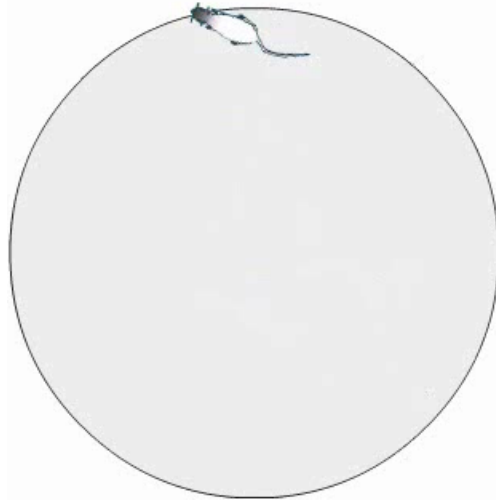
The hippocampus supports memory (e.g. HM), but how does it work?



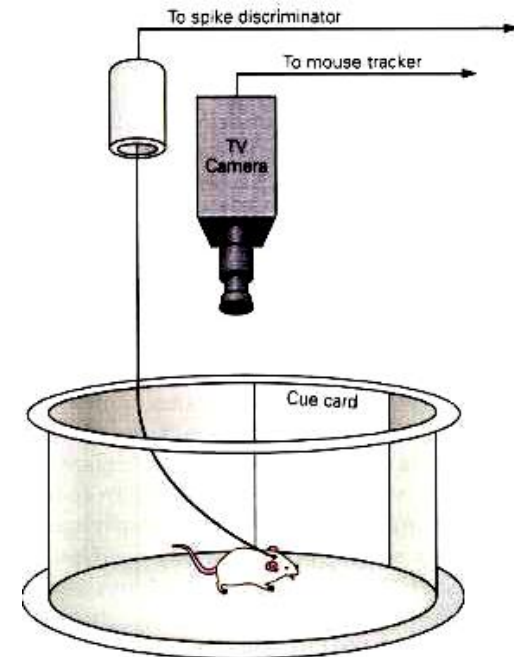
Spatial studies in rodents => likely neural representations.



Place cells- 'allocentric' location



O'Keefe & Dostrovsky, 1971

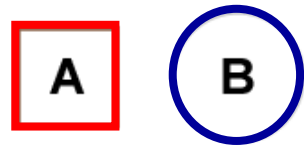


Video by Julija Krupic



Place cells show long term memory and pattern completion

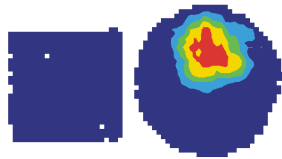
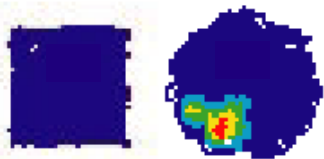
Place cell “remapping:” long-term memory for highly distinct environments.



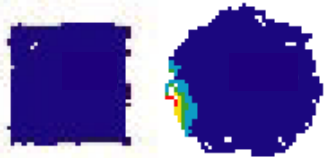
learned distinction remains after 71 days..

CELL

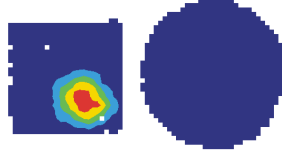
1



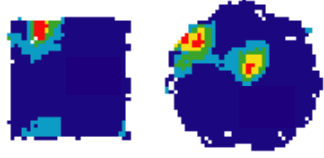
2



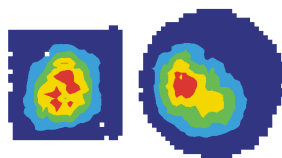
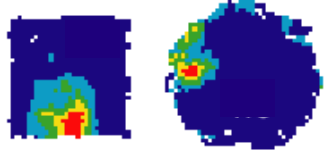
3



4



5



Place cell representation shows attractor dynamics

Wills, Lever, Cacucci, Burgess, O'Keefe, 2005

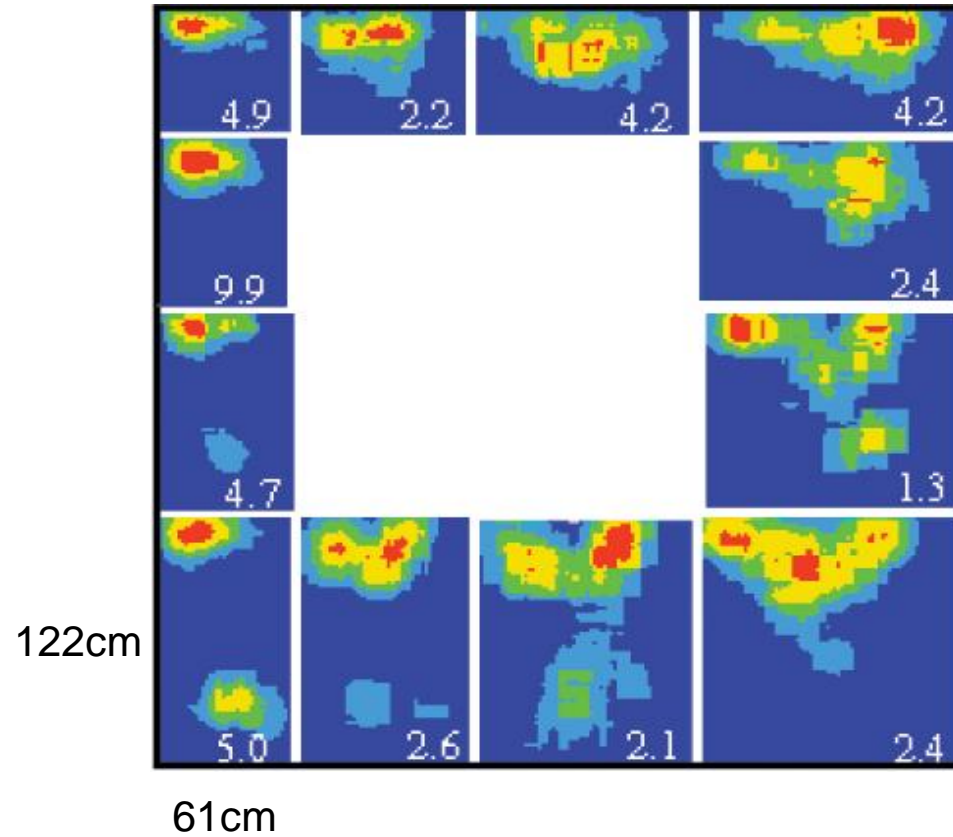
and 'pattern completion' depending on CA3 NMDA receptors

Nakazawa et al., 2002

Lever, Wills, Cacucci, Burgess, O'Keefe, 2002

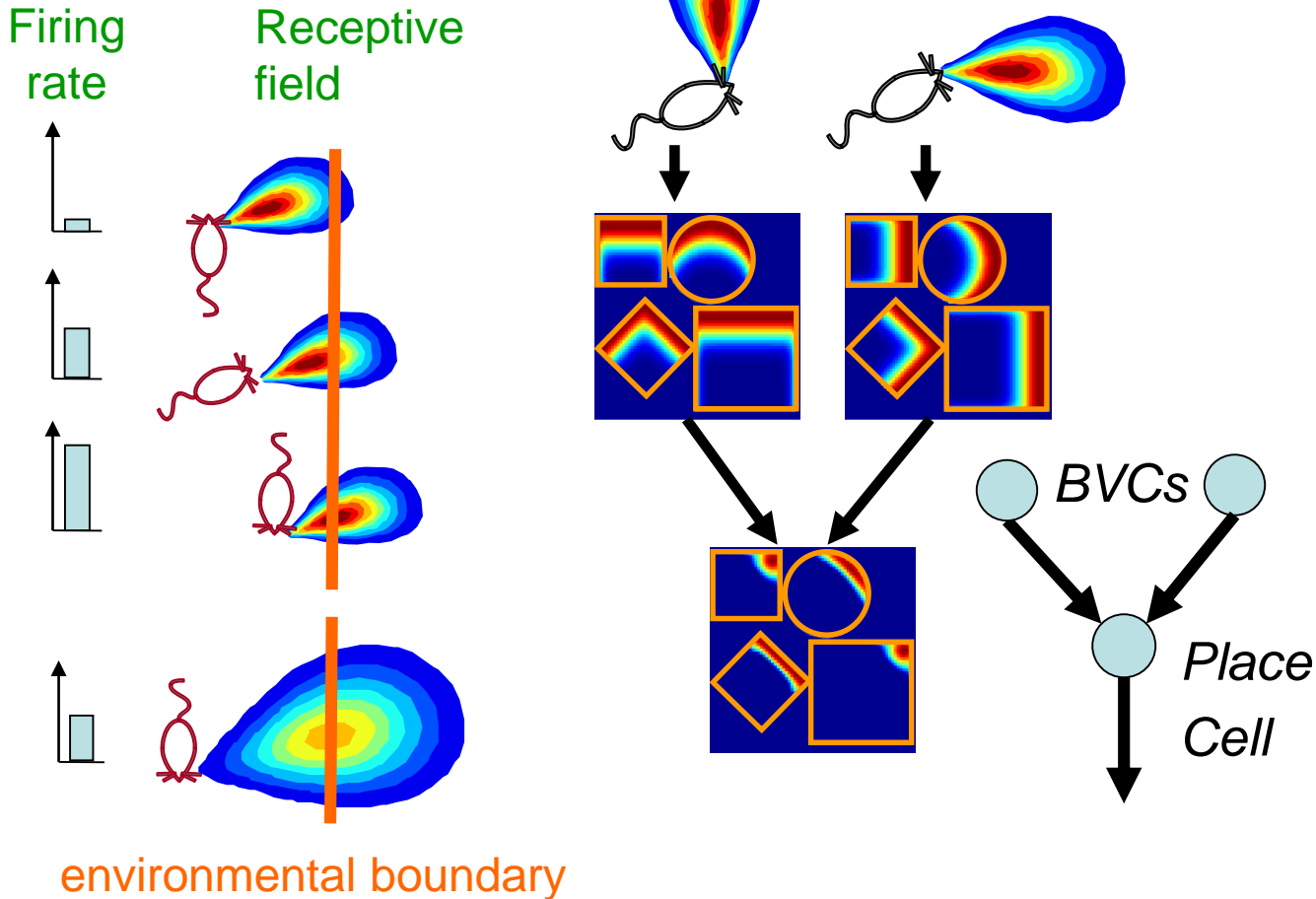


Environmental boundaries particularly influence place cell firing

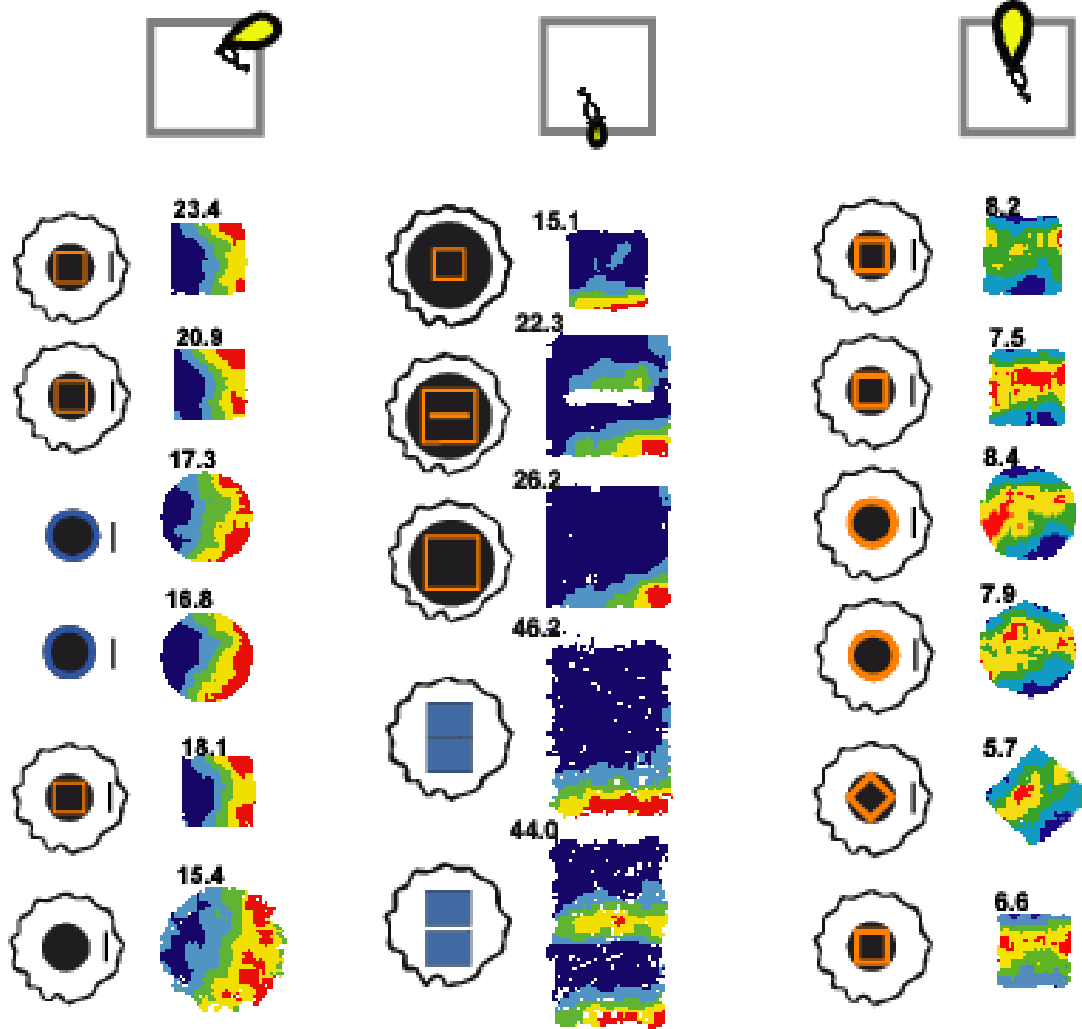


Place Cell firing as a thresholded sum of “Boundary Vector Cell” inputs

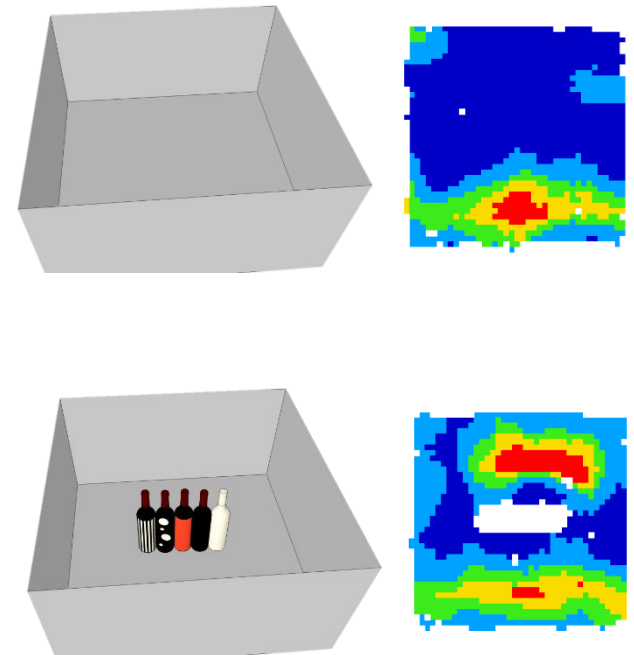
Boundary Vector Cells (BVCs) signal distance to boundary along an *allocentric* direction



BVCs found in subiculum & entorhinal cortex



Including those firing at a distance



Steve Poulter & Colin Lever



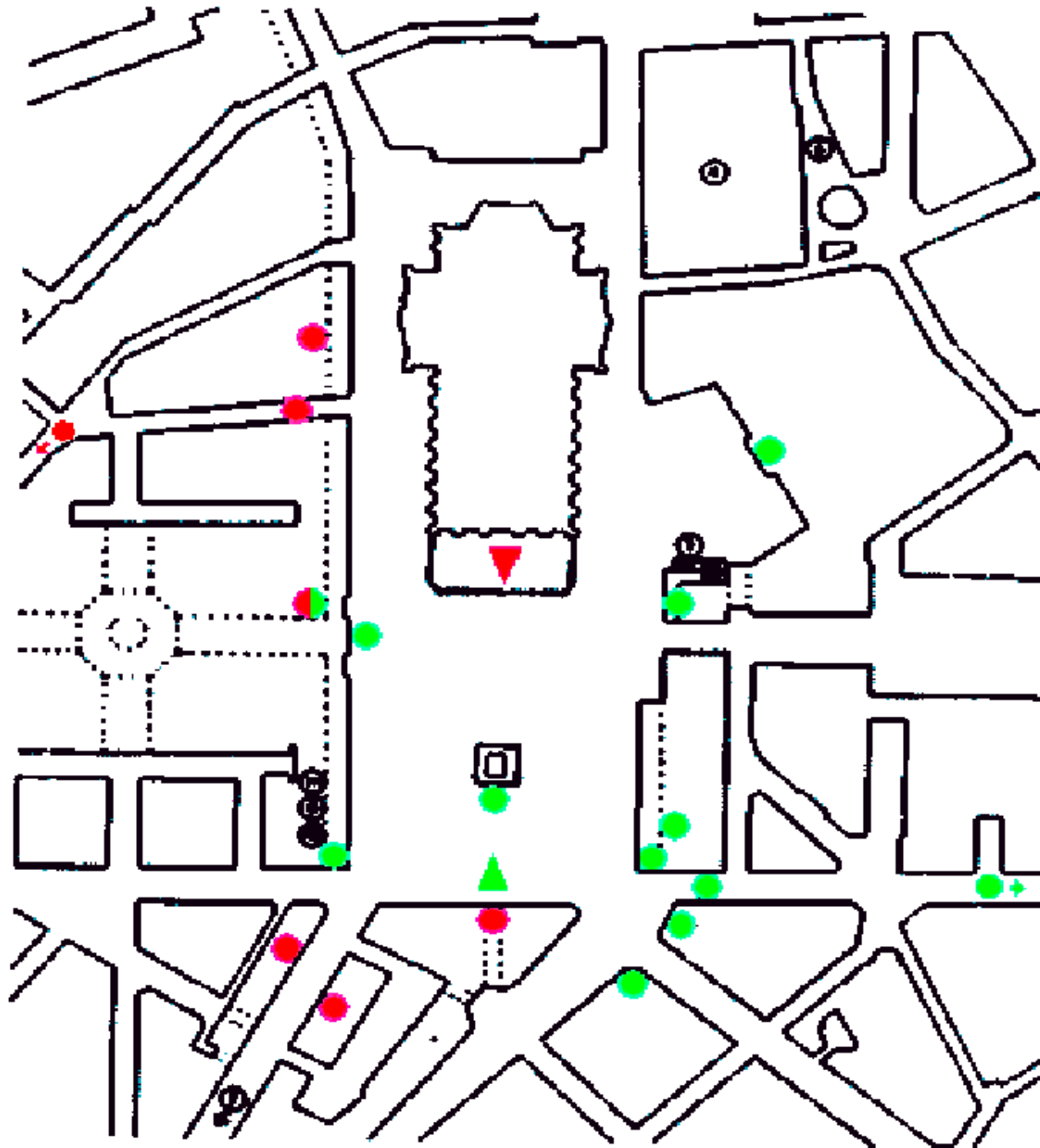
Lever, Burton, Jeewajee, O'Keefe, Burgess, 2009
See also Barry et al, 2006; Solstad et al, 2008

Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) *Place cells & boundary vector cells*
- 3) ***Neural level model of Spatial Memory and Imagery***
- 4) *Place and grid cells, environmental and self-motion inputs?*
- 5) *Grid cells as dynamic imagery?*
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

Hemispatial neglect in memory of Milan square following right parietal damage.

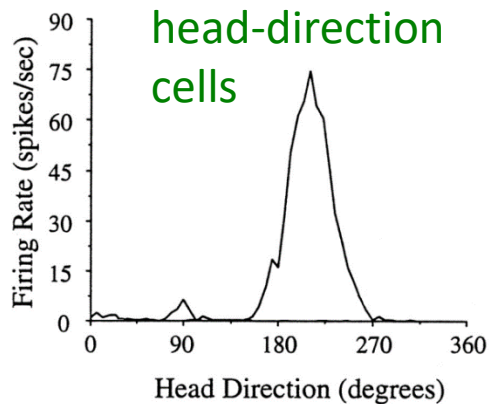
⇒ formation of an egocentric representation in parietal cortex from a stored allocentric representation in medial temporal lobe?



Several identified neural representations support spatial cognition

Hippocampal formation
(allocentric)

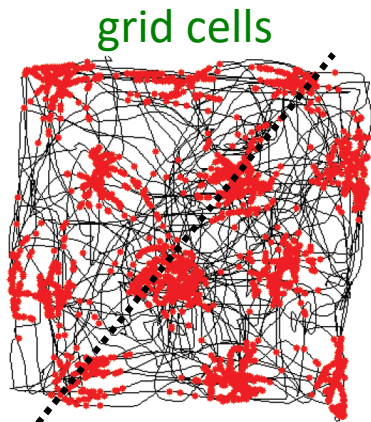
Sensory, **Parietal**, Motor cortices
(egocentric)



Ranck et al, 1984;
Taube et al, 1990

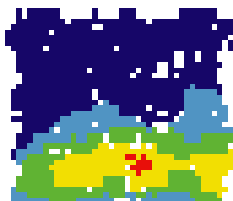


O'Keefe & Dostrovsky, 1971



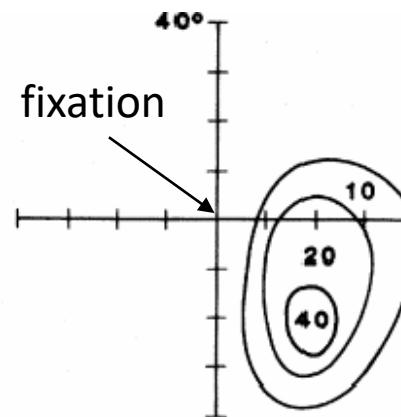
Hafting et al., 2005

boundary cells

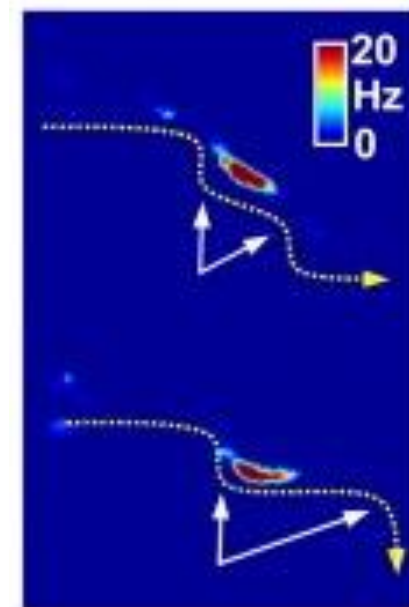


Lever et al, 2009
Solstad et al, 2008

retinal receptive fields



trajectory cells,

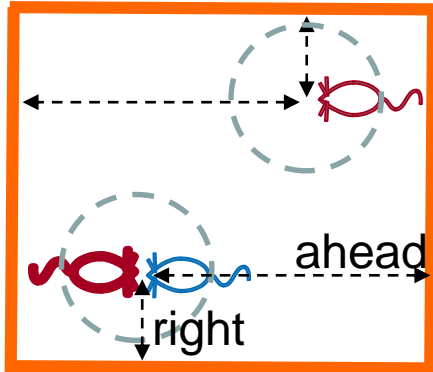


Nitz 2009

Frames of reference for neural coding

'egocentric'

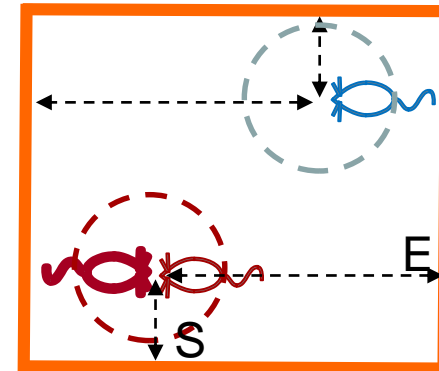
Body-centred location of objects



Perception
Action/Imagery

'allocentric'

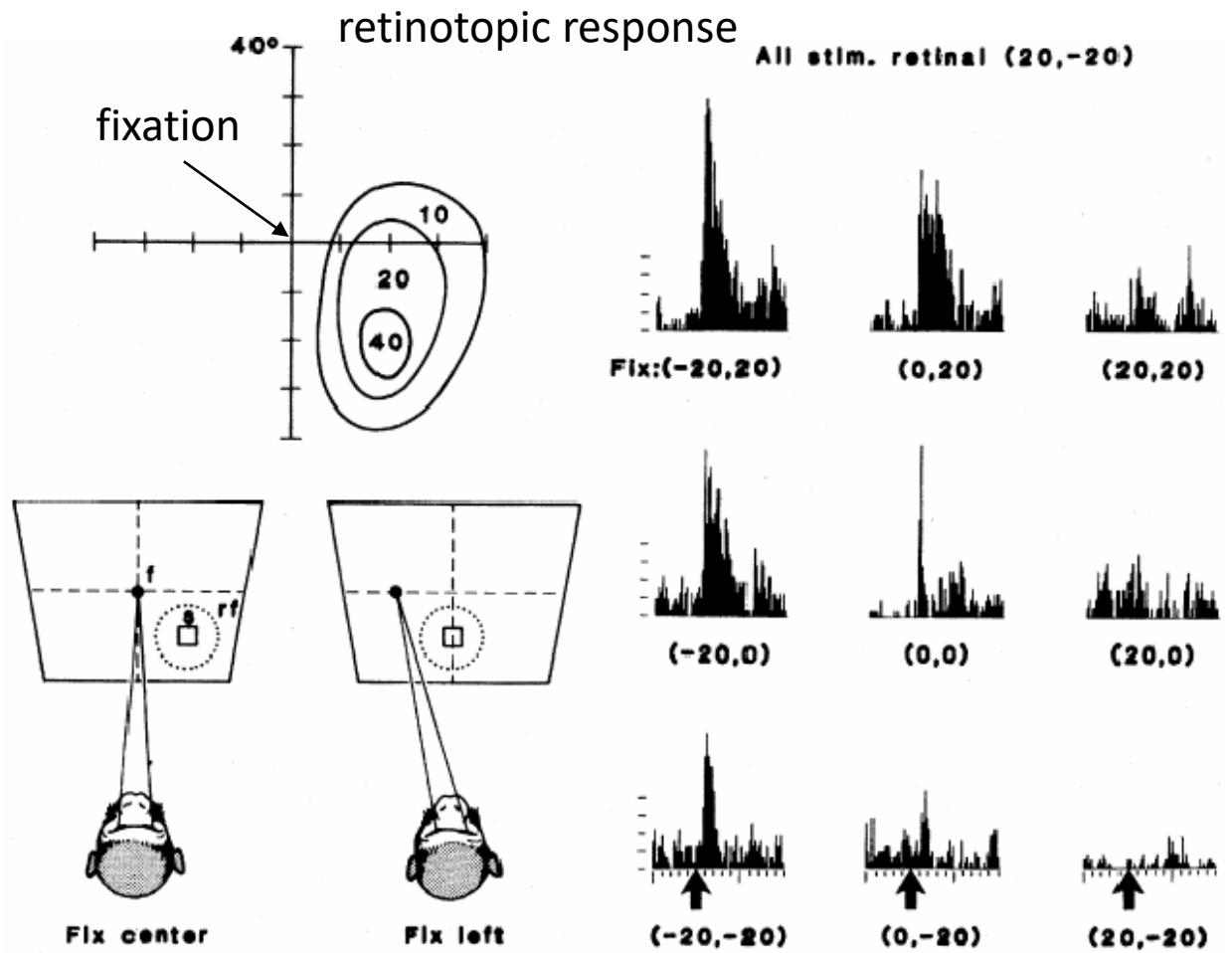
World-centred location of agent



Place cells
Head-direction cells

'Gain field' responses in posterior parietal cortex

i.e. conjunctive responses to (retinotopic) *visual input* x *gaze direction*



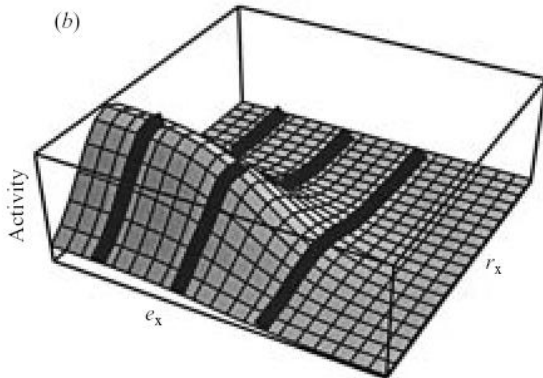
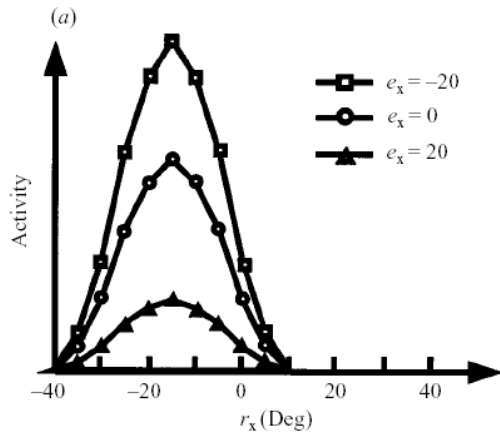
Size of retinotopic visual response is modulated by direction of gaze:

Andersen et al 1985

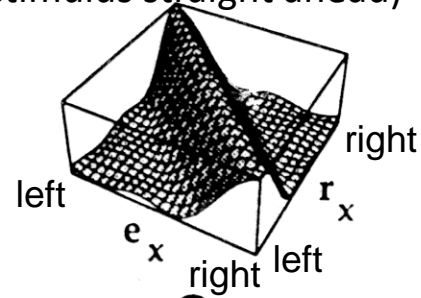
or by direction of the head (Snyder et al 1998).
 Similar responses seen in parieto-occipital ctx (Galletti et al., 1995)

Gain field neurons can produce 'head-centred' or retinotopic representations.

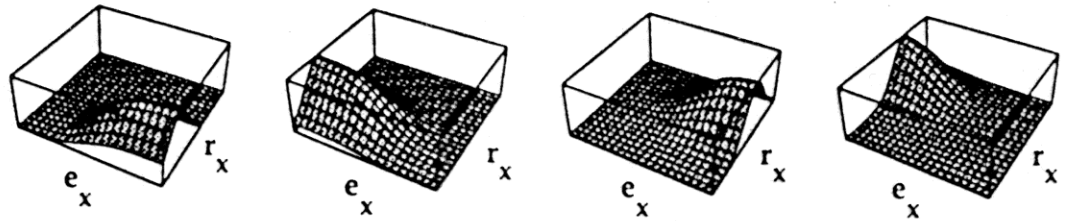
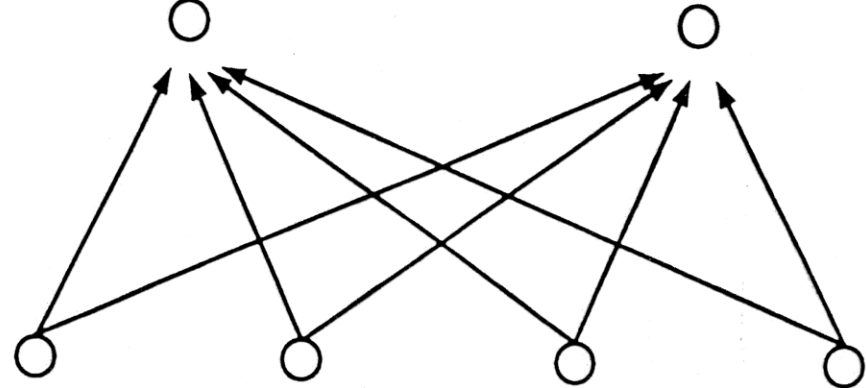
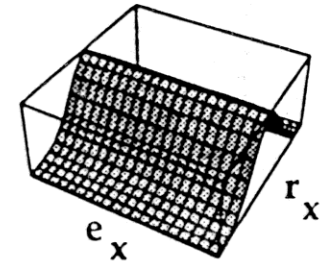
retinal position of stimulus = r_x
eye gaze angle = e_x



Head-Centered
(stimulus straight ahead)

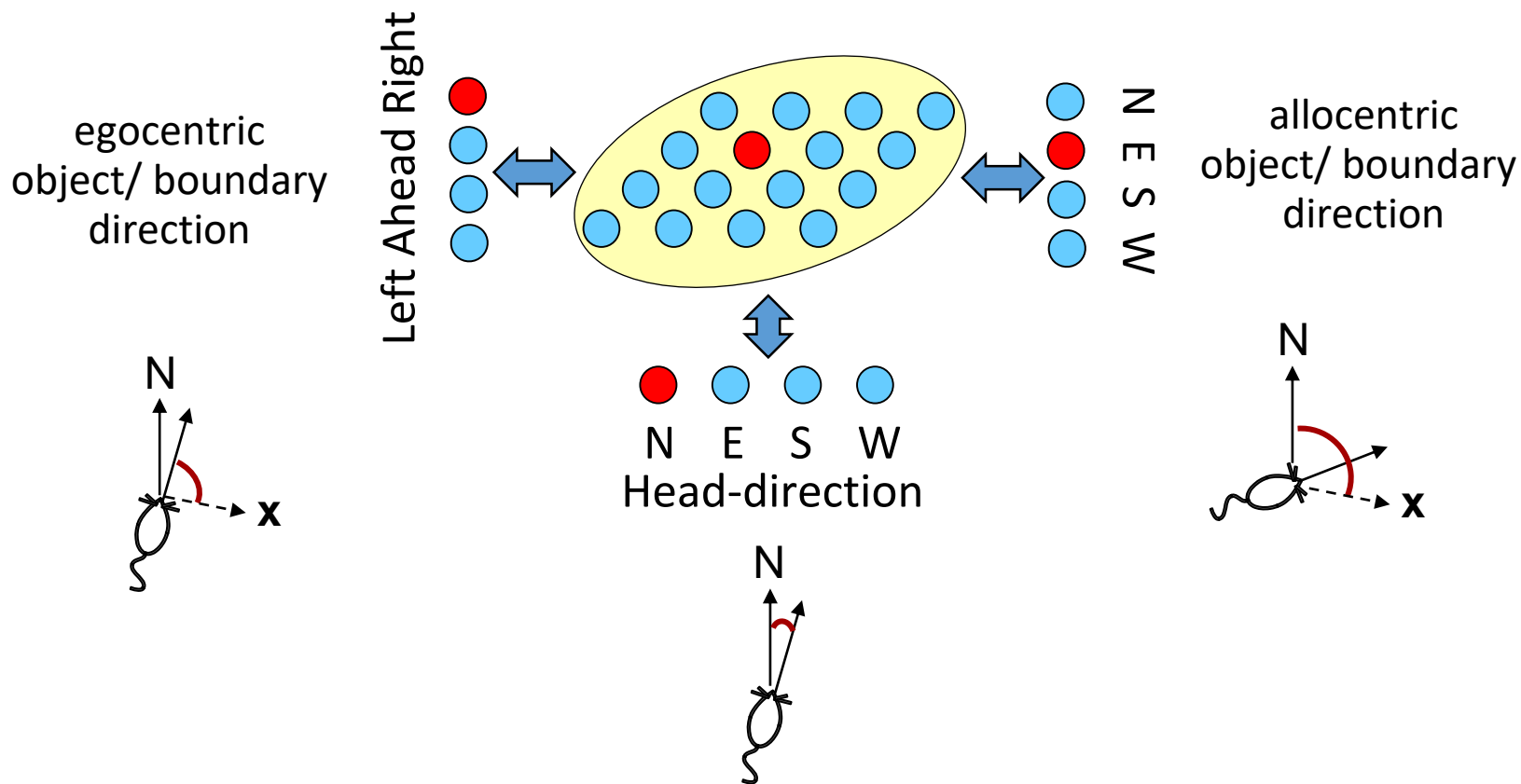


Retinotopic



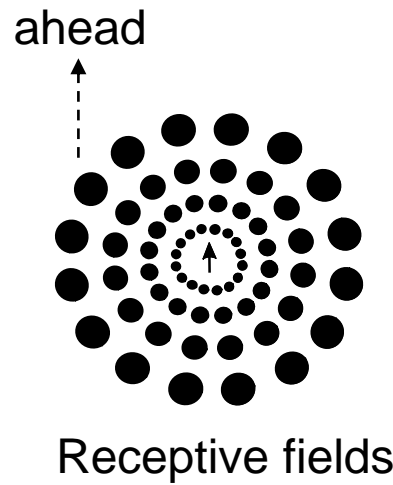
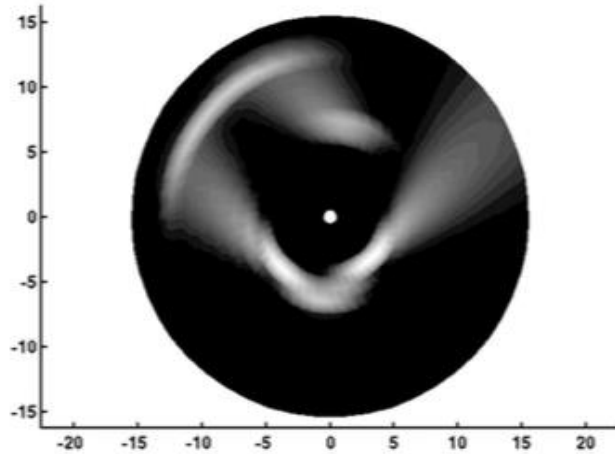
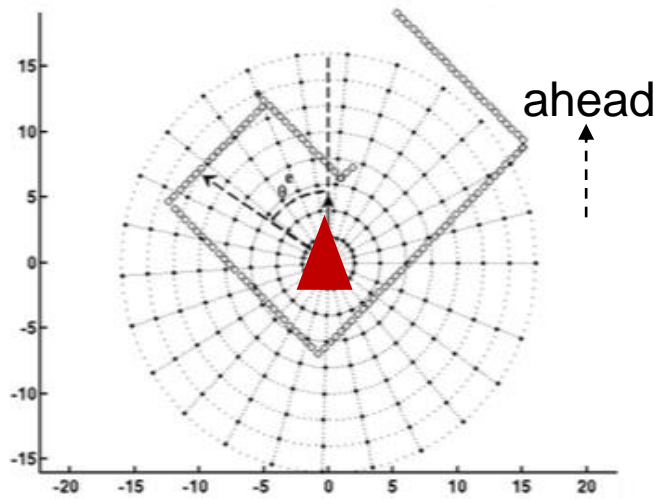
Model of memory & imagery for scenes

Egocentric-allocentric translation by 'gain-field' neurons
(i.e. conjunctive representations of egocentric sensory input x head direction)



*Byrne, Becker, Burgess 2007; Burgess et al., 2001;
See Pouget & Sejnowski, 1997; Deneve et al., 2001.*

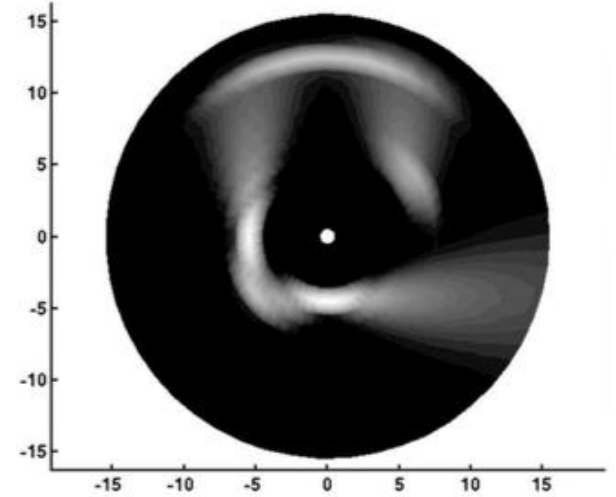
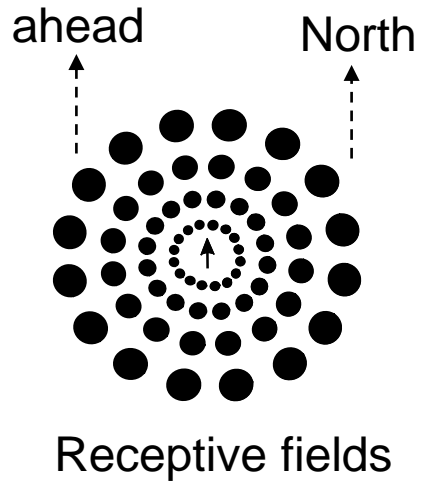
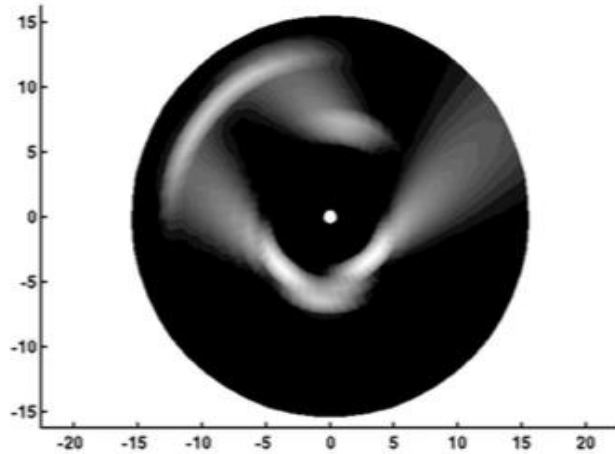
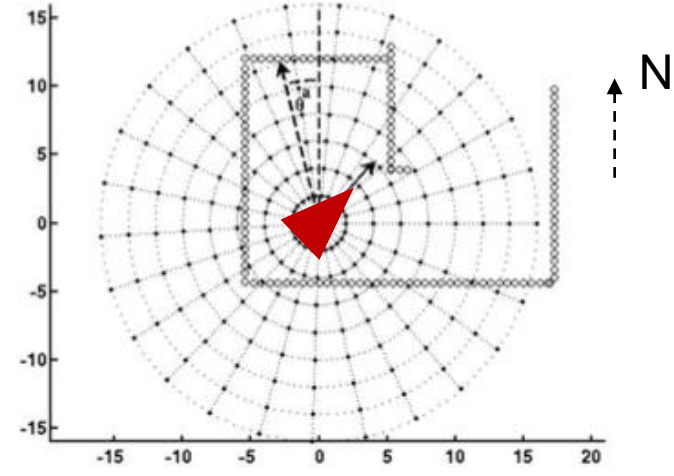
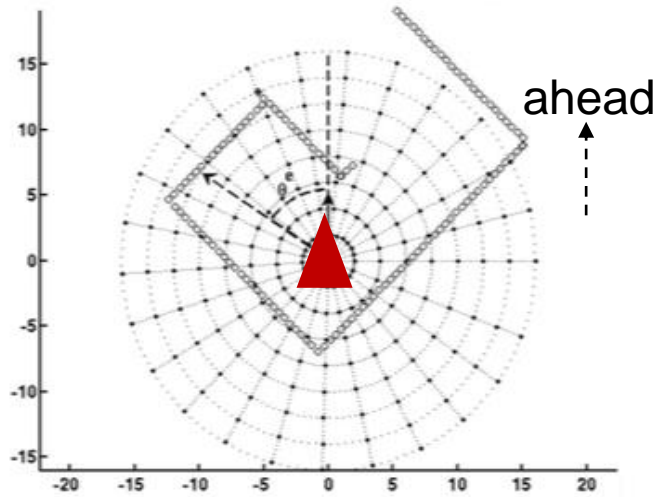
Scene representation by populations of egocentric or allocentric BVCs



Parietal

***egocentric representation
(e.g. visual)***

Scene representation by populations of egocentric or allocentric BVCs



Parietal

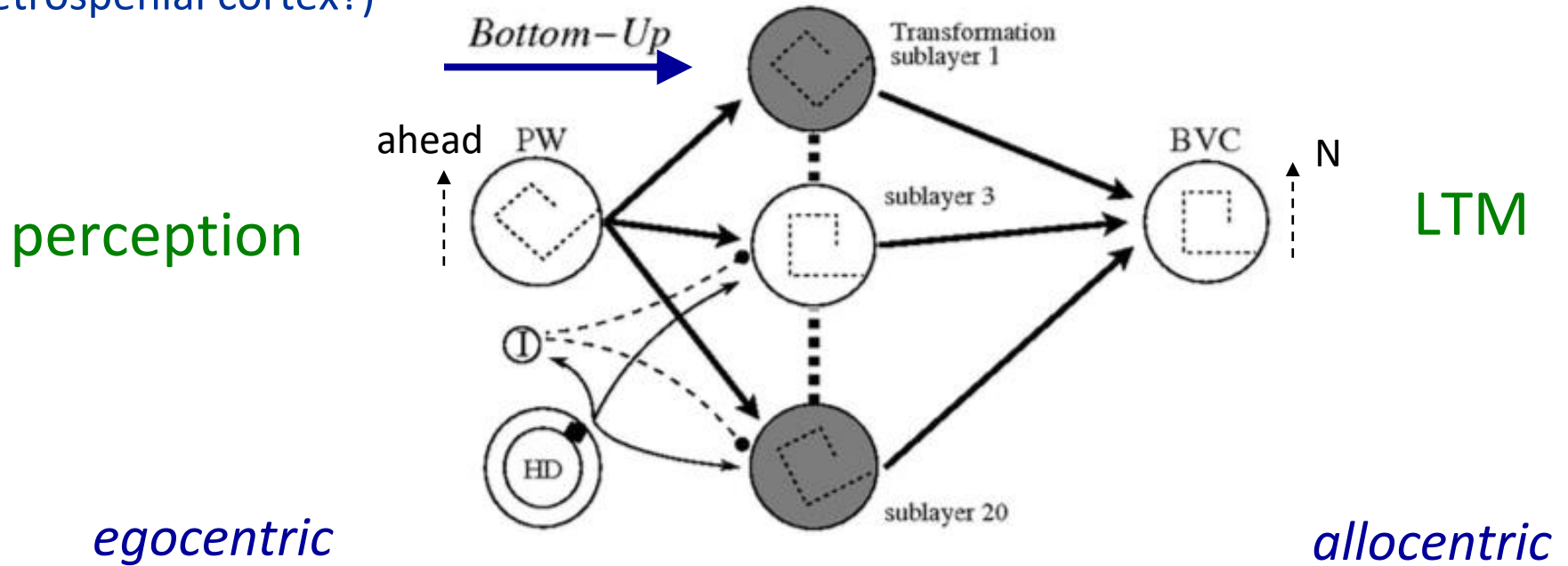
BVCs

***egocentric representation
(e.g. visual)***

allocentric representation

Ego-allo scene translation (retrosplenial cortex?)

'gain field' representation of scene elements x head direction

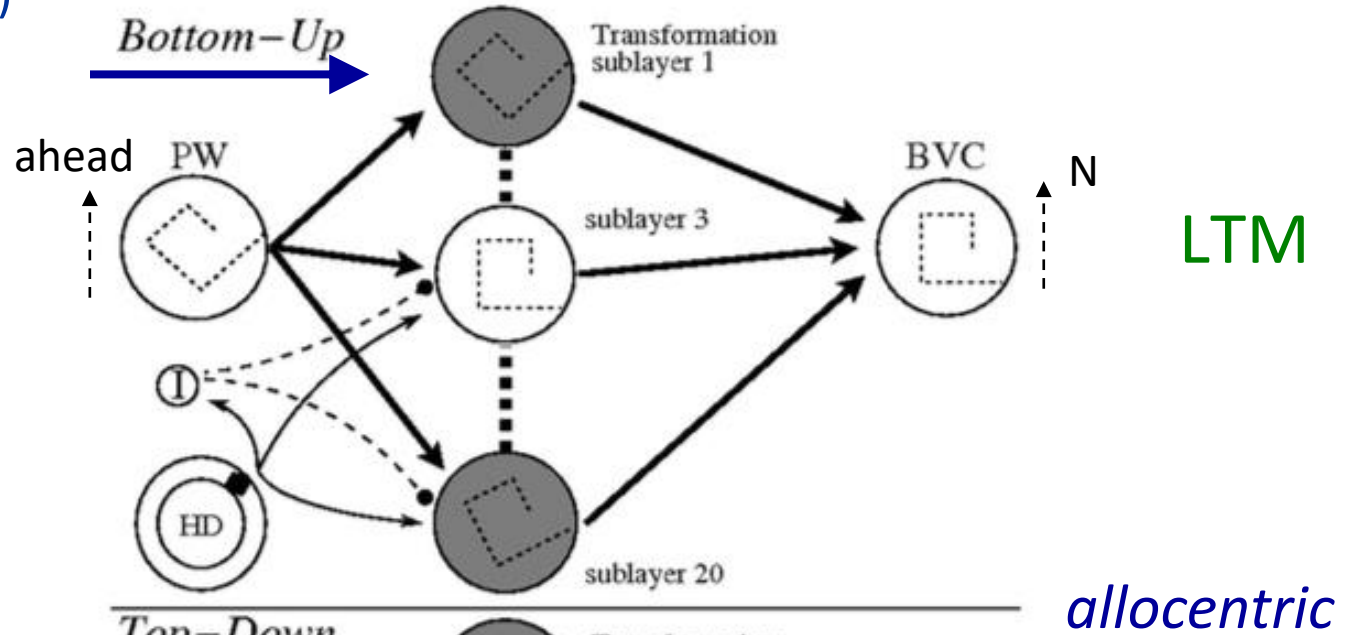


Byrne, Becker, Burgess 2007
Burgess et al., 2001
see also Pouget & Sejnowski 1997

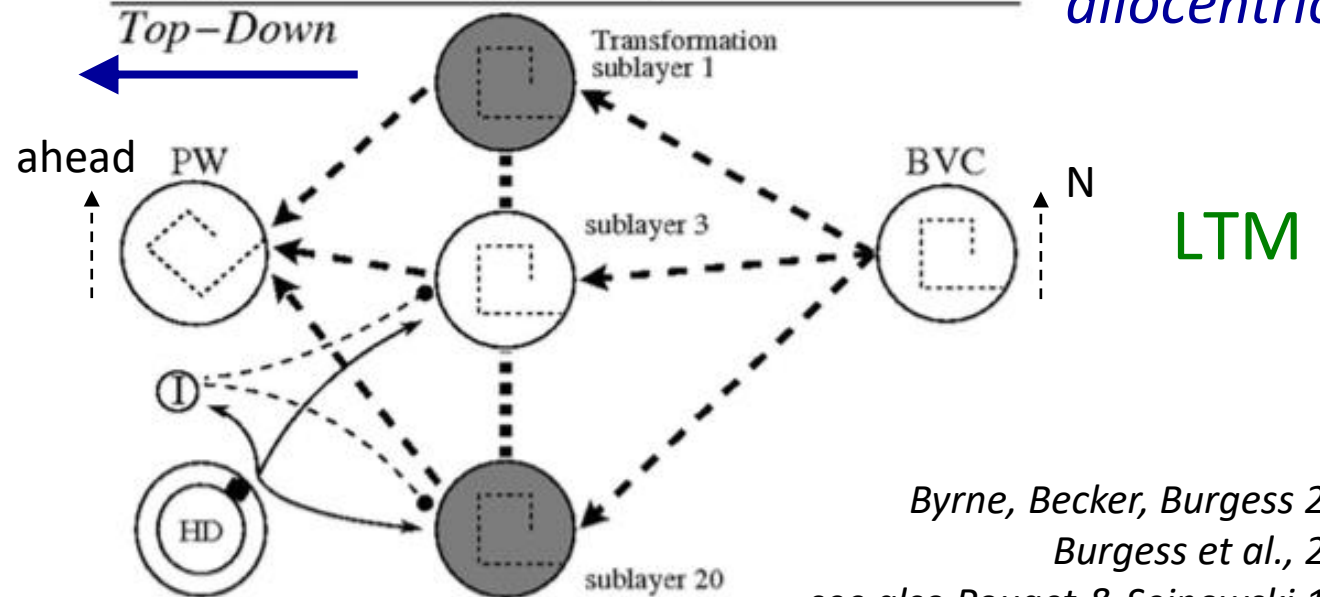
Ego-allo scene translation (retrosplenial cortex?)

'gain field' representation of
scene elements x head direction

perception



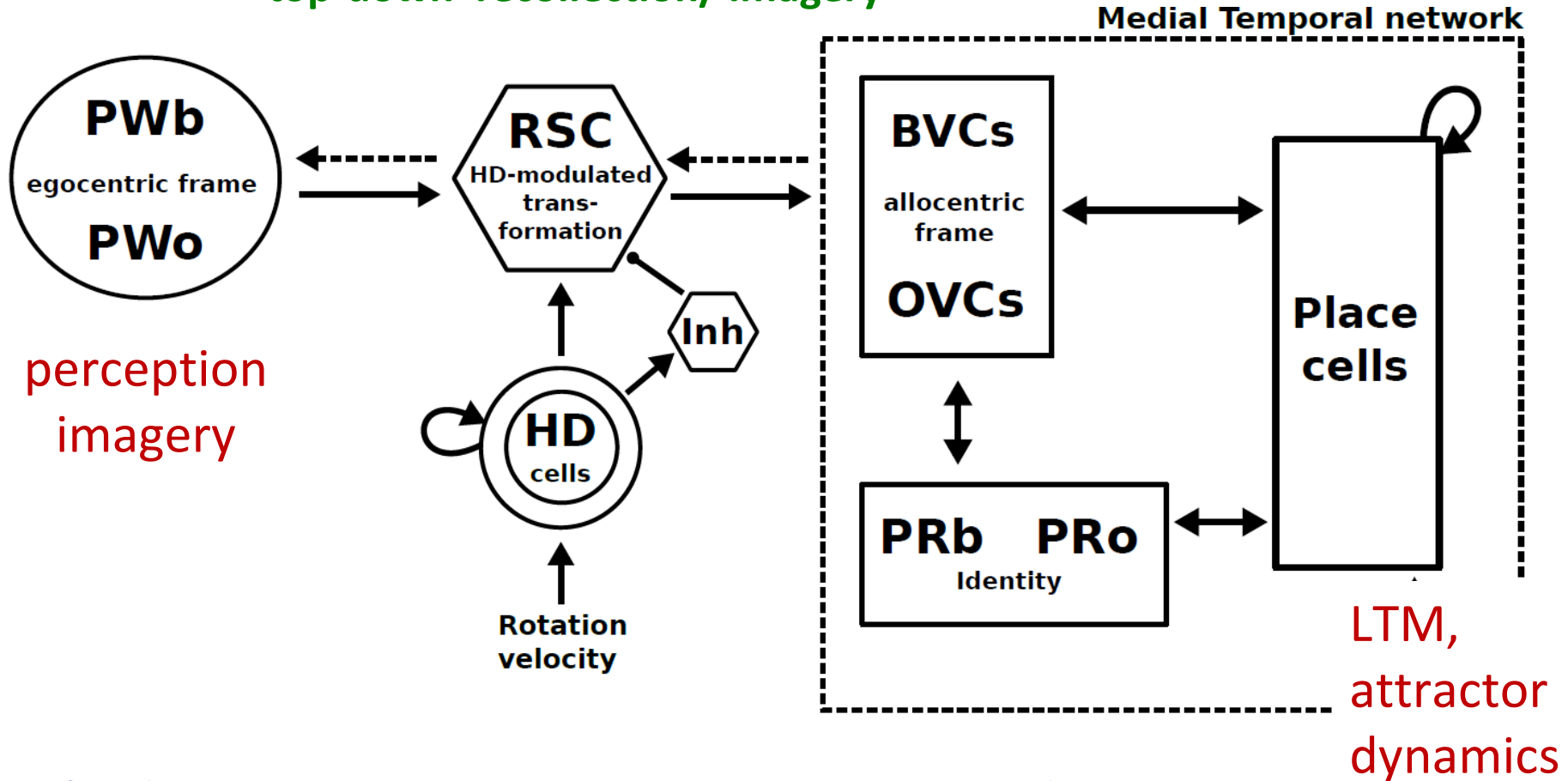
imagery
(& action)



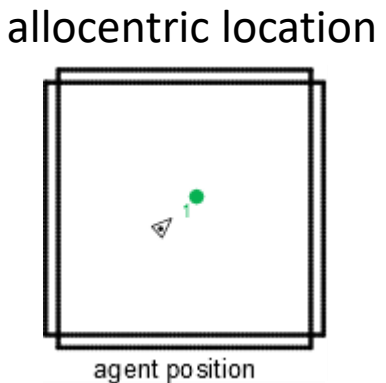
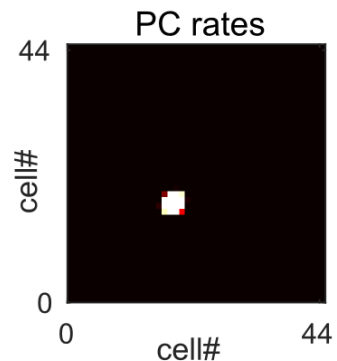
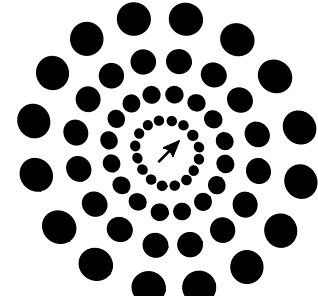
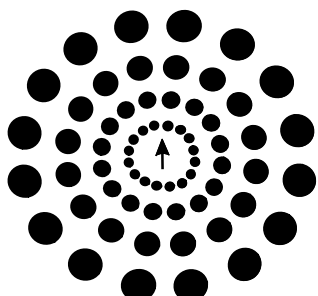
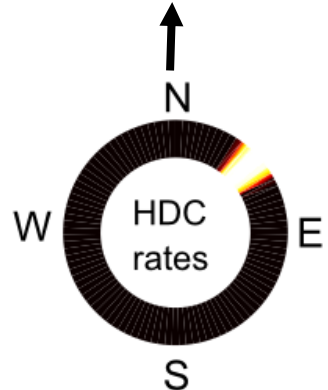
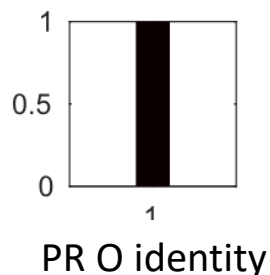
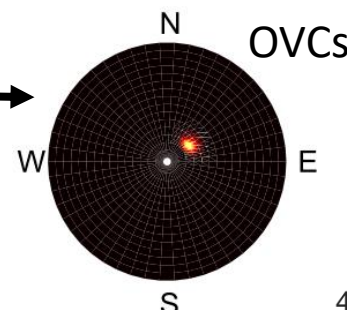
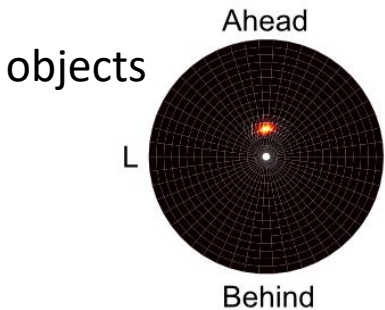
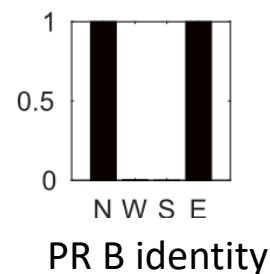
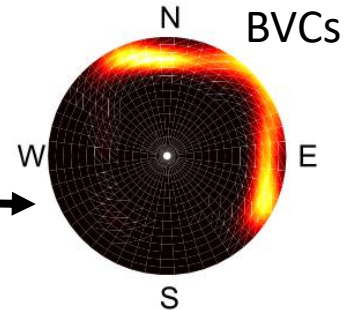
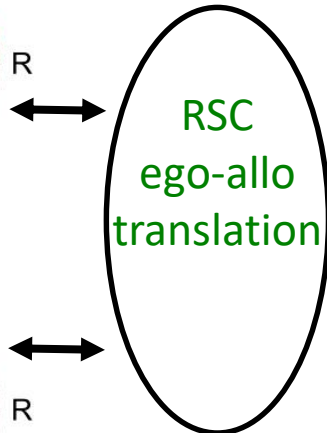
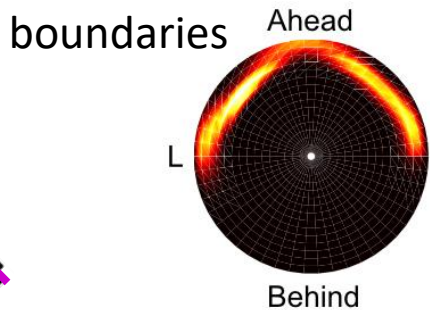
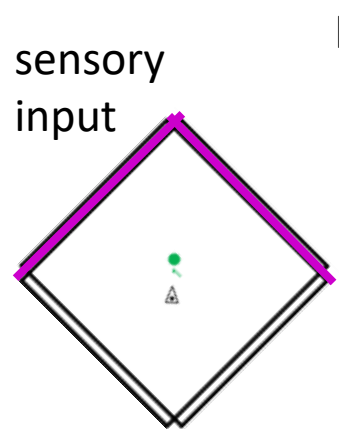
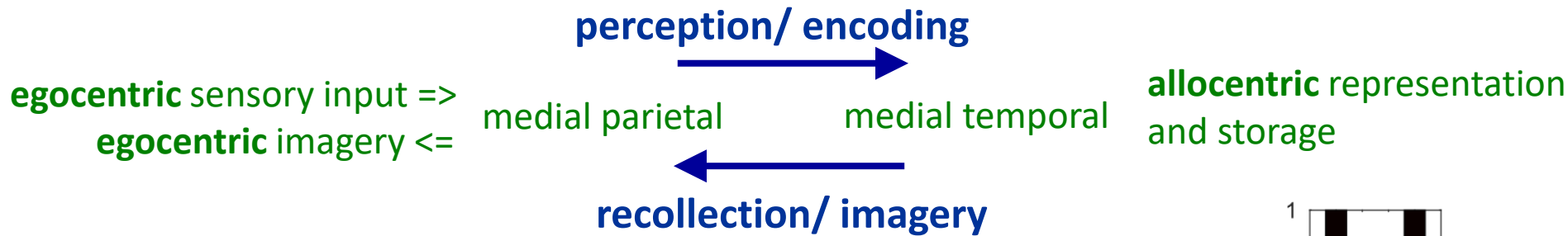
Model of memory & imagery for scenes



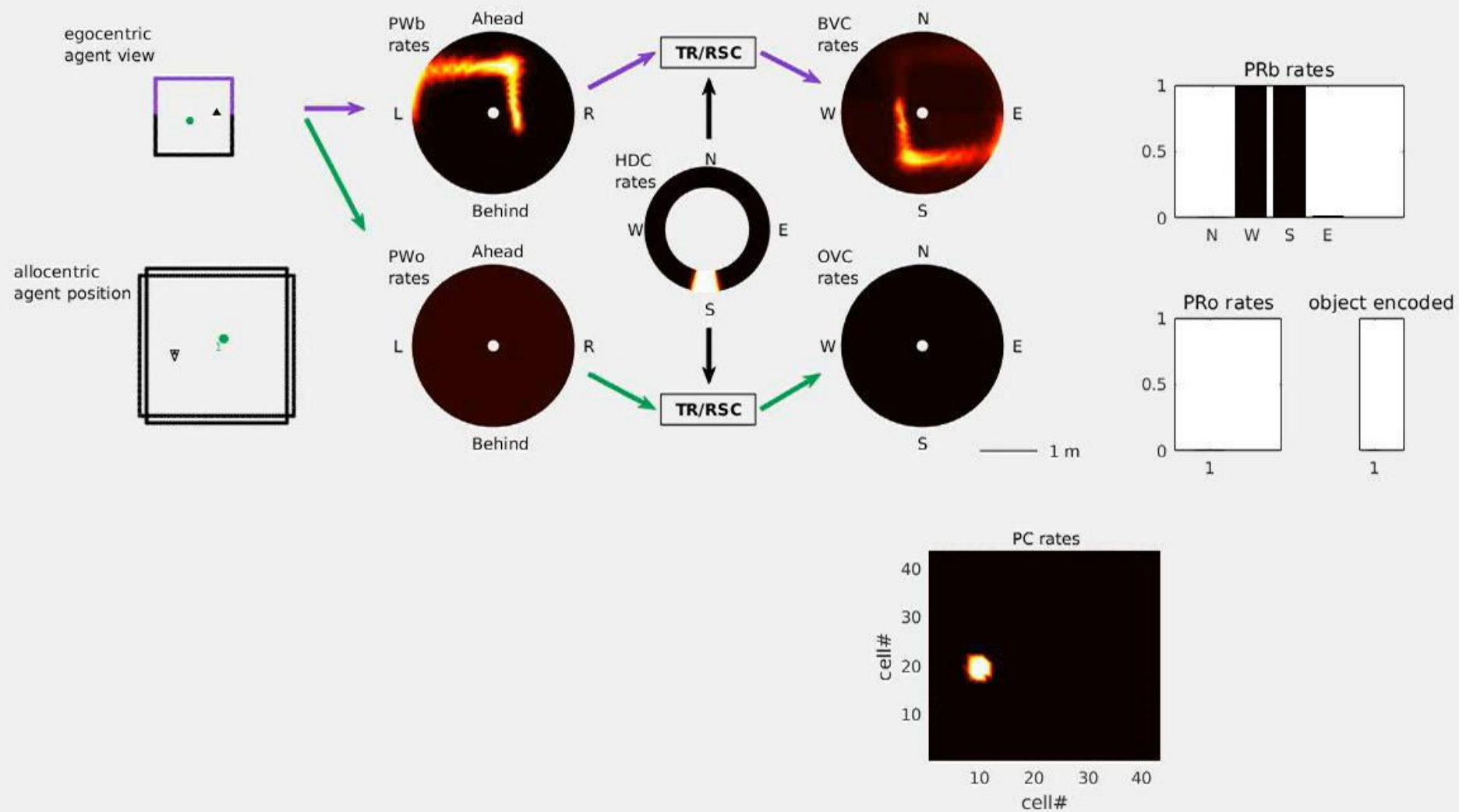
'bottom-up' encoding/ perception
→
←
'top-down' recollection/ imagery



In a familiar environment, MTL connections generate a coherent scene consistent with a single viewpoint (place cells) and direction (HDCs)

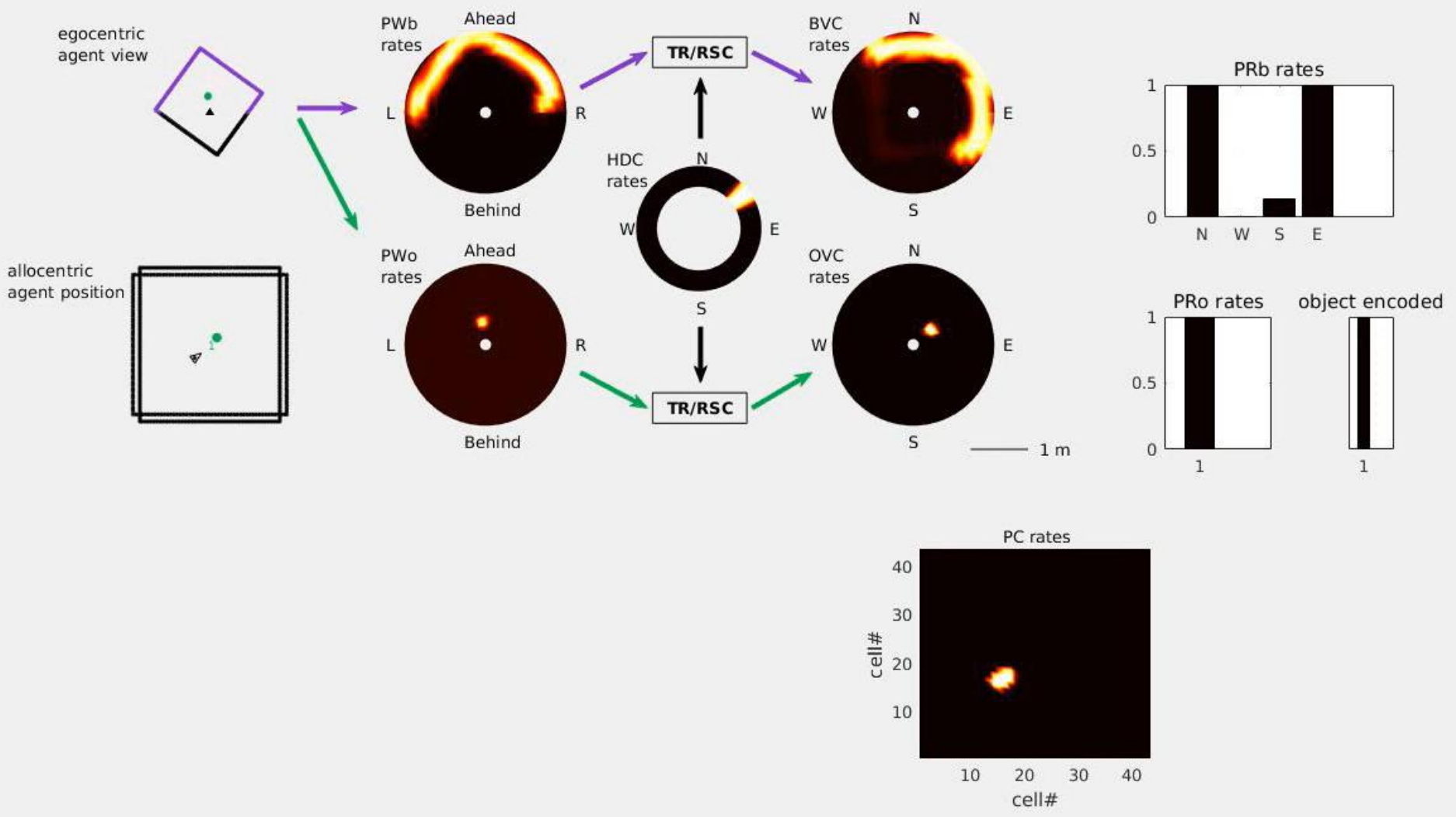


Perceptually driven MTL (bottom-up mode)

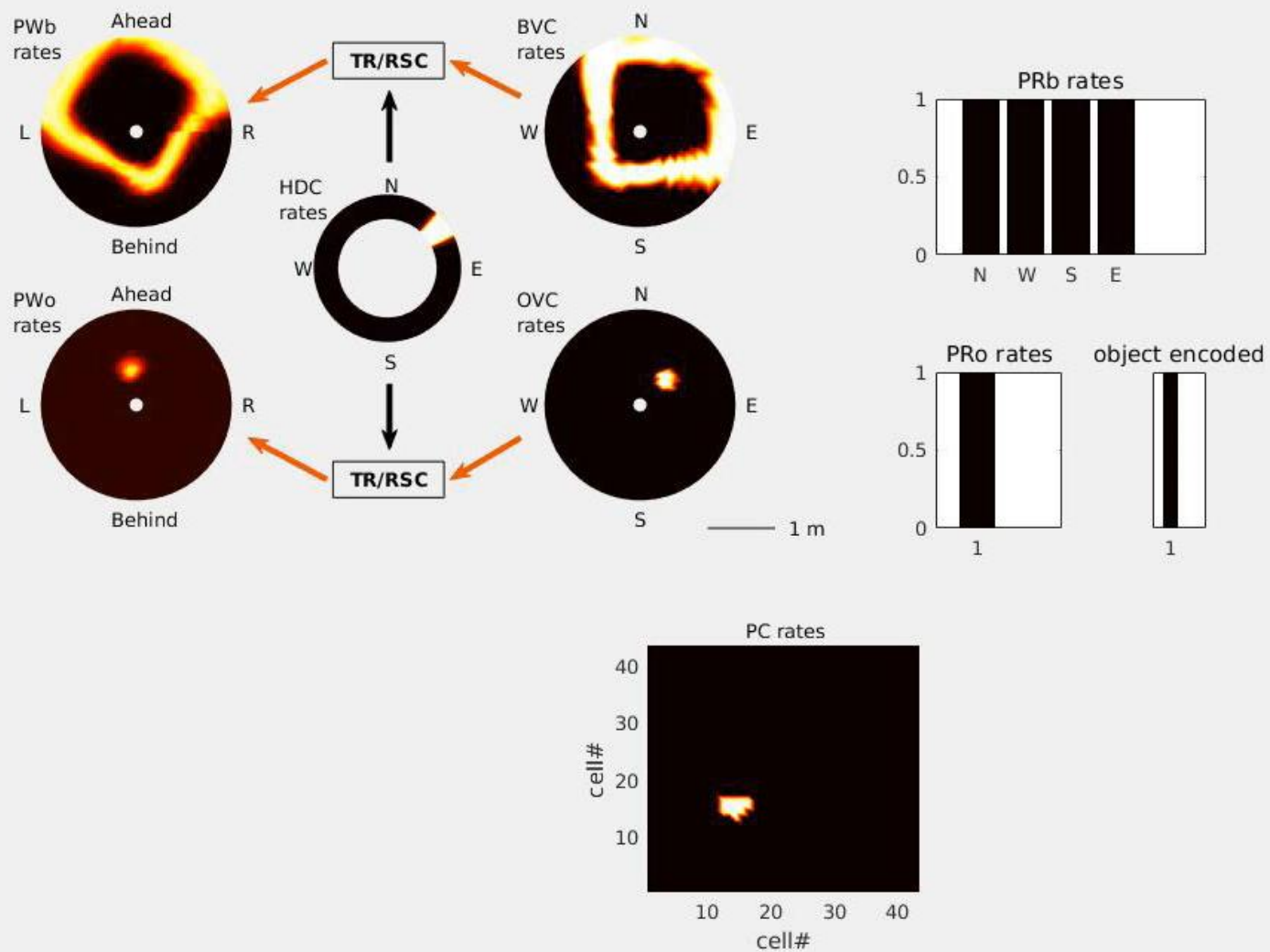


Encountering an object in a familiar environment

Perceptually driven MTL (bottom-up mode)

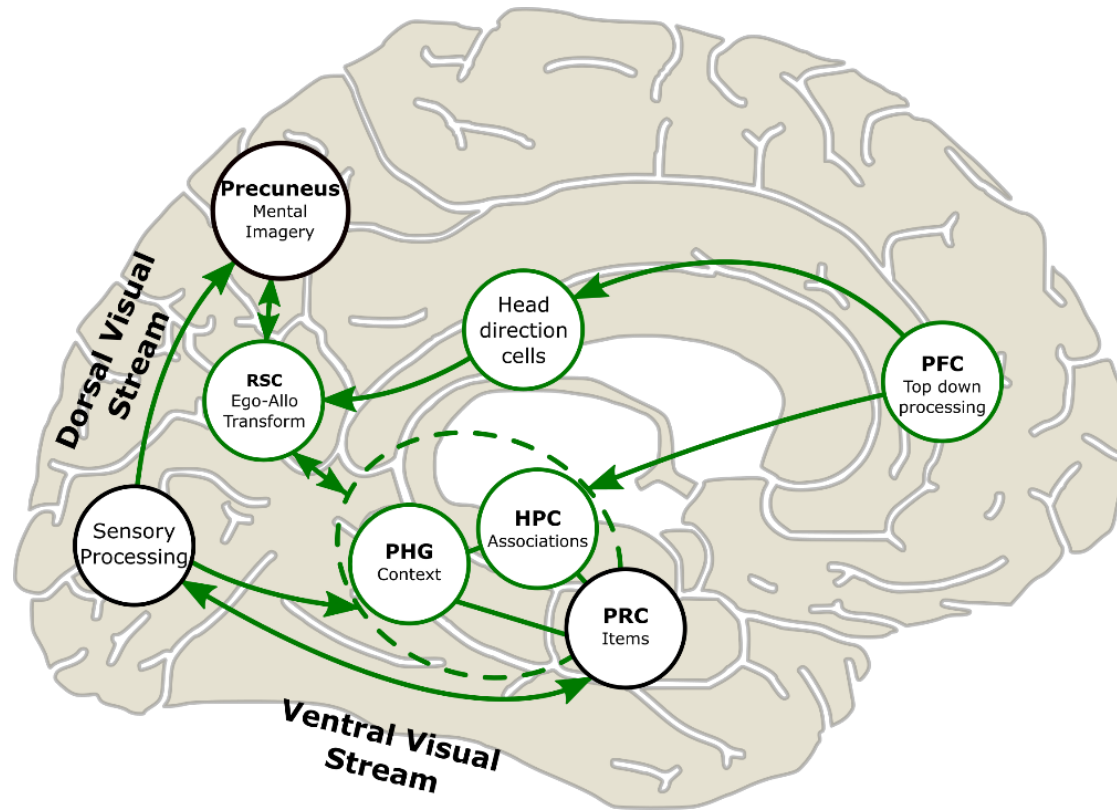


Recollection of encountering the object

egocentric
agent viewallocentric
agent position

Memory enhanced 'perception' of a familiar environment

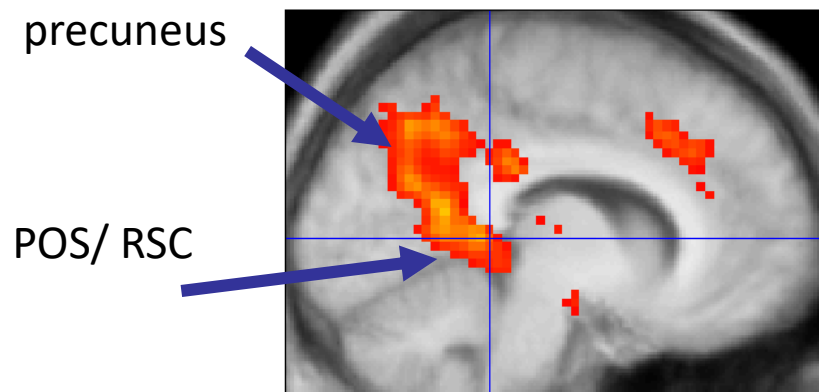
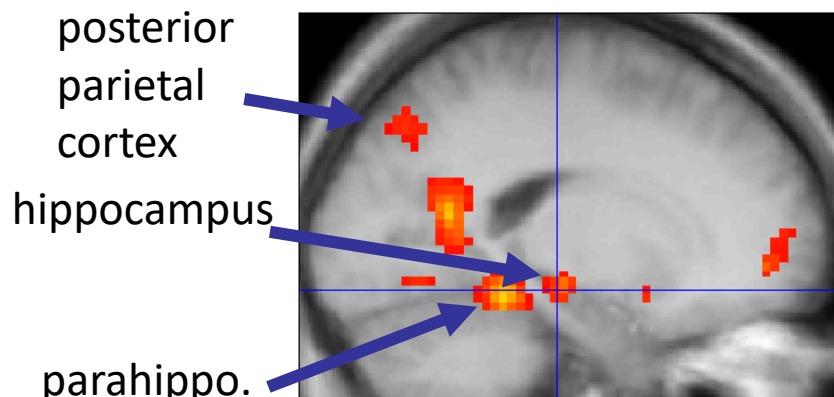
Model allows interpretation of fMRI patterns during recollection/ imagery



In a familiar environment, MTL connections ensure generation of a coherent scene, consistent with a single viewpoint (place cells) and direction (HDCs)

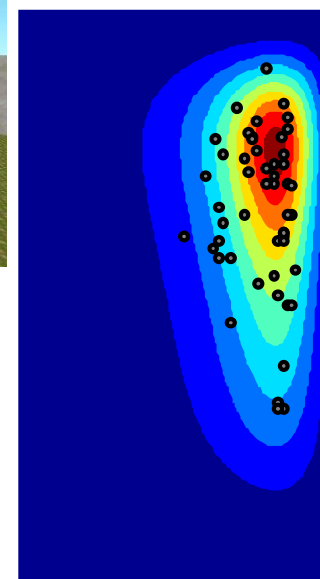
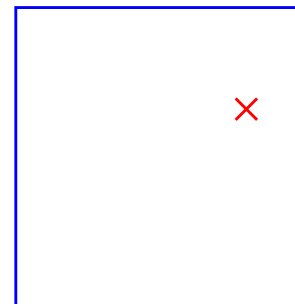
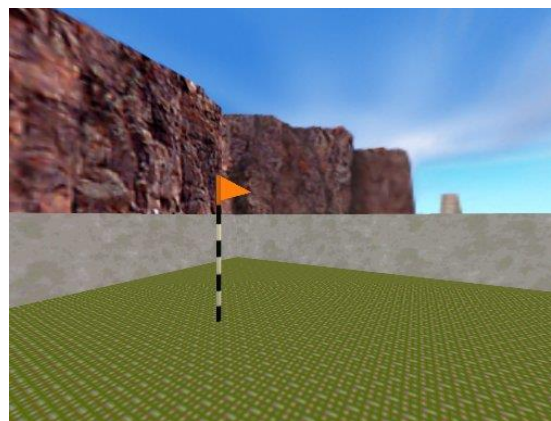
*RSC supports egocentric-allocentric translation, required to associate (allocentric) internal representations with (egocentric) sensory representations
(Egocentric BVCs and OVCs have now been found, Hasselmo & Derdikman labs)*

Model allows interpretation of fMRI patterns during recollection/ imagery



Burgess et al, 2001

& prediction of human search patterns



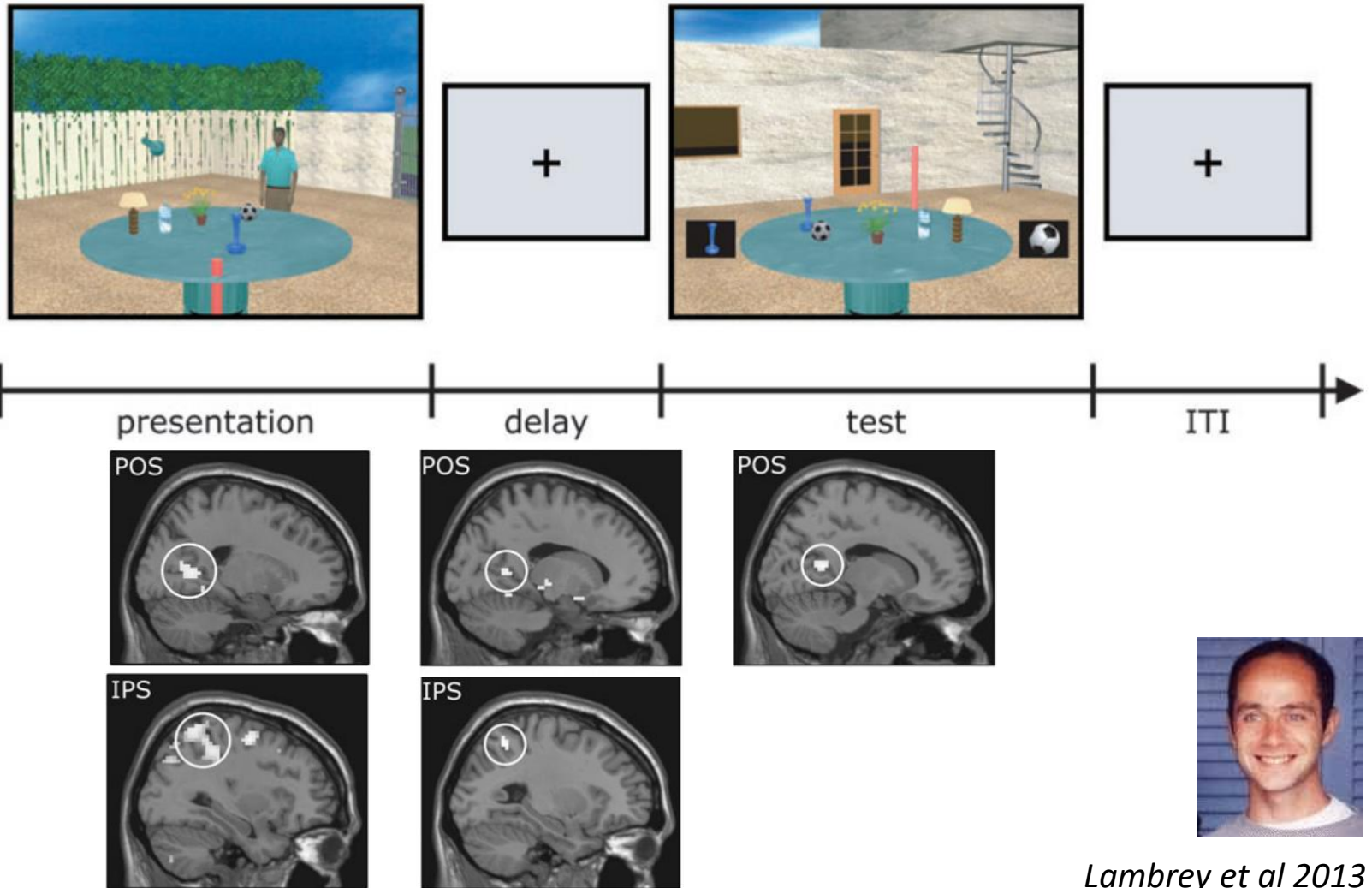
Hartley et al, 2004

The network performs coherent spatial imagery, i.e. related to planning, 'episodic future thinking' and 'scene construction'

*Addis and Schacter, 2007;
Hassabis and Maguire, 2007*

POS/ RSC activity and change of viewpoint in memory

Viewpoint or table will rotate to avatar before test

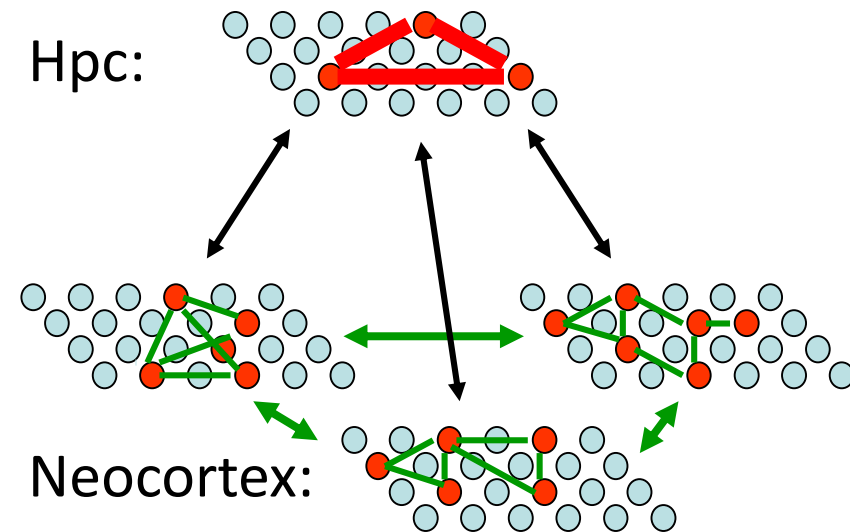


Lambrey et al 2013

RSC associates internal (allocentric) representations to (egocentric) sensory inputs
- strong associations form to stable sensory features (e.g. Auger et al., 2012)

Relation to pattern completion and models of Episodic Memory

- Pattern completion is seen in reconstruction of location-object-identity in scene.
- Consistent with Marr's model of hippocampus & Tulving's idea of holistic episodic recollection/ re-experience.
- Consistent with measures of pattern completion in Episodic memory see *Horner et al (2015)*.



Marr, 1971; Gardner-Medwin, McNaughton, Alvarez, Squire, McClelland, O'Reilly, Treves, Rolls, Teyler & DiScenna; Damasio;

Functional roles for Papez's circuit?

Hippocampus (place cells):
imposing a common viewpoint on
retrieval/ imagery.

Fornix:

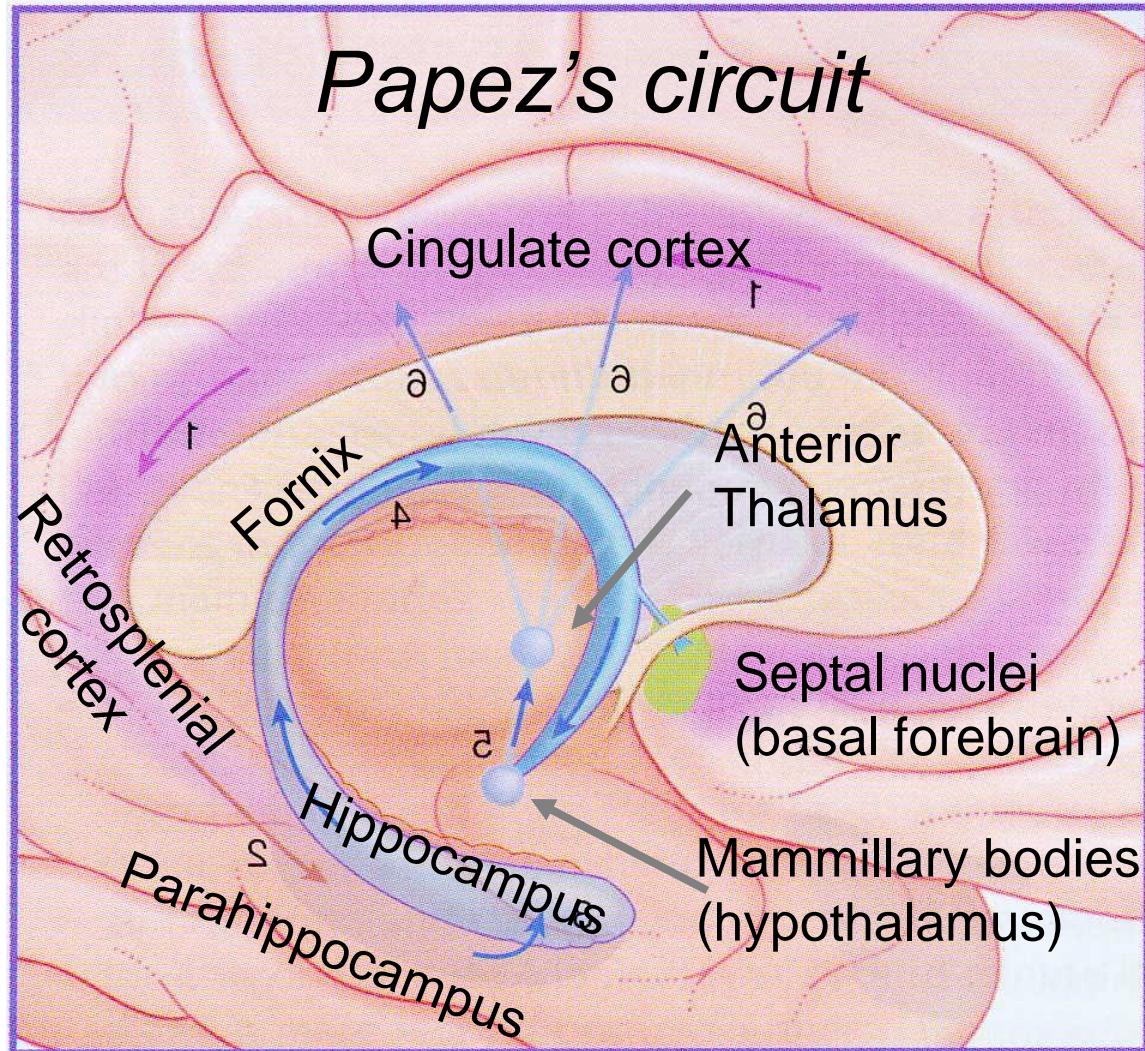
Head-direction cells: imposing a
viewing direction

Theta cells/VCOs: grid cells, path
integration, moving viewpoint in
imagery.

ACh/novelty/learning

Diencephalic amnesia

(Aggleton & Brown, 1999; Gaffan;
Delay & Brion 1969). E.g.,
patient NA (Squire & Slater,
1978), Korsakoff's syndrome.

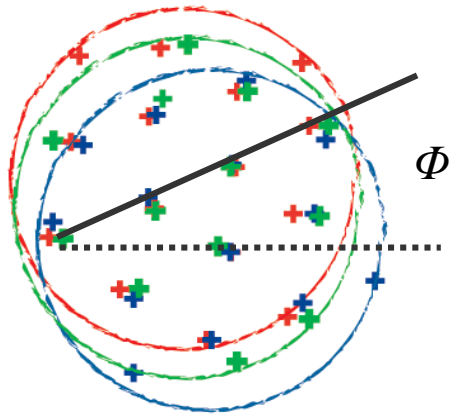
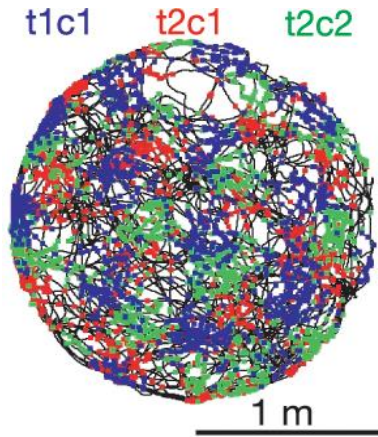


Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) *Place cells & boundary vector cells*
- 3) *Neural level model of Spatial Memory and Imagery*
- 4) ***Place and grid cells, environmental and self-motion inputs?***
- 5) *Grid cells as dynamic imagery?*
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

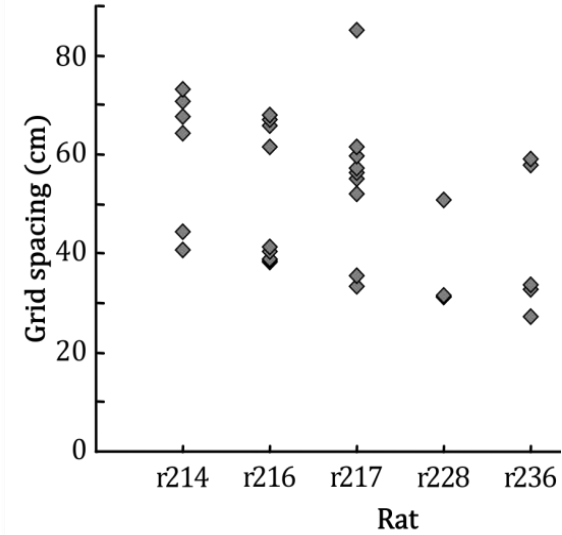
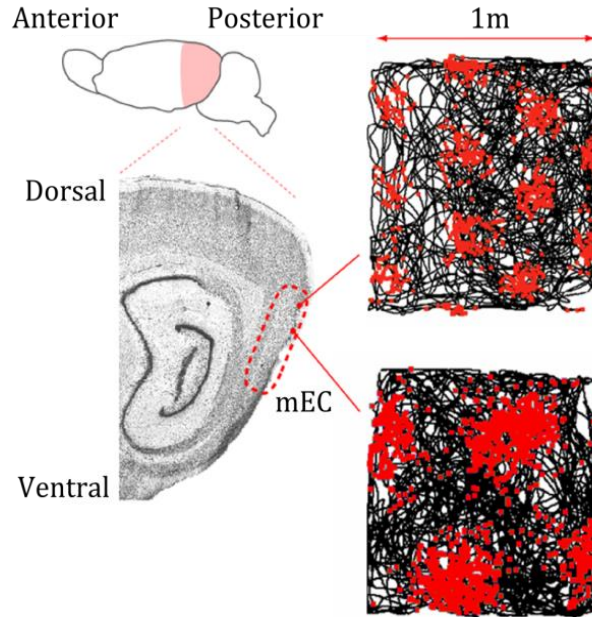
Grid cells – thought to represent location by integrating self-motion.

The grids of nearby cells share orientation & scale



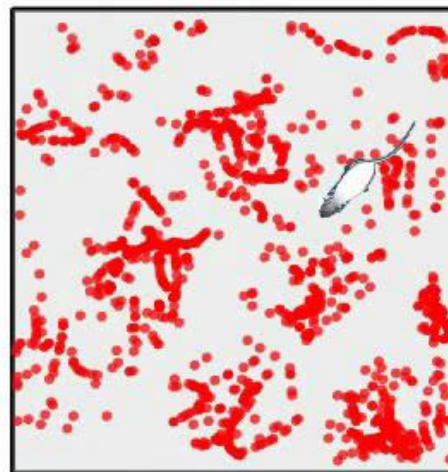
Hafting et al., 2005

Grid cells occur in modules with discrete scales



Barry et al, 2007;

see also Stensola et al., 2012



Video by Julija Krupic

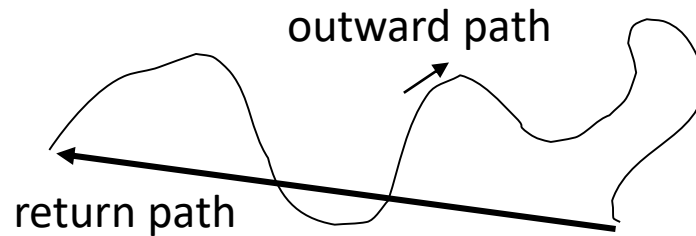


Two ways to know where you are:

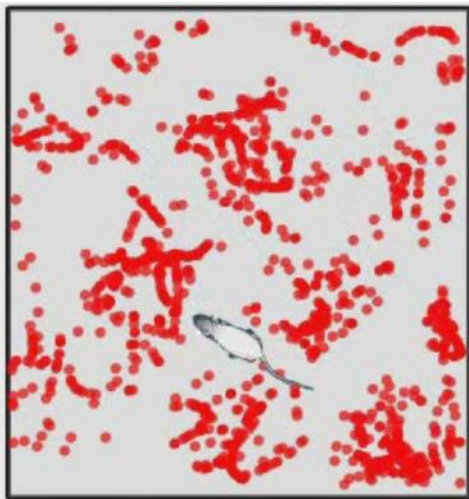
1. Environmental information

(Environmental boundaries particularly influence place cells)

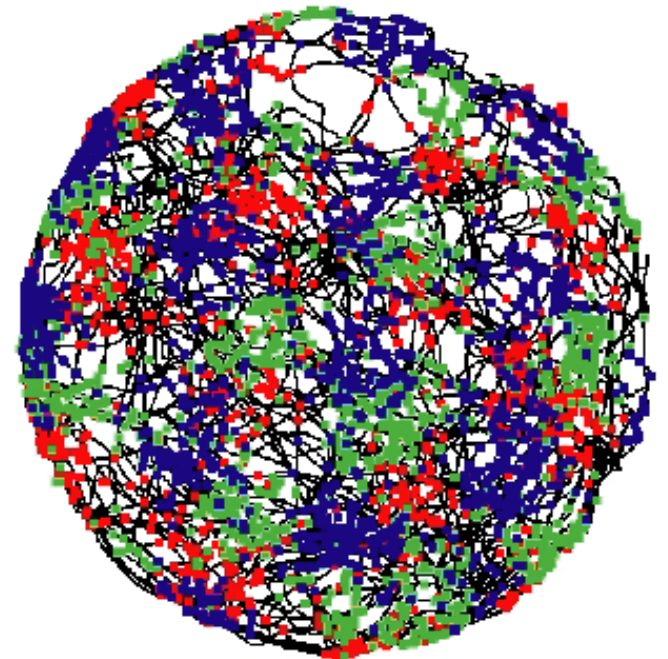
2. Path integration



Grid cells



Hafting et al., 2005



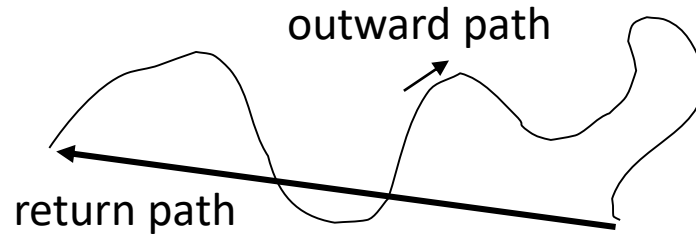
Video by Julija Krupic

Two ways to know where you are:

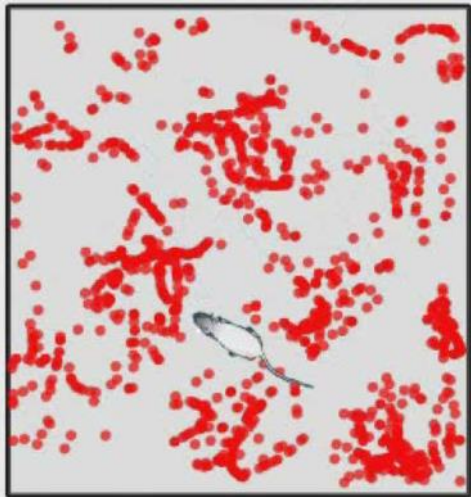
1. Environmental information

(Environmental boundaries particularly influence place cells)

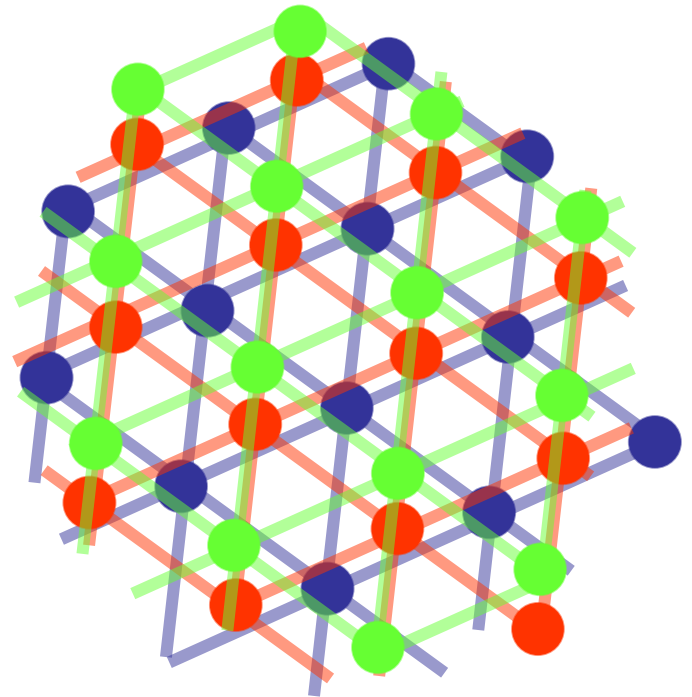
2. Path integration



Grid cells



Hafting et al., 2005



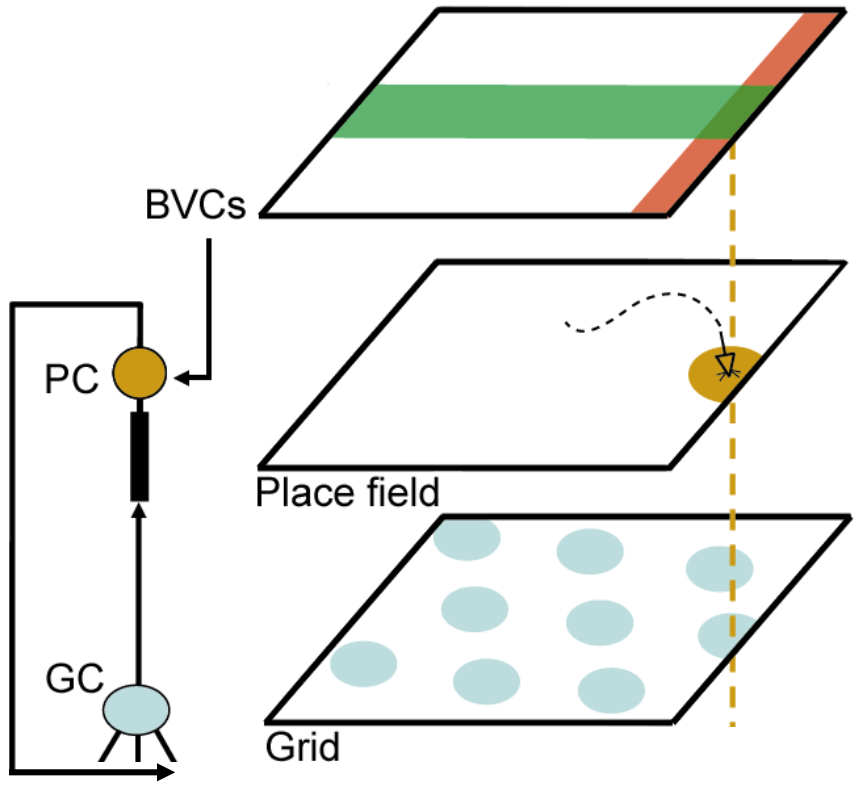
Video by Julija Krupic

Interactions between place cells and grid cells

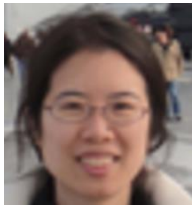
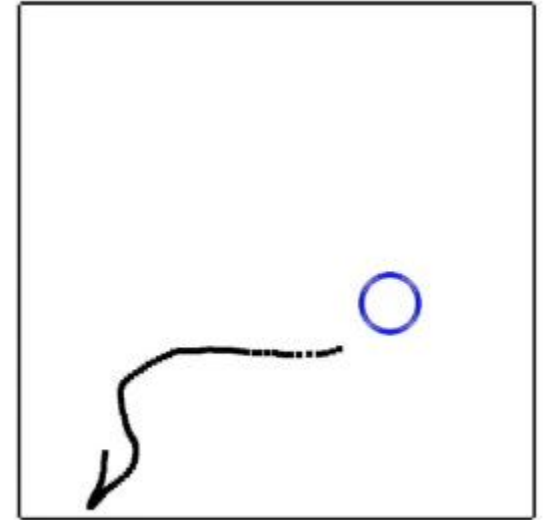
Estimating self-location combines environmental & self-motion information

Environmental information
(\Leftrightarrow Boundary Vector Cells)

Self- motion

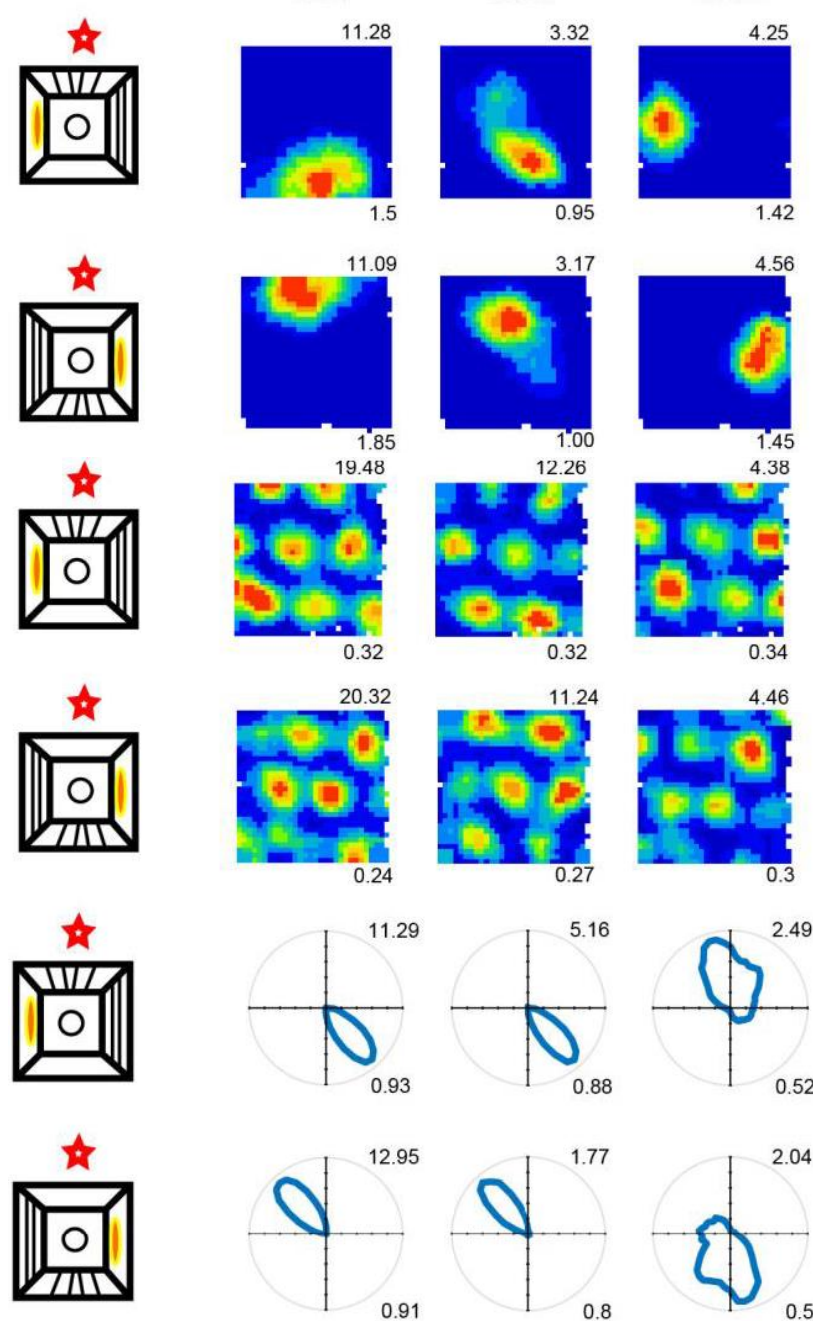


2D VR for mice (invisible reward task)

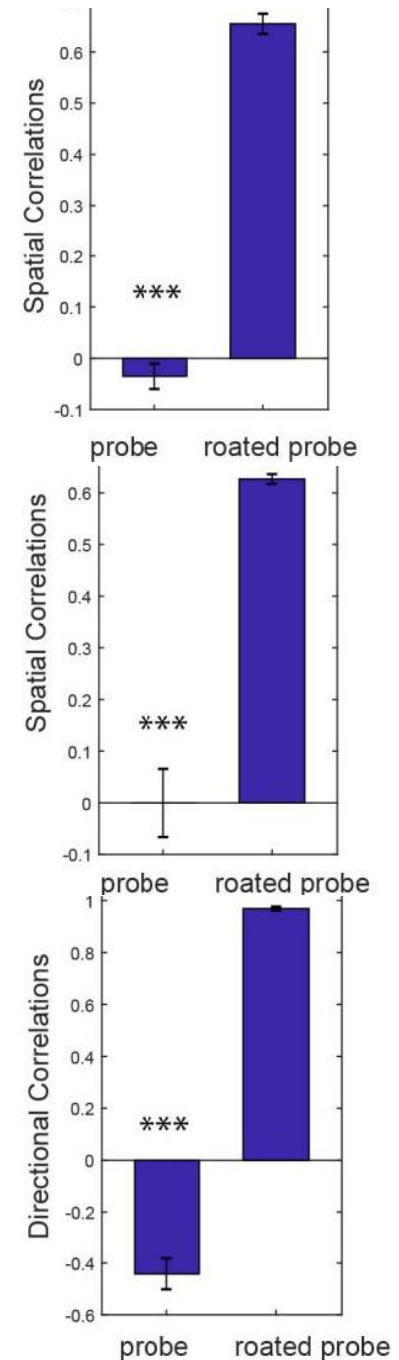


Guifen Chen, John King, Yi Lu, Francesca Cacucci, Neil Burgess, eLife 2018

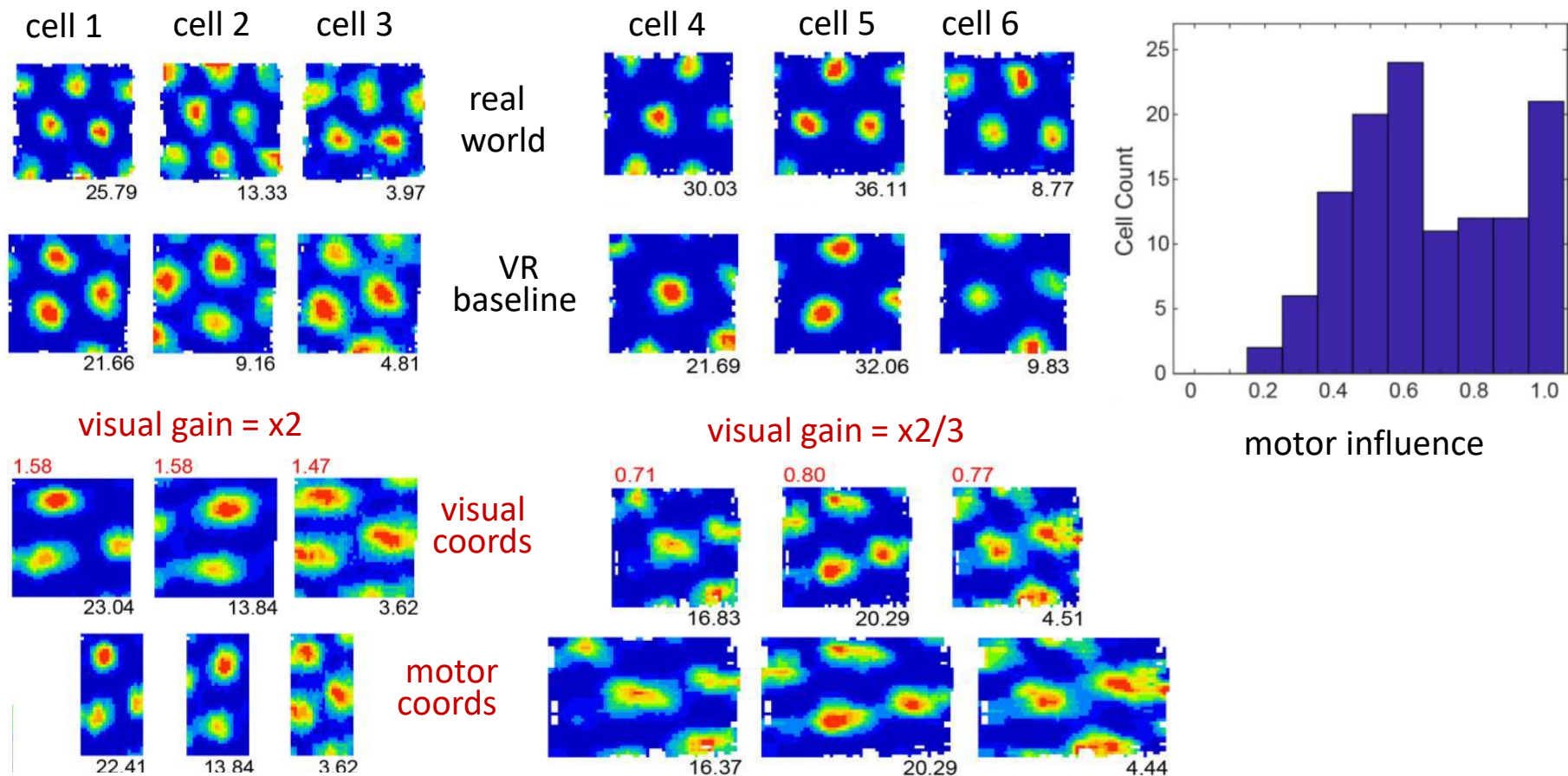
2d VR allows expression of normal place, grid & head-direction firing patterns, controlled by virtual cues (e.g. 180° rotation of VR and entry point)



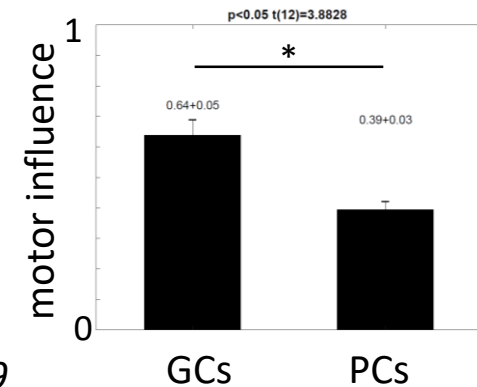
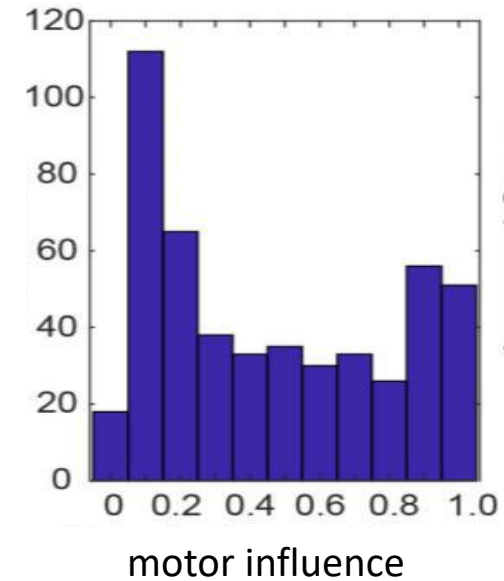
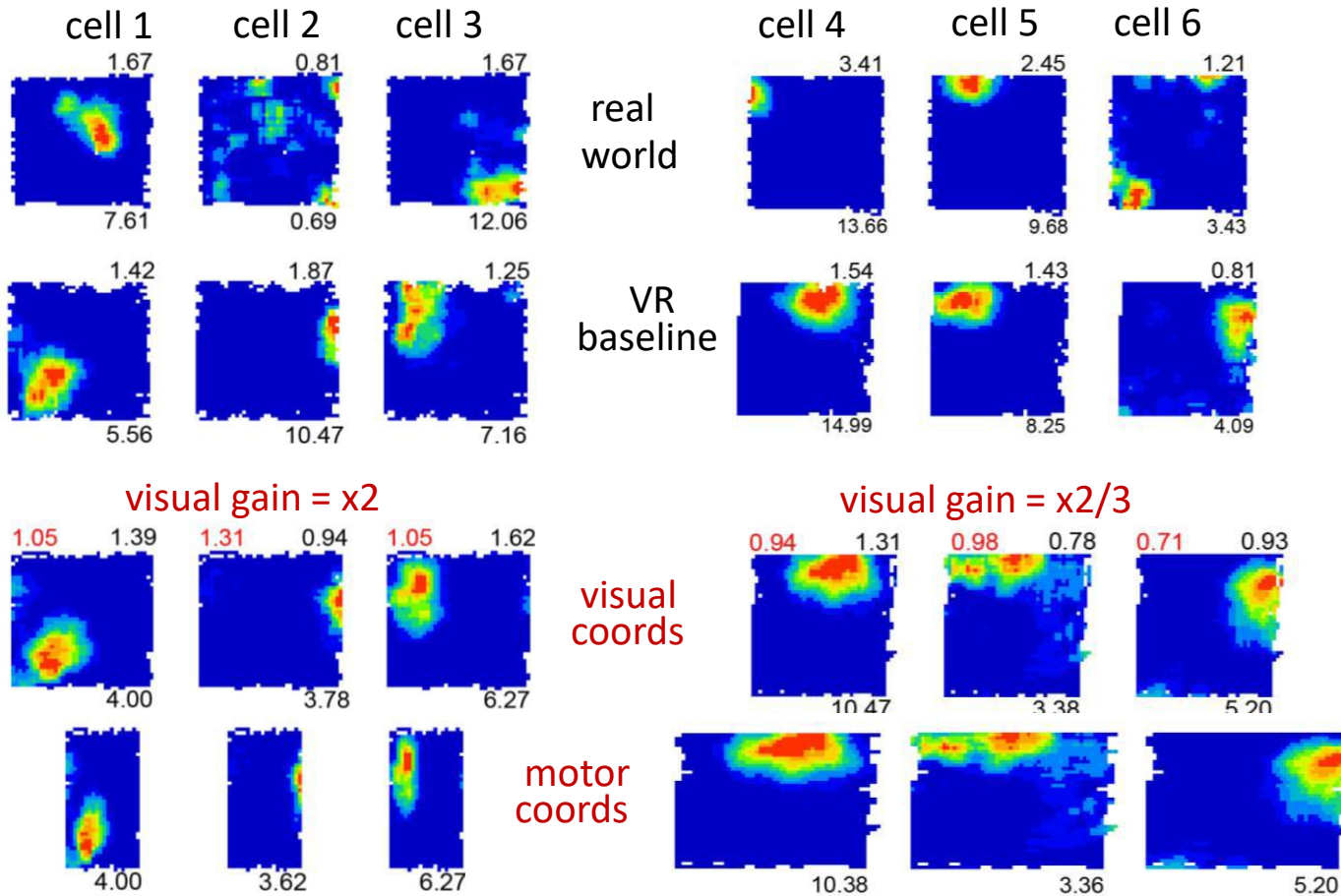
Correlation with baseline



Grid cell firing patterns reflect self-motion more than vision



Place cell firing patterns reflect vision more than self-motion

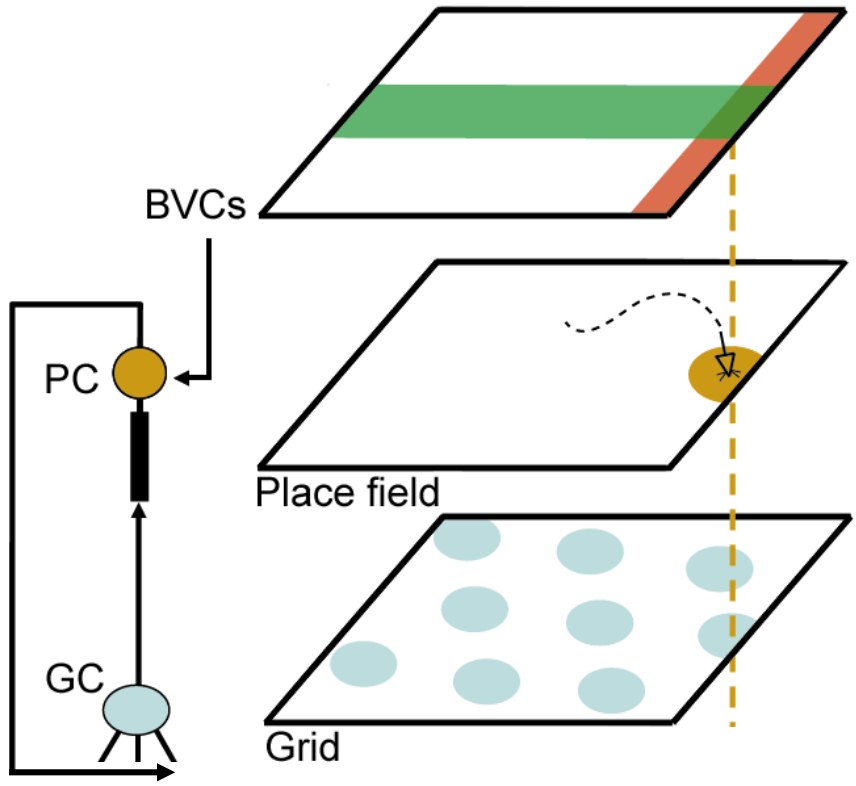


Interactions between place cells and grid cells

Estimating self-location combines environmental & self-motion information.

Environmental information
(\Leftrightarrow Boundary Vector Cells)

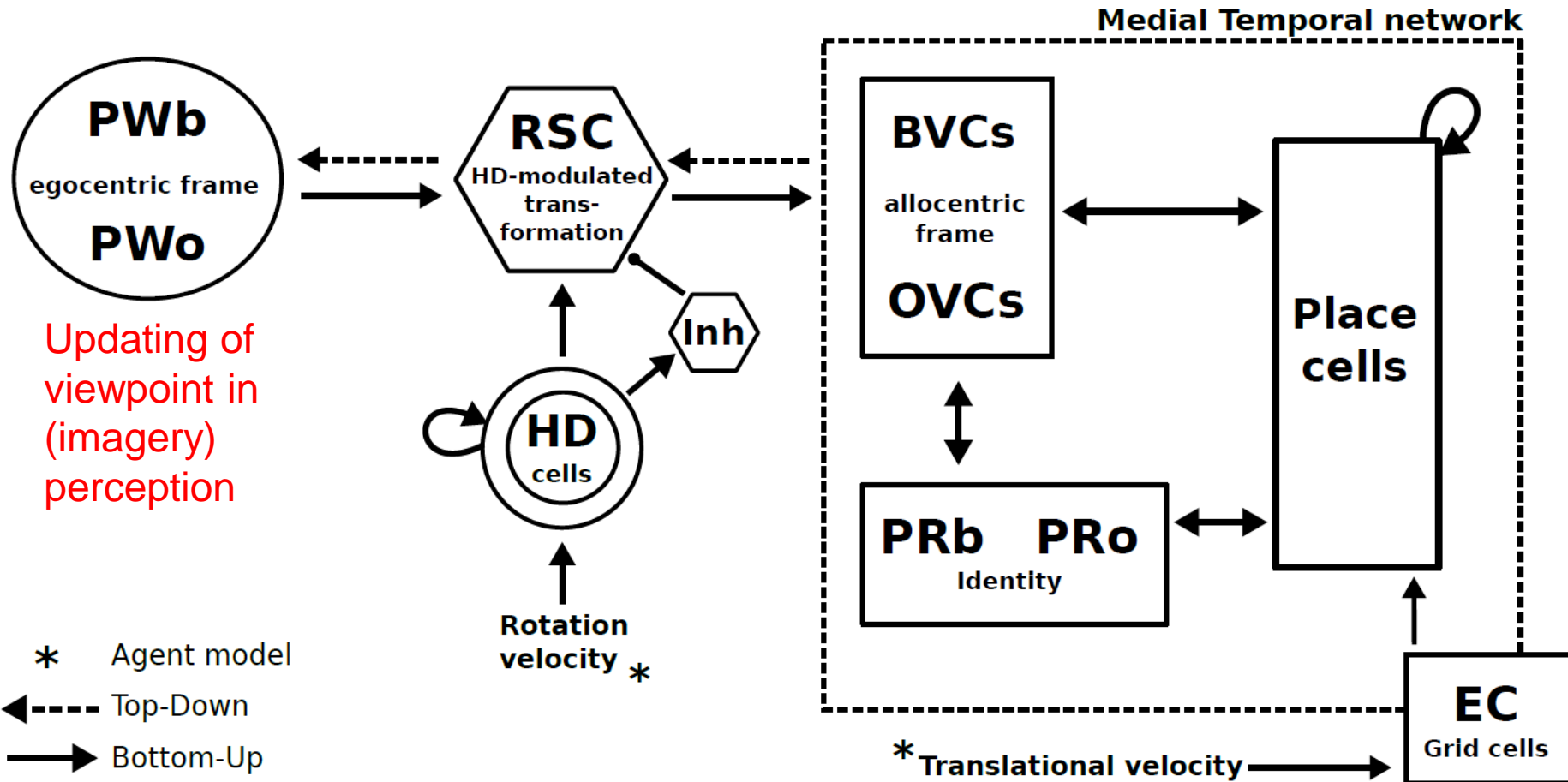
Self- motion



Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) *Place cells & boundary vector cells*
- 3) *Neural level model of Spatial Memory and Imagery*
- 4) *Place and grid cells, environmental and self-motion inputs?*
- 5) ***Grid cells as dynamic imagery?***
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

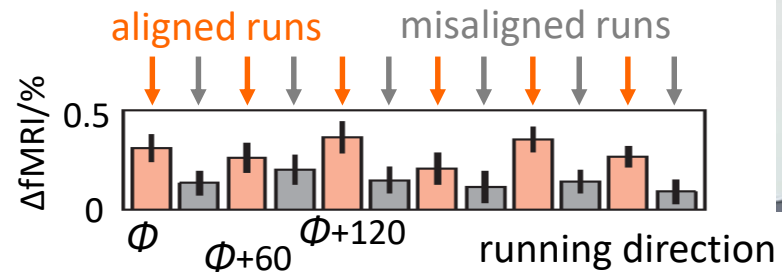
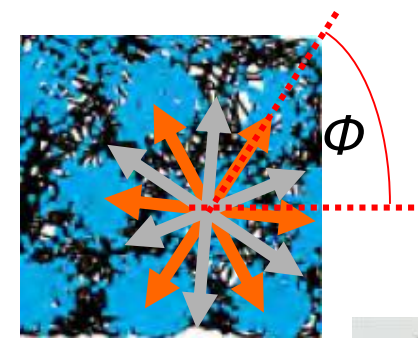
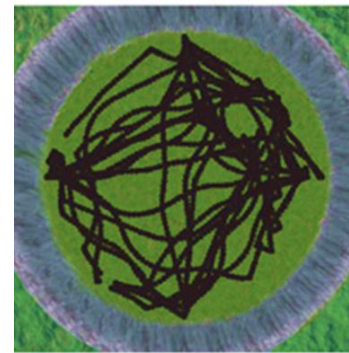
Grid cells and memory/imagery



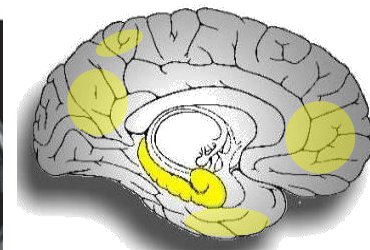
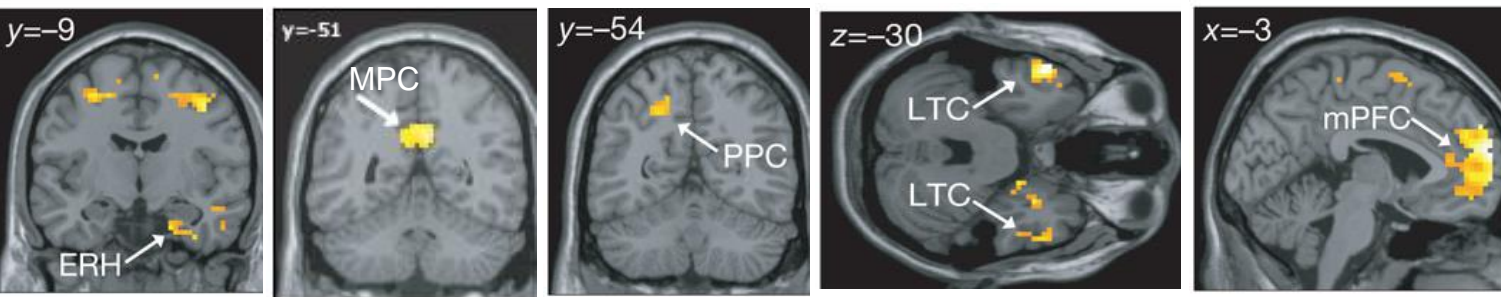
Grid cells in the human autobiographical memory system?

Doeller, Barry, Burgess, 2010

populations of *aligned* grids (modules) => changes in fMRI signal with virtual running direction



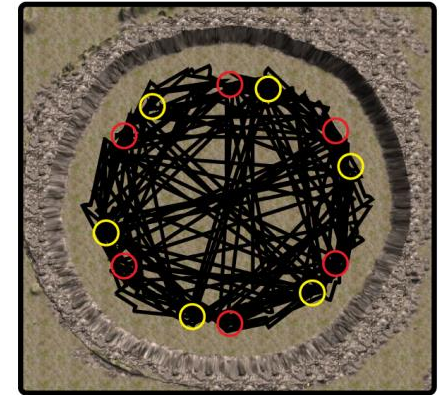
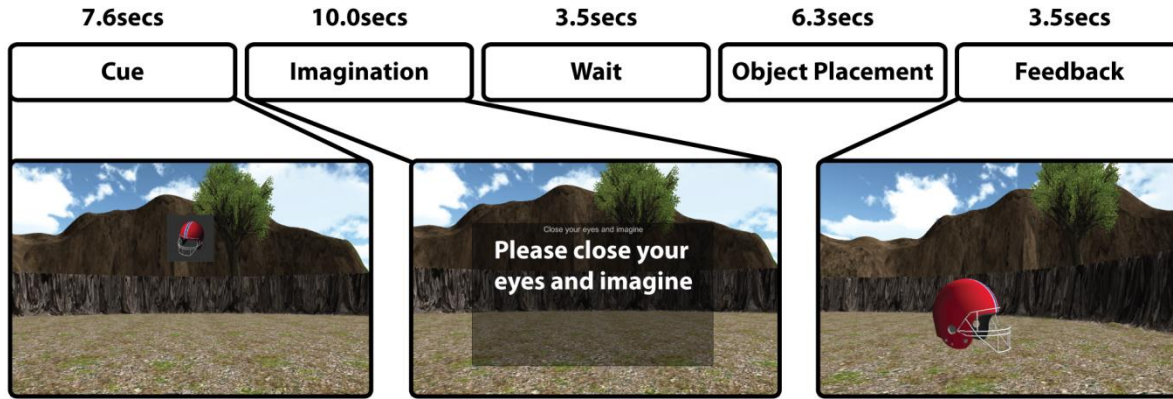
Task designed by John King



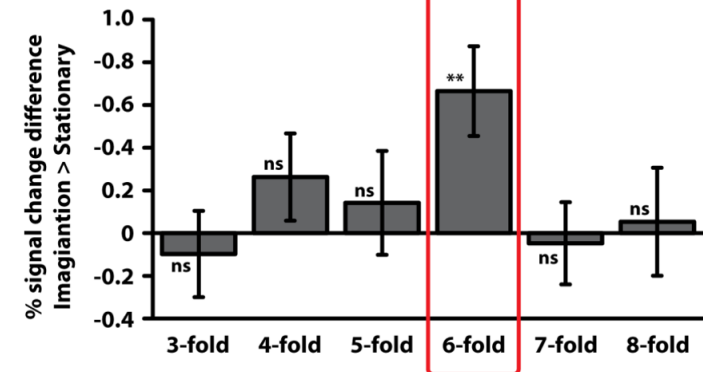
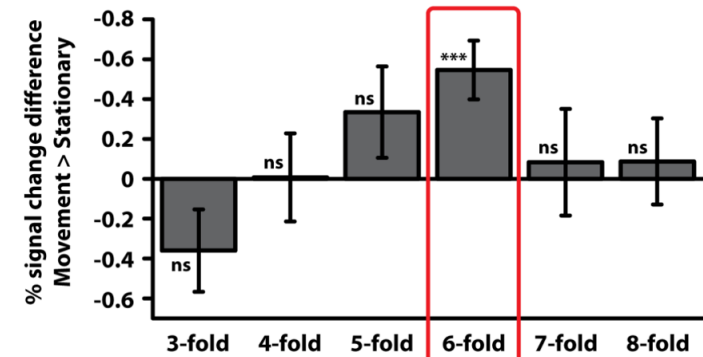
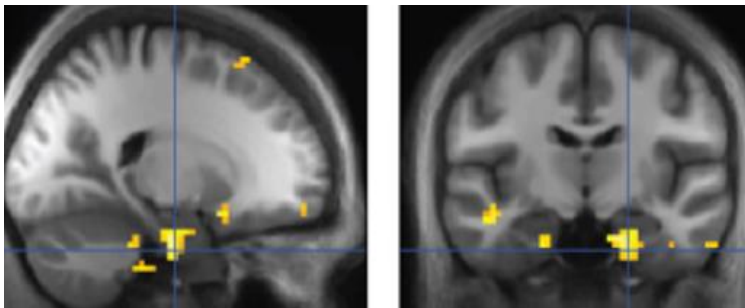
Autobiographical memory system

=> Grid cells allow path integration, and movement of viewpoint in imagery?

Grid-like processing of movement of viewpoint in imagery



60° symmetry in fMRI signal with *imagined* running direction in Entorhinal cortex (aligned with that in virtual movement)



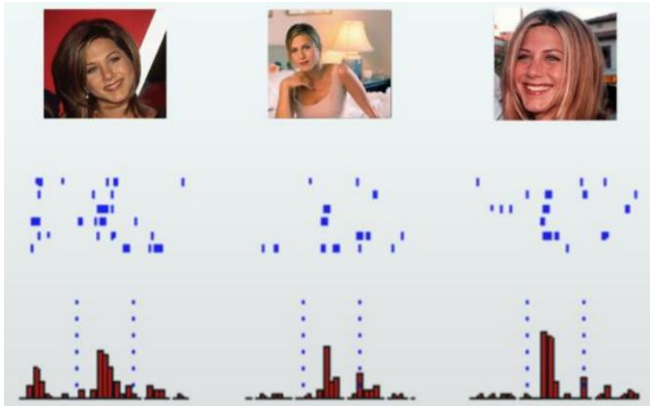
Horner et al., 2016



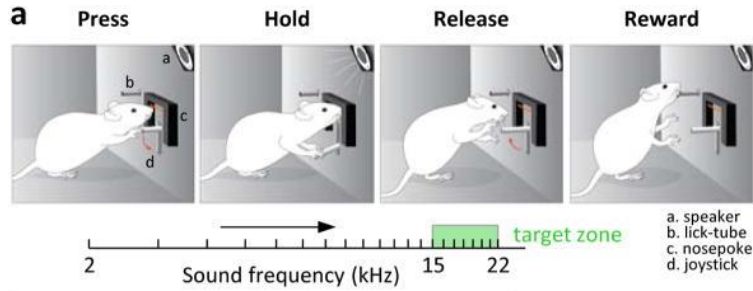
Abstract neural representations

- 1) *Frames of reference for spatial representation*
- 2) *Place cells & boundary vector cells*
- 3) *Neural level model of Spatial Memory and Imagery*
- 4) *Place and grid cells, environmental and self-motion inputs?*
- 5) ***Grid cells as dynamic imagery?***
- 6) *Place and grid cells, representing states and transitions for planning?*
 - A. *Hippocampus & striatum: Model-based versus model-free RL?*
 - B. *Dual representations theory, PTSD and intrusive imagery*

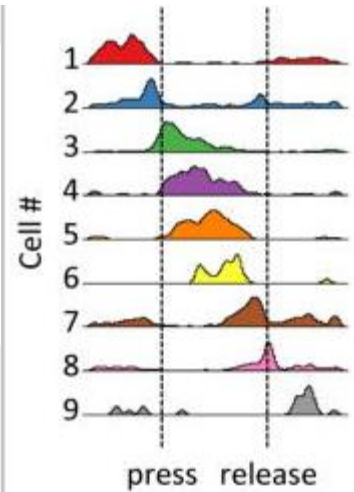
Hippocampal cells can represent abstract concepts, such as 'place' but also, e.g., personal identity or sound frequency?



Quiroga et al., (2005)

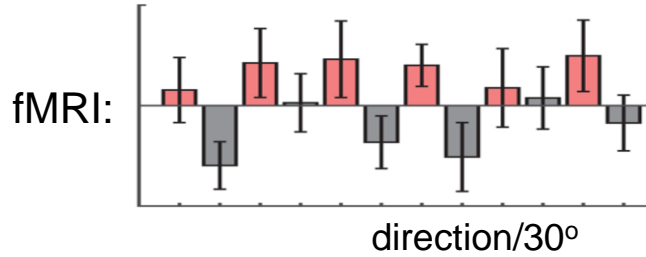


Aronov, Nevers, Tank (2017)

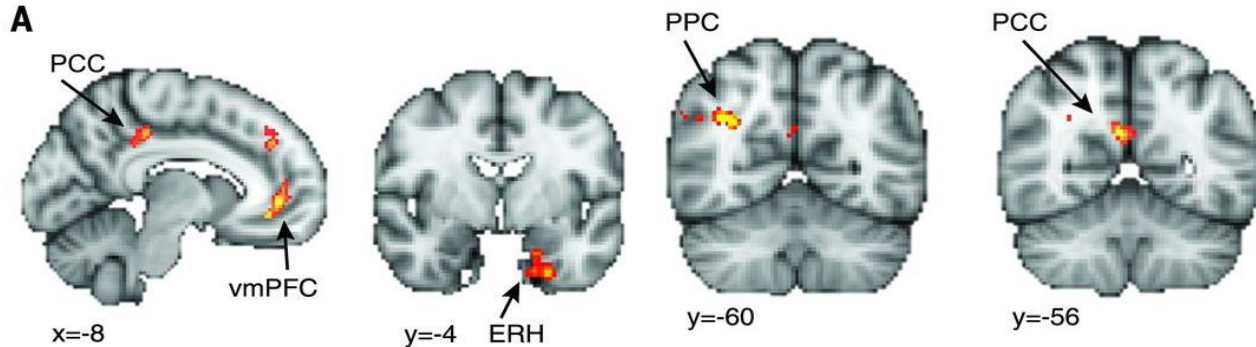


Grid cell firing patterns reflect the transition structure of learned conceptual spaces?

Navigation in space of bird neck & leg length



Constantinescu, O'Reilly, Behrens 2016

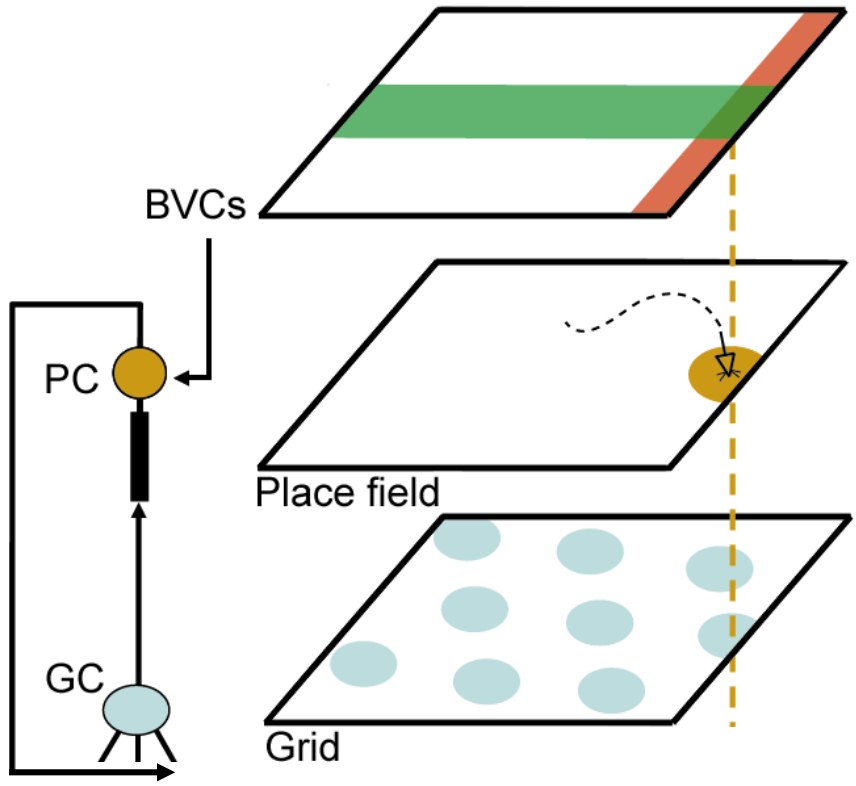


Interactions between place cells and grid cells

Representing bodies of conceptual knowledge (states) and transitions between them?

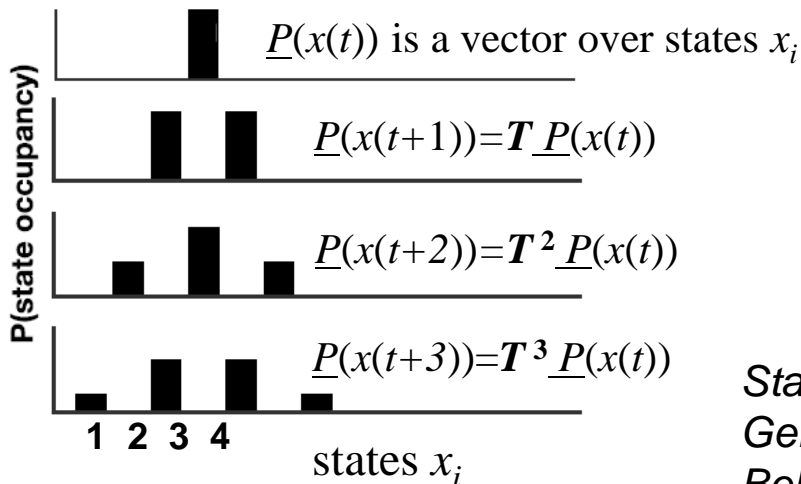
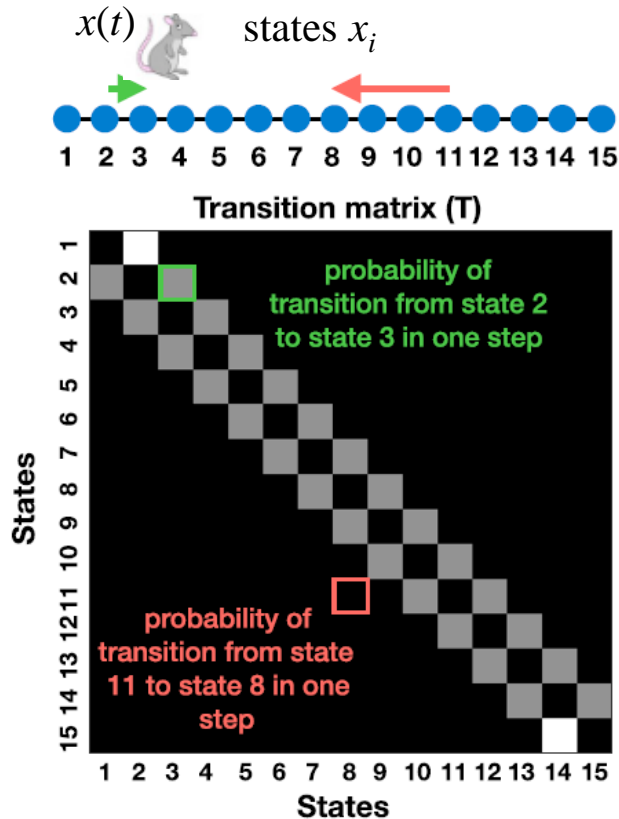
State information (place)
(\Leftrightarrow Feature Vector Cells?)

Transition structure
(self- motion)



'Intuitive Planning..'

..with neural populations



$\underline{P}(x(t))$ is a vector over states x_i :

PC_i firing profile is \underline{F}_i
firing rate is $f_i(x(t))$

$$\underline{P}(x(t)) \sim \sum_j f_j(x(t)) \underline{F}_j$$

$$\underline{P}(x(t+1)) \sim \sum_j f_j(x(t)) T \underline{F}_j$$

GC_i firing profile = \underline{G}_i
firing rate = $g_i(x(t))$

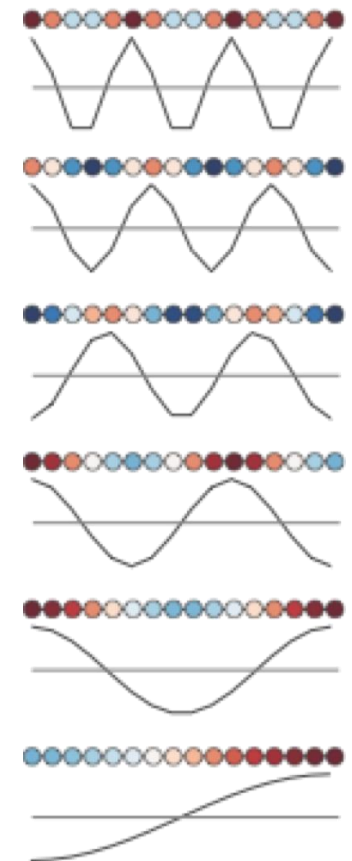
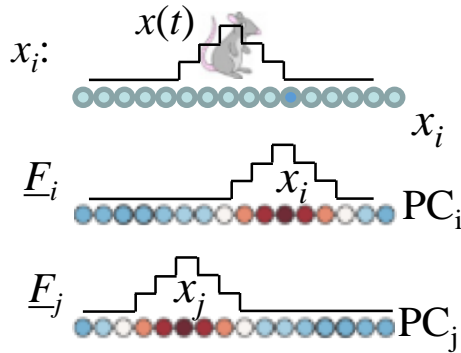
$$\underline{P}(x(t)) \sim \sum_j g_j(x(t)) \underline{G}_j$$

$$\underline{P}(x(t+1)) \sim \sum_j g_j(x(t)) T \underline{G}_j$$

$$\text{If } T \underline{G}_j(x) = \lambda_j \underline{G}_j(x)$$

$$\underline{P}(x(t+1)) \sim \sum_j \lambda_j g_j(x(t)) \underline{G}_j$$

$$\begin{aligned} \underline{P}(x(\tau \geq t) = x_i) \\ \sim \sum_j (\gamma \lambda_i + \gamma^2 \lambda_i^2 + \dots) g_j(x(t)) \underline{G}_j \\ \sim \sum_j g_j(x(t)) / (1 - \gamma \lambda_i) \underline{G}_j \end{aligned}$$



Stachenfeld, Botvinick,
Gershman, Gerstner, Baram..
Behrens

Place cell read-out of GCs

PC_i firing profile is \underline{F}_i , firing rate is $f_i(x(t))$ driven by GCs?

If $f_i(x(t)) \sim \sum_j w_{ij} g_j(x(t))$

[e.g. Hebbian $w_{ij} \sim \underline{F}_i \cdot \underline{G}_j$]

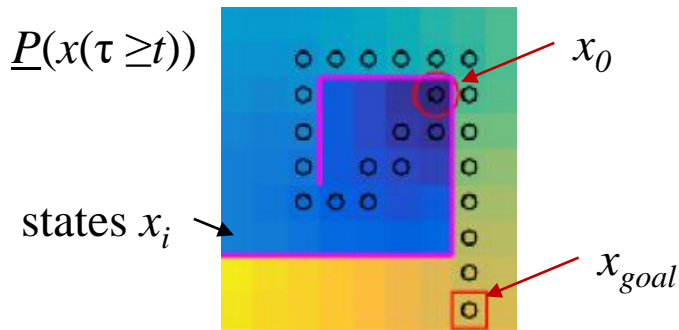
then $f_i(x(t)) \sim P(x(t) = x_i)$

If $f_i(x(t)) \sim \sum_j \lambda_j w_{ij} g_j(x(t))$

then $f_i(x(t)) \sim P(x(t+1) = x_i)$

If $f_i(x(t)) \sim \sum_j w_{ij} g_j(x(t)) / (1 - \gamma \lambda_j)$

Then $f_i(x(t)) \sim P(x(\tau \geq t) = x_i)$



$\underline{P}(x(t))$ is a vector over states x_i :

PC_i firing profile is \underline{F}_i
firing rate is $f_i(x(t))$

$\underline{P}(x(t)) \sim \sum_j f_j(x(t)) \underline{F}_j$

$\underline{P}(x(t+1)) \sim \sum_j f_j(x(t)) \mathbf{T} \underline{F}_j$

GC_i firing profile = \underline{G}_i
firing rate = $g_i(x(t))$

$\underline{P}(x(t)) \sim \sum_j g_j(x(t)) \underline{G}_j$

$\underline{P}(x(t+1)) \sim \sum_j g_j(x(t)) \mathbf{T} \underline{G}_j$

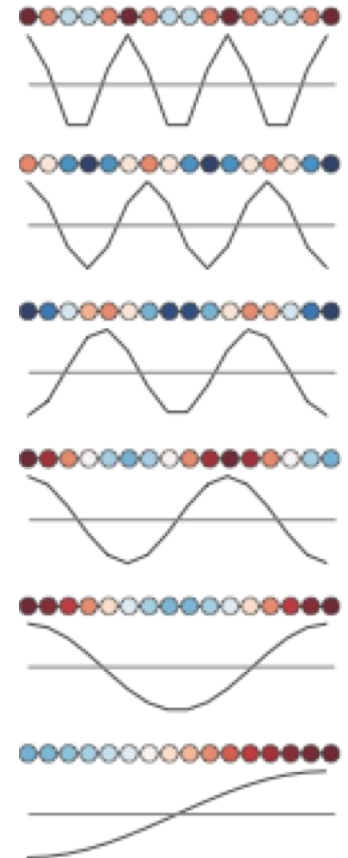
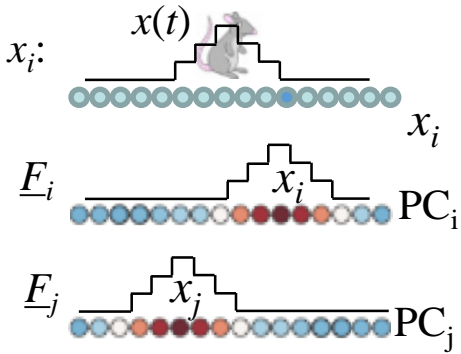
If $\mathbf{T} \underline{G}_j(x) = \lambda_j \underline{G}_j(x)$

$\underline{P}(x(t+1)) \sim \sum_j \lambda_j g_j(x(t)) \underline{G}_j$

$\underline{P}(x(\tau \geq t) = x_i)$

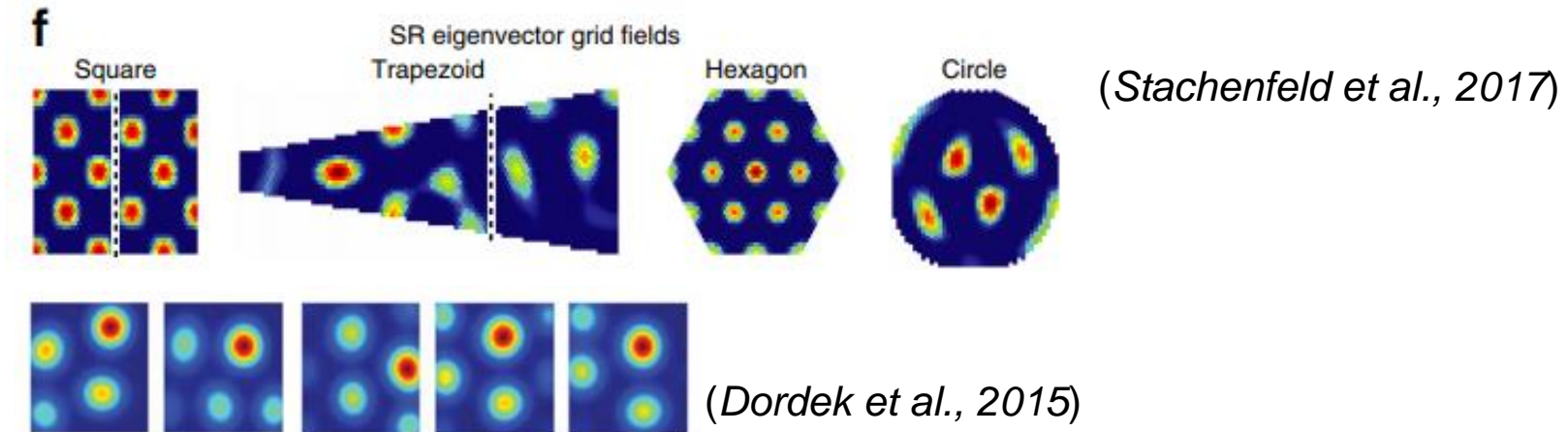
$\sim \sum_j (\gamma \lambda_j + \gamma^2 \lambda_j^2 + \dots) g_j(x(t)) \underline{G}_j$

$\sim \sum_j g_j(x(t)) / (1 - \gamma \lambda_j) \underline{G}_j$



So.. If you want a set of basis vectors to represent where you are in state space, choosing Eigenvectors of the task Transition Matrix makes planning easy.

And.. Grid firing profiles might be Eigenvectors of a diffusive transition matrix \mathbf{T} (i.e. $\mathbf{T} \underline{G}_i(x) = \lambda_i \underline{G}_i(x)$), or of the covariance matrix of PC firing (e.g. learned via Oja's rule)



Implications

- *Place cell read-out shifts from current location to future locations by re-weighting GC inputs, can give Successor Representation (SR)*
- *Gradient ascent on SR allows navigation to any other state*
- *Local changes to Transitions require re-learning of eigenvectors (GCs): via replay?*
- *Common transition structure across tasks captured by GCs, while PCs 'remap' to specific stimuli, allows generalisation to new tasks (aka 'schemas' & 'consolidation' of statistical structure), see 'T.E.M.' (Whittington et al BioRxiv, 2019)*

Conclusions

- Considerable progress has been made in understanding how environmental and self-motion information combine in neural representations of location and orientation in rodents.
- We can use this to create a neural-level understanding of spatial memory, learning and imagination in humans, and begin to apply it to conceptual knowledge?

Thanks to:

*Andrej Bičanski
John King
Guifen Chen
Yi Lu
John O'Keefe
Francesca Cacucci
Lone Hørlyck
James Bisby
Tim Behrens*

*Caswell Barry
Dan Bush
Christian Doeller
Aidan Horner
Colin Lever
Hugo Spiers
Suzanna Becker
Tom Hartley
Chris Brewin*

