UCL NEUROSCIENCE

UCL

# Neural representations of spaces and concepts

Neil Burgess
Institute of Cognitive Neuroscience
University College London

SWC PhD program, October 2018

## *Abstract neural representations*

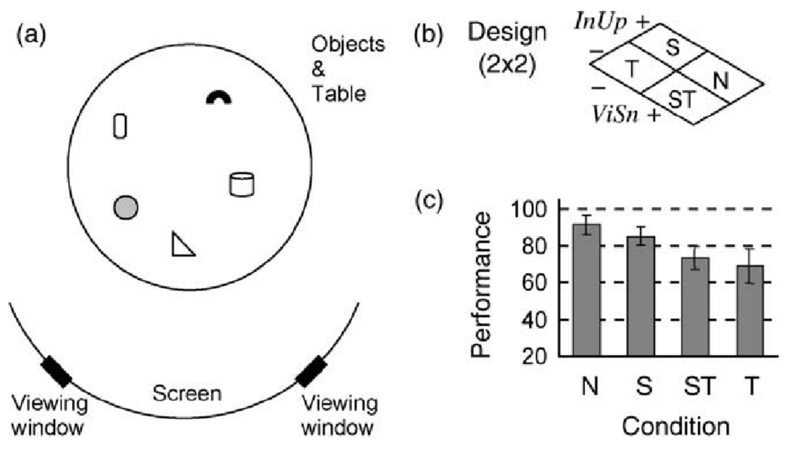
1) *Frames of reference for spatial representation*
2) *Place cells & boundary vector cells*
3) *Neural level model of Spatial Memory and Imagery*
4) *Grid cells and place cells*
5) *Grid cells as dynamic imagery, a general model for planning?*

A. *Hippocampus & striatum: Model-based versus model-free RL?*
B. *Dual representations theory, PTSD and intrusive imagery*

## Multiple parallel representations in spatial memory.

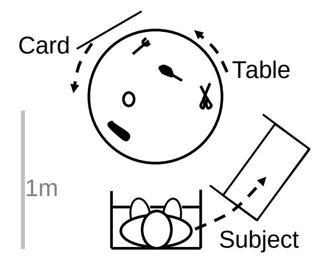
Effects of consistency with 'Visual Snapshots' & Internal 'Spatial Updating'

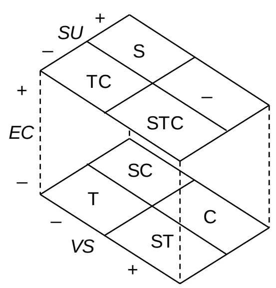
*Wang & Simons 1999*

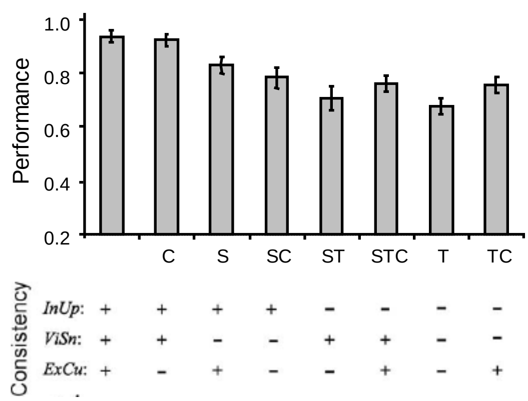



## Multiple parallel representations in spatial memory.

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



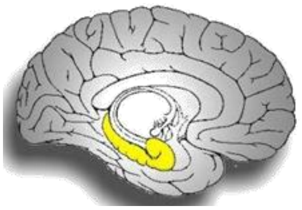

*Burgess, Spiers, Paleologou, 2004*

## *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) *Grid cells and place cells*
5) *Grid cells as dynamic imagery, a general model for planning?*

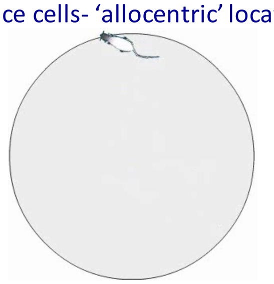
A. *Hippocampus & striatum: Model-based versus model-free RL?*
B. *Dual representations theory, PTSD and intrusive imagery*

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

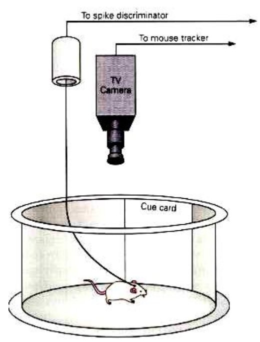


Spatial studies in rodents => likely neural representations.

Place cells- ‘allocentric’ location



*O’Keefe & Dostrovsky, 1971*

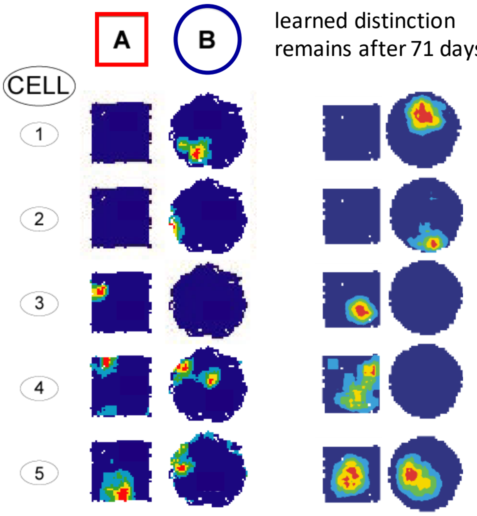



*Video by Julija Krupic*

## Place cells show long term memory and pattern completion

Place cell "remapping:" long-term memory for highy distinct environments.

learned distinction remains after 71 days..




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

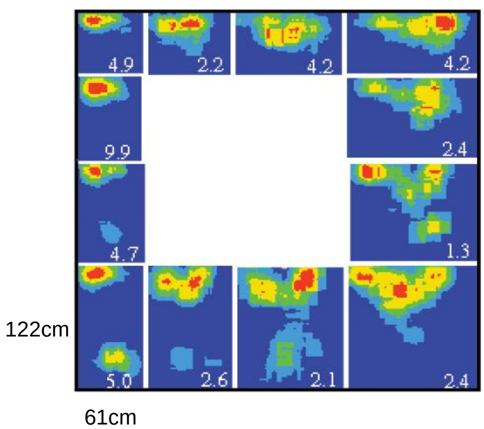
Place cell representation shows attractor dynamics

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

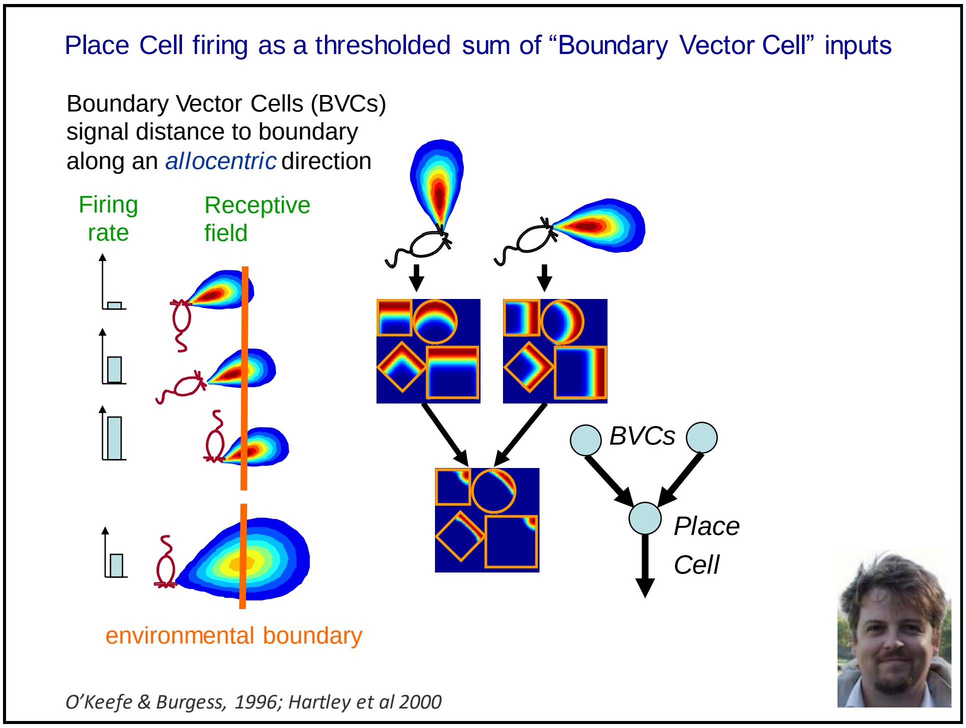
and 'pattern completion' depending on CA3 NMDA receptors

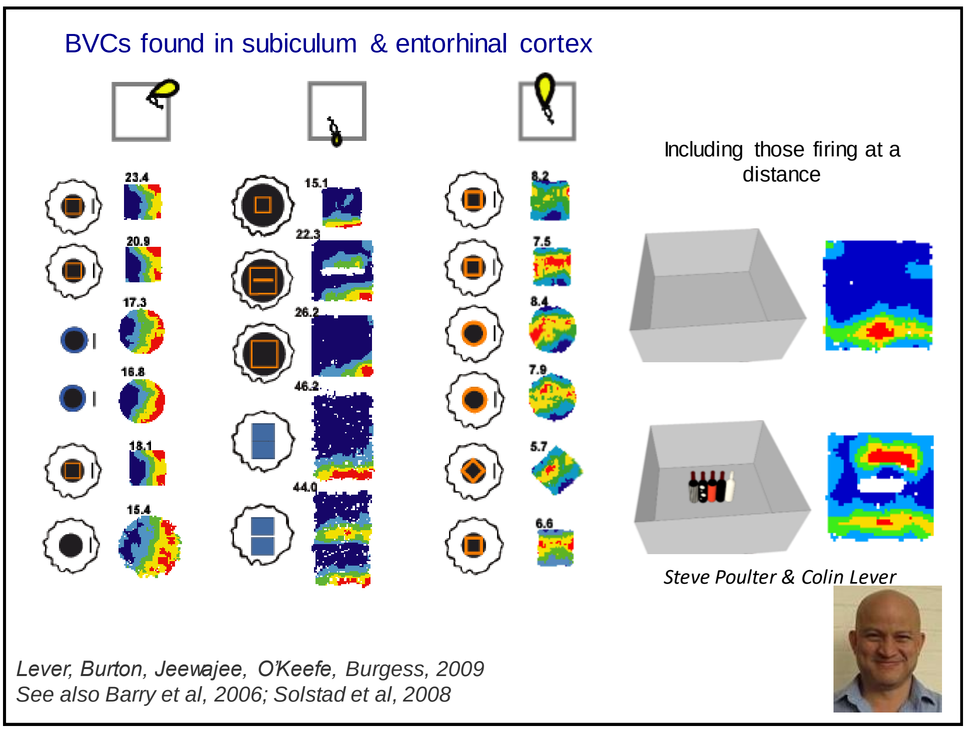
*Nakazawa et al., 2002*

## Environmental boundaries particularly influence place cell firing

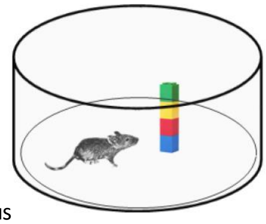



*O'Keefe & Burgess (1996)*

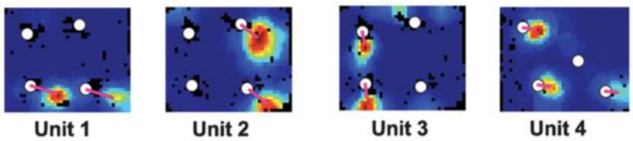

Place Cell firing as a thresholded sum of "Boundary Vector Cell" inputs
Boundary Vector Cells (BVCs)
signal distance to boundary
along an allocentric direction
Firing rate
Receptive field
environmental boundary
BVCs
Place Cell
O'Keefe & Burgess, 1996; Hartley et al 2000


BVCs found in subiculum & entorhinal cortex
23.4
20.9
17.3
16.8
18.1
15.4
15.1
22.3
26.2
46.2
44.0
8.2
7.5
8.4
7.9
5.7
6.6
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

## Object Vector Cells

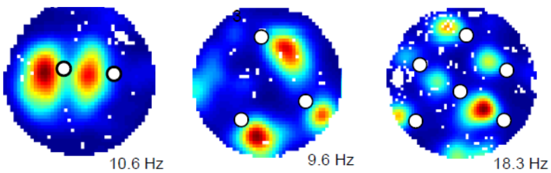


Recently found, in hippocampus




Desmukh & Knierim, 2013

and medial entorhinal cortex




Moser et al., BiorXiv 2018

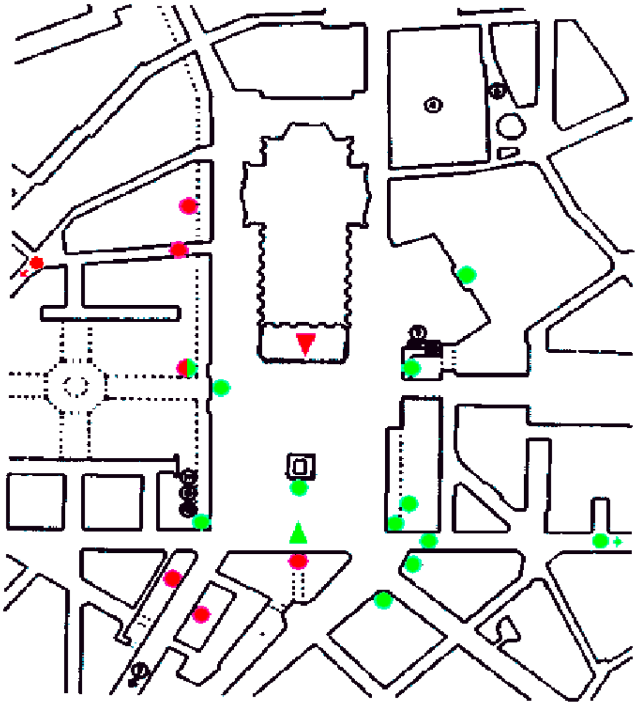
## *Abstract neural representations*

1) *Frames of reference for spatial representation*
2) *Place cells & boundary vector cells*
3) ***Neural level model of Spatial Memory and Imagery***
4) *Grid cells and place cells*
5) *Grid cells as dynamic imagery, a general model 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?



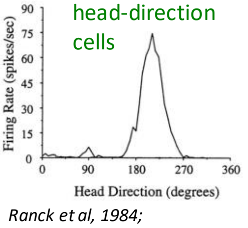
*Bisiach & Luzzatti(1978)*

## Several identified neural representations support spatial cognition

Hippocampal formation (allocentric)

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

head-direction cells

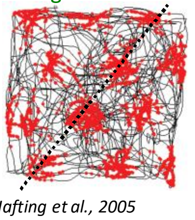



*Ranck et al, 1984; Taube et al, 1990*

place cells

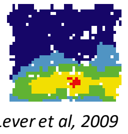
*O'Keefe & Dostrovsky, 1971*

grid cells



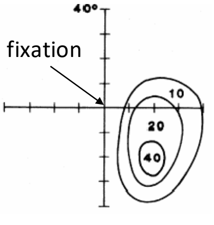
*Hafting et al., 2005*

boundary cells

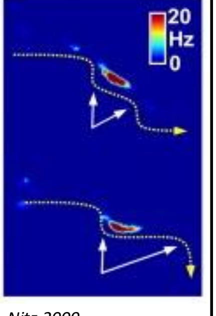


*Lever et al, 2009*
*Solstad et al, 2008*

retinal receptive fields




trajectory cells,

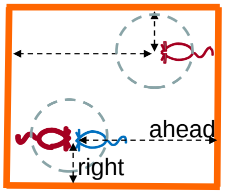



*Nitz 2009*

## Frames of reference for neural coding

*'egocentric'*

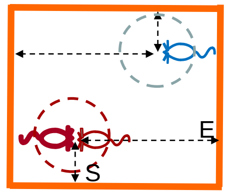
Body-centred location of objects




Perception
Action/Imagery

*'allocentric'*

World-centred location of agent



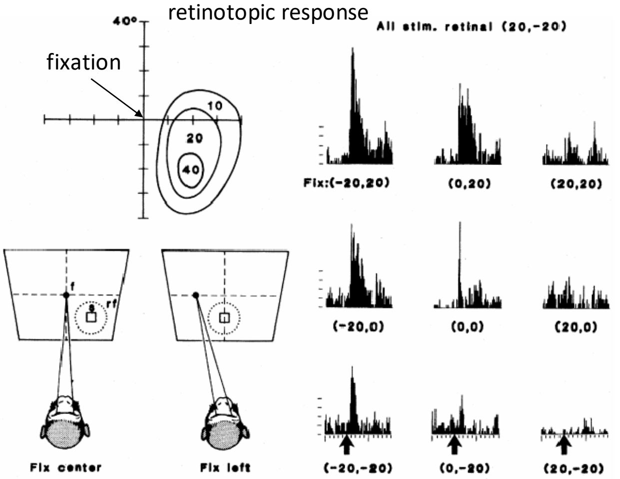

Place cells
Head-direction cells

*Burgess et al 2001*

## '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:



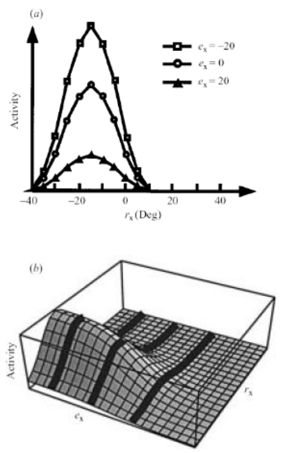

*Andersen et al 1985*

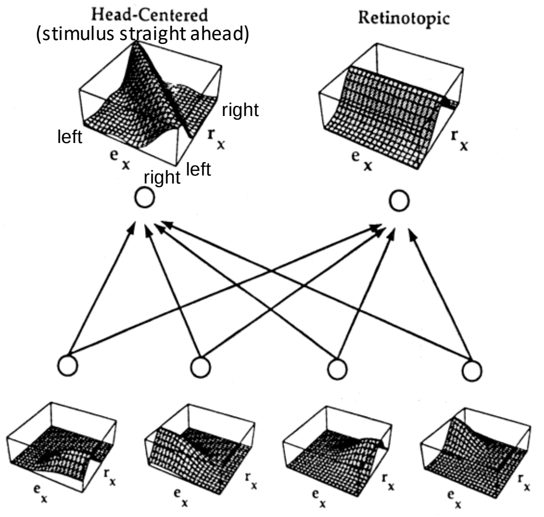
or by *direction of the head (Snyder et al 1998).*
Similar responses seen in parieto-occipital ctx (Galletti et al., 1995)

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

retinal position of stimulus = $r_x$

eye gaze angle = $e_x$

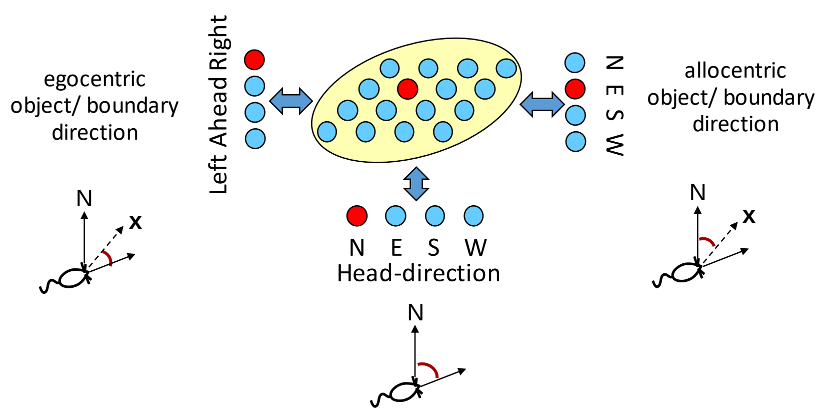



Pouget & Sejnowski, 1997

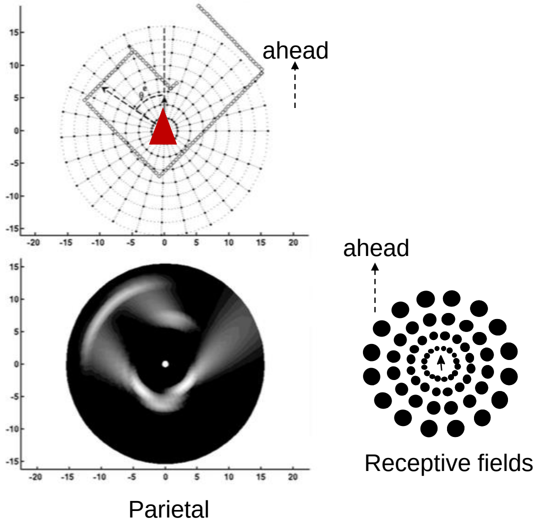
## Model of memory & imagery for scenes

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



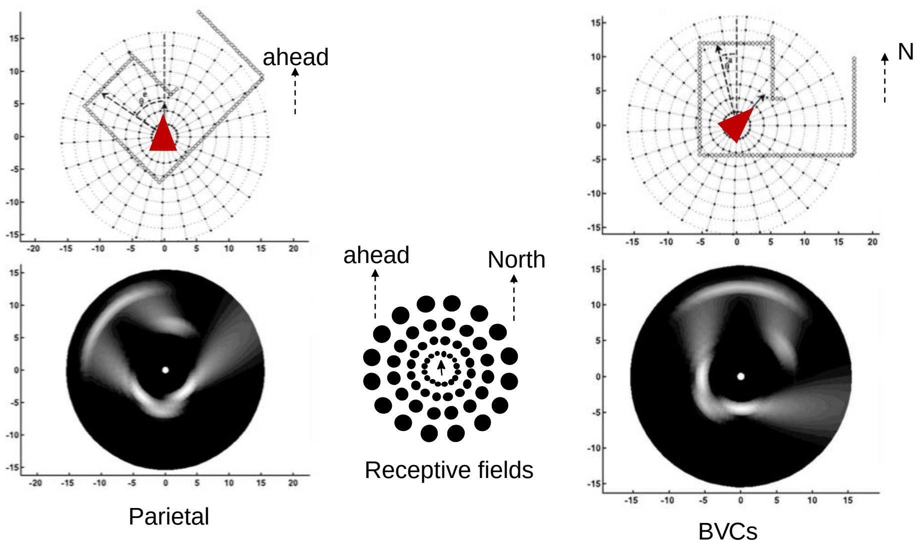

*Byrne, Becker, Burgess 2007; Burgess et al., 2001;*
*See Pouget & Sejnowski, 1997; Deneve et al., 2001.*

Scene representation by populations of egocentric or allocentric BVCs




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

Scene representation by populations of egocentric or allocentric BVCs



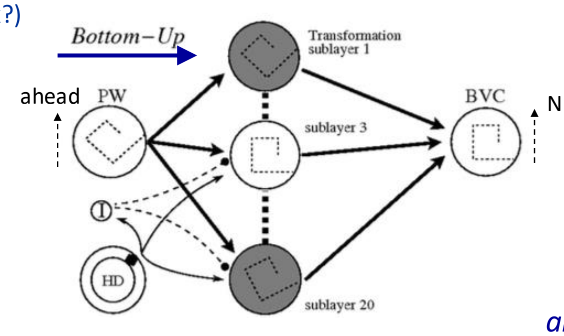

***egocentric representation (e.g. visual)***

***allocentric representation***

*Becker & Burgess 2001; Burgess et al., 2001; Byrne, Becker, Burgess 2007*

## Ego-allo scene translation
(retrospenial cortex?)

'gain field' representation of
scene elements x head direction




perception

LTM

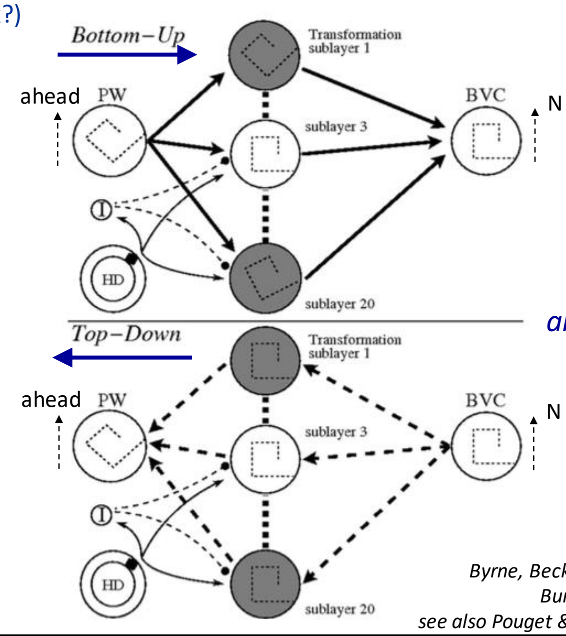
*egocentric*

*allocentric*

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

## Ego-allo scene translation
(retrospenial cortex?)

'gain field' representation of
scene elements x head direction




perception

LTM

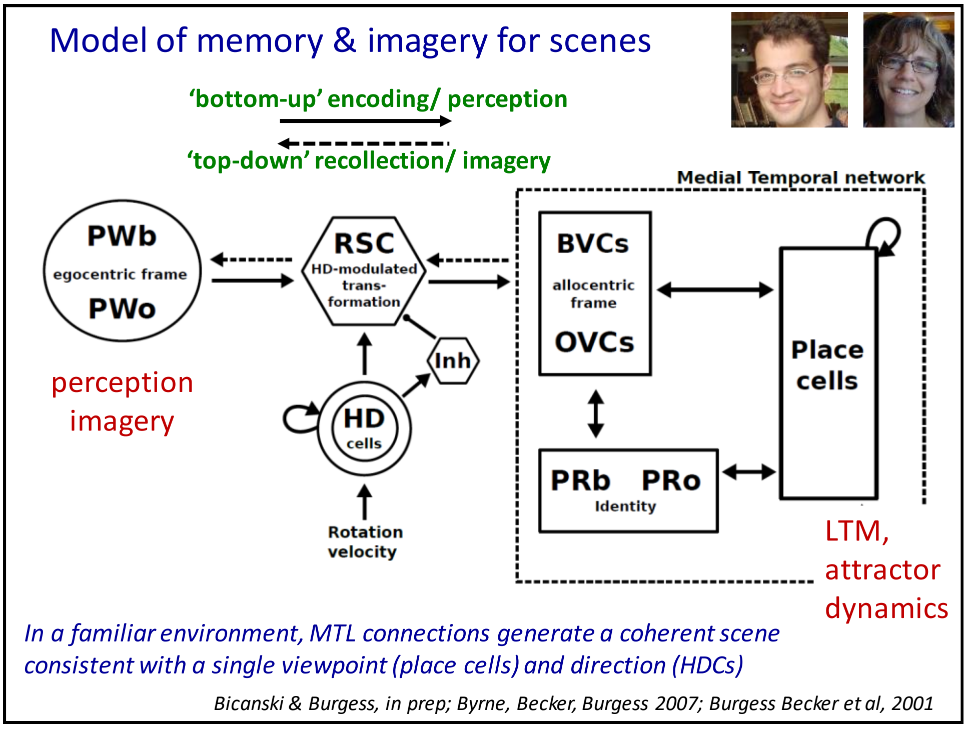
*egocentric*

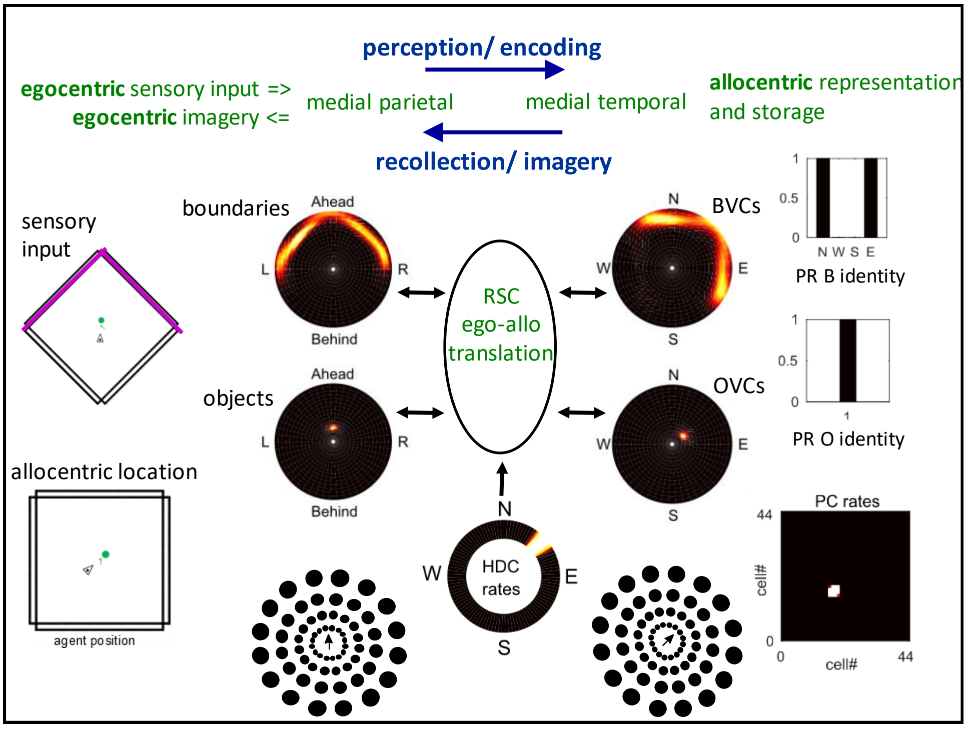
*allocentric*

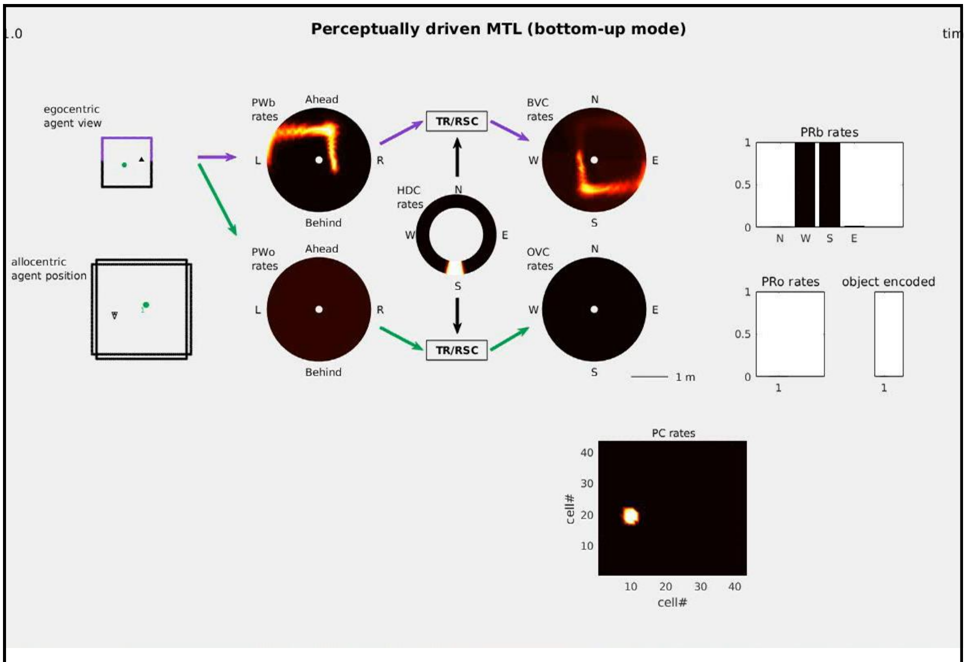
imagery
(& action)

LTM

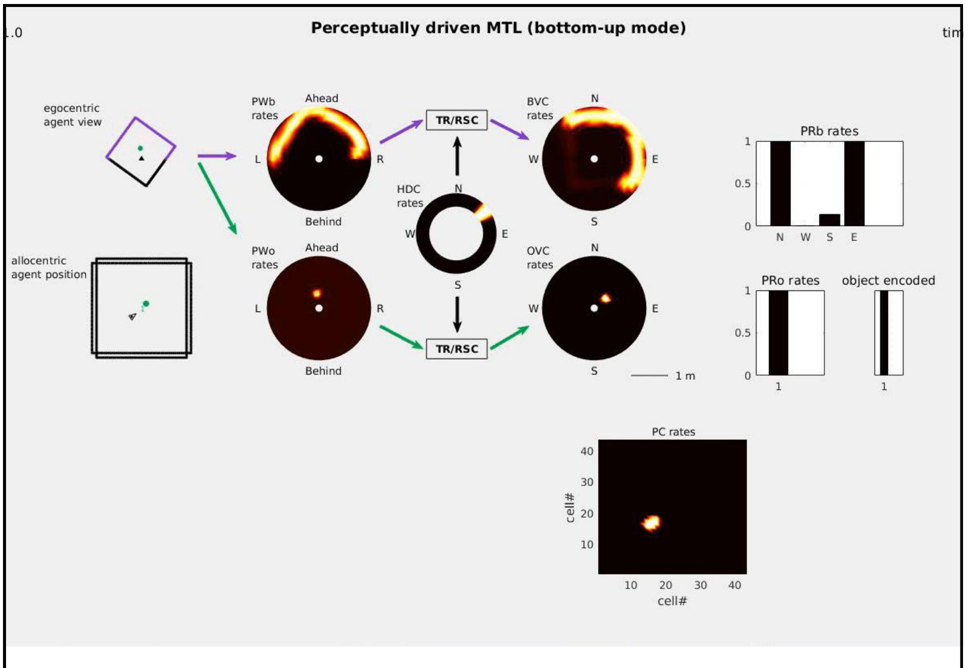
*Byrne, Becker, Burgess 2007*
*Burgess et al., 2001*
*see also Pouget & Sejnowski 1997*


Model of memory & imagery for scenes
'bottom-up' encoding/ perception
'top-down' recollection/ imagery
Medial Temporal network
PWb
egocentric frame
PWo
RSC
HD-modulated trans-formation
BVCs
allocentric frame
OVCs
Place cells
Inh
HD cells
Rotation velocity
PRb PRo
Identity
perception
imagery
LTM, attractor dynamics
In a familiar environment, MTL connections generate a coherent scene consistent with a single viewpoint (place cells) and direction (HDCs)
Bicanski & Burgess, in prep; Byrne, Becker, Burgess 2007; Burgess Becker et al, 2001

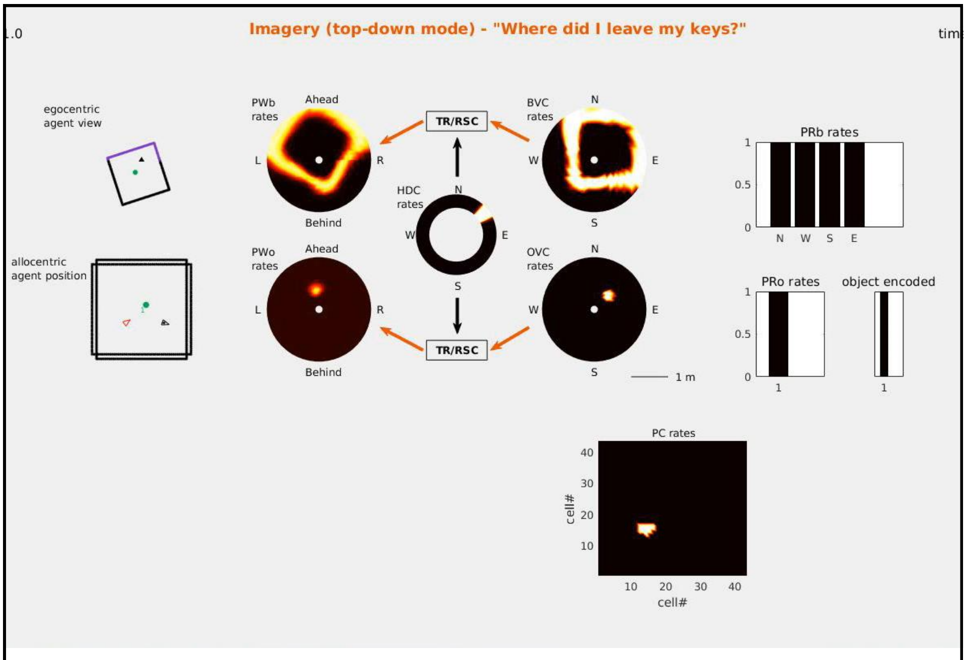

perception/ encoding
egocentric sensory input =>
egocentric imagery <=
medial parietal
medial temporal
allocentric representation and storage
recollection/ imagery
sensory input
boundaries
Ahead
L
R
Behind
RSC ego-allo translation
N
W
E
S
BVCs
PR B identity
objects
OVCs
PR O identity
allocentric location
agent position
HDC rates
PC rates
cell#




Encountering an object in a familiar environment

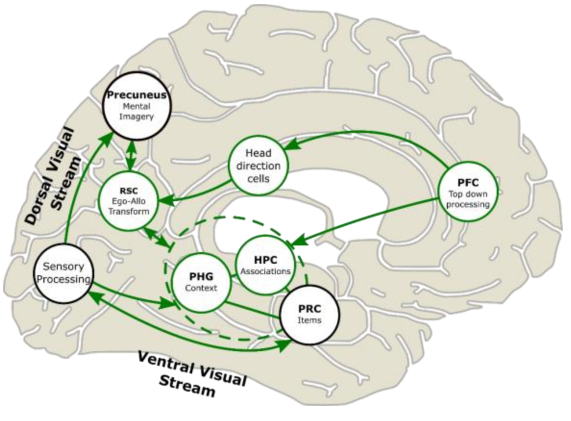



Recollection of encountering the object




Memory enhanced 'perception' of a familiar environment

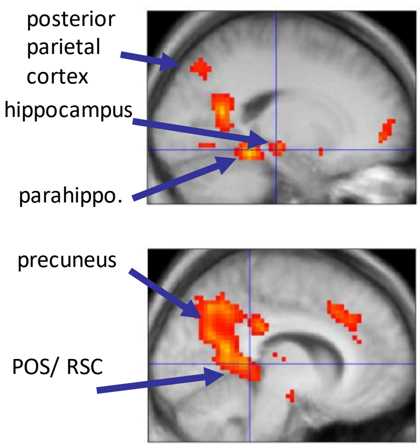
Model allows interpretation of fMRI patterns during recollection/ imagery




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

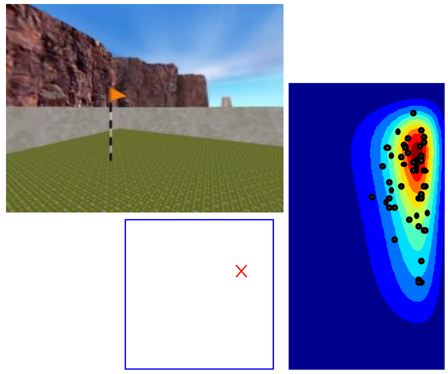
*RSC /POS supports egocentric-allocentric translation, required to associate (allocentric) internal representations with (egocentric) sensory representations*
*- e.g. stronger associations will form to stable sensory features,* see Auger et al., *2012*

## Model allows interpretation of fMRI patterns during recollection/ imagery




*Burgess et al, 2001*

& prediction of human search patterns

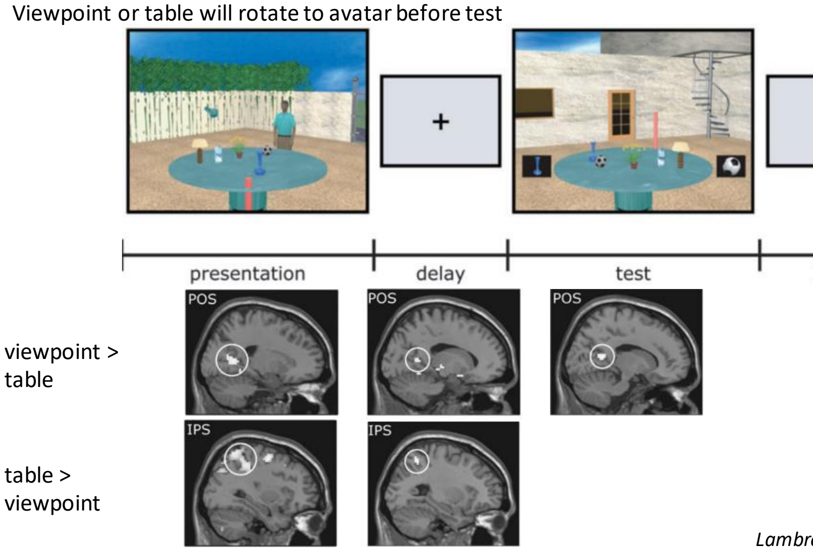


*Hartley et al, 2004*

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

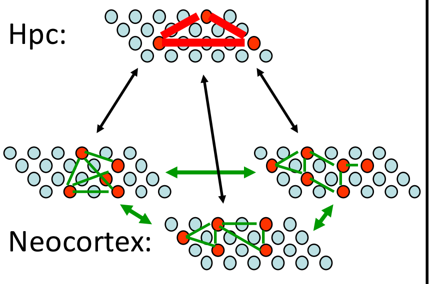



*Lambrey et al 2013*

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 Bardur Joensen 10:00 Friday; Horner & Burgess (2013, 2014) Horner et al (2015).*



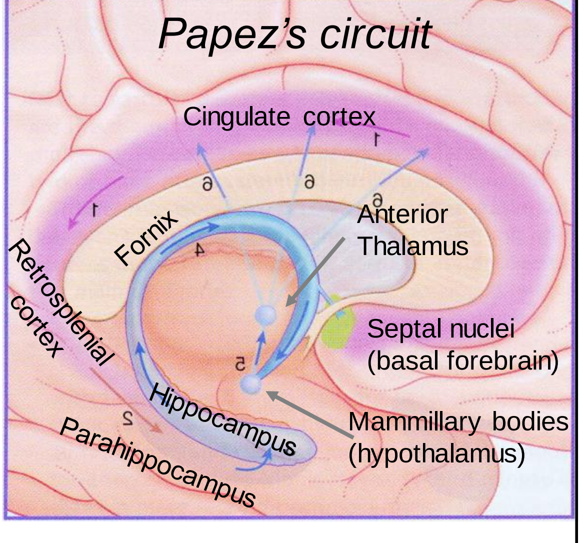

***Marr, 1971****; Gardner-Medwin, McNaughton, Alvarez, Squire, McClelland, O'Reilly, Treves, Rolls, Teyler & DiScenna; Damasio;*

## 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

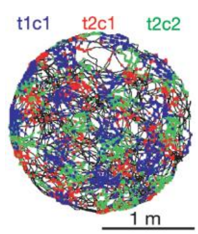
1) *Frames of reference for spatial representation*
2) *Place cells & boundary vector cells*
3) *Neural level model of Spatial Memory and Imagery*
4) ***Grid cells and place cells***
5) *Grid cells as dynamic imagery, a general model for planning?*

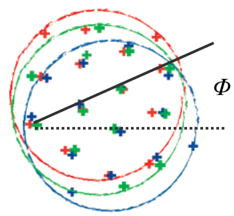
A. *Hippocampus & striatum: Model-based versus model-free RL?*

B. *Dual representations theory, PTSD and intrusive imagery*

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

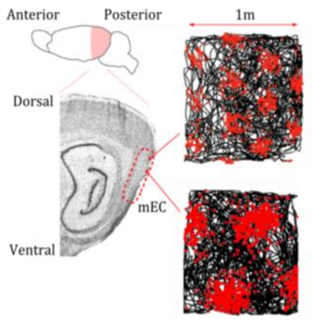
The grids of nearby cells share orientation & scale

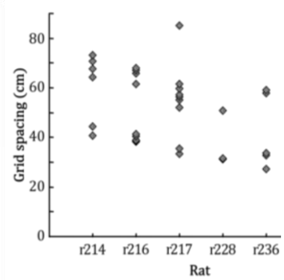


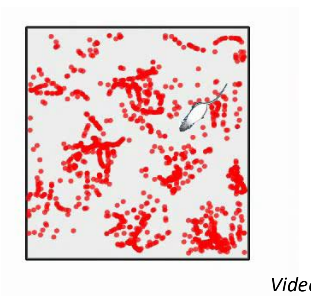

*Hafting et al., 2005*

Grid cells occur in modules with discrete scales




*Barry et al, 2007;*
*see also Stensola et al., 2012*



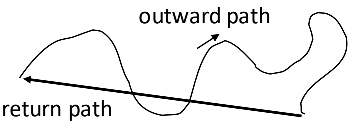
*Video by Julija Krupic*

## Two ways to know where you are:

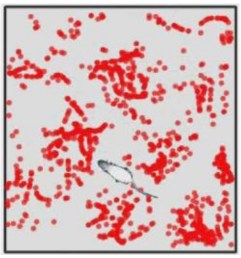
1. Environmental information

(Environmental boundaries particularly influence place cells)

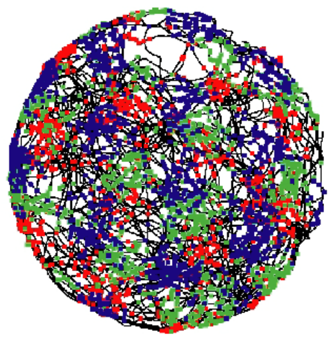
2. Path integration




Grid cells



*Hafting et al., 2005*



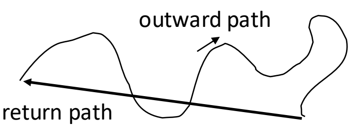
*Video by Julija Krupic*

## Two ways to know where you are:

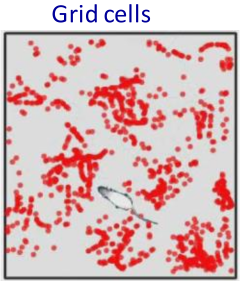
1. Environmental information

(Environmental boundaries particularly influence place cells)

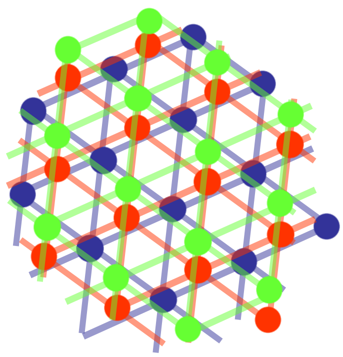
2. Path integration




Grid cells



*Hafting et al., 2005*



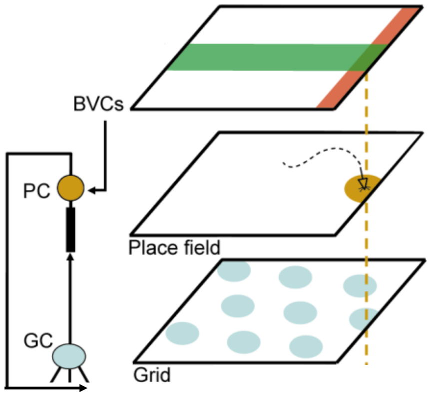
*Video by Julija Krupic*

## Interactions between place cells and grid cells

Estimating self-location combines environmental & self-motion information

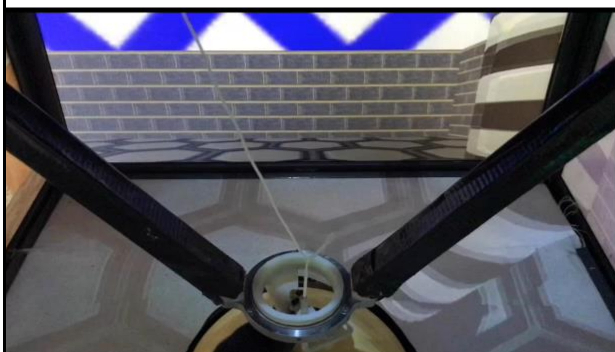
Environmental information
(⇔ Boundary Vector Cells)

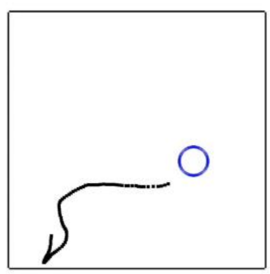
Self- motion

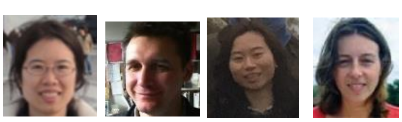



*Burgess et al, 2007*

## 2D VR for mice (invisible reward task)

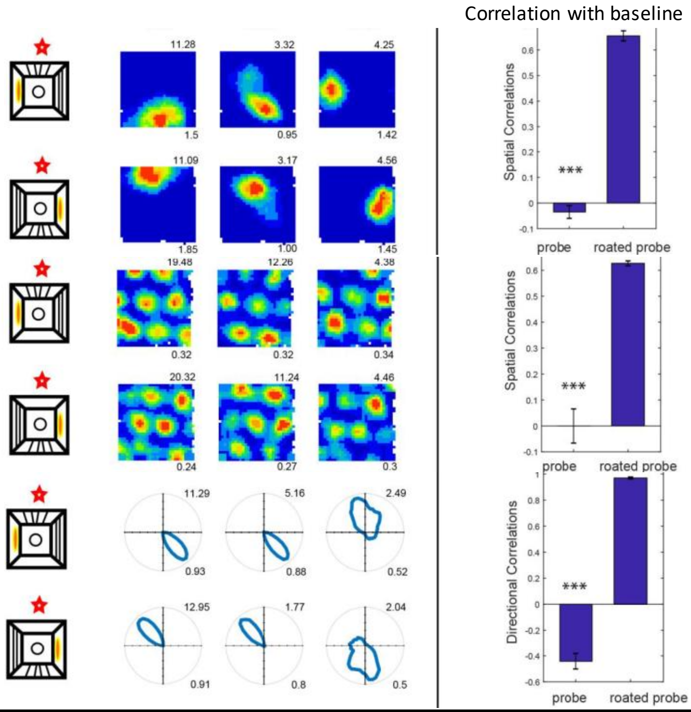






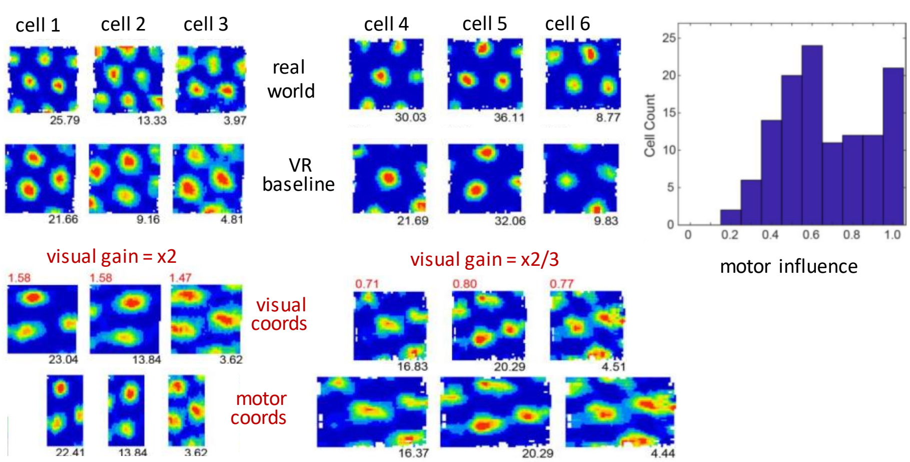
*Guifen Chen, John King, Yi Lu, Francesca Cacucci, Neil Burgess, bioRxiv 2018*

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)



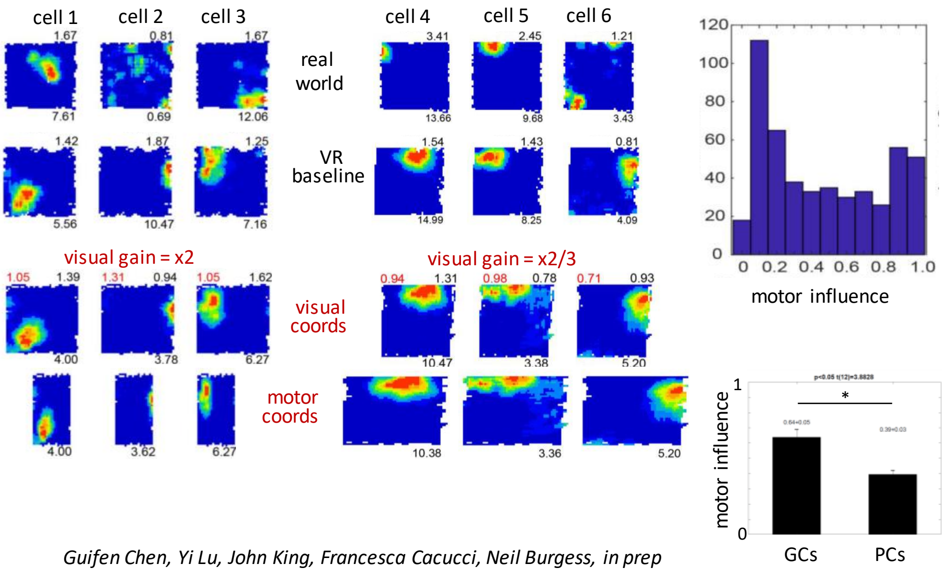

*Chen et al, bioRxiv 2018*

## Grid cell firing patterns reflect self-motion more than vision




*Guifen Chen, Yi Lu, John King, Francesca Cacucci, Neil Burgess, in prep*

## Place cell firing patterns reflect vision more than self-motion



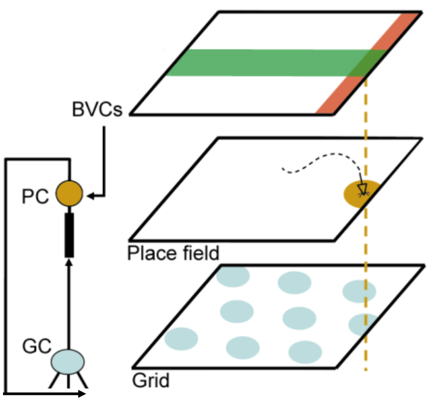

*Guifen Chen, Yi Lu, John King, Francesca Cacucci, Neil Burgess, in prep*

## Interactions between place cells and grid cells

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

Environmental information
(⇔ Boundary Vector Cells)

Self- motion



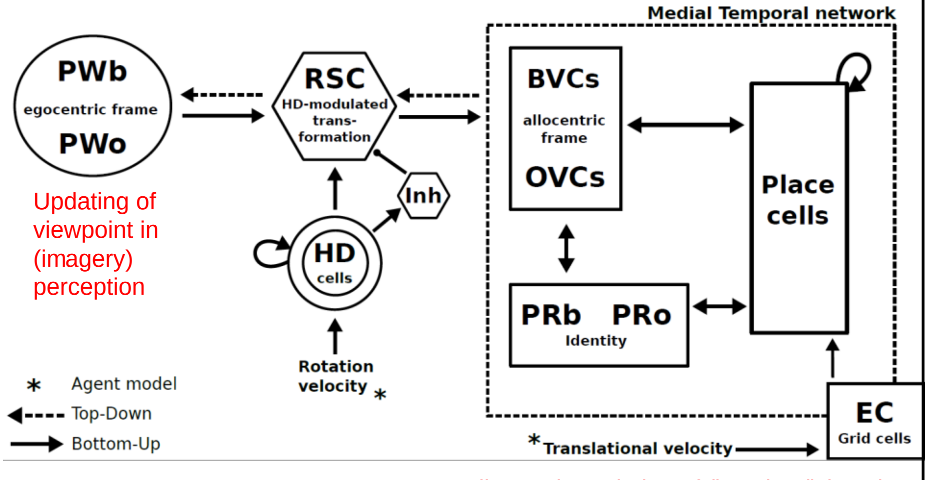

*Burgess et al, 2007*

## *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) *Grid cells and place cells*
5) ***Grid cells as dynamic imagery, a general model for planning?***

A. *Hippocampus & striatum: Model-based versus model-free RL?*
B. *Dual representations theory, PTSD and intrusive imagery*

## Grid cells and memory/imagery




Updating of viewpoint in (imagery) perception

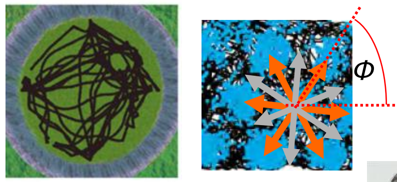
Allocentric updating of (imagined) location

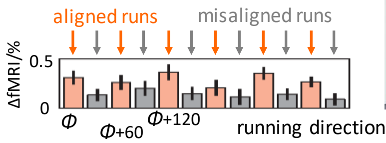
*Bicanski & Burgess, in prep*

## Grid cells in the human autobiographical memory system?

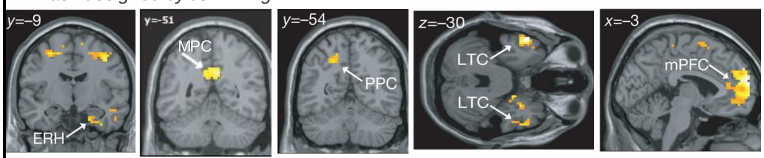
*Doeller, Barry, Burgess, 2010*

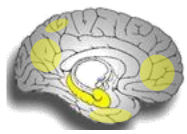
populations of *aligned* grids (modules) => changes in fMRI signal with virtual running direction


*Task designed by John King*

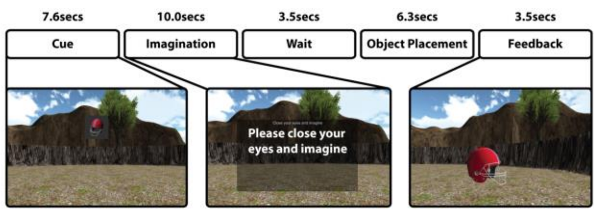


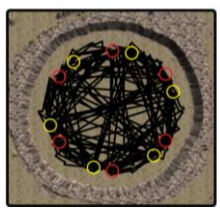



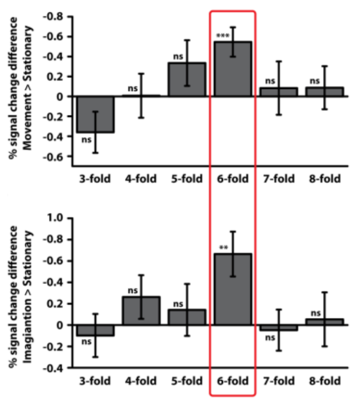
Autobiographical memory system

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

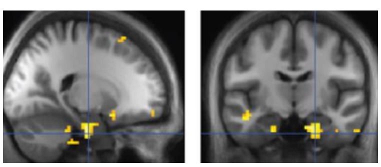
## Grid-like processing of movement of viewpoint in imagery



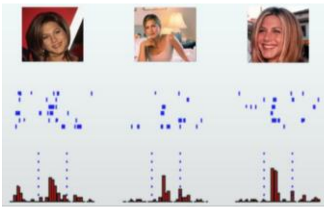



60° symmetry in fMRI signal with *imagined* running direction in Entorhinal cortex (aligned with that in virtual movement)

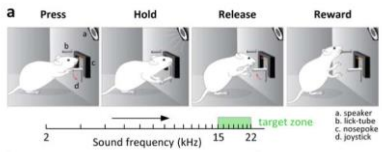


*Horner et al., 2016*

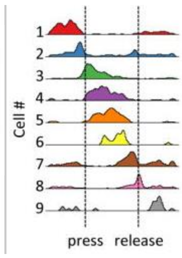
## Hippocampal cells can represent abstract concepts, such as 'place' but also, e.g., personal identity or sound frequency?



*Quiroga et al., (2005)*



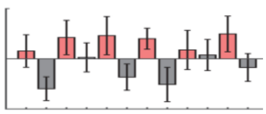

*Aronov, Nevers, Tank (2017)*



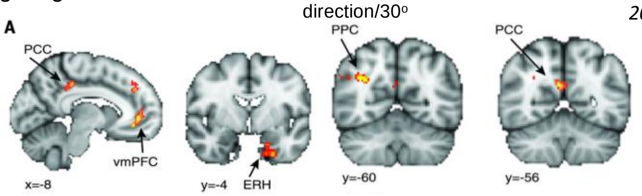

## Grid cell firing patterns reflect the transition structure of learned conceptual spaces?

Navigation in space of bird neck & leg length

fMRI:




*Constantinescu, O'Reilly, Behrens 2016*



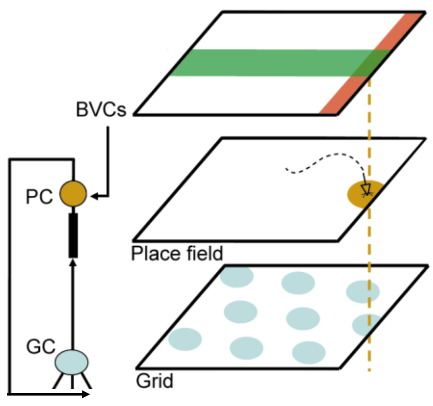

## Interactions between place cells and grid cells

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

State information (place)

(⇔ Feature Vector Cells?)

Transition structure (self- motion)



*Abstract neural representations*

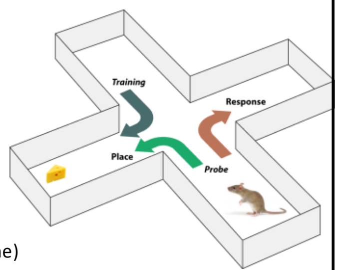
1) *Frames of reference for spatial representation*
2) *Place cells & boundary vector cells*
3) *Neural level model of Spatial Memory and Imagery*
4) *Grid cells and place cells*
5) *Grid cells as dynamic imagery, a general model for planning?*

A. ***Hippocampus & striatum: Model-based versus model-free RL?***

B. *Dual representations theory, PTSD and intrusive imagery*

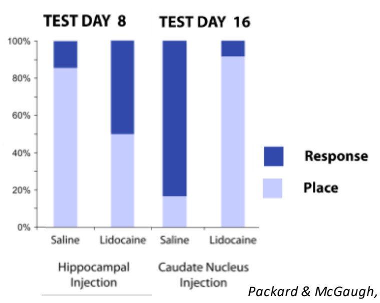
## Hippocampo-striatal model of navigation

Packard & McGaugh Task



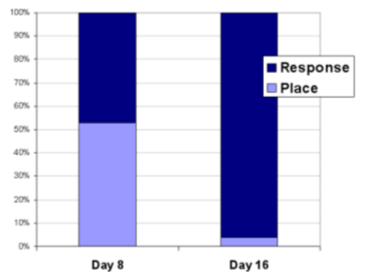

Switch from hippocampal 'place' navigation
to striatal 'response' navigation
during T-maze learning

Experimental data




*Packard & McGaugh, 1996*

Simulation results (saline)

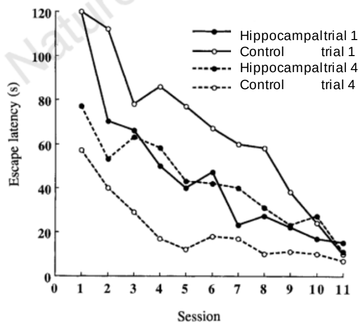


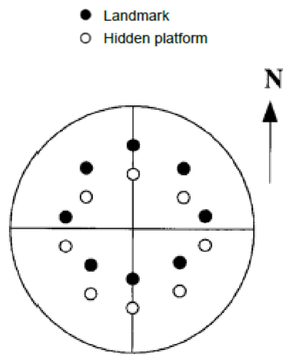
## Hippocampo-striatal model of navigation

Pearce et al Task

Learning water maze with local landmark, including effects of hippocampal lesions

Experimental data *Pearce et al., 1998*

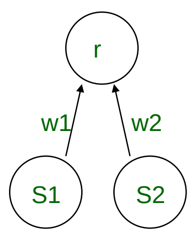



## Rescorla-Wagner rule (reward prediction error) and multiple stimuli

What about when multiple stimuli are present? e.g. S1, S2→r

How would animals respond to S1 or S2?
How should the model be modified?




$w_i \rightarrow w_i + \varepsilon S_i \delta_i$

(a) $\delta_i = r - w_i S_i$

i.e. separate error terms for each $S_i$

(b) $\delta_i = \delta = r - V$; $V = \Sigma_i w_i S_i$

V is expected reinforcement r given all stimuli

i.e. single error term for all stimuli $S_i$: $\delta$

the difference between actual r and V (expected r)

## Experiments with multiple stimuli

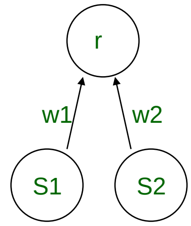
| Experimental terms | Phase 1: | Phase 2: | Test: |
|---|---|---|---|
| **Overshadowing:** | S1, S2 → r | | S1? weak resp |
| **Blocking:** | S1 → r | S1, S2 → r | S2? – |

Which model is favoured?

$w_i \rightarrow w_i + \varepsilon S_i \delta_i$

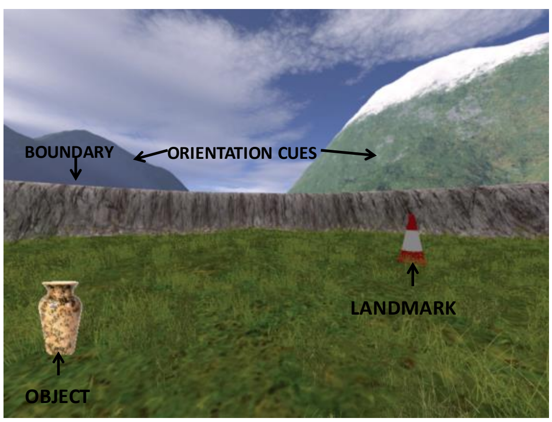
*Blocking* (Kamin, 1969) and *overshadowing* (Kamin, 1969; Pavlov, 1927) imply:

(b) $\delta_i = \delta = r - V;\ V = \Sigma_j w_j S_j$

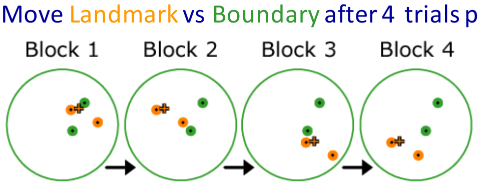



i.e. single error term for all stimuli (the Rescorla-Wagner rule)
= difference between reinforcement and expected reinforcement given all stimuli

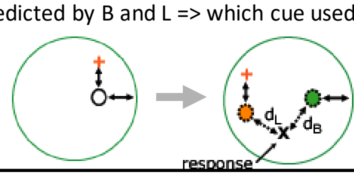
## Boundaries versus landmarks in human spatial memory




Move Landmark vs Boundary after 4 trials per object

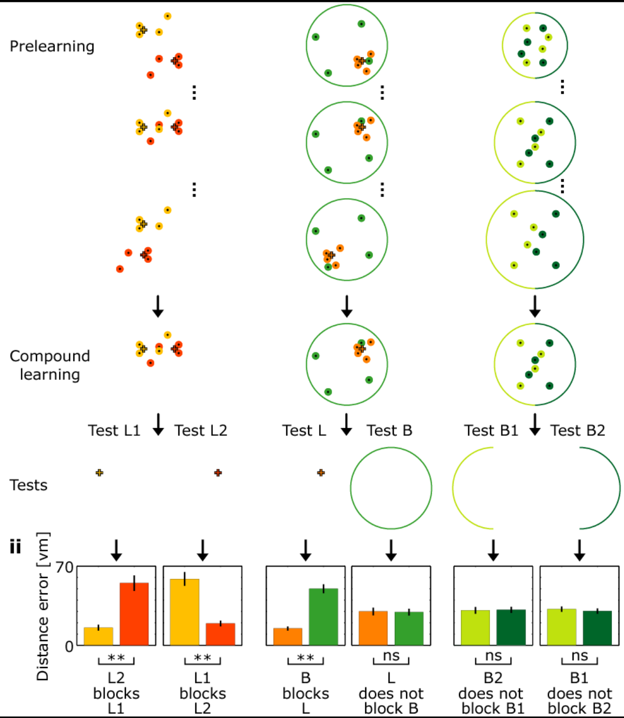



Proximity of response to the locations predicted by B and L => which cue used

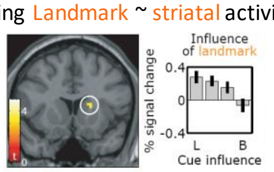


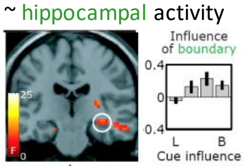
Learning to landmarks obeys *associative reinforcement* ("blocking"); learning to boundaries is *incidental* (no "blocking").

*Striatum* = reinf learning
*Hippocampus* = incidental Hebbian association.

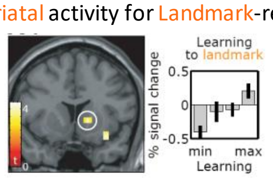


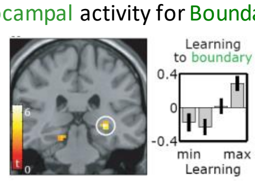

Replacing objects:
using Landmark ~ striatal activity;   using Boundary ~ hippocampal activity




Learning from feedback (improvement on next trial with same object) ~
striatal activity for Landmark-related objects; hippocampal activity for Boundary-related objects



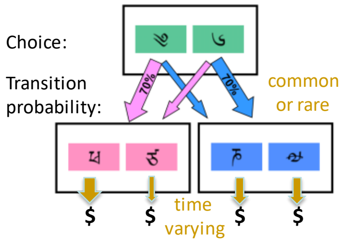




Learning locations relative to landmarks obeys *associative reinforcement* (shows blocking and overshadowing).

Learning locations relative to boundaries is *incidental* (no blocking or overshadowing)

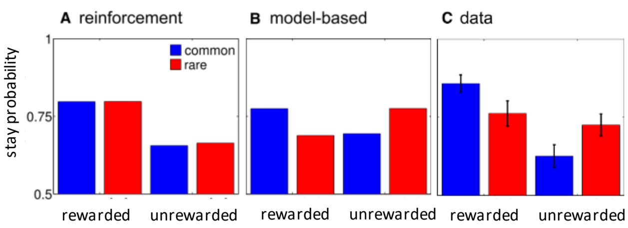
*Doeller, King, Burgess (2008); Doeller, Burgess (2008)*

## Probabilistic choice task separates model-based vs simple reinforcement learning




*Daw et al., 2011*

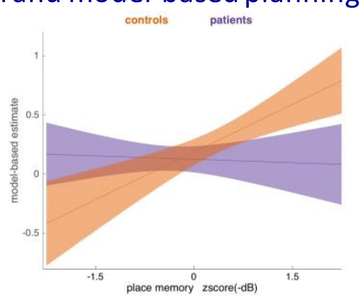
Choice on next trial as function of reward and transition on previuos trial



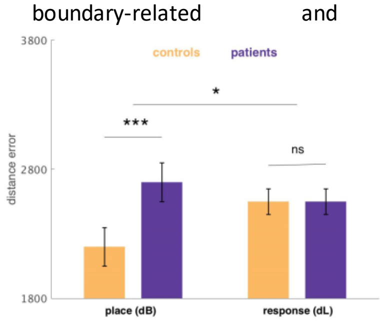

## Hippocampus supports spatial navigation and model-based planning?

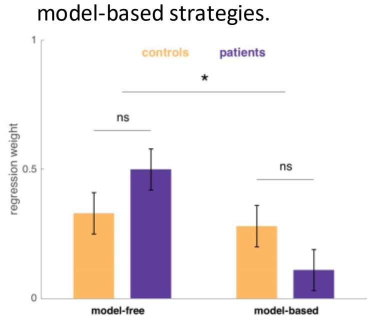
*Vikbladh, Meager, King, Blackmon, Devinsky, Shohamy, Burgess, Daw, biorXiv 2018*

Boundary-related (place) strategy ~ model-based strategy in healthy controls



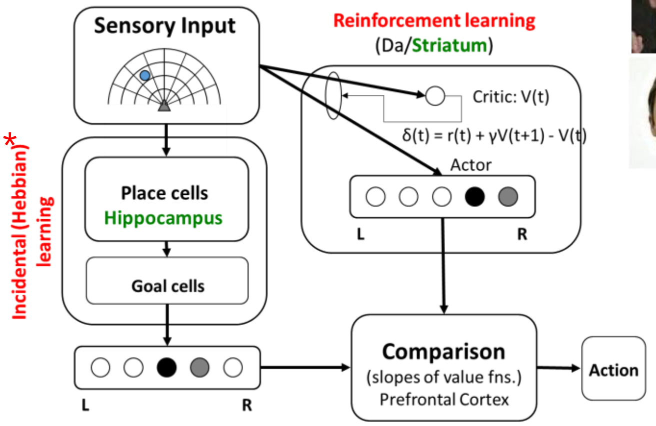

Anterior temporal lobectomy biases away from both boundary-related and model-based strategies.



## Hippocampo-striatal model of navigation

Architecture:



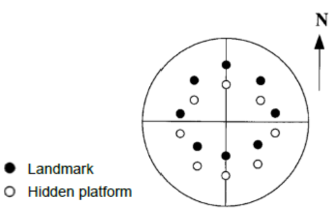

* pattern completion among place cells + delta rule tracks presence of goal (could be reward, or any other object) without cue competition

Chersi F, Burgess N 2015 Cognitive architecture of spatial navigation: Hippocampal & Striatal contributions. *Neuron* 88:64-77. Geerts et al in prep. *See also Dollé et al 2010; Sheynikhovich et al 2009.*

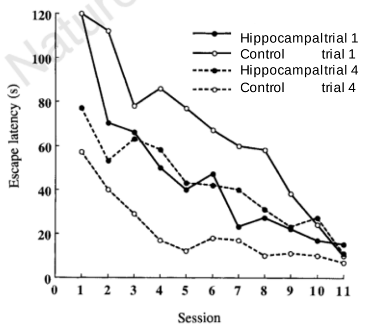
## Hippocampo-striatal model of navigation

Pearce et al Task

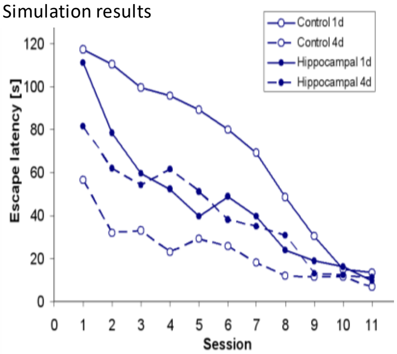
Learning water maze with local landmark, including effects of hippocampal lesions




Experimental data *Pearce et al., 1998*




Simulation results



## *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) *Grid cells and place cells*
5) *Grid cells as dynamic imagery, a general model for planning?*

A. *Hippocampus & striatum: Model-based versus model-free RL?*

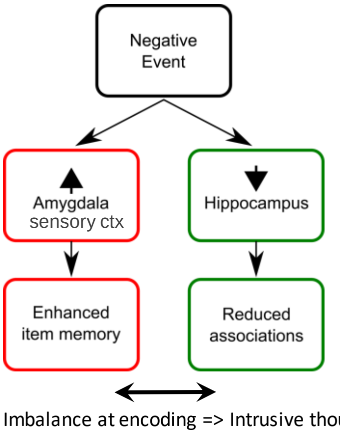
B. ***Dual representations theory, PTSD and intrusive imagery***

## A dual representation account of intrusive memories

Negative experiences affect distinct representations in different ways:

↑ Strengthens sensory/affective representations through amygdala up-regulation

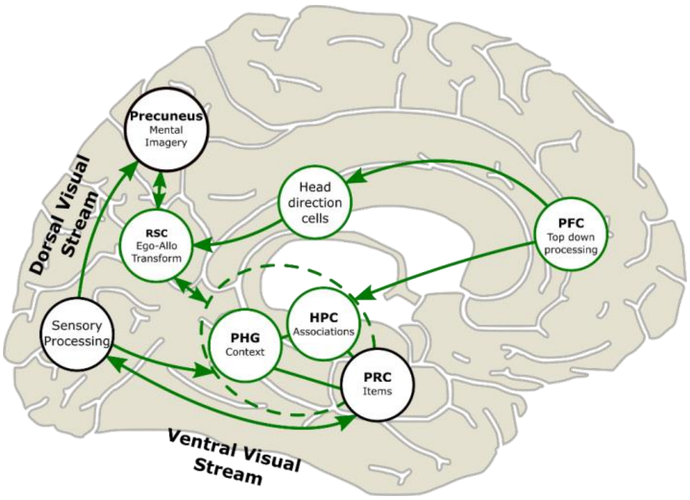
↓ Weakens associative/contextual representations through down-regulation of the hippocampus

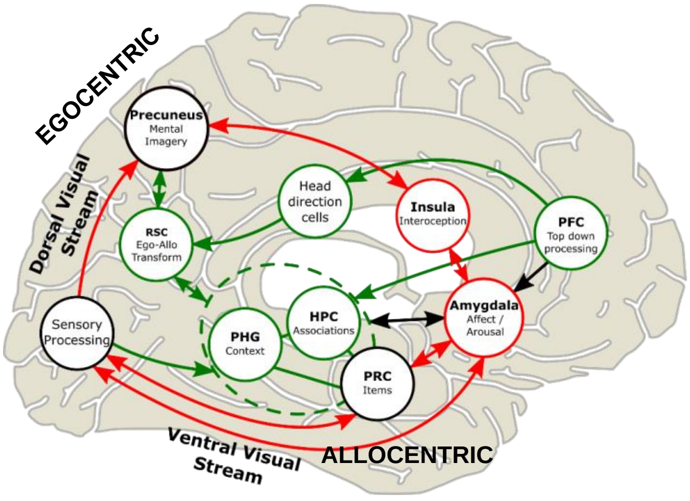



Imbalance at encoding => Intrusive thoughts
Therapy strengthens association between negative content and appropriate context

cf Unitary model: intrusive traumatic memories are just very strong autobiographical memories (e.g. Rubin)

*Brewin, Gregory, Lipton, Burgess 2010*
*following Jacobs & Nadel (1998)*



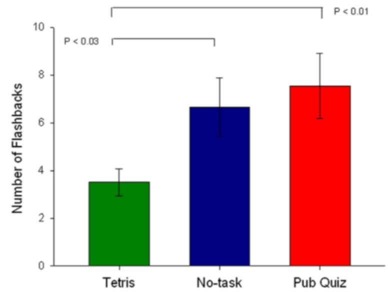




*Bisby & Burgess (2017)*

## Two options for reducing intrusive thoughts following a traumatic event?

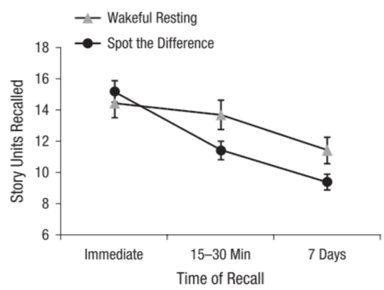
*Disrupting consolidation of sensory representations* can reduce intrusive thoughts after watching a traumatic video (e.g. by playing Tetris)

*Holmes et al., 2004; Holmes et al., 2010*




Brief wakeful rest can facilitate consolidation of neutral episodic memories *e.g. Dewar et al., 2012*

**Will wakeful rest *enhance* consolidation of *hippocampal* representations and so *reduce* intrusive thoughts?**



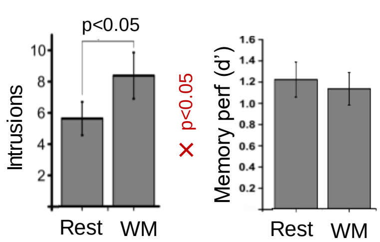

## Brief wakeful rest and memory intrusions

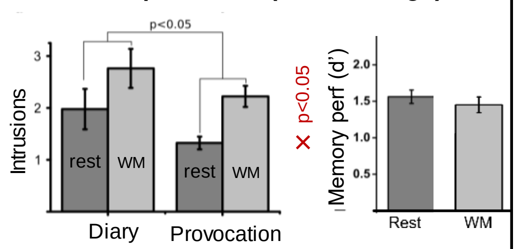
Day 1 | Day 1-7 | Day 8

Encoding | Intervention | Intrusion Diary | Memory Test | Provocation Task


20 clips | 10mins

**Experiment 2 (within-design)**




Brief wakeful rest reduced *intrusions* but not *deliberate memory*, supporting a *dual representation* account of intrusive thoughts. *Horlyck, Bisby, Burgess, in prep*

## *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:*

*Graham Hitch*
*Andrej Bičanski*
*John King*
*Guifen Chen*
*Yi Lu*
*John O'Keefe*
*Francesca Cacucci*
*Lone Hørlyck*
*James Bisby*
*Tom Wills*

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