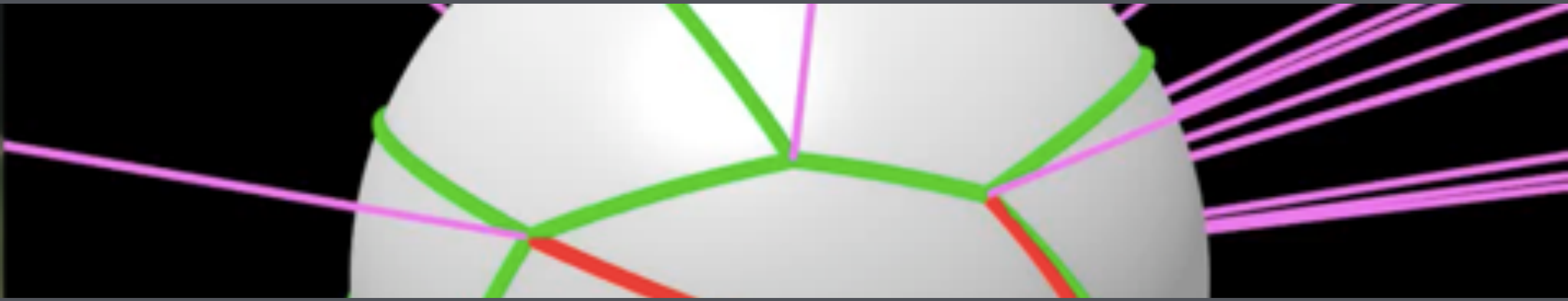
School of Psychology and
Clinical Language Sciences

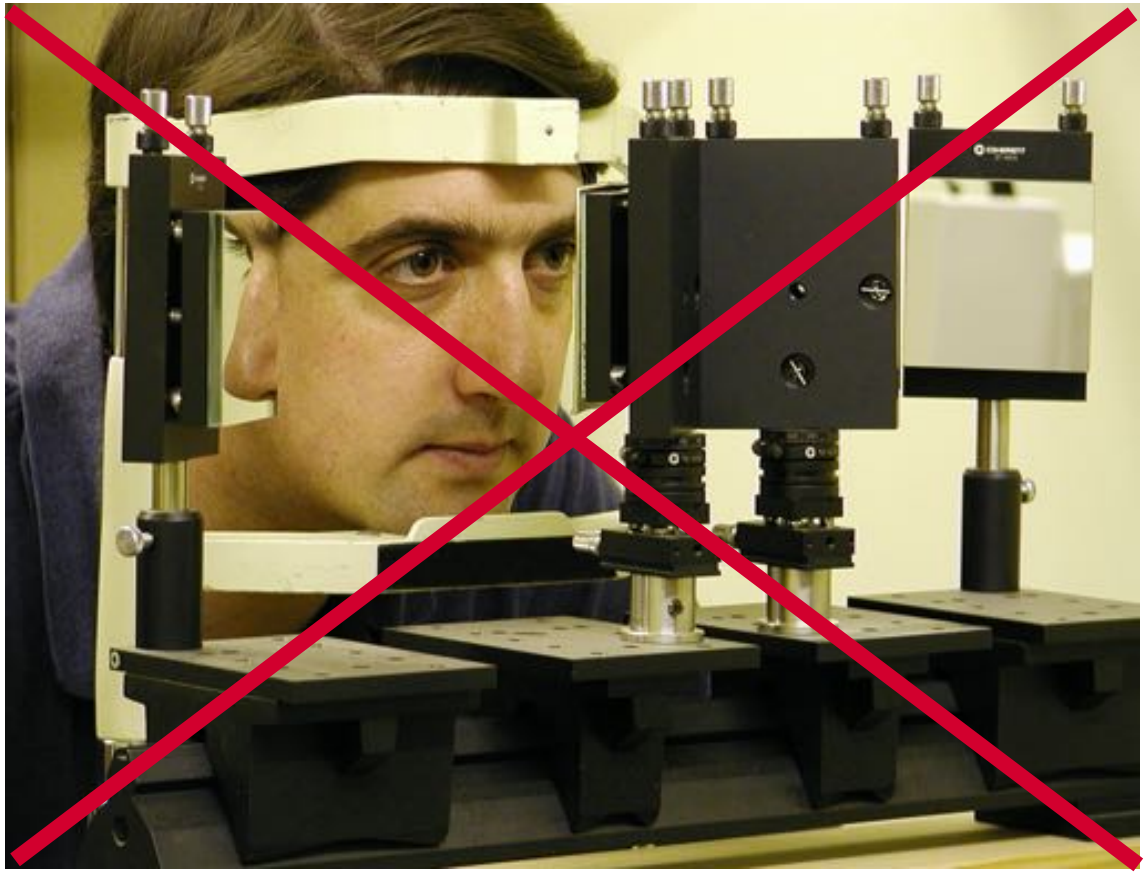
University of Reading

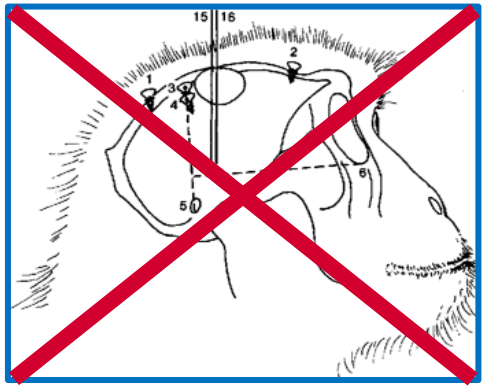
USING VR TO STUDY NAVIGATION AND 3D VISUAL PERCEPTION IN FREELY MOVING OBSERVERS

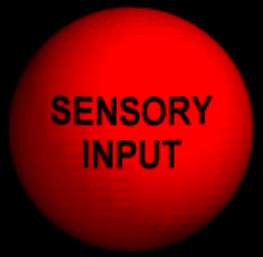


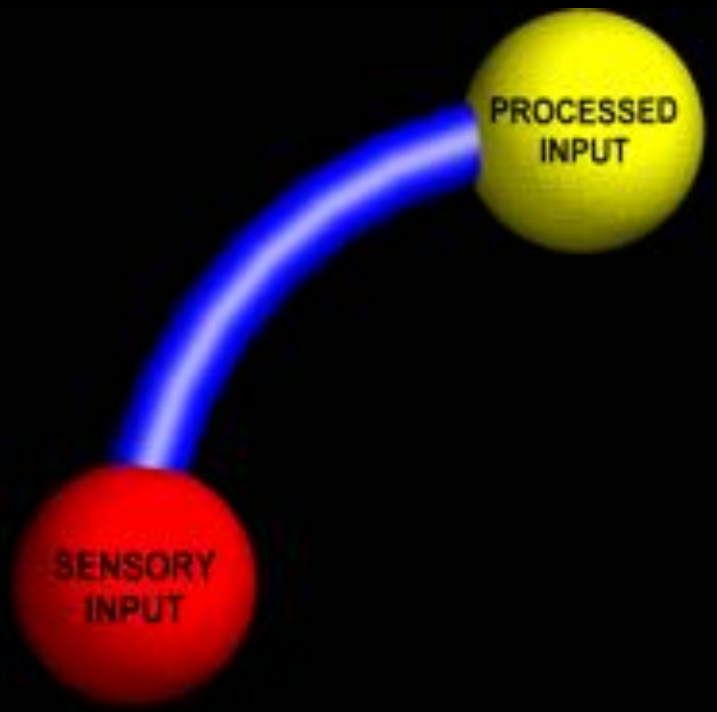
Andrew Glennerster

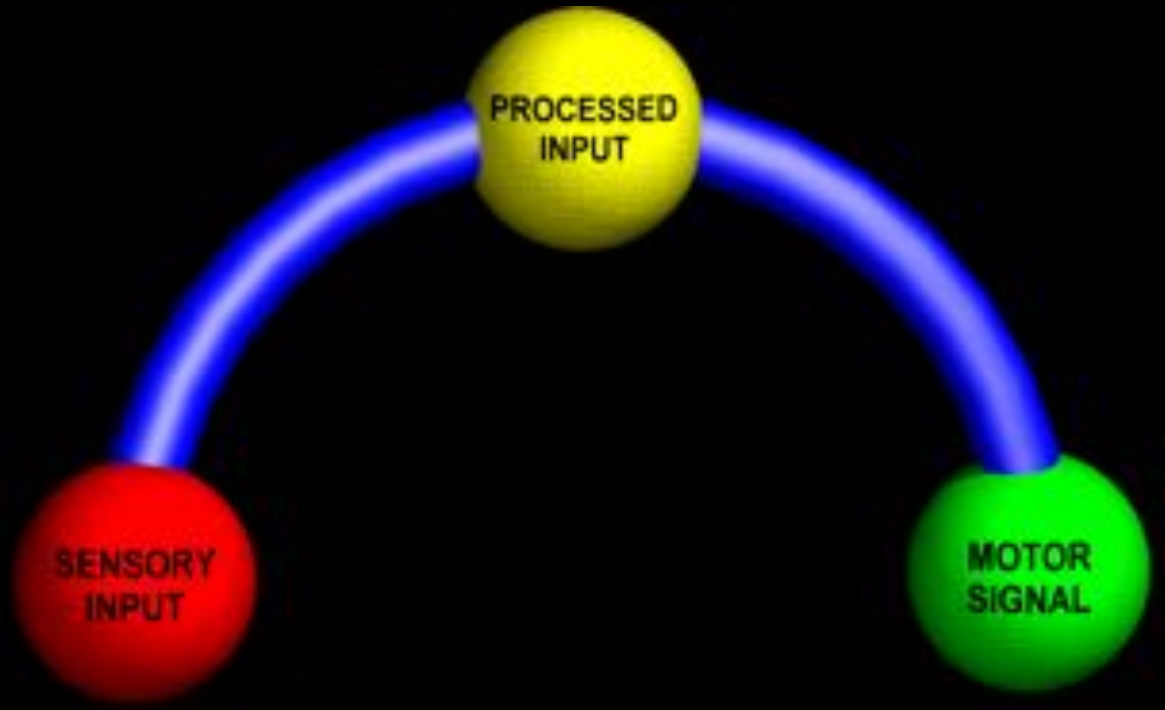
How NOT to do experiments

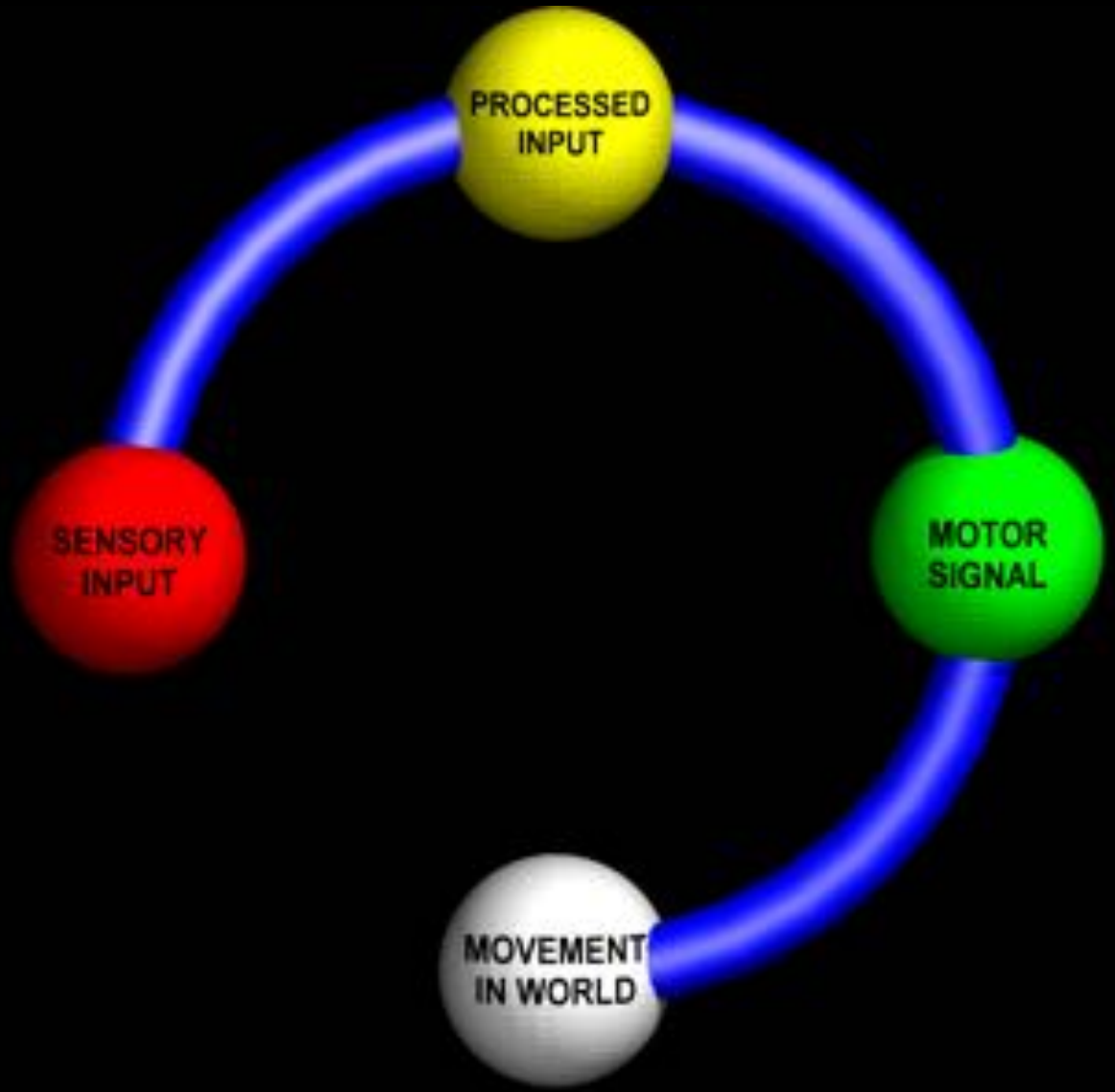


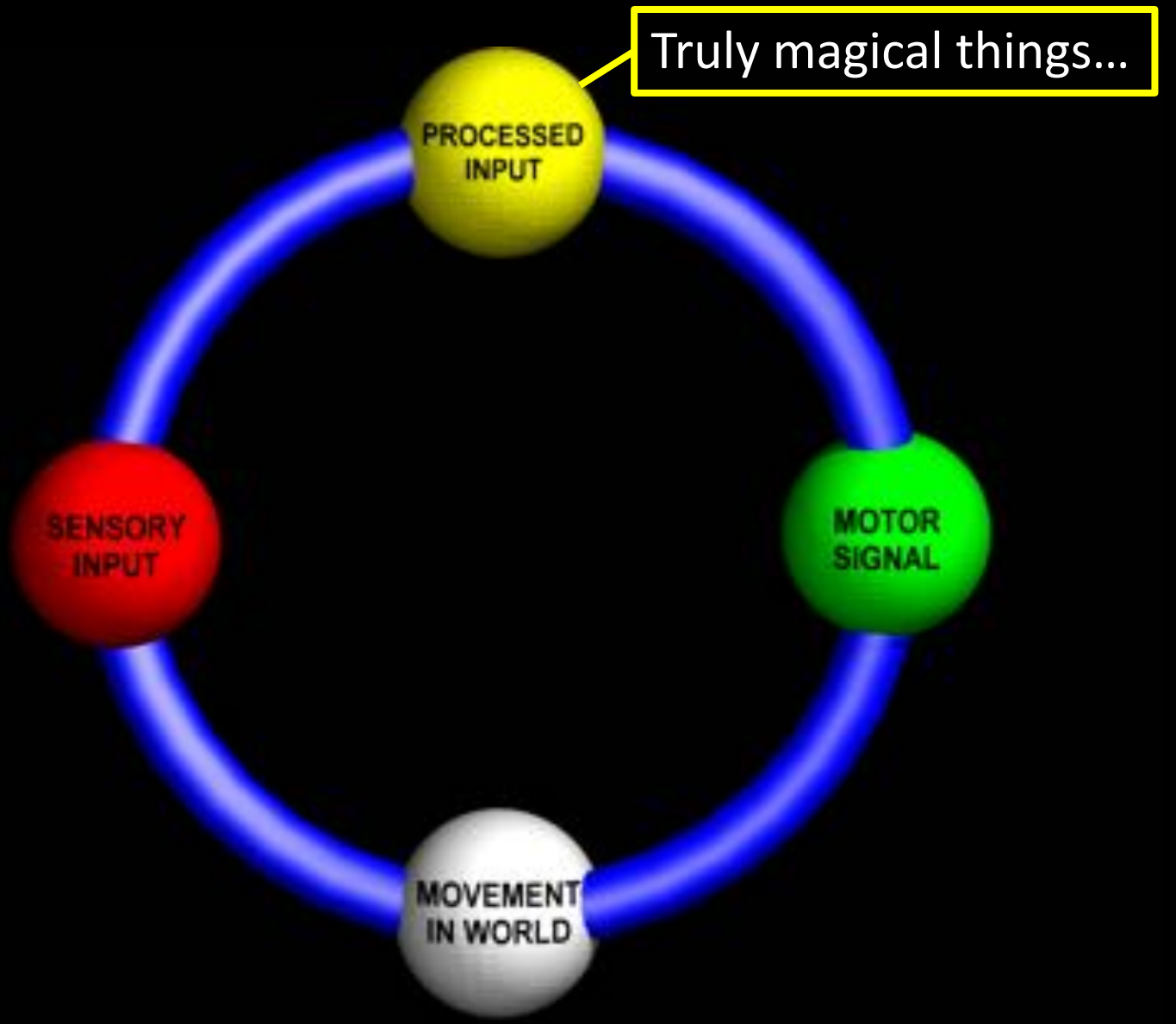



SENSORY
INPUT


PROCESSED
INPUT
SENSORY
INPUT

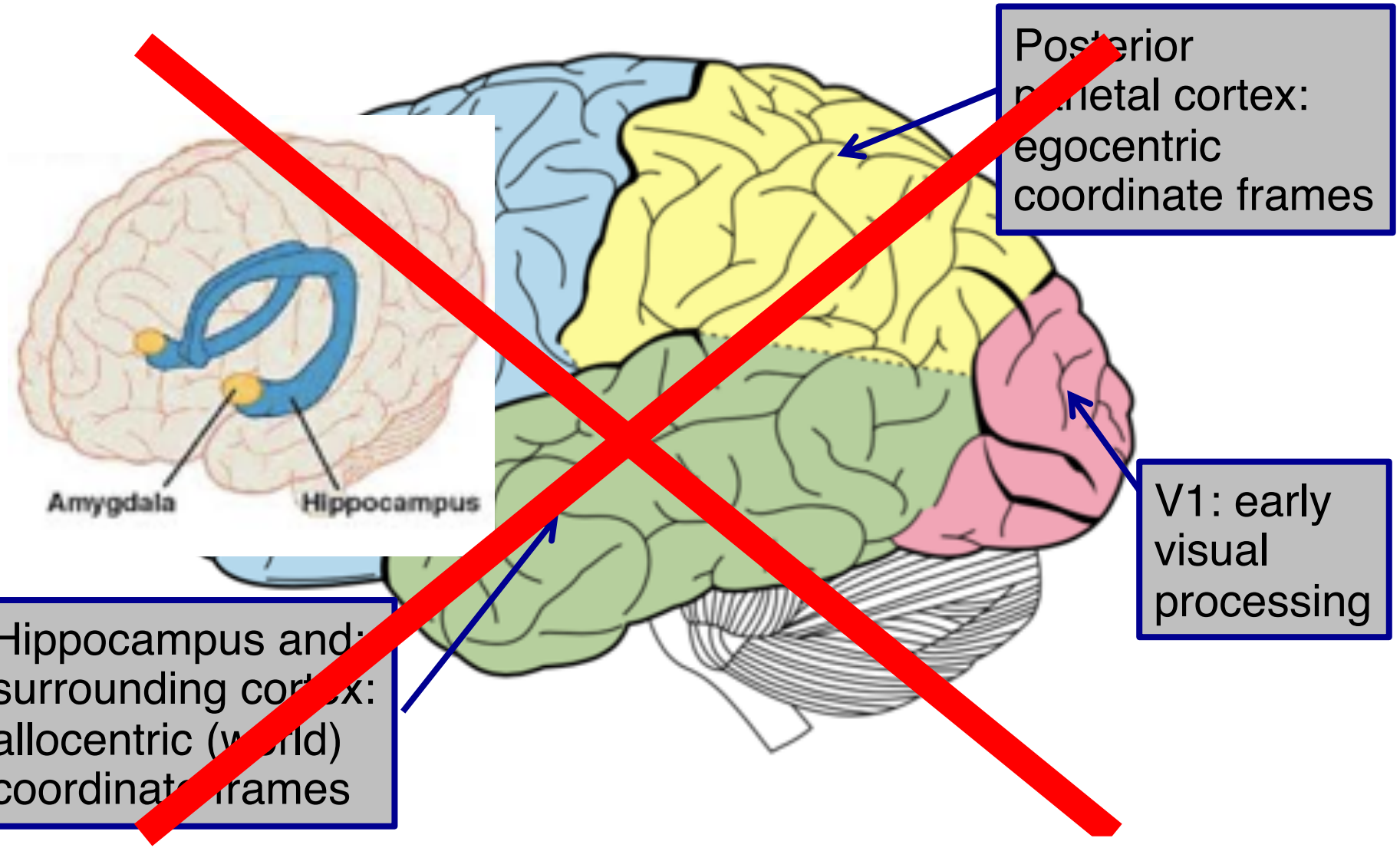

PROCESSED
INPUT
SENSORY
INPUT
MOTOR
SIGNAL

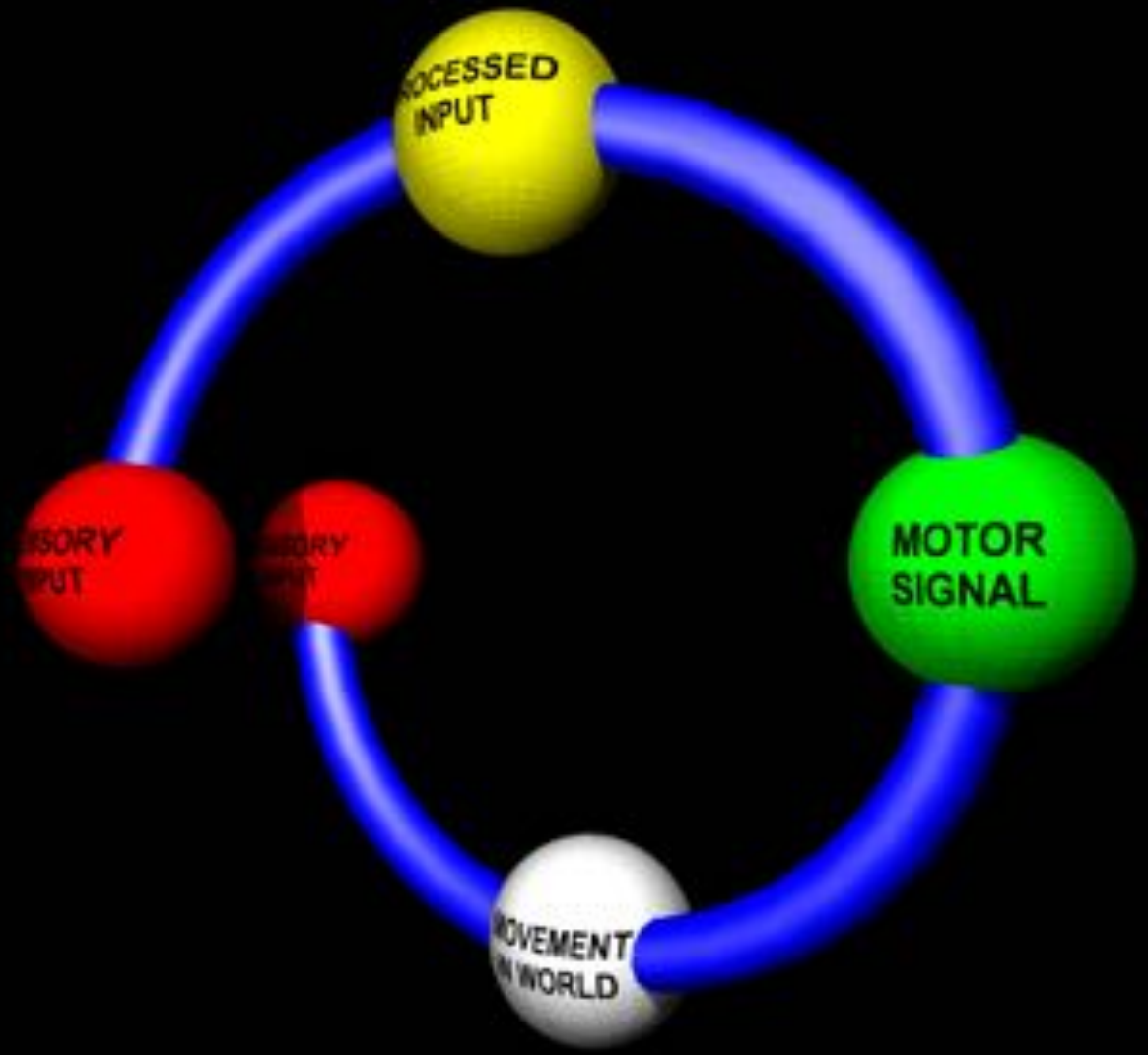

PROCESSED INPUT
MOTOR SIGNAL
MOVEMENT IN WORLD
SENSORY INPUT

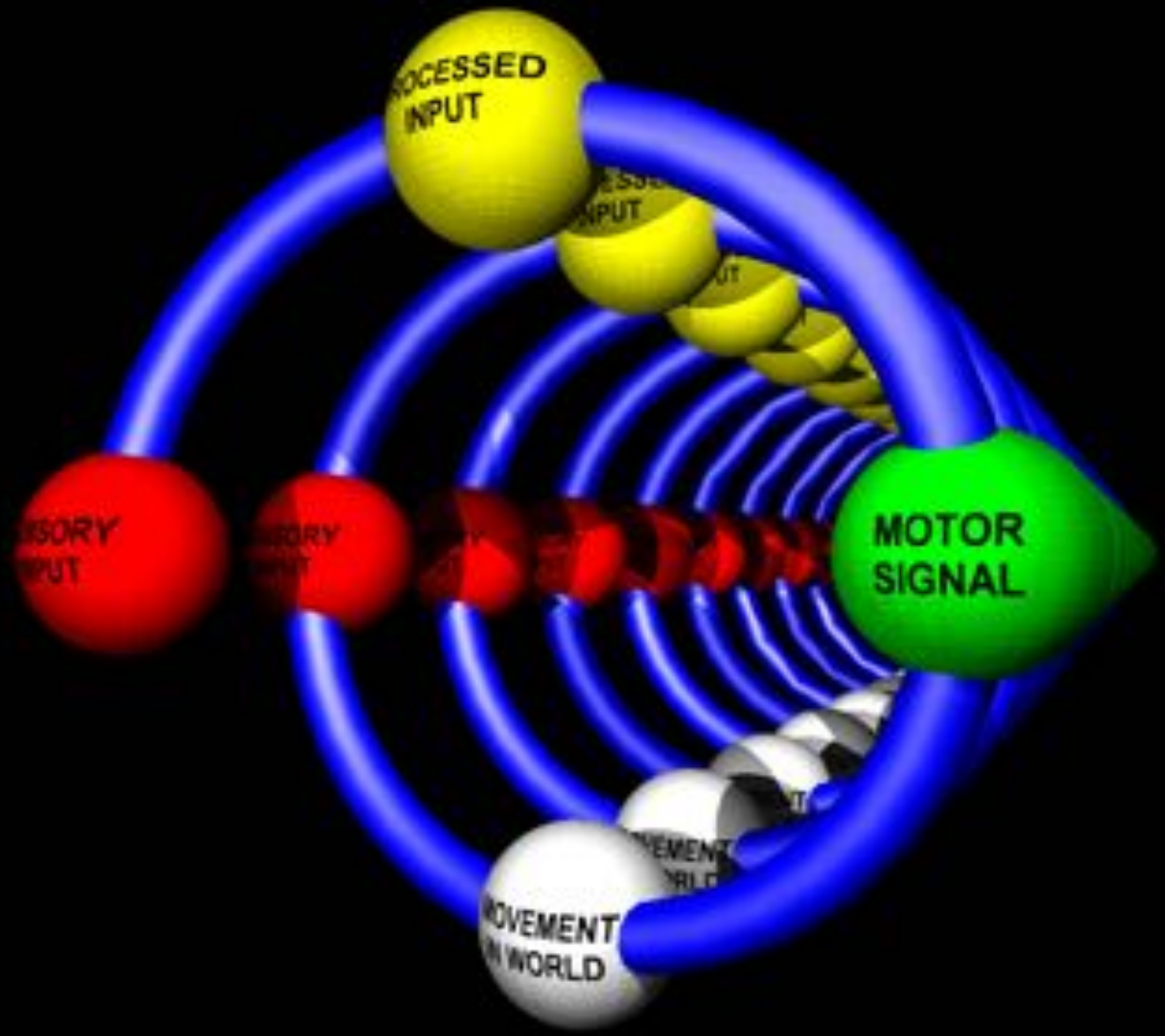

Truly magical things…
PROCESSED INPUT
MOTOR SIGNAL
MOVEMENT IN WORLD
SENSORY INPUT

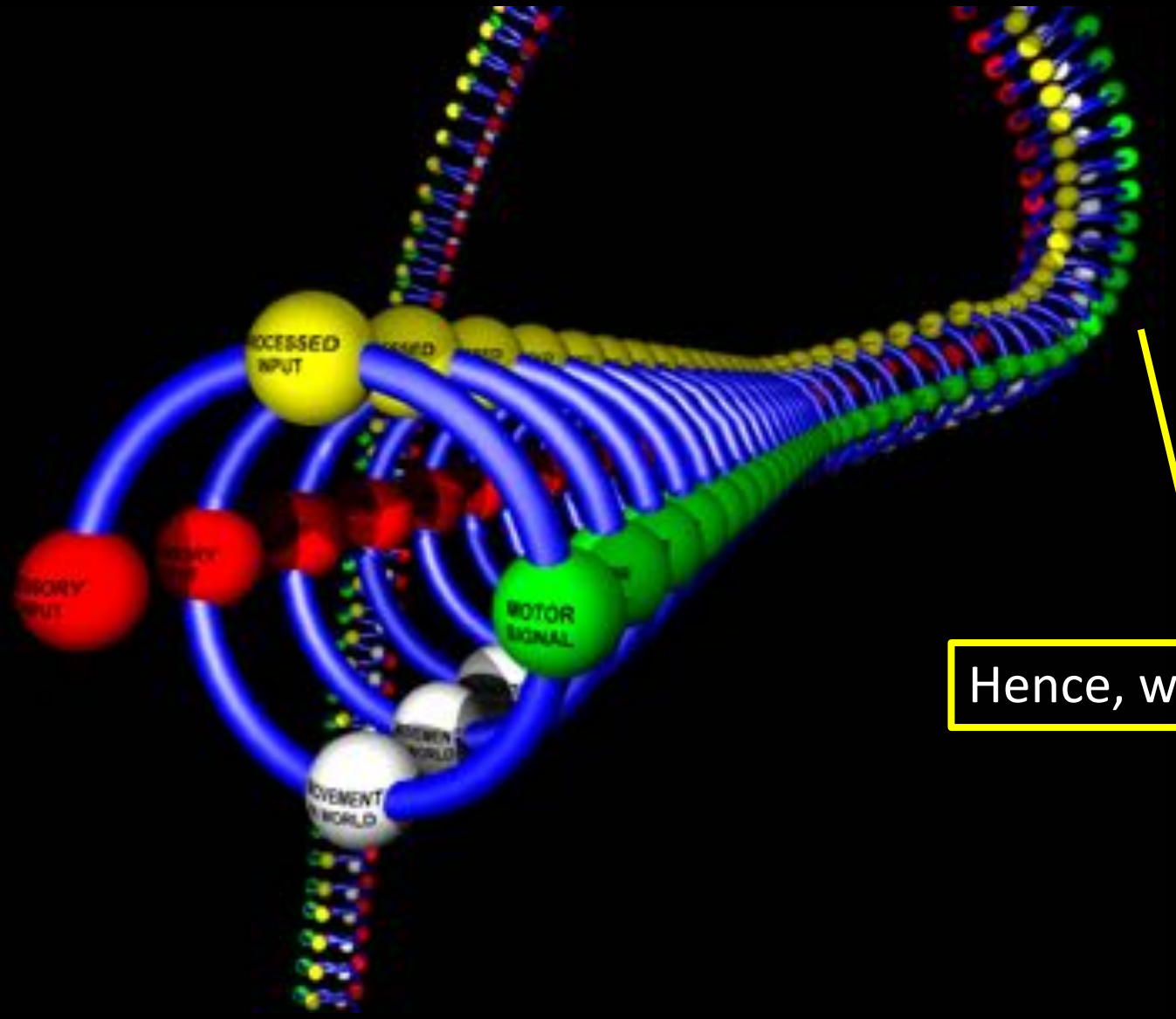
Current hypothesis

University of Reading




OCESSED
INPUT
MOTOR
SIGNAL
OVEMENT
WORLD

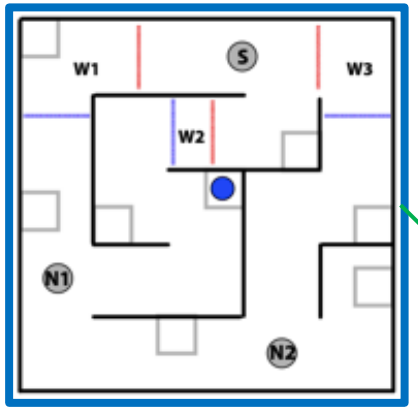

OCESSED
INPUT
MOTOR
SIGNAL
OVEMENT
WORLD



Hence, we need VR…

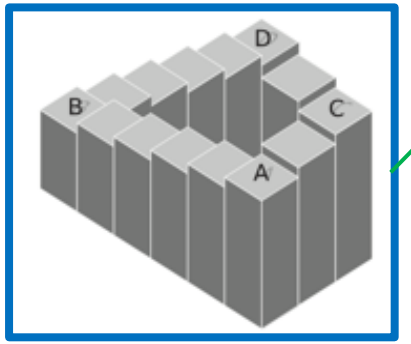
Three experiments:

Alex Muryy



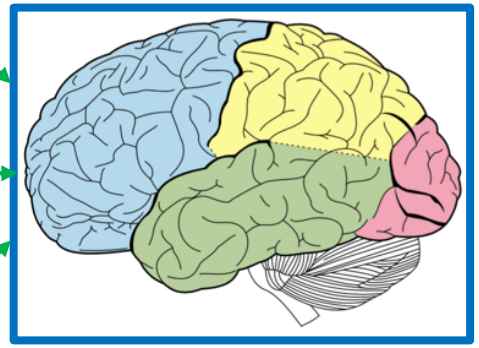
Jenny Vuong

Ellen Svarverud



Are there 3D representations in the brain?

Criterion:
Can a single 3D representation explain performance across multiple tasks?

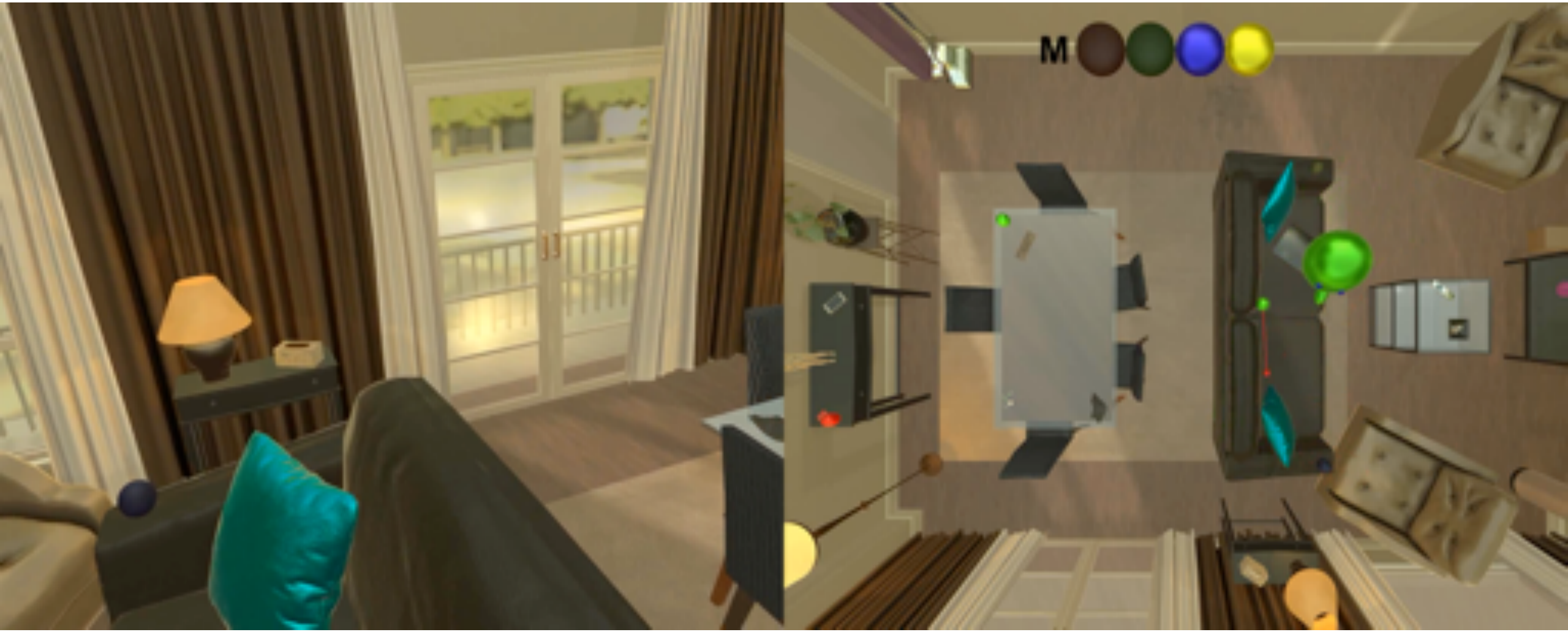


Only possible to test alternative hypotheses if the observer moves: i.e. VR is essential

Others using VR to argue a similar case:

- Warren (2019), *Journal of Experimental Biology*
- Strickrodt, Bülthoff, Meilinger (2018), *Journal of Experimental Psychology: LMC*
- Gillner and Mallot (1998), *Journal of Cognitive Neuroscience*

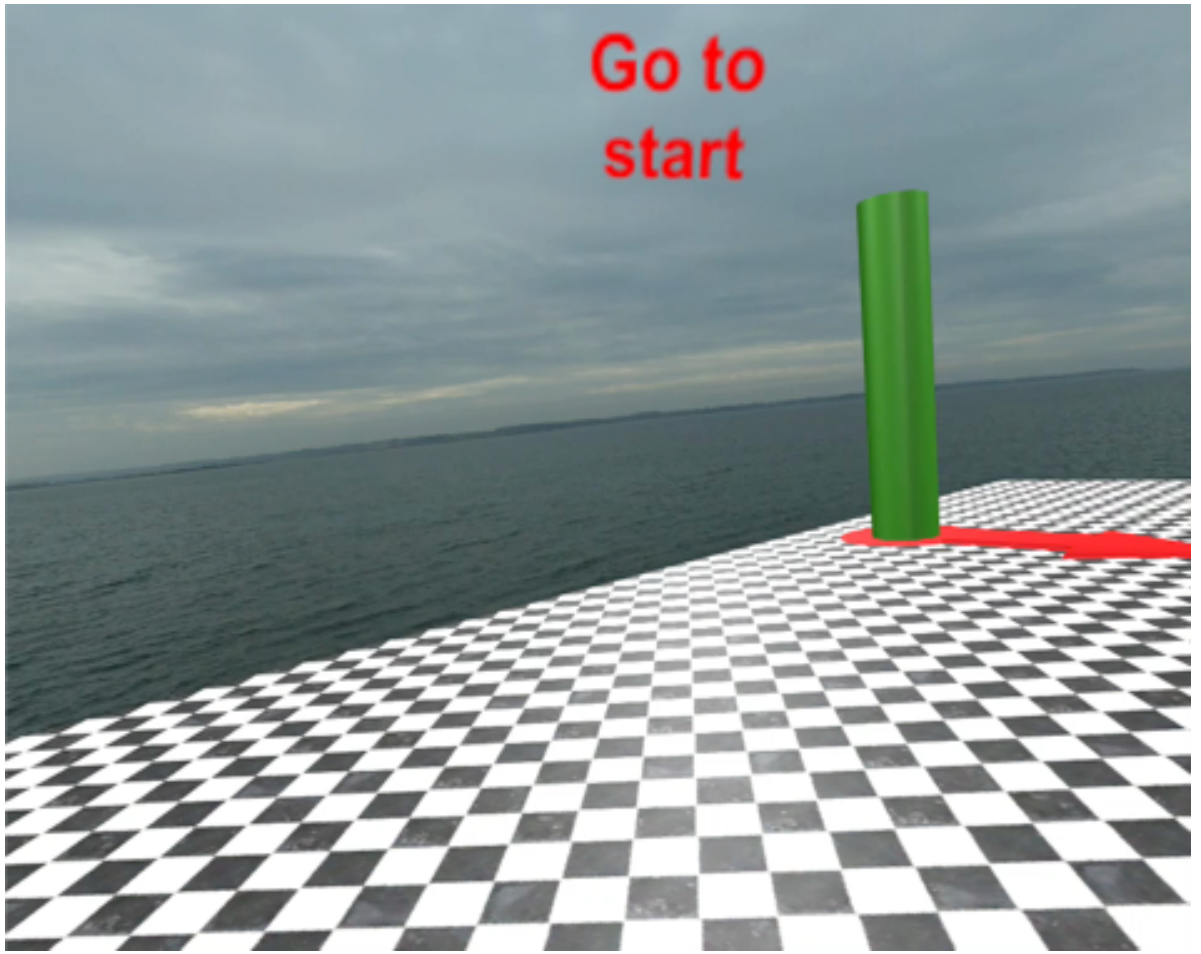
University of Reading



Alex Muryy

University of Reading

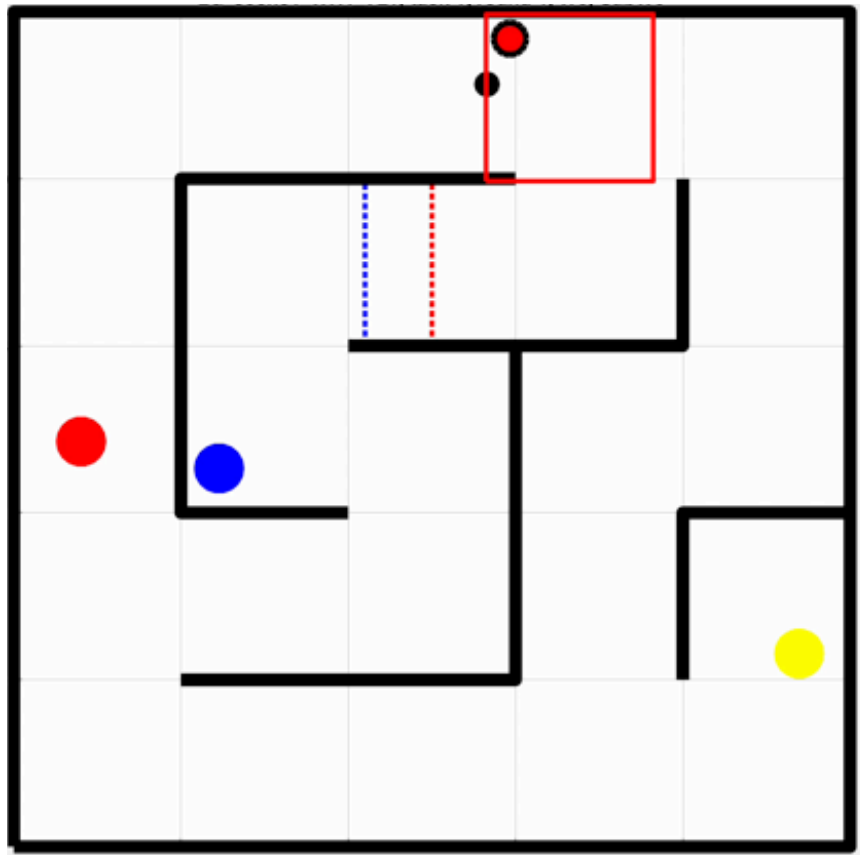
Learning to point to targets in a maze



Tasks:
(i) find targets in specified order and
(ii) point to them…

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

Learning to point to targets in a maze



Life gets harder…

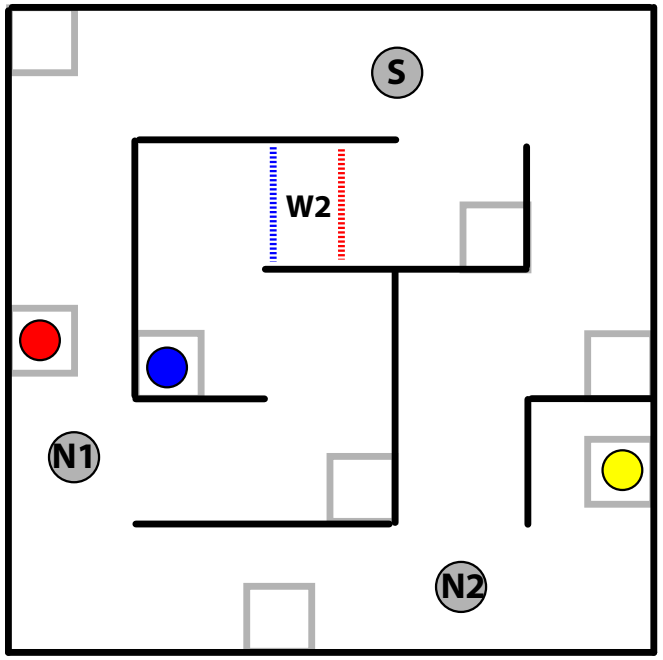
Learning phase (repeat x5):
a) Navigation: go Start-R-G-B-Y
b) Pointing: from Y point to S, R, G, B

Test phase (x3):
a) Random sequences
b) Point to all targets

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

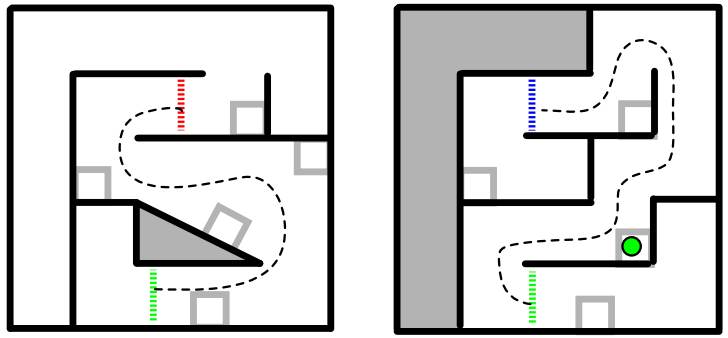
University of Reading

Non-metric scene: 1 wormhole

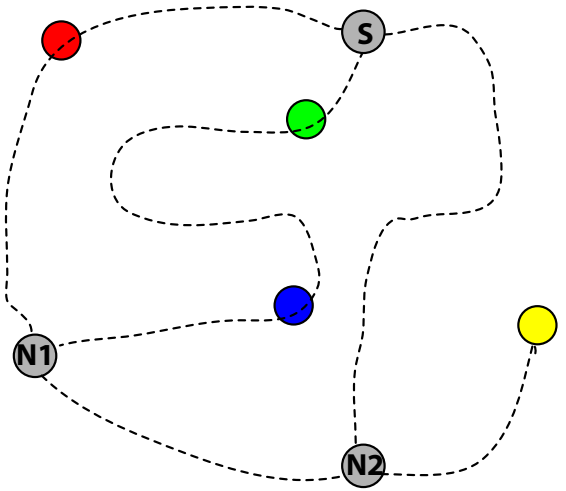


- topological node
- small walls
- wormhole

wormhole

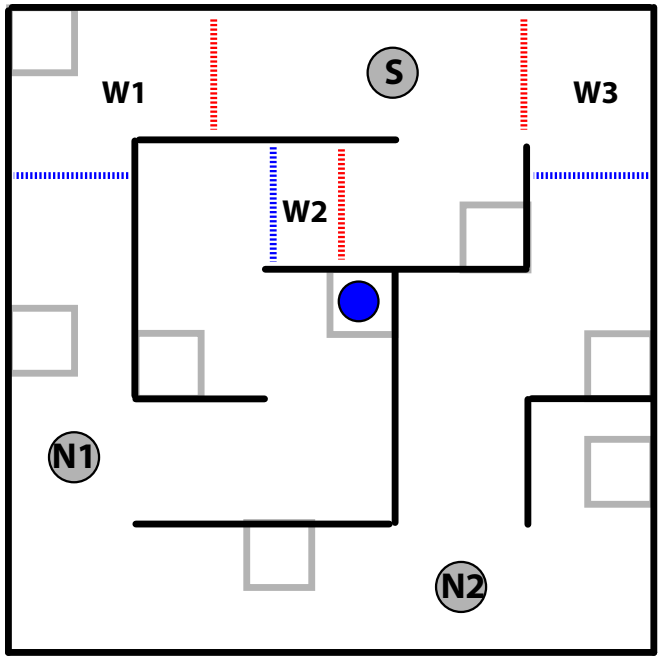


topological graph

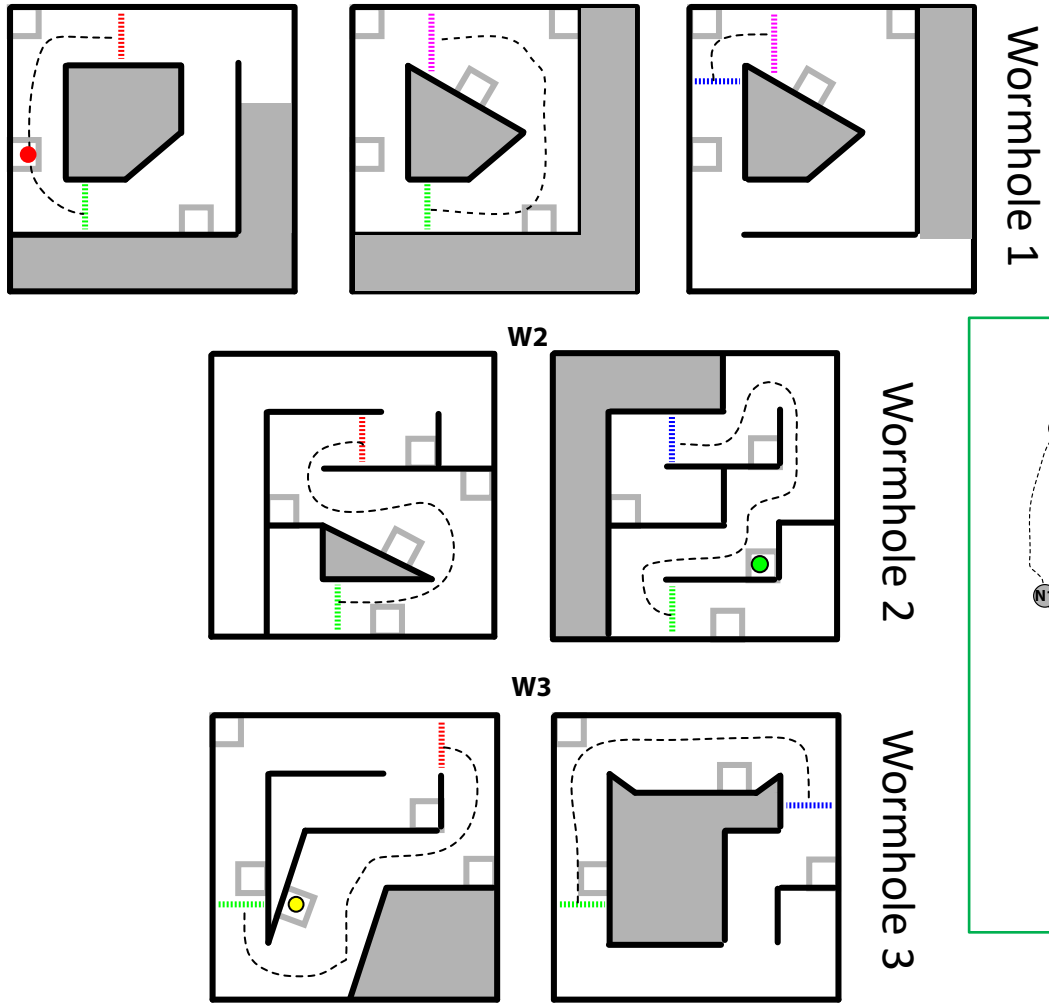


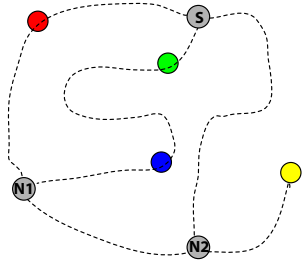
Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

Non-metric scene: 3 wormholes



- topological node
- small walls
- wormhole

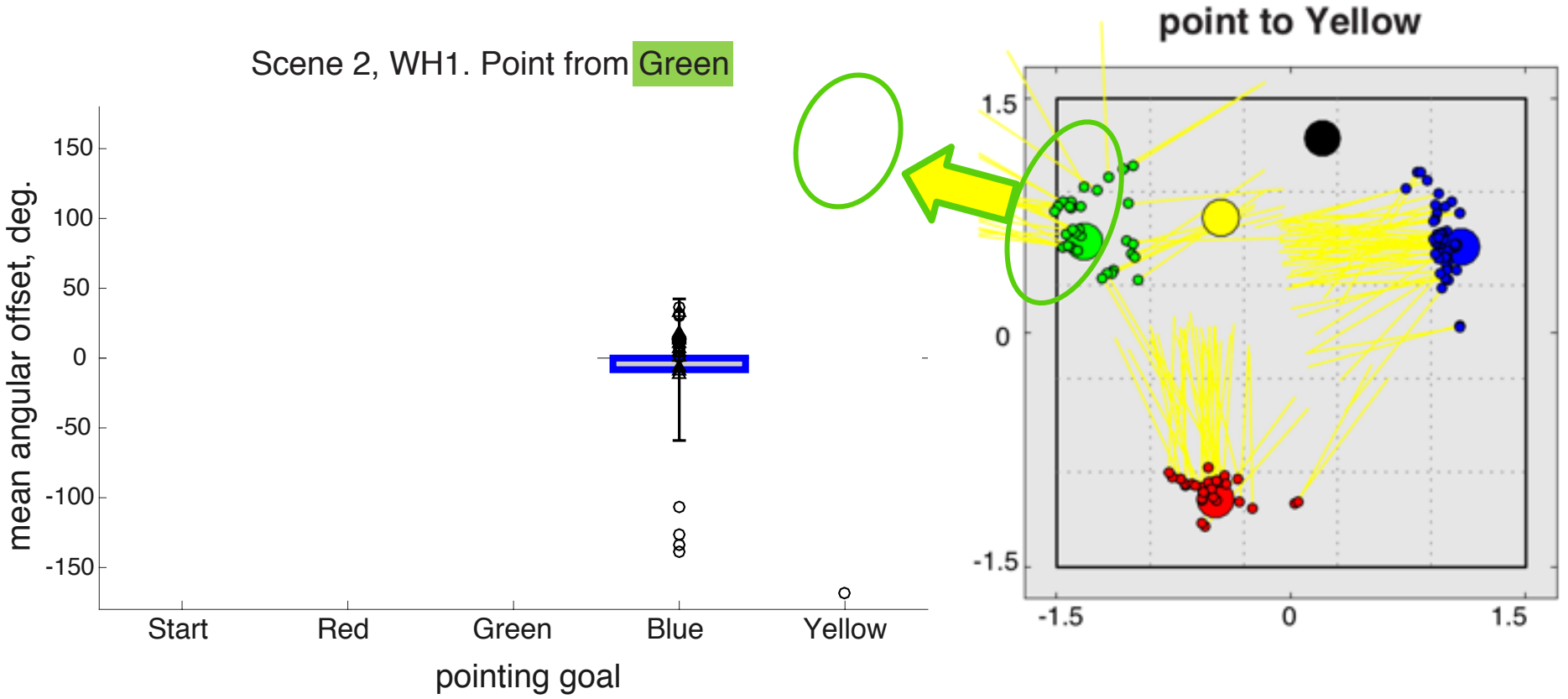




… but same topological structure

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

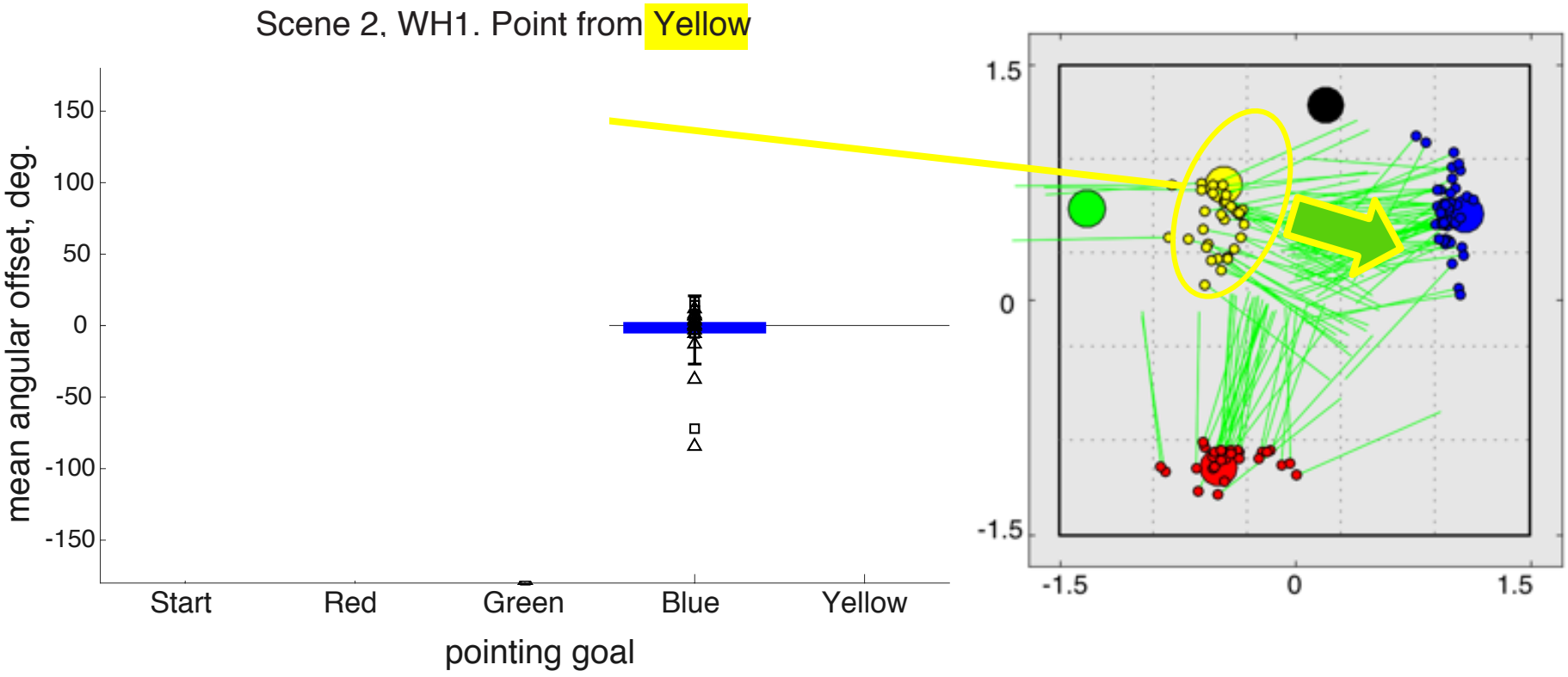
Pointing: very large errors



Pointing to some targets leads to very large, systematic errors.

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

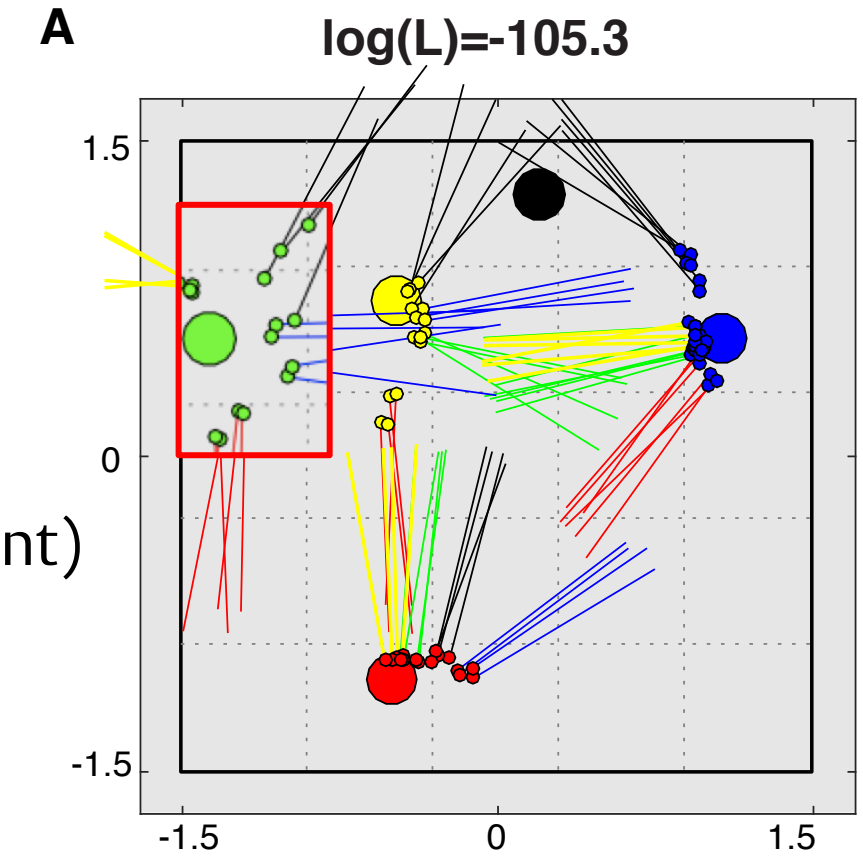
Pointing: very large errors



Pointing to some targets leads to very large, systematic errors.

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

Pointing: very large errors



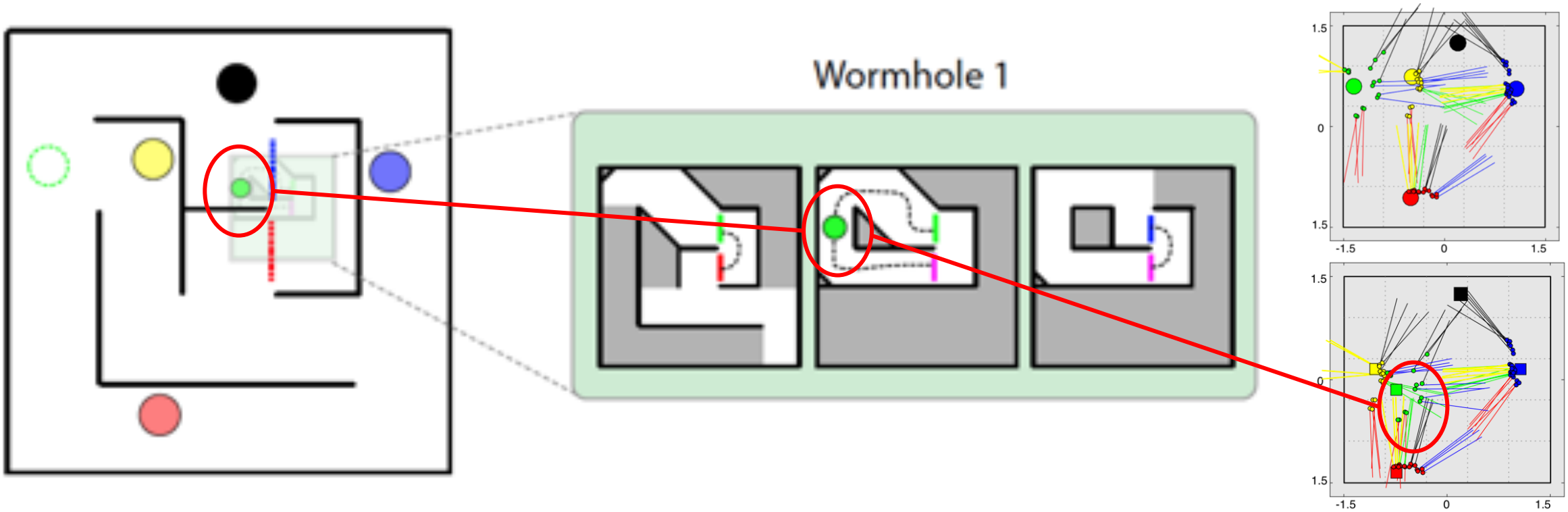
Original (one participant)

Maximum likelihood configuration

In the most likely configurations, green is to the east of yellow.

Muryy and Glennerster (2018) *Lecture Notes in Computer Science*

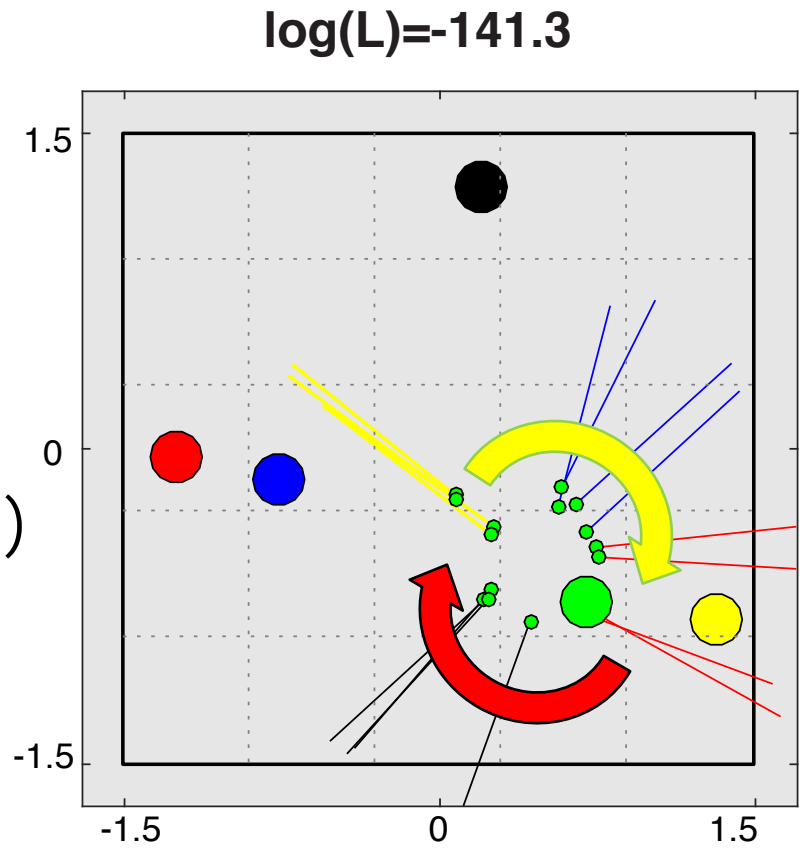
Pointing: very large errors



It seems as if participants ‘squash’ the wormhole corridors into a smaller region than they actually occupy .

Adding in rotation

Original (one participant)



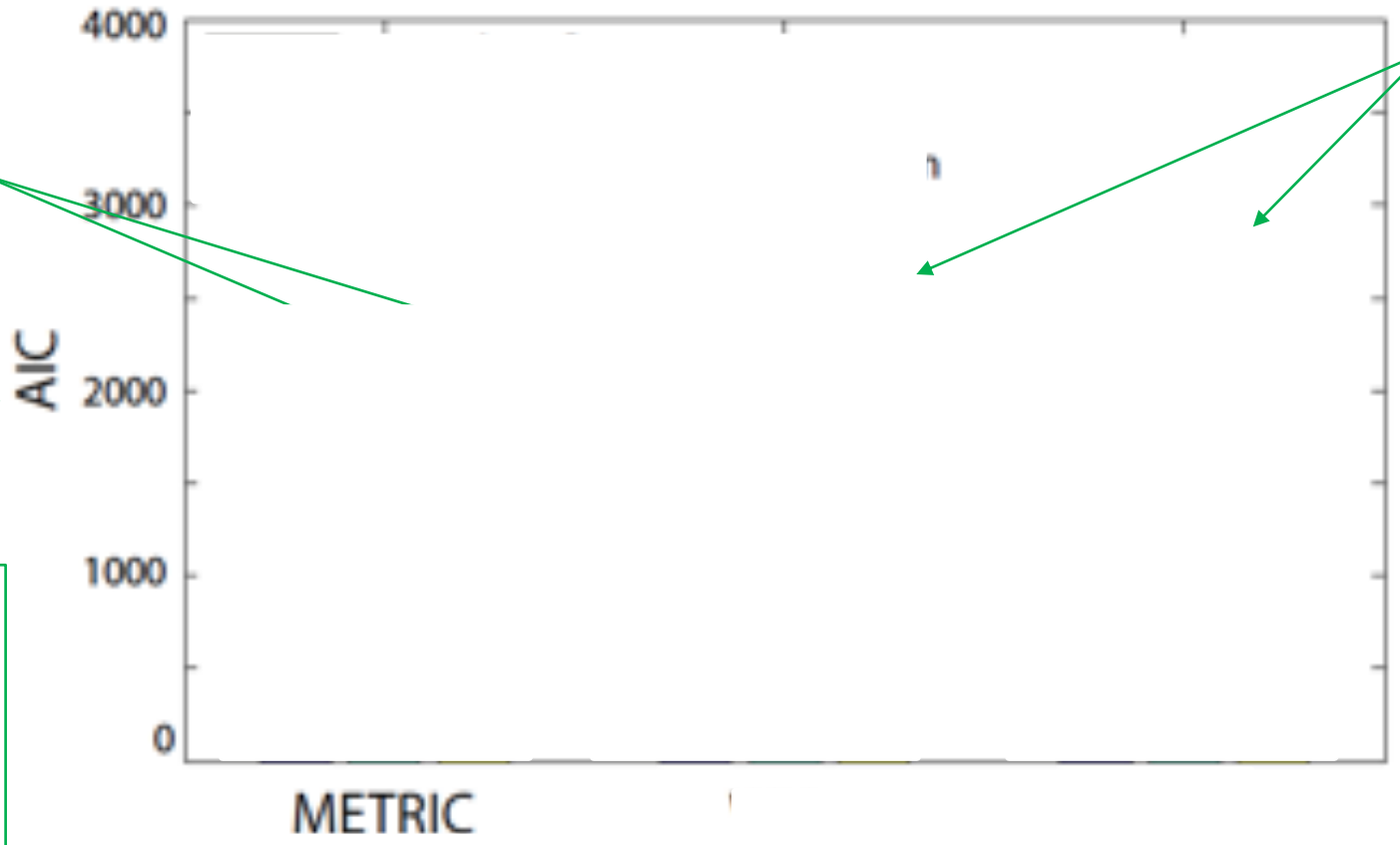
180° rotation of all pointing directions

This takes into account the possibility that people are disoriented.
But it is not compatible with a single, consistent 3D representation.

Model comparison

In the METRIC condition, optimised translation and rotation are not better models than the original configuration (when penalized for the extra parameters in the models)

Akaike information criterion, a measure of likelihood (low is more likely)

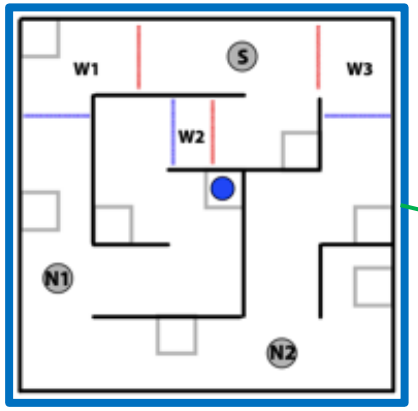


In the WORMHOLE conditions, the best model is one that optimizes the location of the targets (i.e. a distorted world) *and* optimizes the rotation of the observer independently at each pointing zone

This is not a consistent, metric, 3D representation

Three experiments:

Alex Muryy



Are there 3D representations in the brain?

Criterion:
Can a single 3D representation explain performance across multiple tasks?

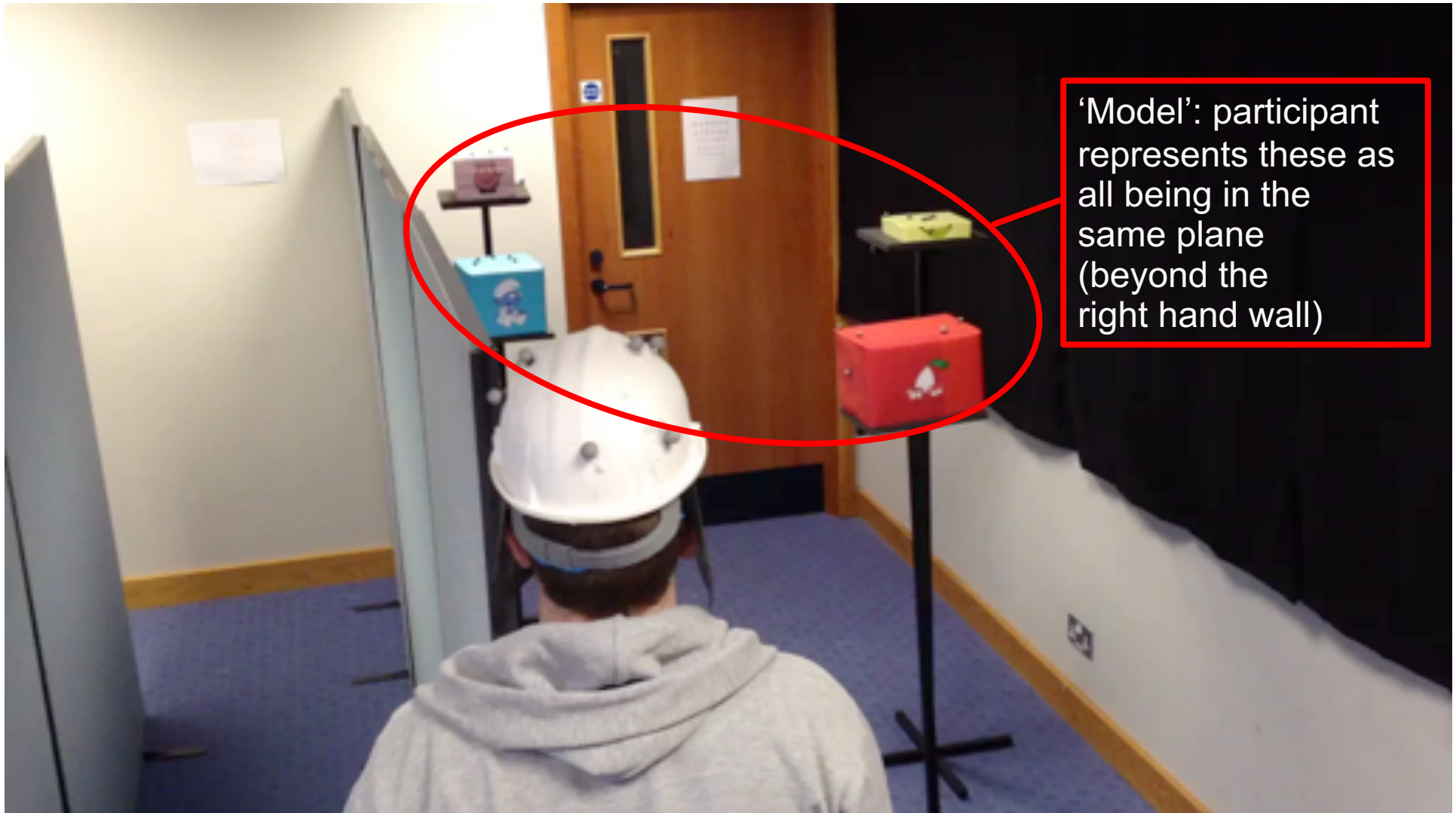
No

Can we update the visual direction of unseen objects as we move?

Jenny Vuong

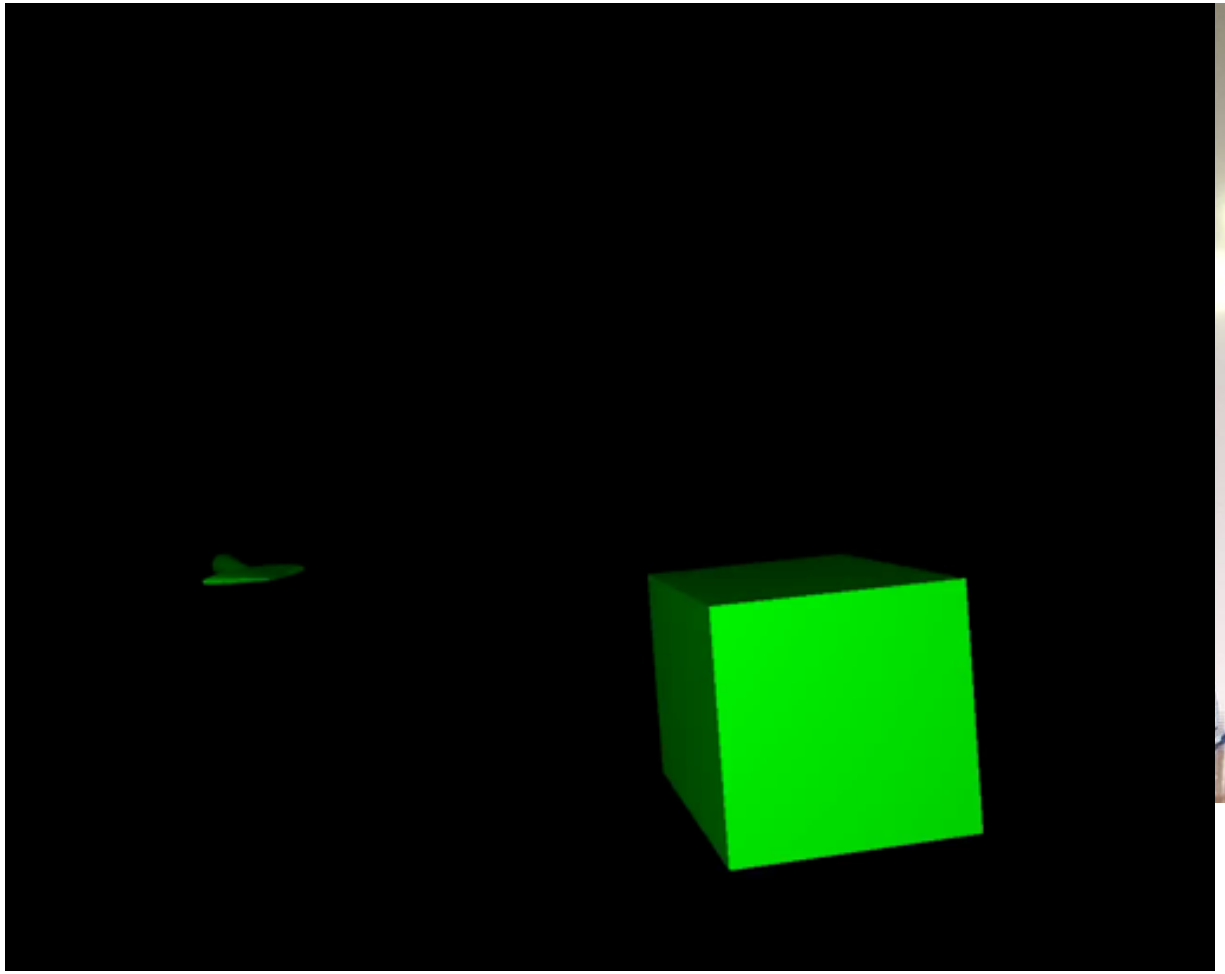
Vuong et al (2019) *Scientific Reports*

Can we update the visual direction of unseen objects as we move?



Vuong et al (2019) *Scientific Reports*

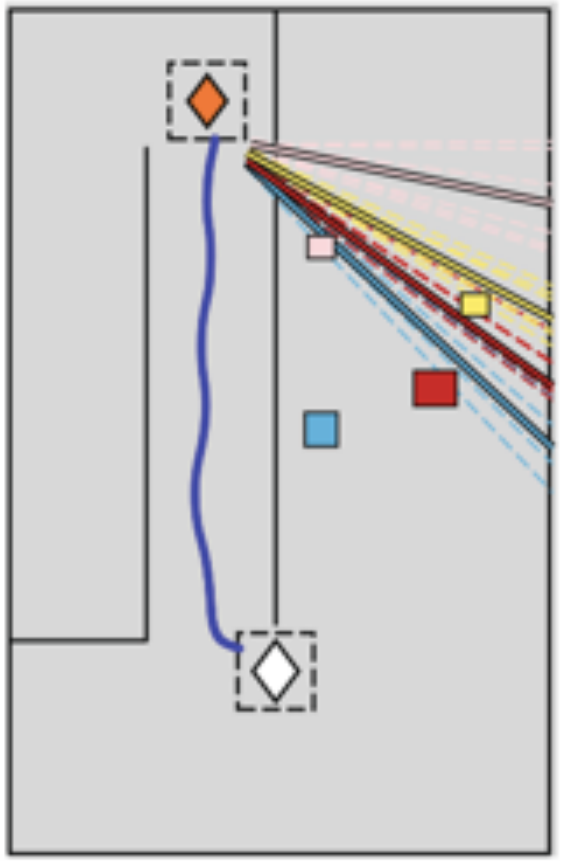
Can we update the visual direction of unseen objects as we move?



nVis SX111 HMD
Vicon tracking

Vuong et al (2019) *Scientific Reports*

People show large, consistent biases

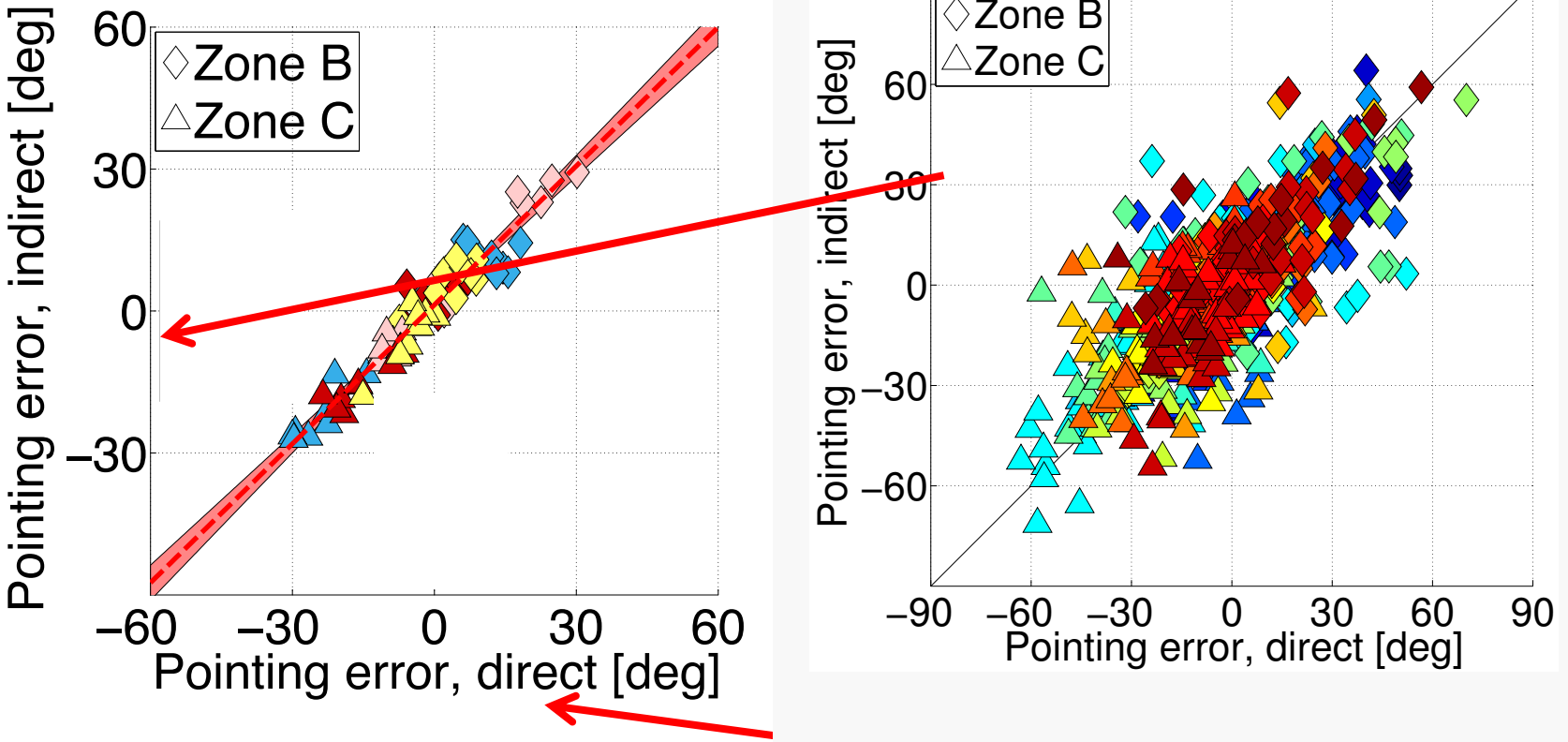


Task:

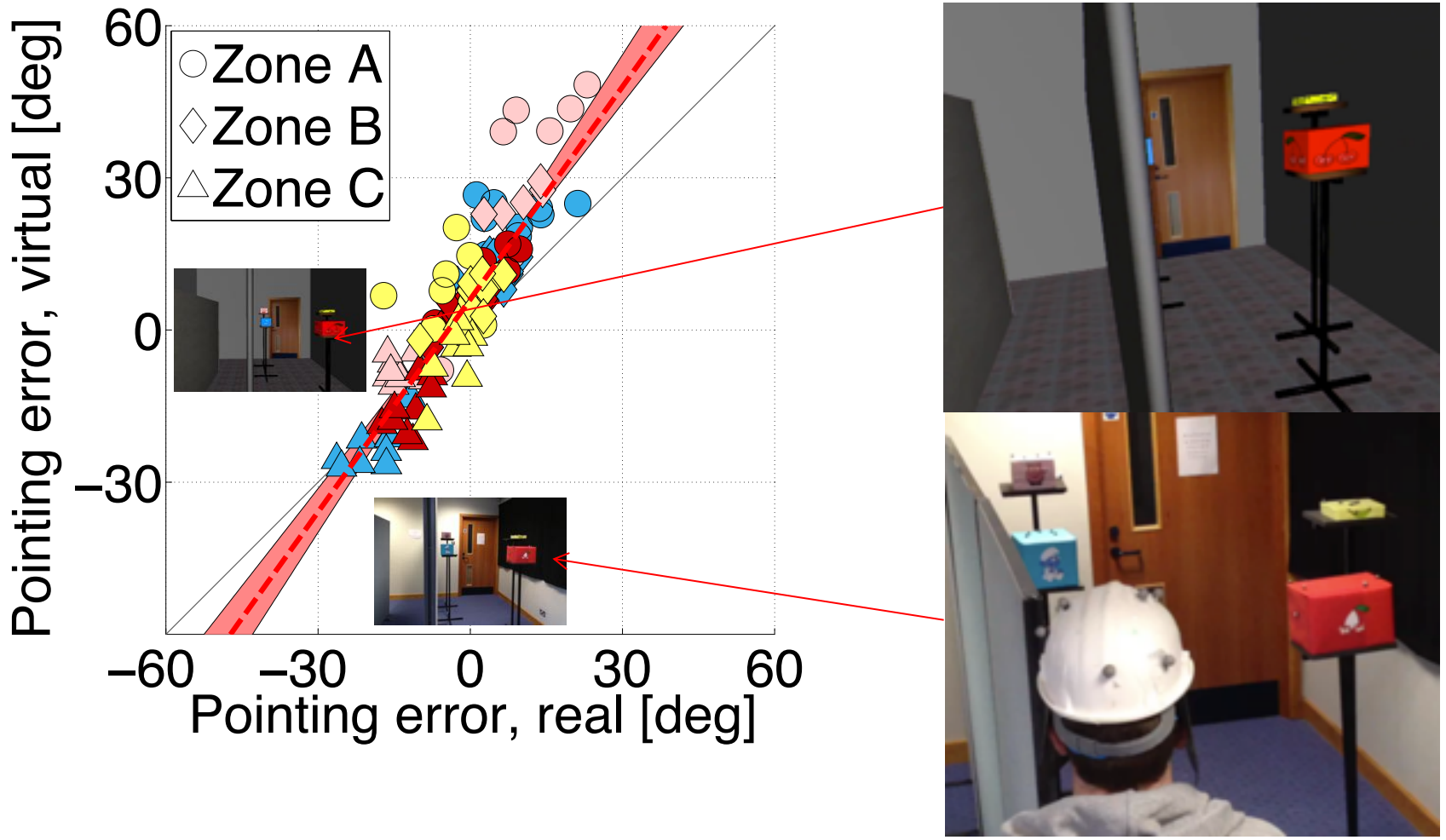
- view a scene
- walk without any further view of the objects
- point to the objects
- easy to do if we update our location in a 3D reconstruction

http://www.jennyvuong.net/dataWebsite/rawdata_main.html

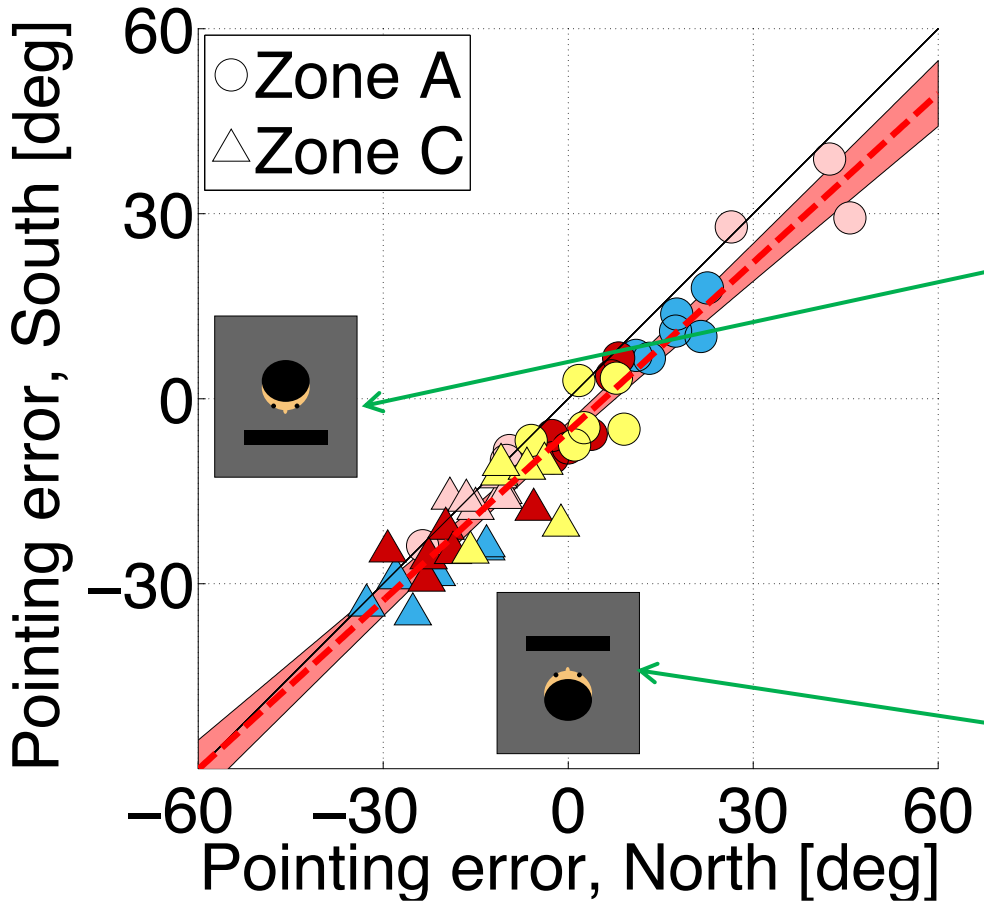
… independent of the route they take …



… similar biases in real and virtual worlds …



… whether looking ‘north’ or ‘south’ …



Facing ‘south’

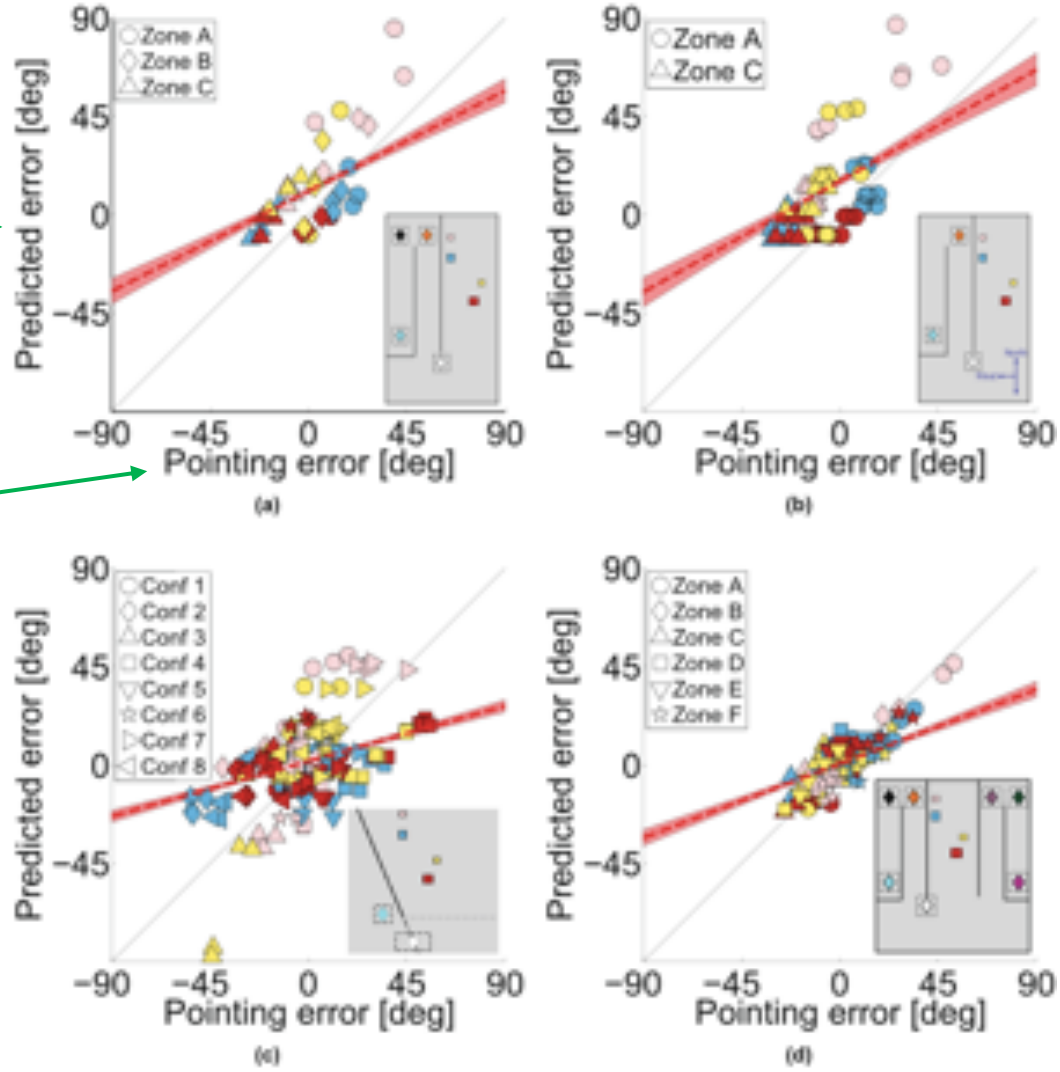
… just another example of pointing errors being correlated across conditions

Facing ‘north’

If you assume the brain builds a 3D model…

Prediction

Data

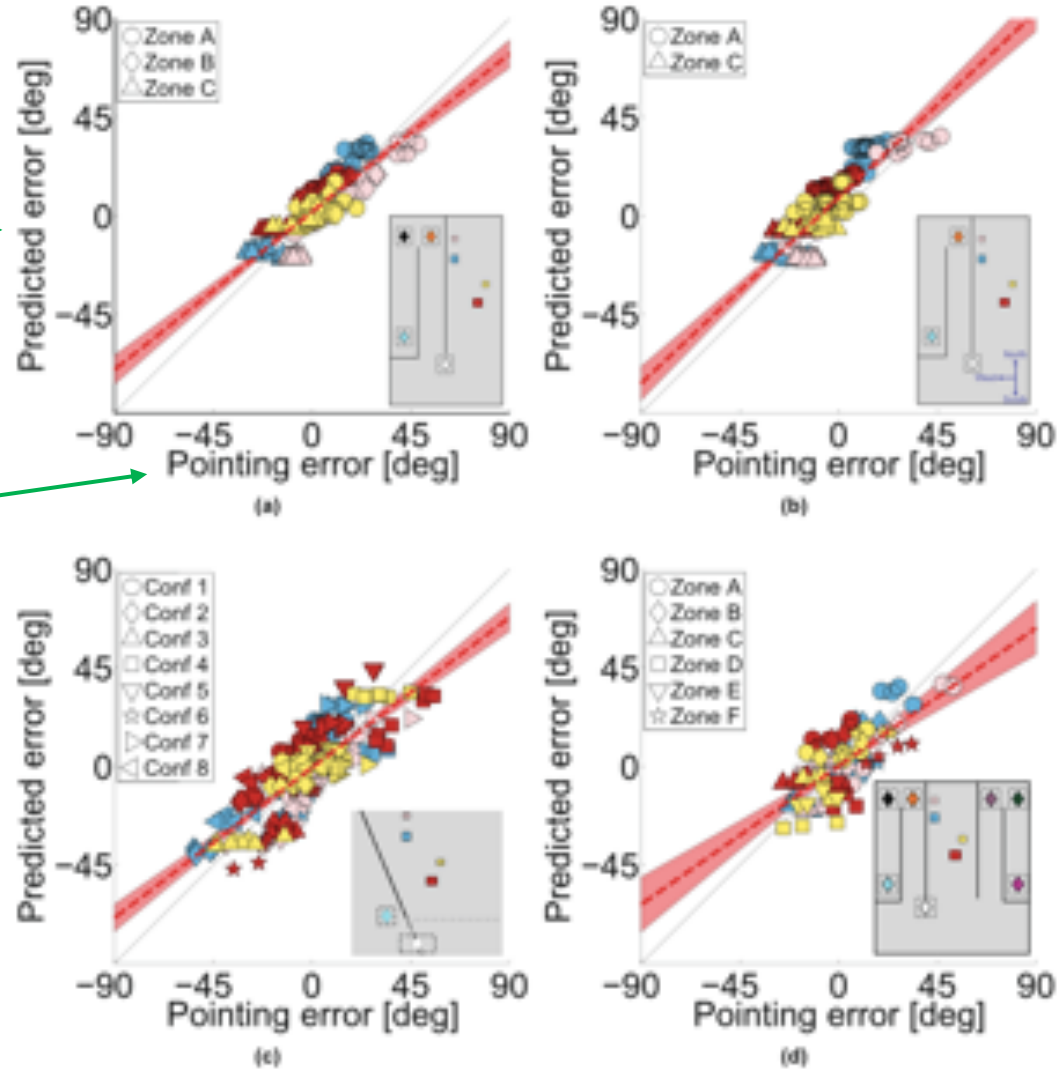


This model is being as generous as possible to the idea that the brain builds a 3D model (either correct or distorted)

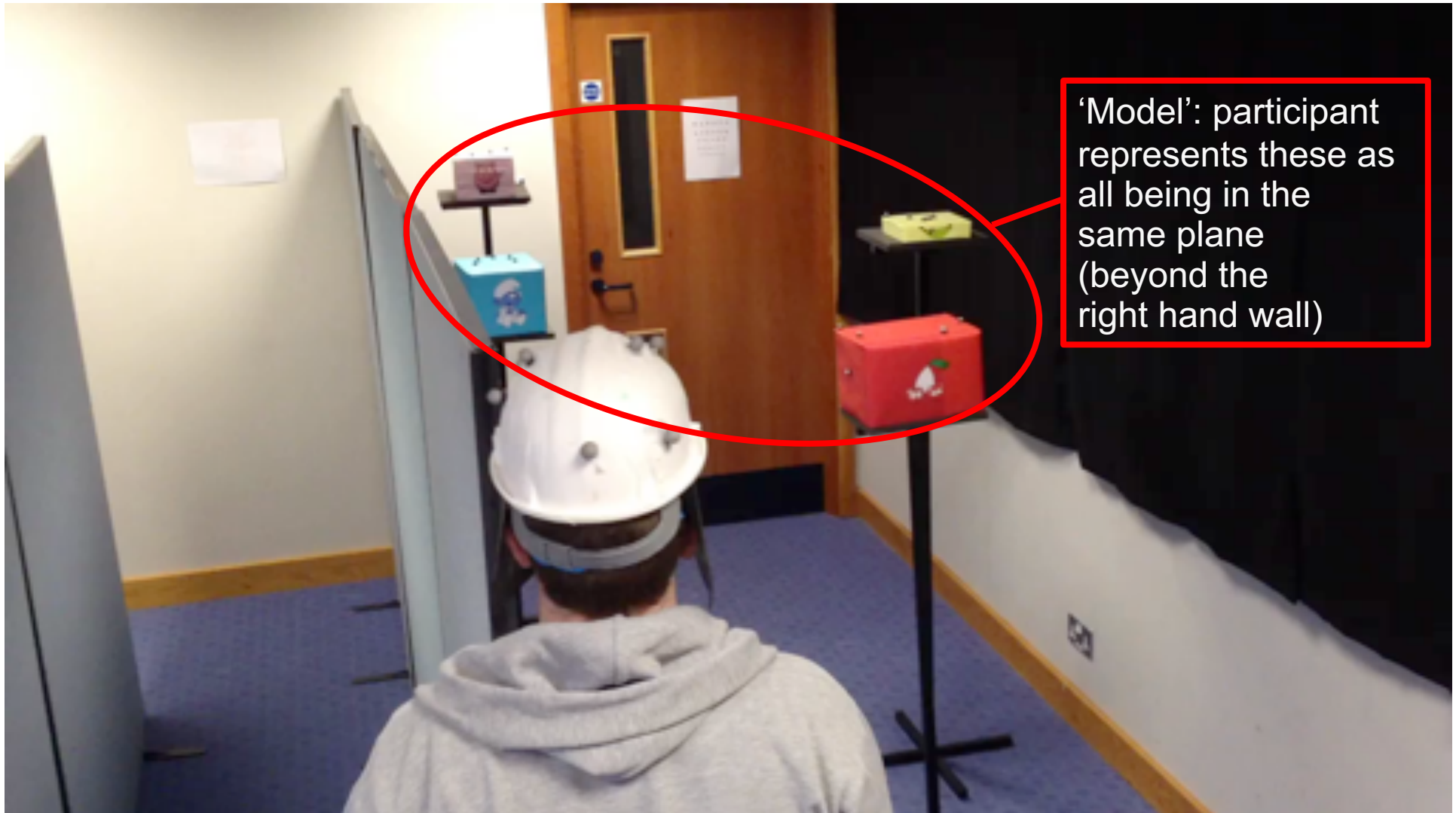
If you assume it does NOT build a 3D model…

Prediction

Data

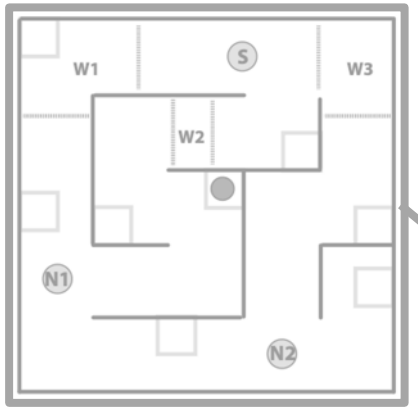


Can we update the visual direction of unseen objects as we move?



Three experiments:

Alex Muryy



Jenny Vuong

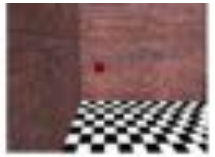
Are there 3D representations in the brain?

Criterion:
Can a single 3D representation explain performance across multiple tasks?

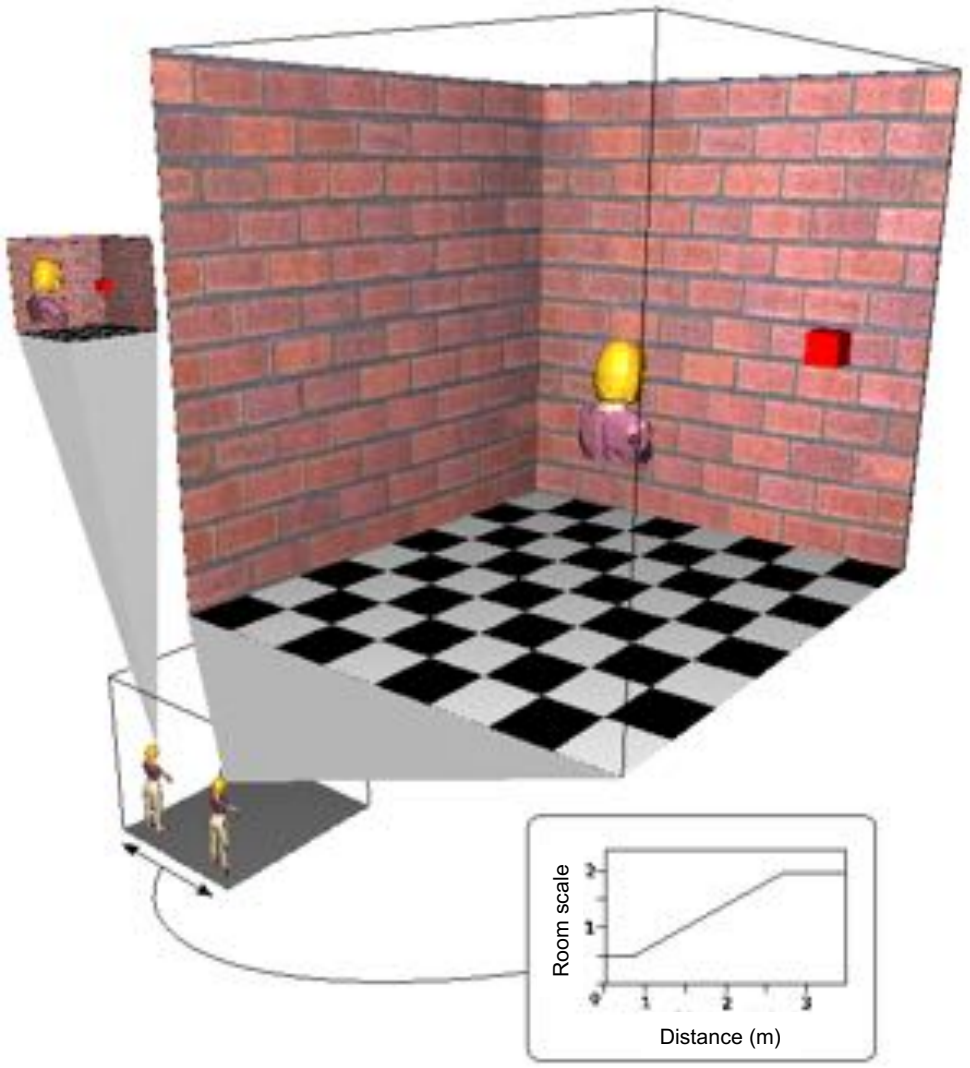
No

Relative depth judgements

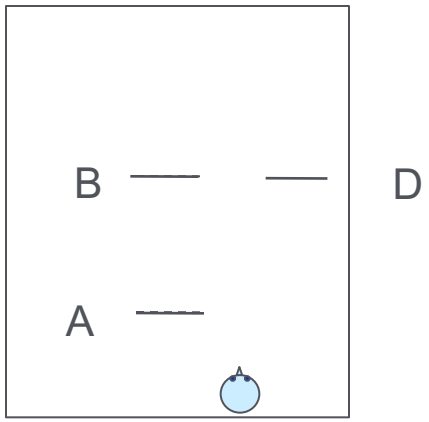
View in small room



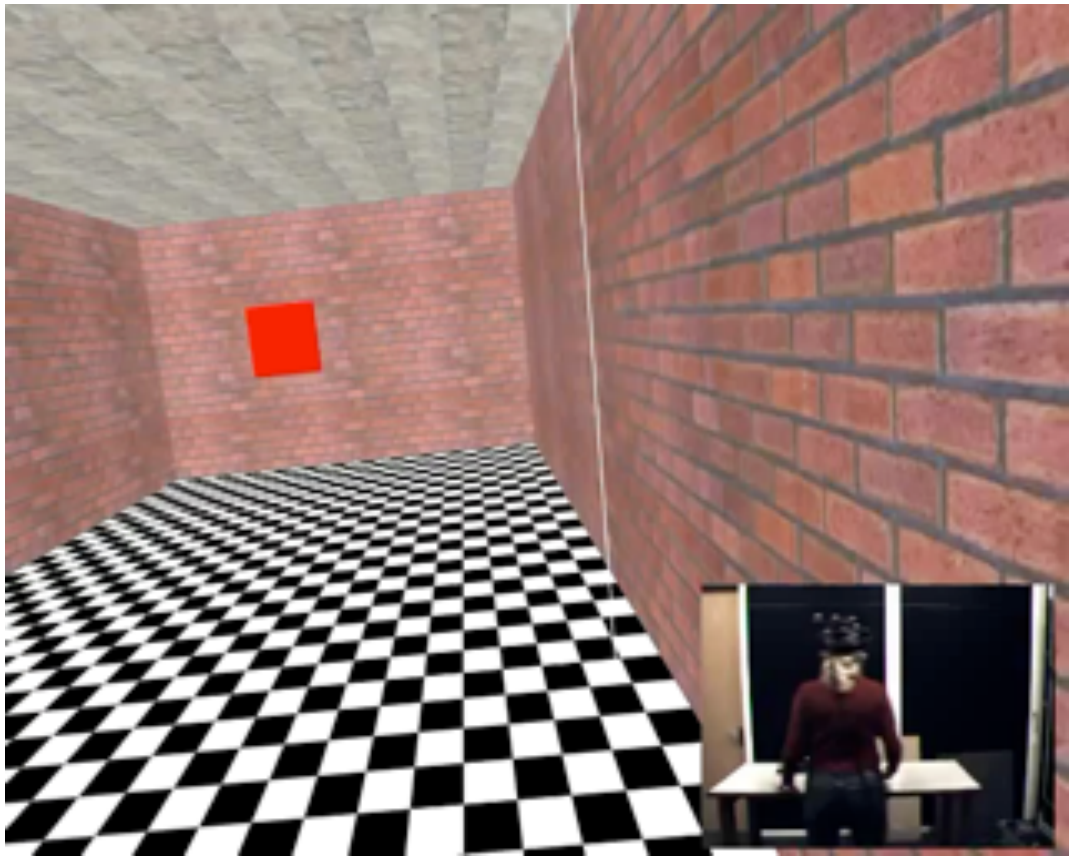
View in large room



Relative depth judgements

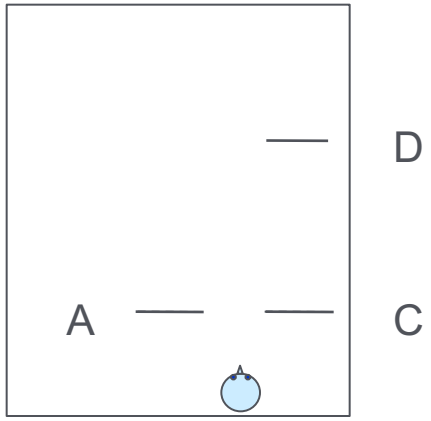


A, B and D are at the same **perceived distance**

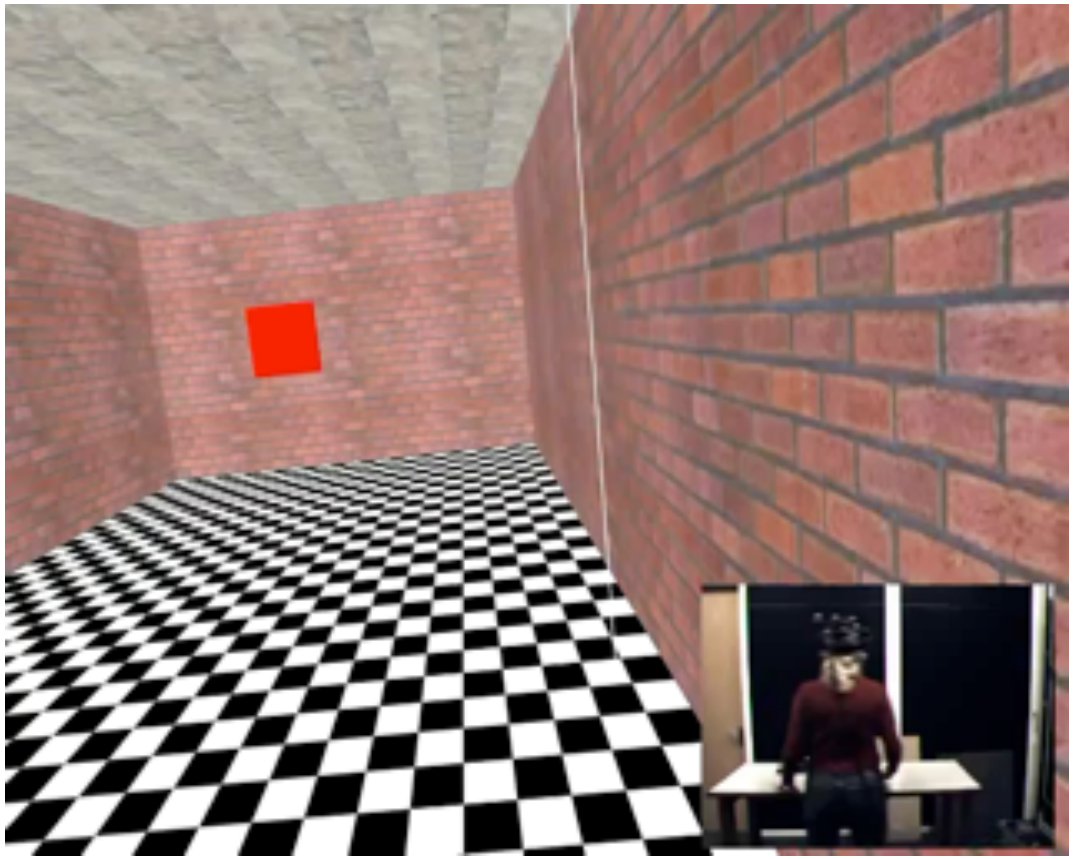


Svarverud et al (2012) *PLoS ONE*

Relative depth judgements

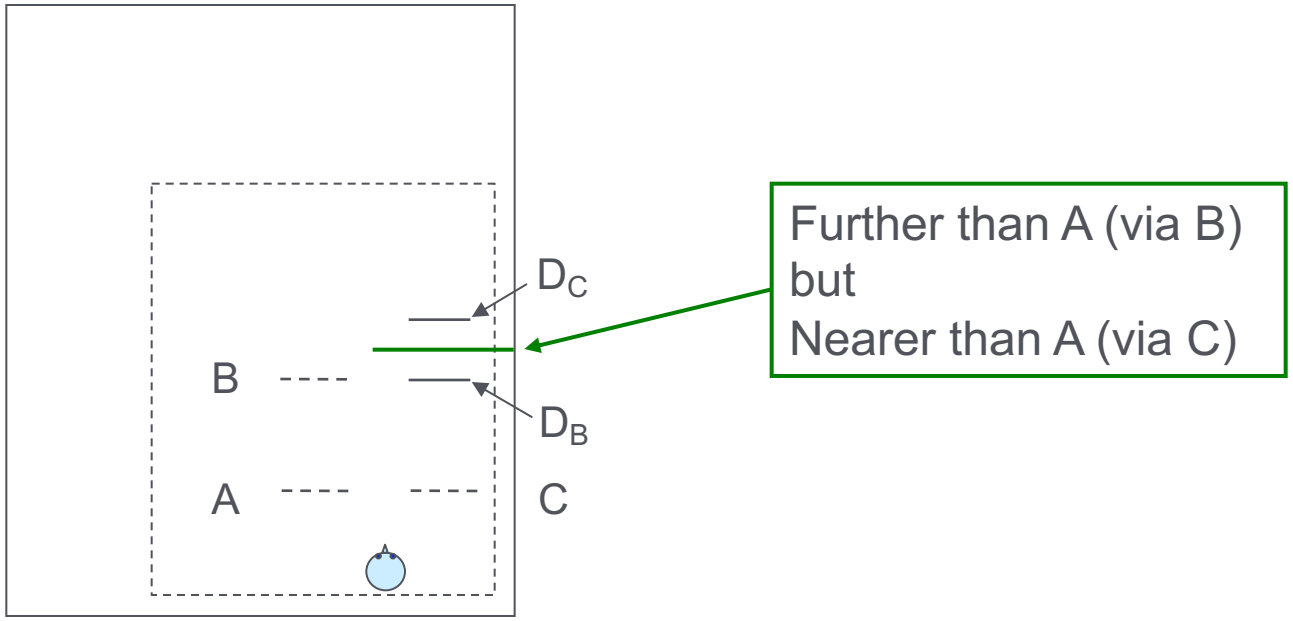


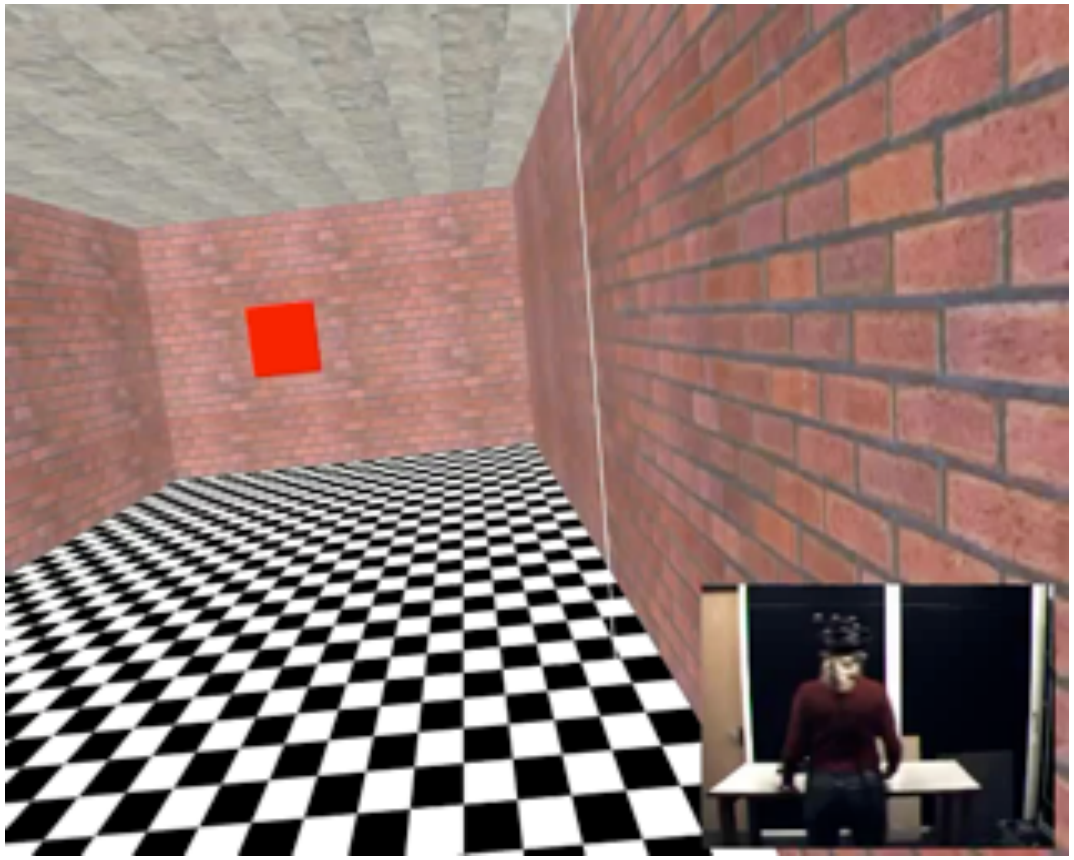
A, C and D are at the same **perceived distance**



Svarverud et al (2012) *PLoS ONE*

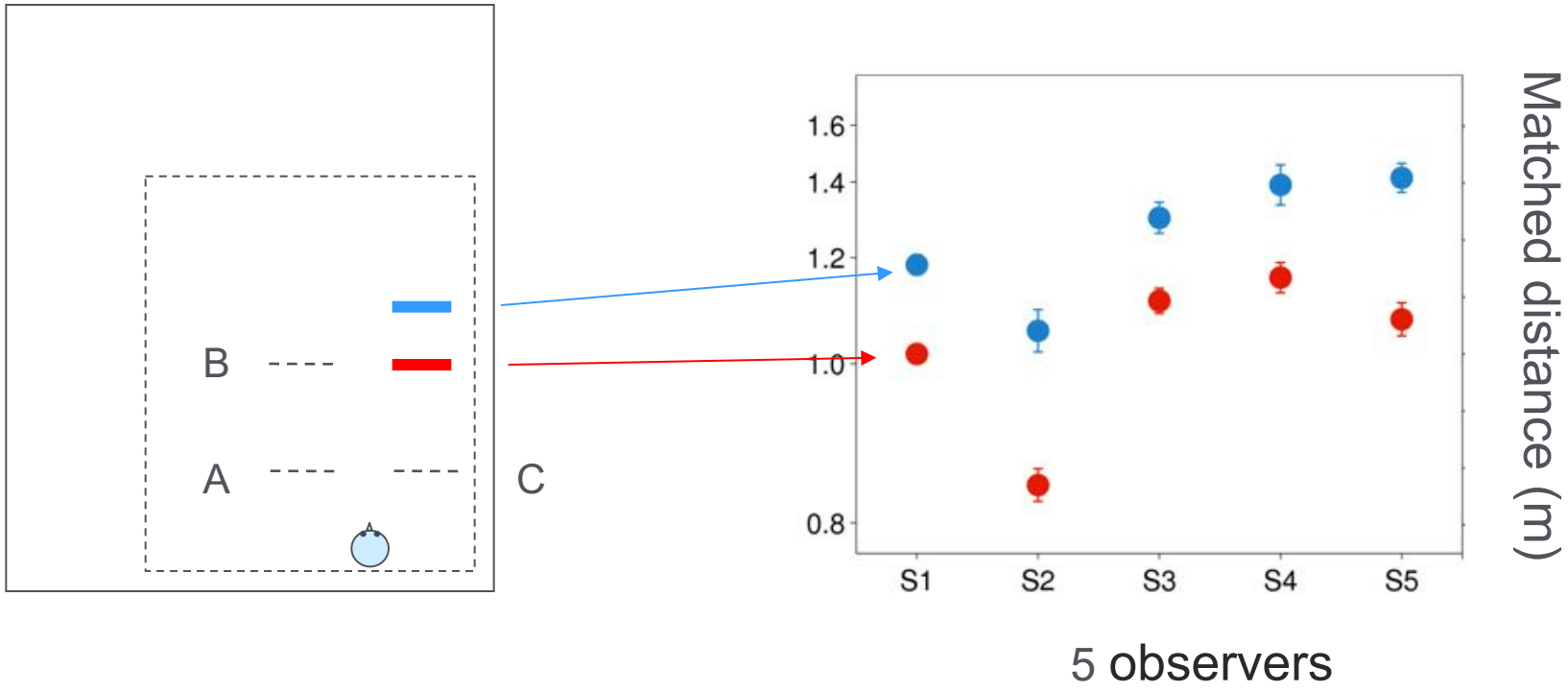
Relative depth judgements





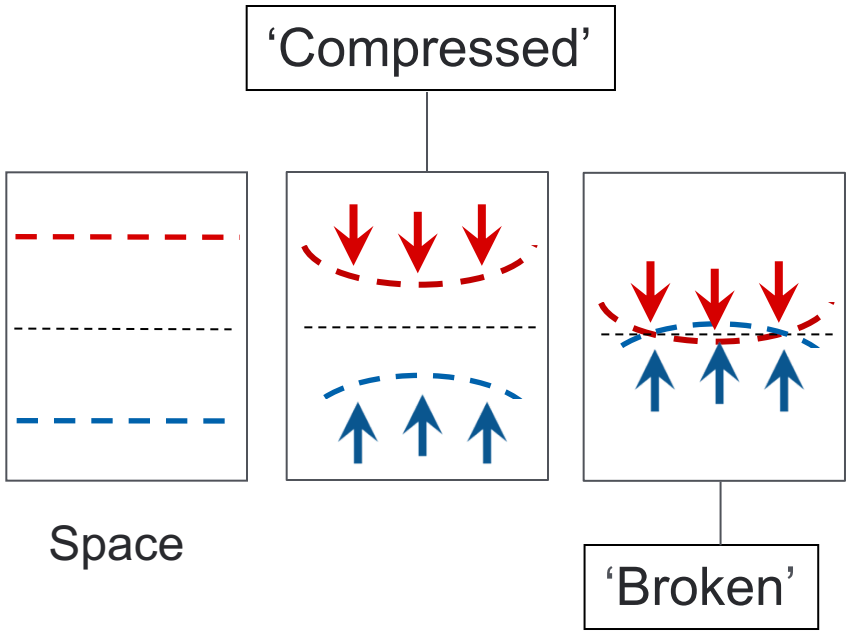
Svarverud et al (2012) *PLoS ONE*

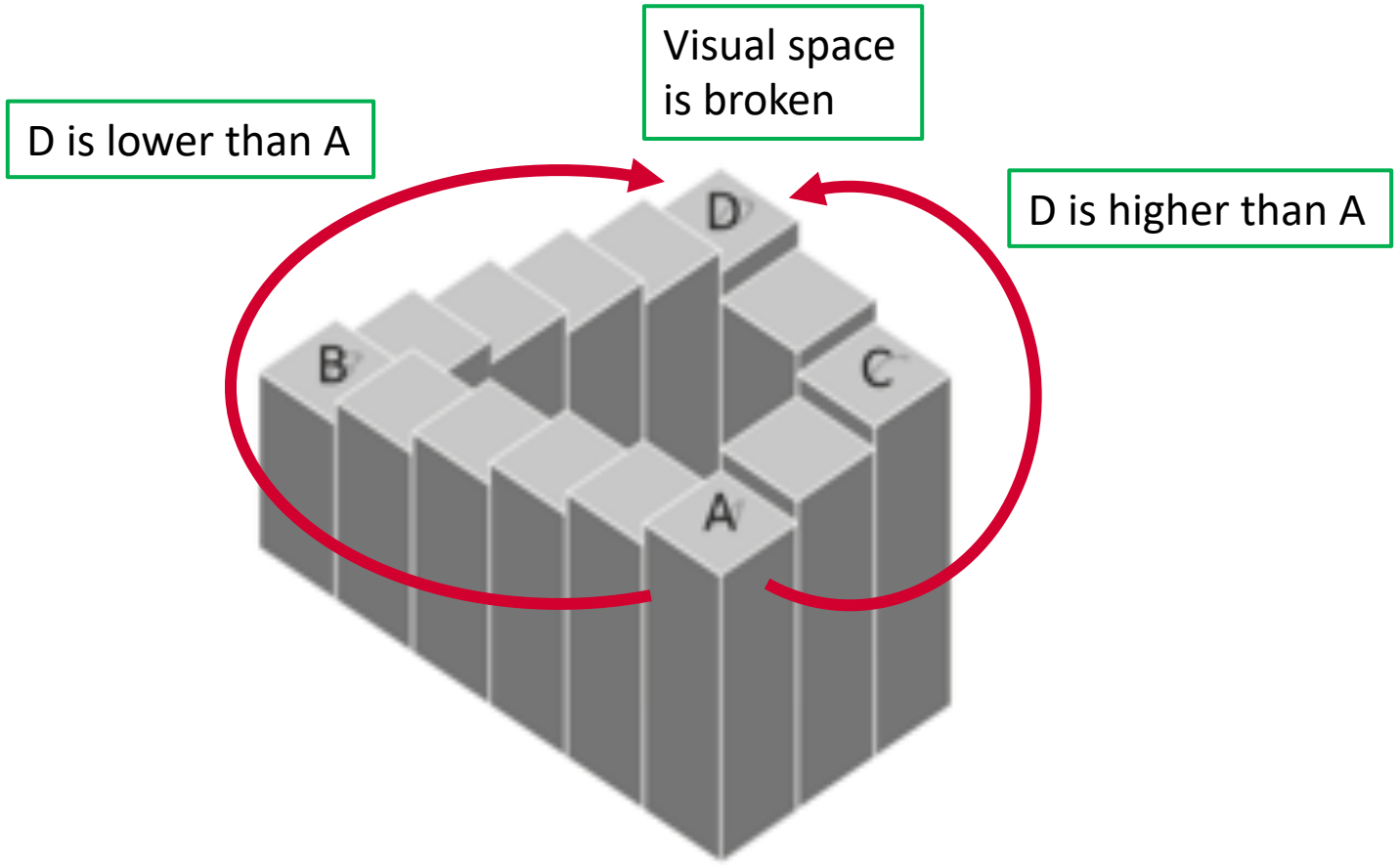
Relative depth judgements



Svarverud et al (2012) *PLoS ONE*

Relative depth judgements

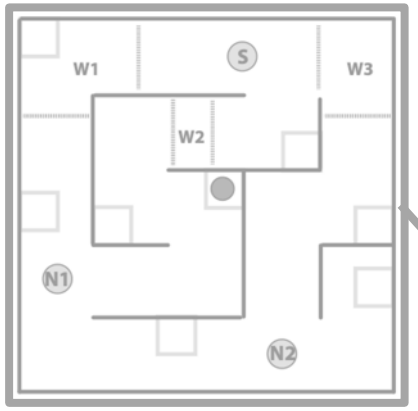




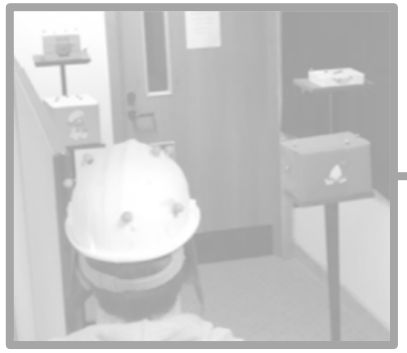
Svarverud et al (2012) *PLoS ONE*

Three experiments:

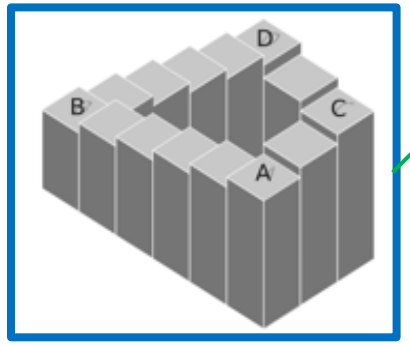
Alex Muryy



Jenny Vuong



Ellen Svarverud



Are there 3D representations in the brain?

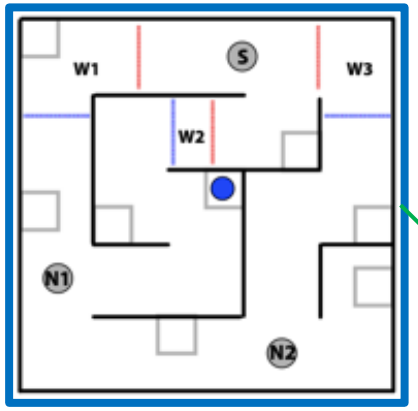
Criterion:
Can a single 3D representation explain performance across multiple tasks?

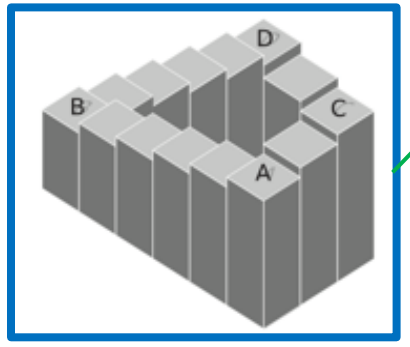
No

‘Neural rendering’ without a 3D reconstruction

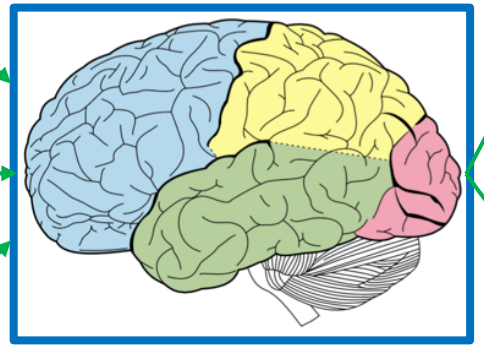
Eslami et al (2018) https://www.youtube.com/watch?v=G-kWNQJ4idw

Take home message:





Are there 3D representations in the brain?



Only possible to test alternative hypotheses if the observer moves: i.e. VR is essential

Thanks…

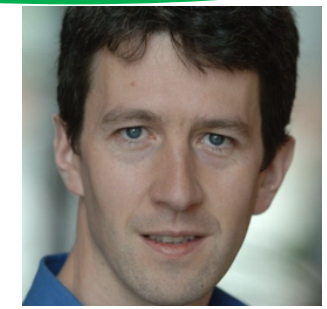
Alex Muryy

Jenny Vuong

Ellen Svarverud

Peter Scarfe

Lyndsey Pickup



Andrew Fitzgibbon

Stuart Gilson

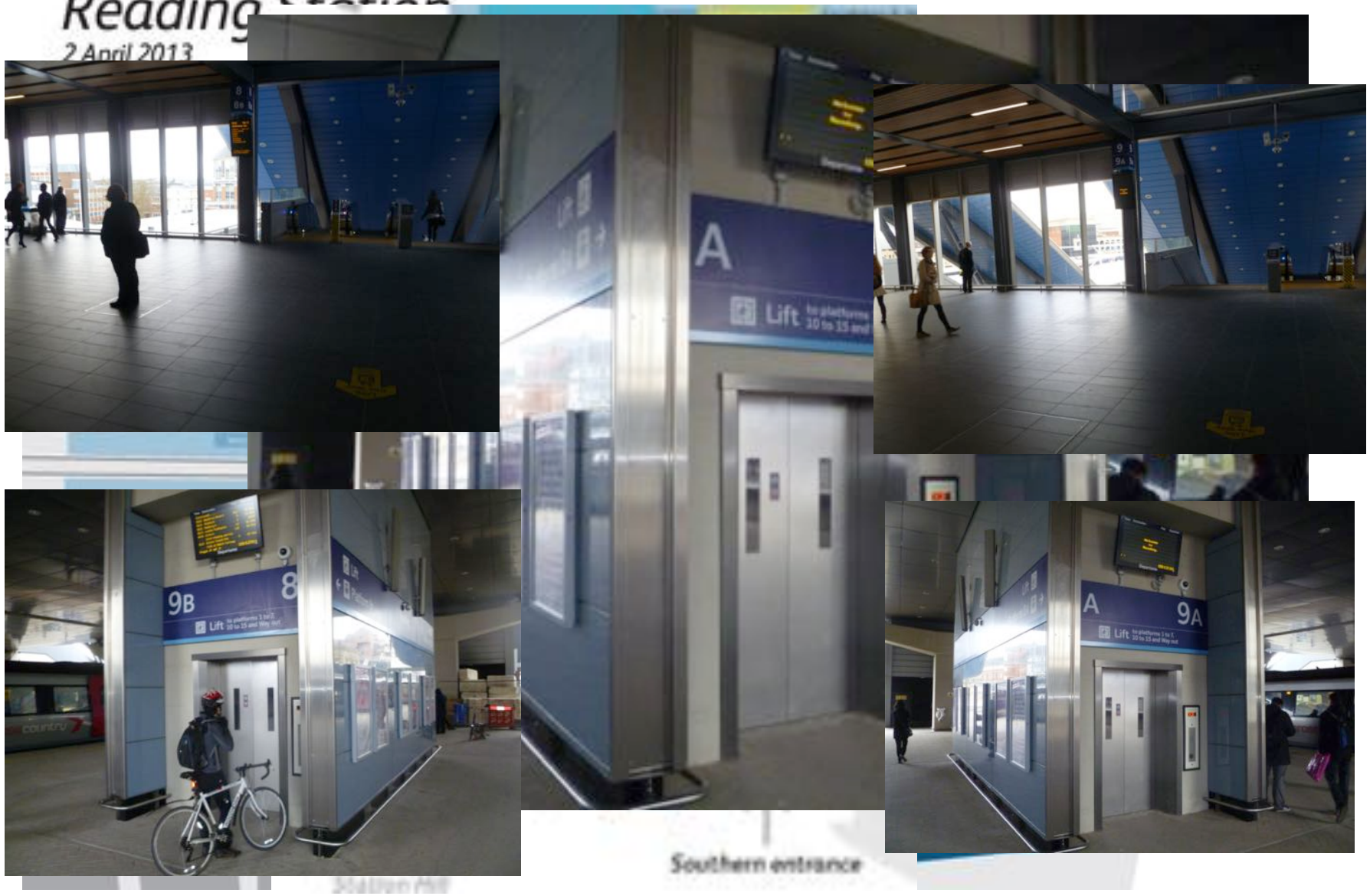
University of Reading

Microsoft Research

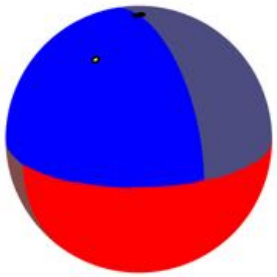
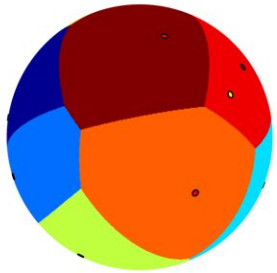
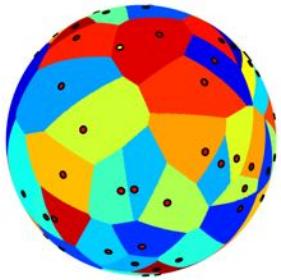
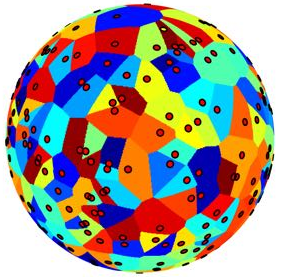
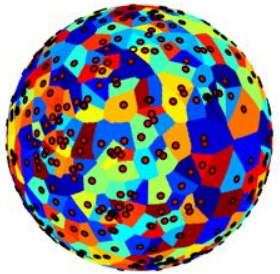
EPSRC

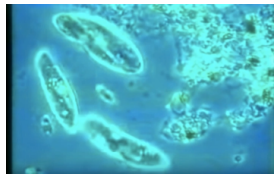
[dstl]

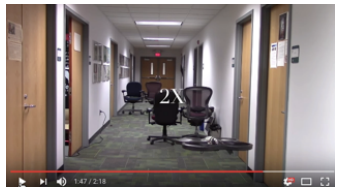
No need for VR....



Evolution of policy networks

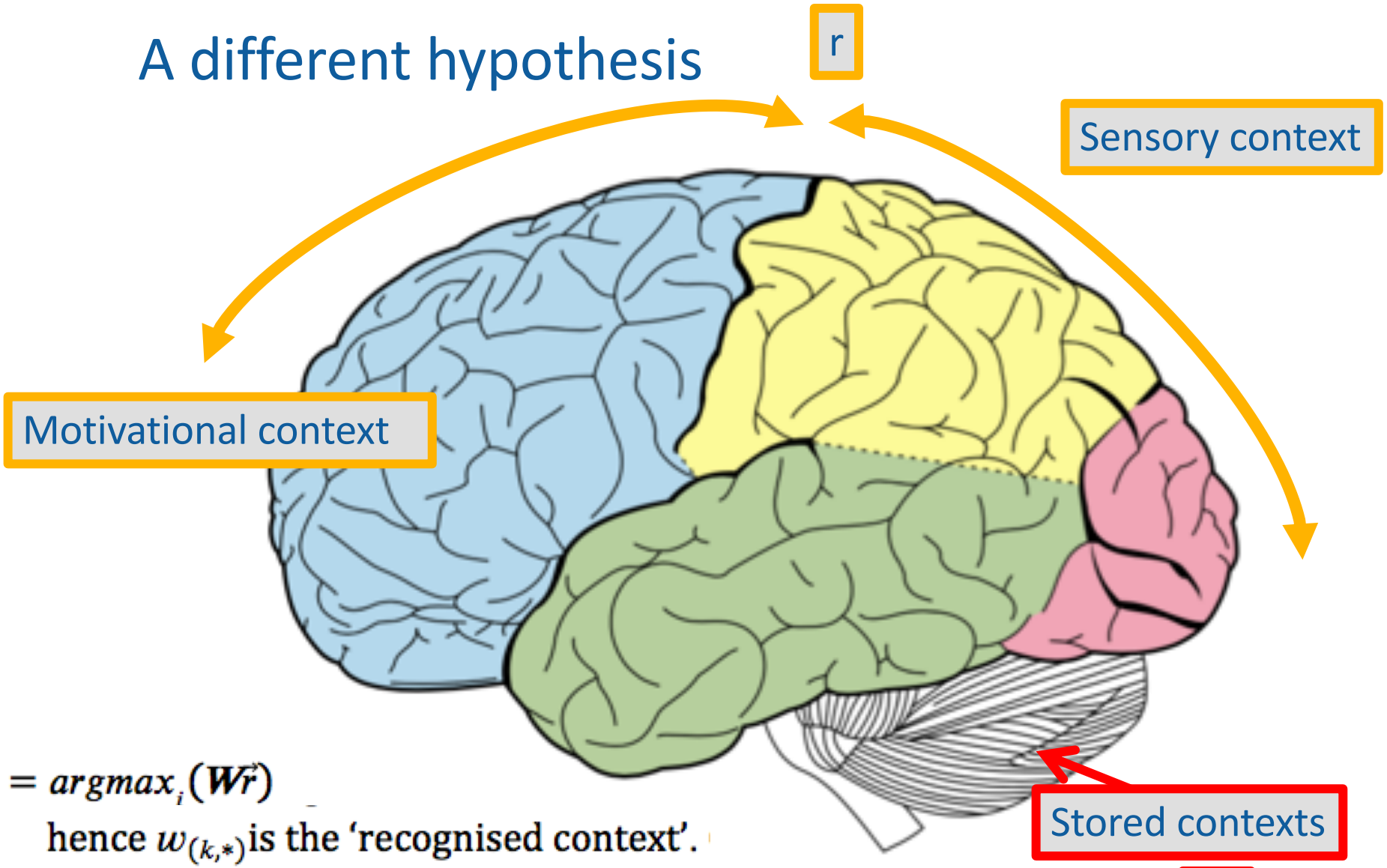




?

- bacterium
- plants
- RL
- humans
- higher intelligence

A different hypothesis



$k = argmax_i(\vec{Wr})$

hence $w_{(k,*)}$ is the 'recognised context'.