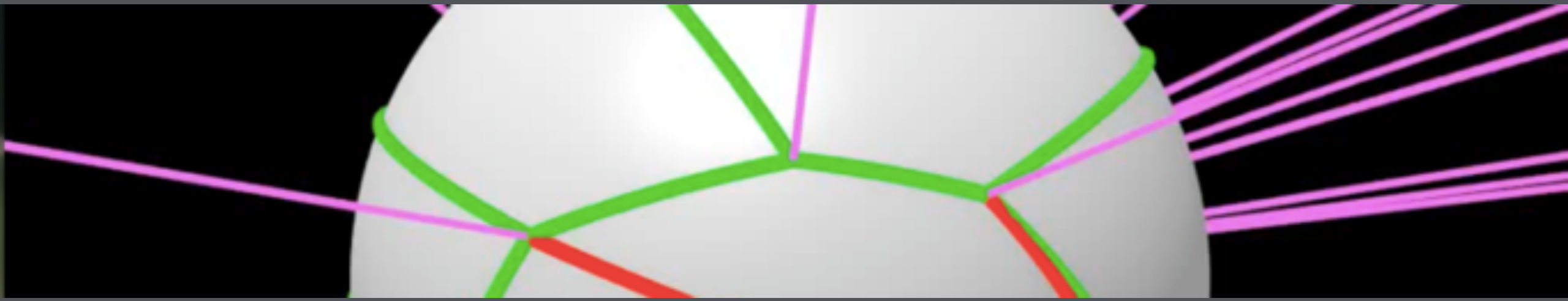
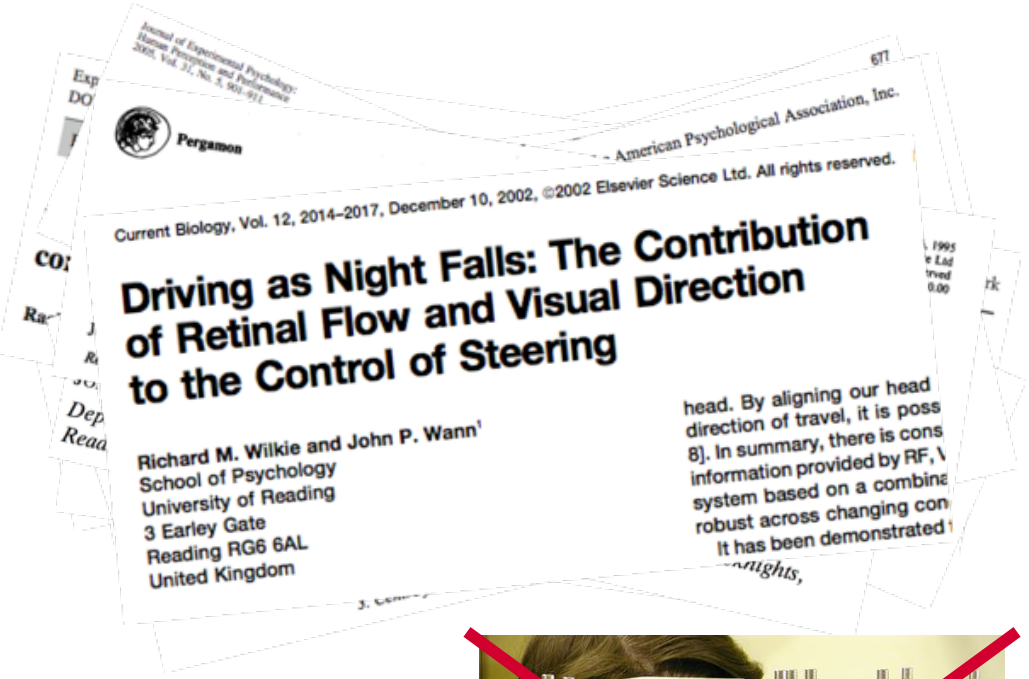
School of Psychology and
Clinical Language Sciences

University of Reading

TWO SPHERES AND 3D VISION



Andrew Glennerster

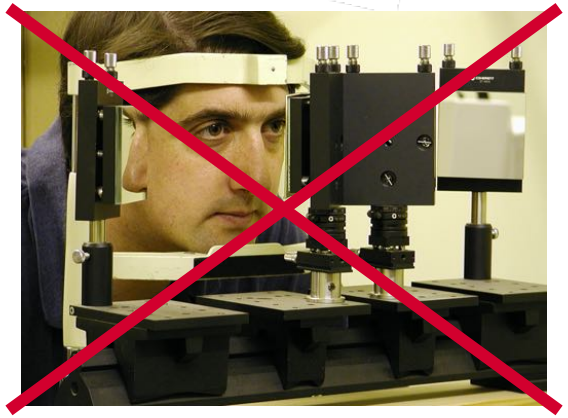


Pergamon

Current Biology, Vol. 12, 2014–2017, December 10, 2002, ©2002 Elsevier Science Ltd. All rights reserved.

Driving as Night Falls: The Contribution of Retinal Flow and Visual Direction to the Control of Steering

Richard M. Wilkie and John P. Wann[1]
School of Psychology
University of Reading
3 Earley Gate
Reading RG6 6AL
United Kingdom



University of Reading

Jenny Vuong

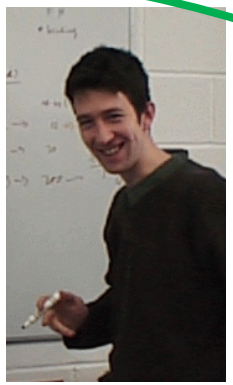
Alex Muryy

Luise Gootjes-Dreesbach

Peter Scarfe

James Stazicker

Miles Hansard



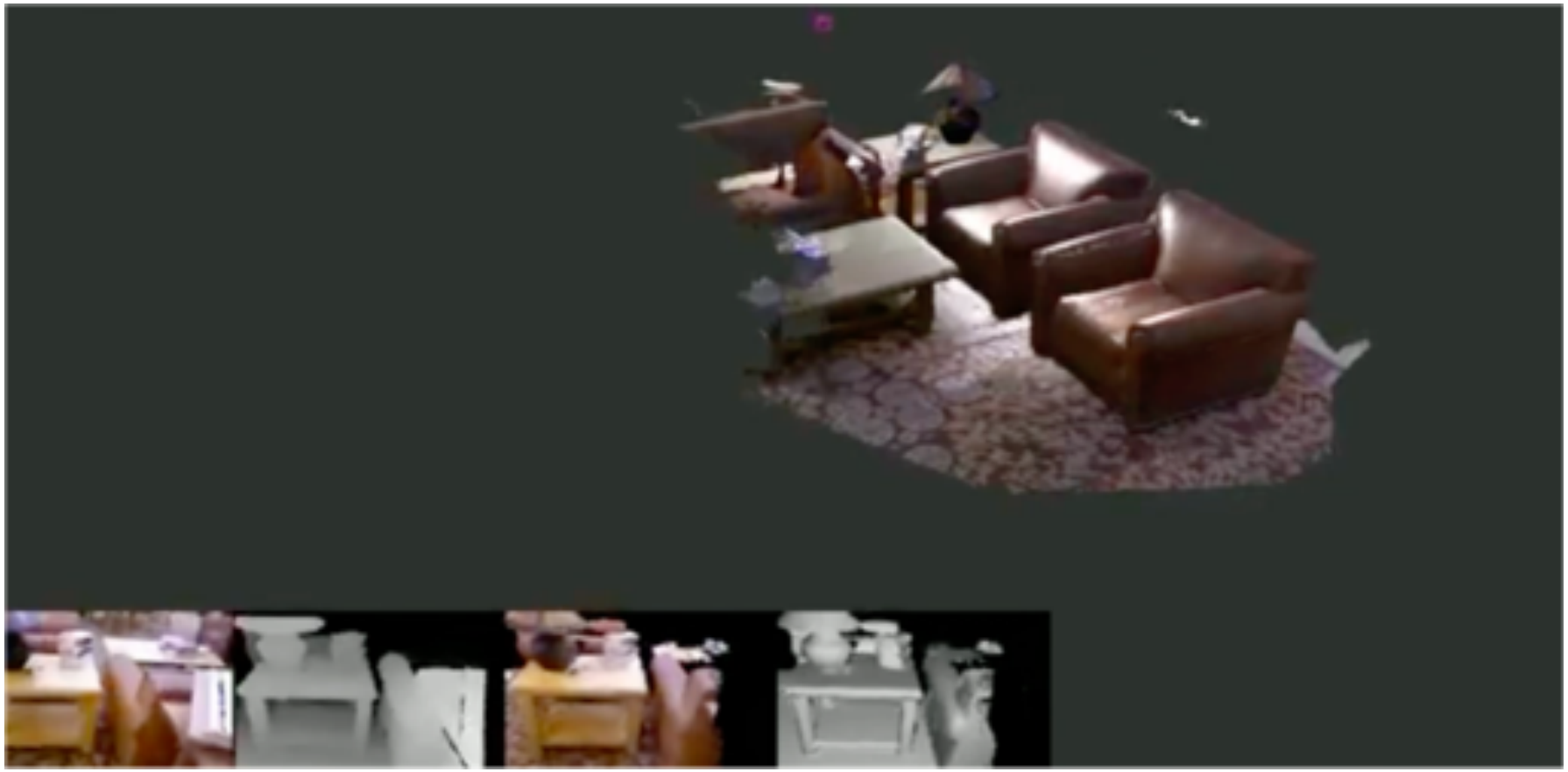
Andrew Fitzgibbon

Microsoft Research

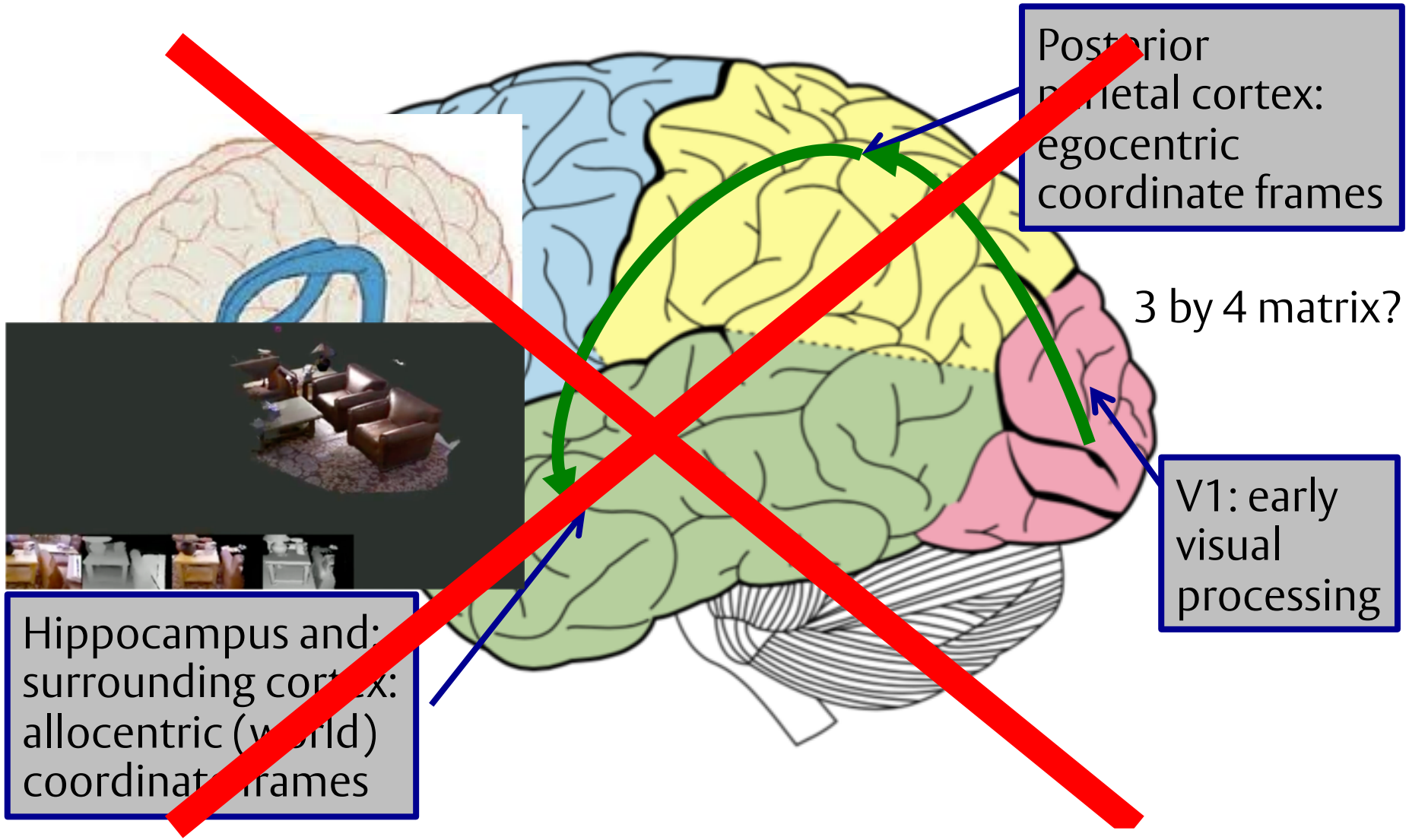
EPSRC

[dstl]

University of Reading



Current hypothesis

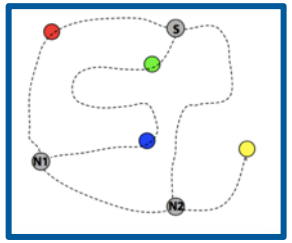


Outline

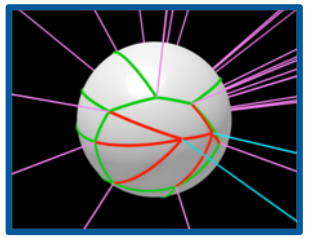
- Updating visual direction
  - some evidence and a ‘model’

A task that would seem to require the brain to build a 3D model

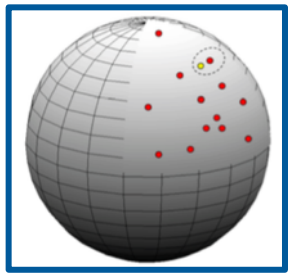
- Navigating through wormholes
  - a 3D model is not the best explanation



- A sphere of visual directions
  - information about viewing distance
  - A 2½ -D sketch



- A sphere of sensory+motivational states
  - a gradual increase in dimensionality

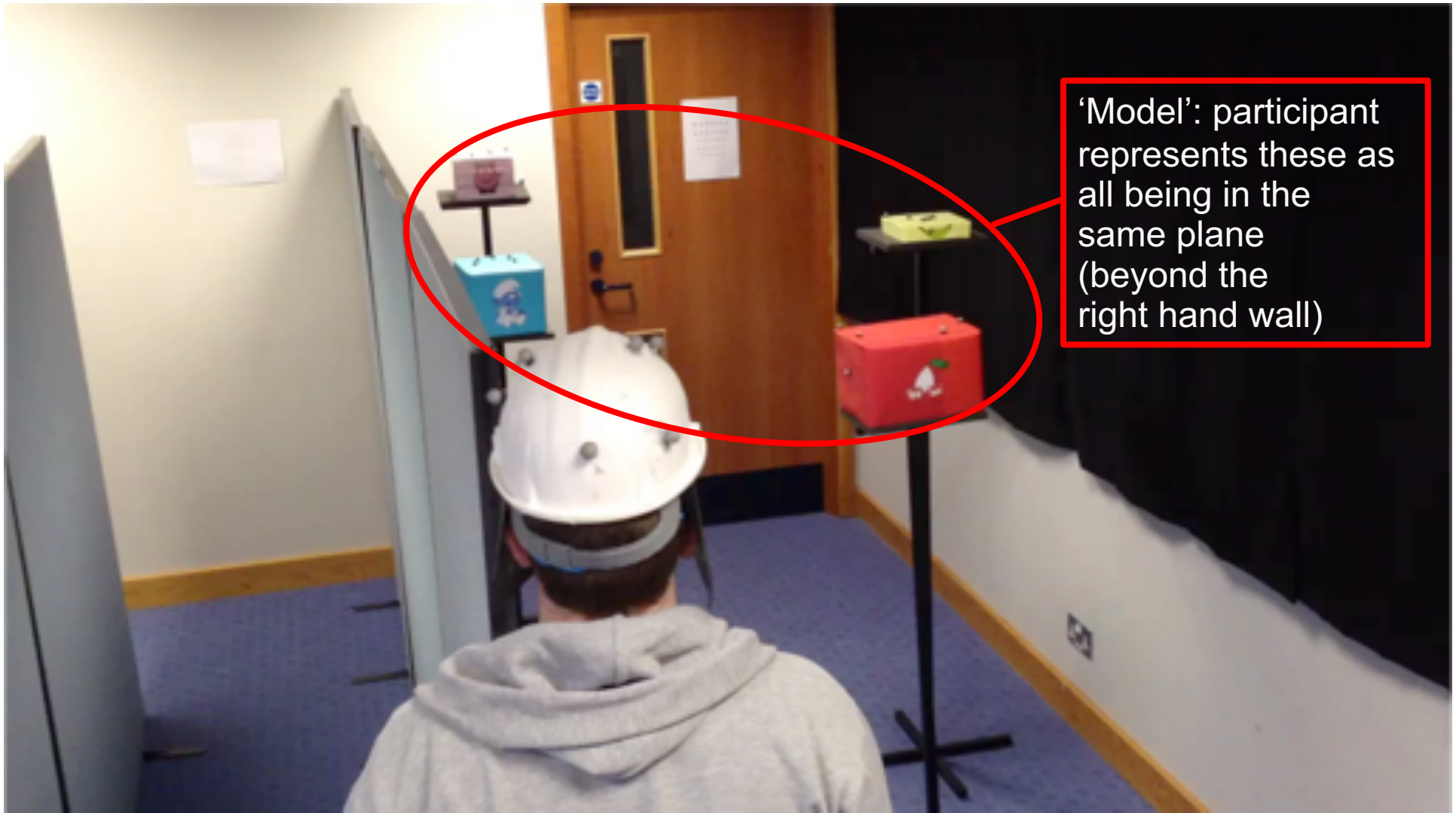


Can we update the visual direction of unseen objects as we move?

Jenny Vuong

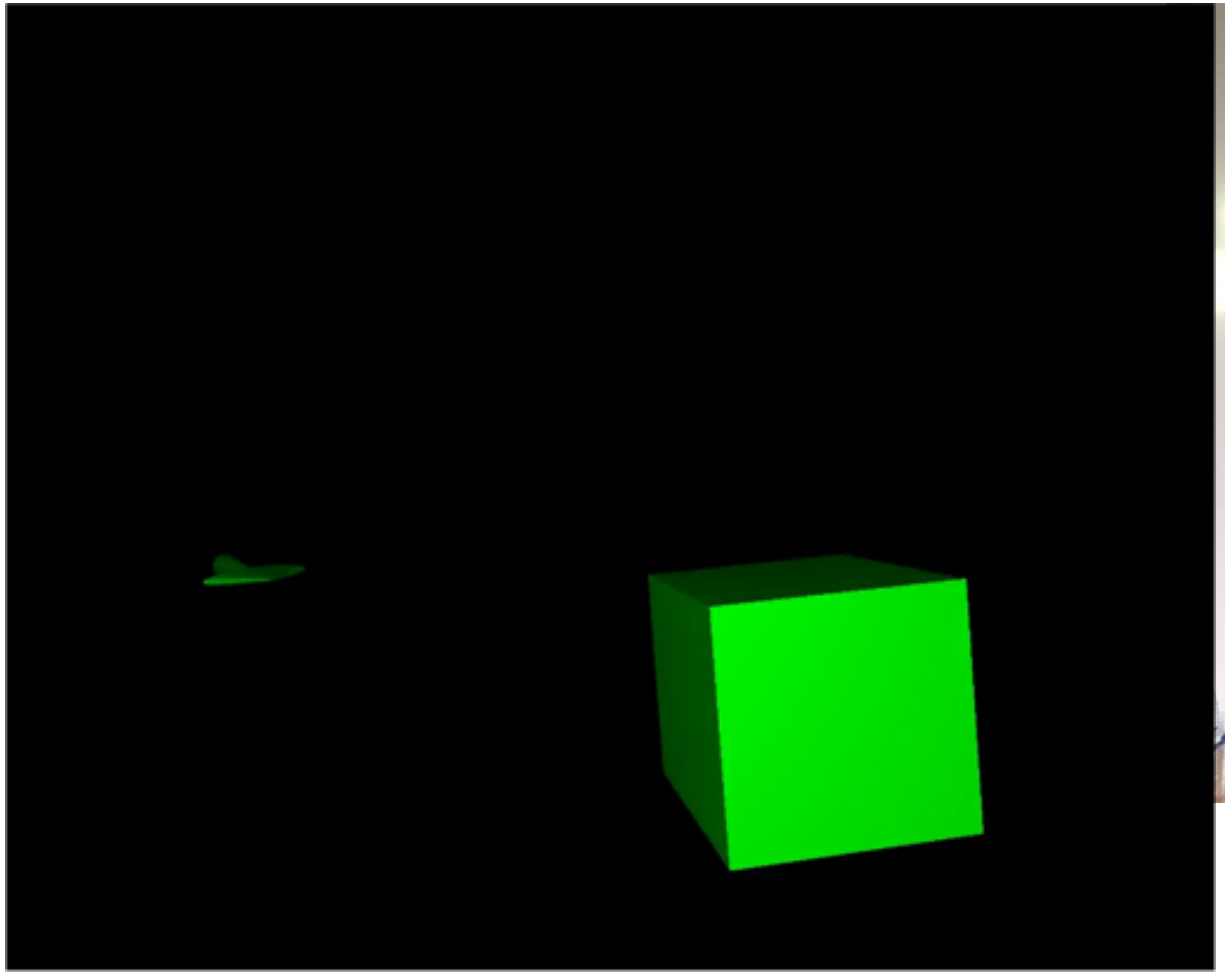
Can we update the visual direction of unseen objects as we move?

Jenny Vuong



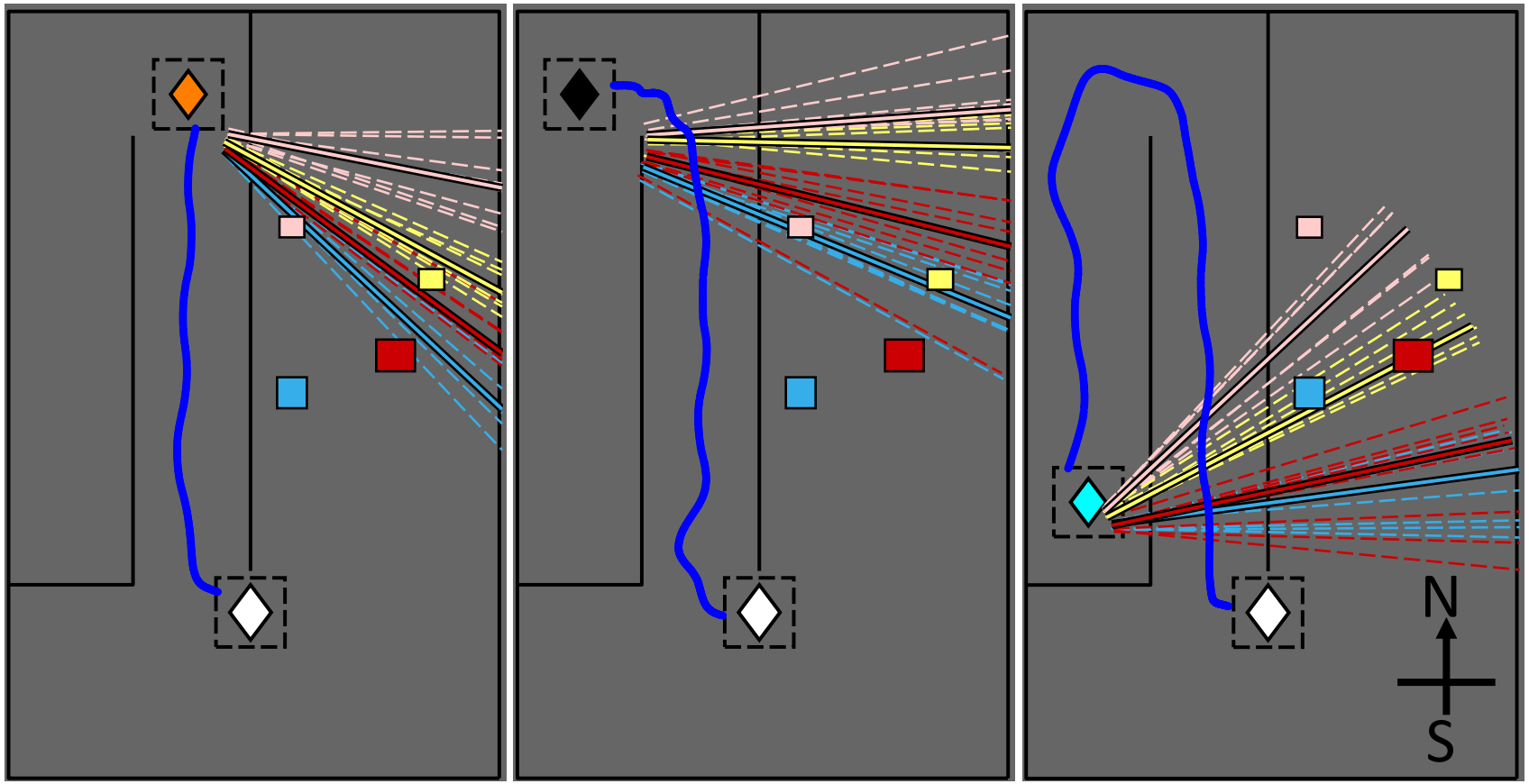
Can we update the visual direction of unseen objects as we move?

Jenny Vuong



nVis SX111 HMD
Vicon tracking

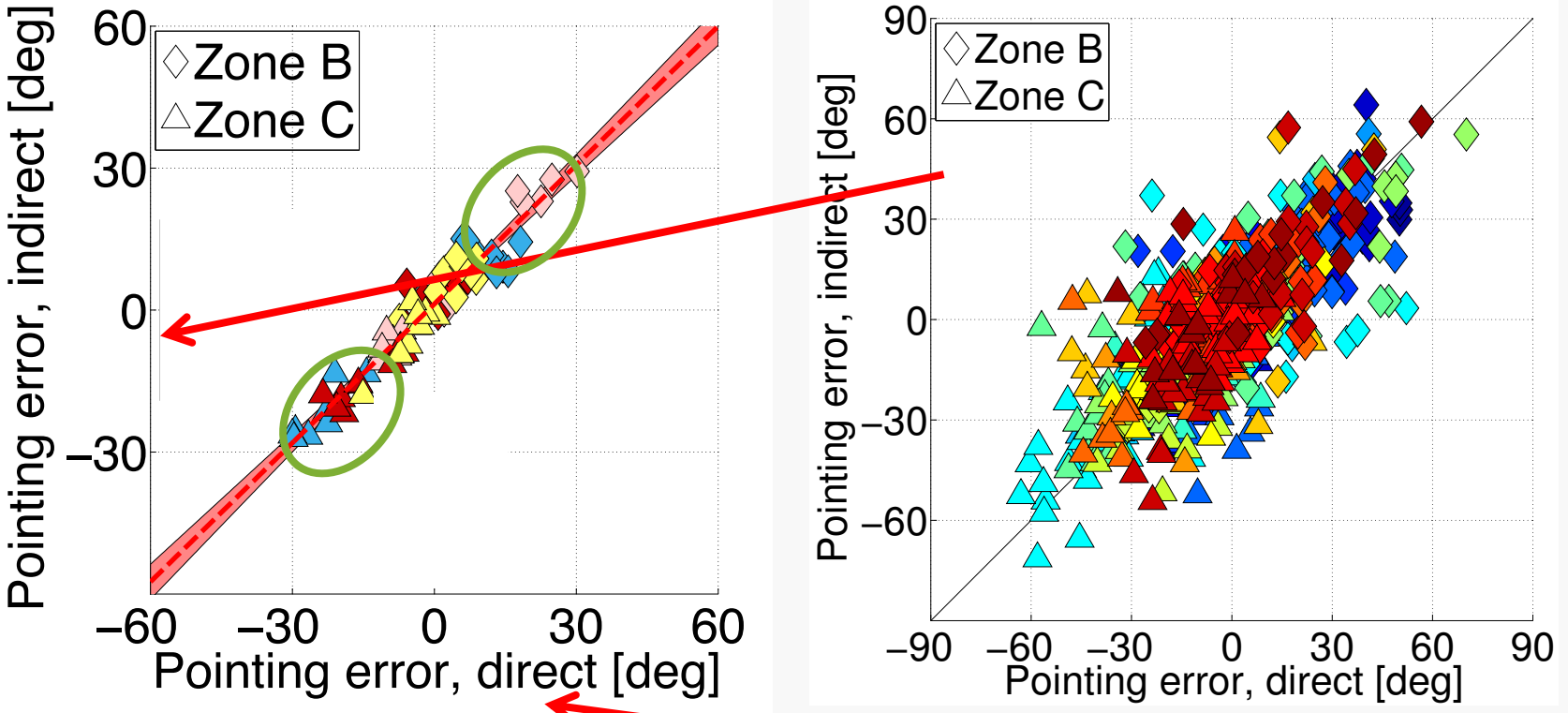
People show large, consistent biases



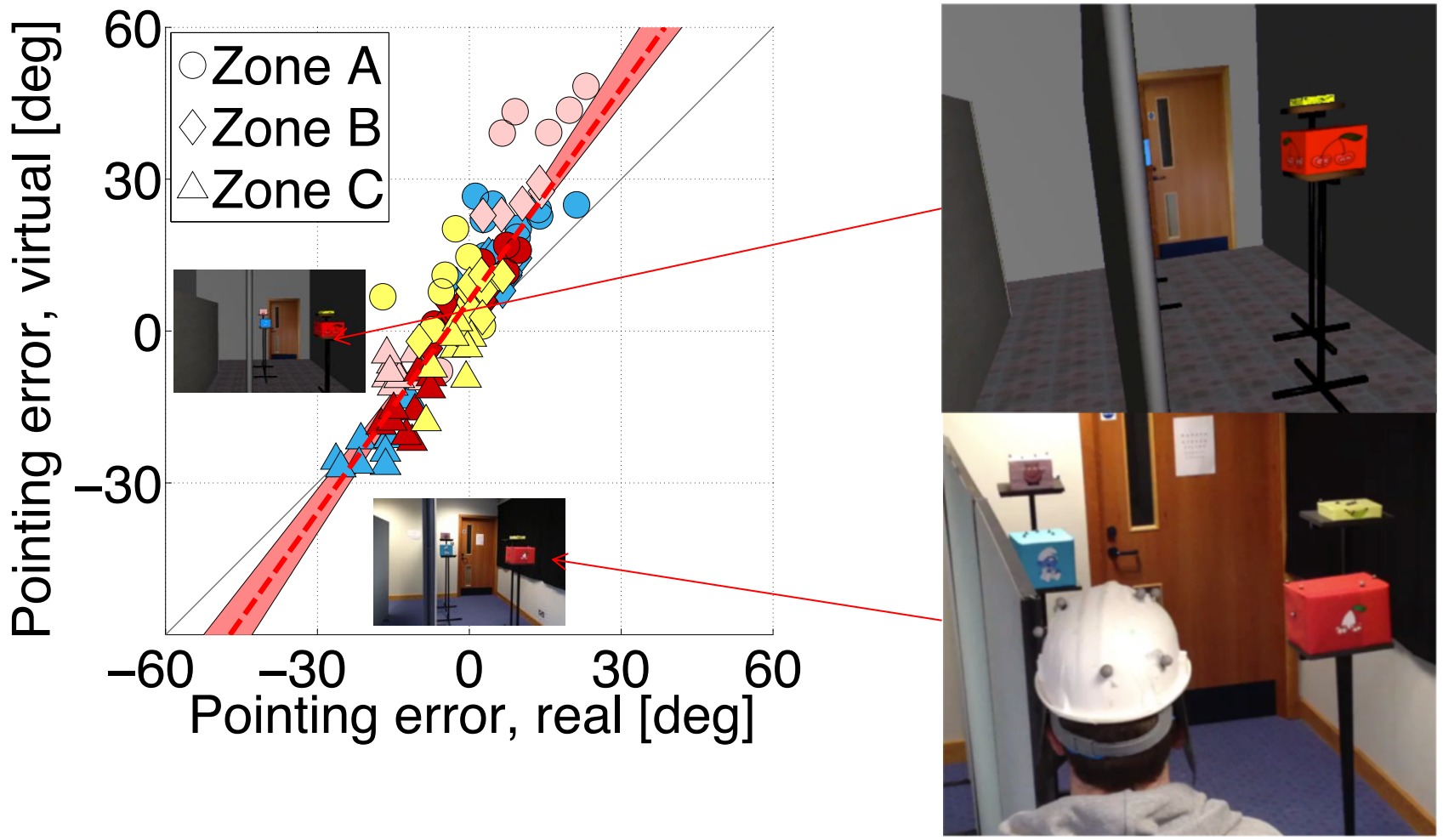
- Task:
  - view a scene
  - walk without any further view of the objects
  - point to the objects
  - easy to do if we update our location in a 3D reconstruction (SLAM)

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

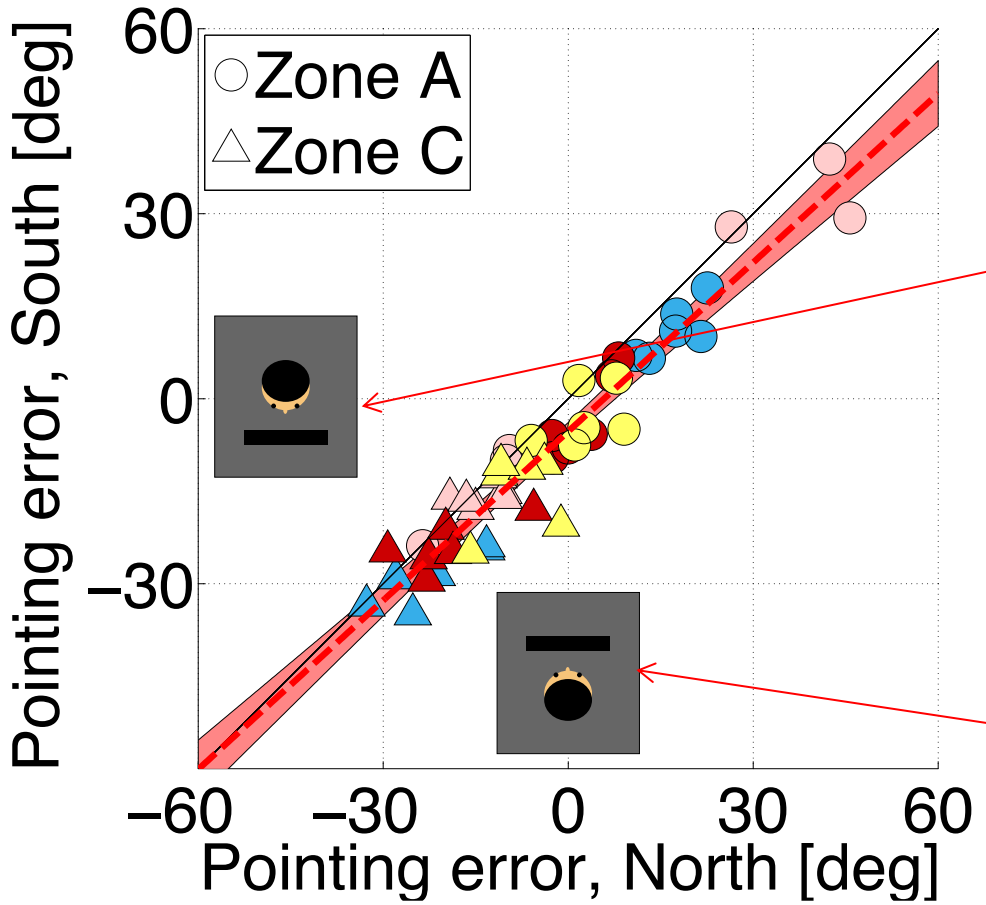
… independent of the route they take …



… similar biases in real and virtual worlds …



… whether looking ‘north’ or ‘south’ …

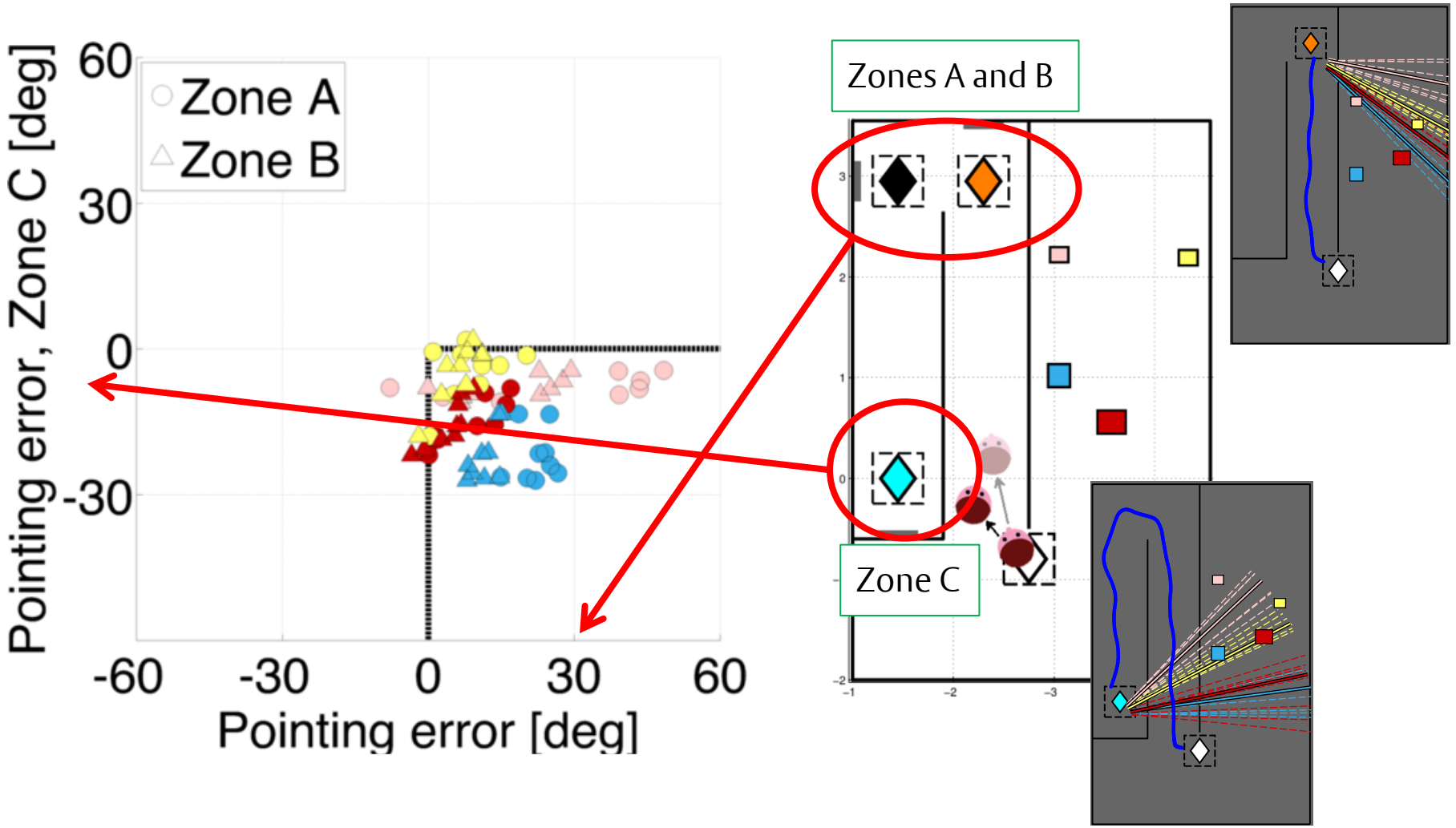


Facing ‘south’

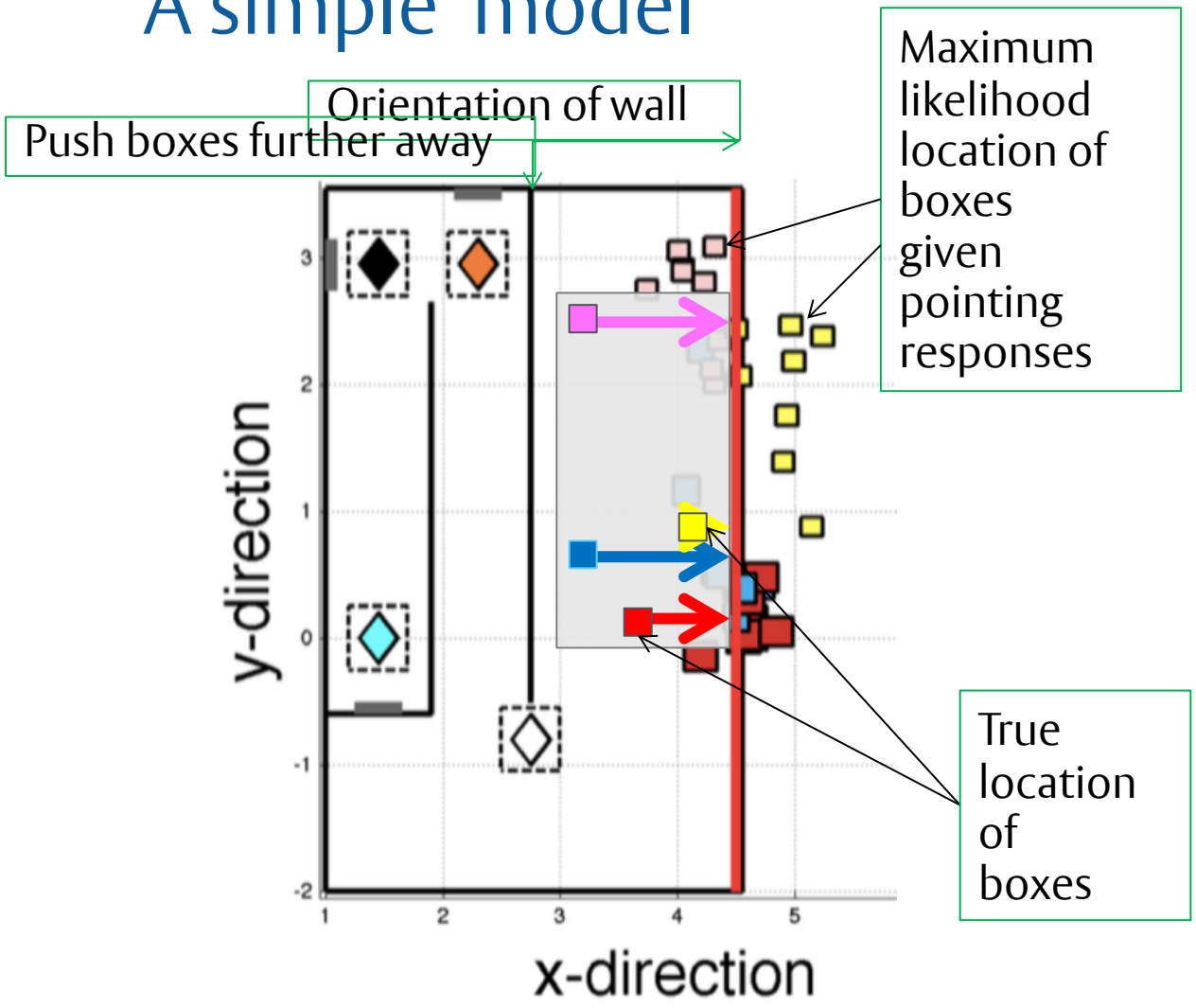
… just another example of pointing errors being correlated across conditions

Facing ‘north’

… but heavily dependent on pointing zone

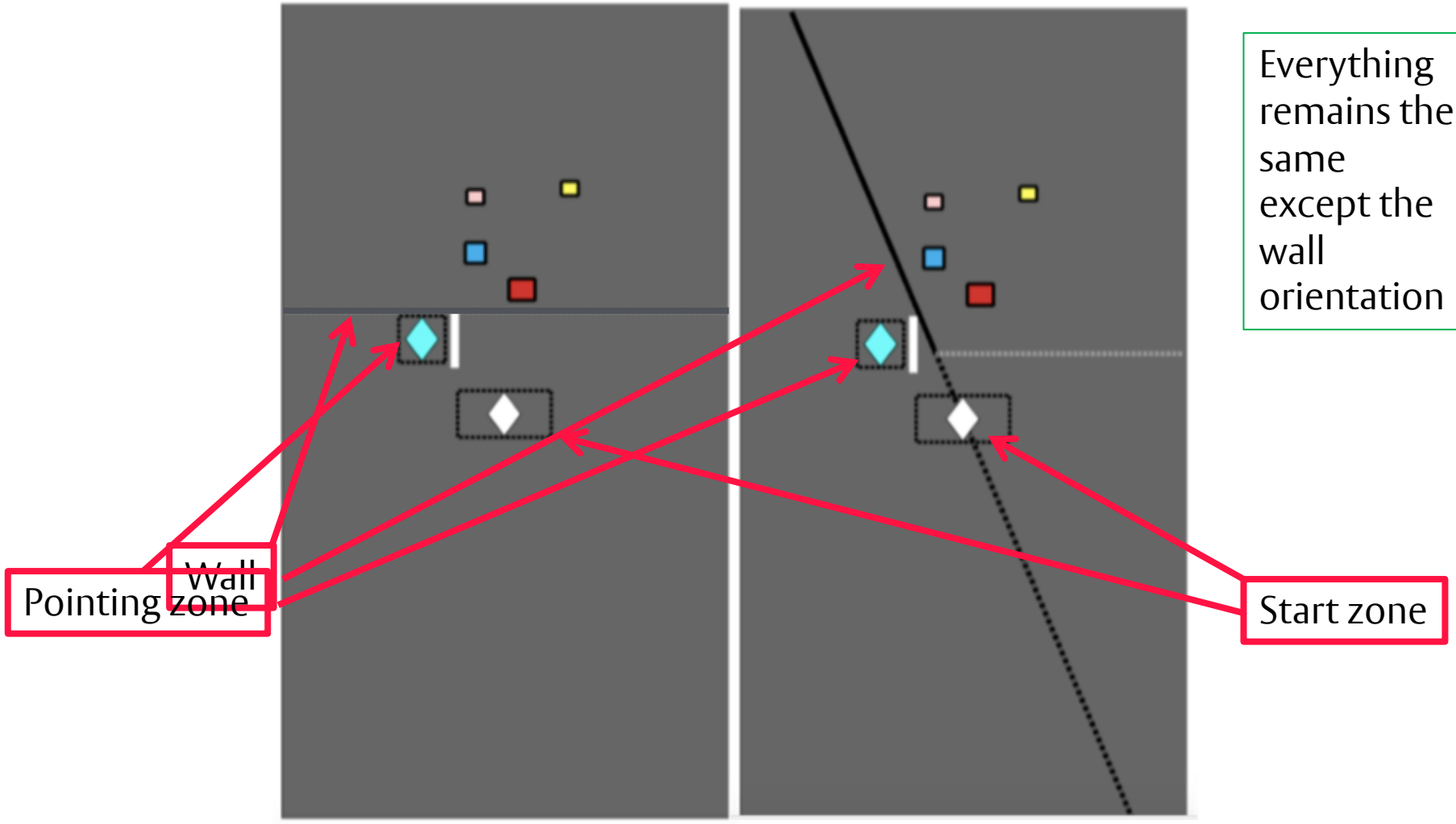


A simple 'model'



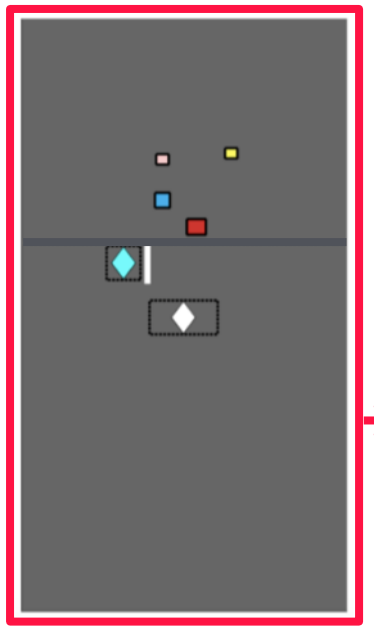
- Participants behave as if they ignore crucial aspects of the geometry of the scene
  - pointing responses suggest they assume objects lie in a plane (or something close to this)

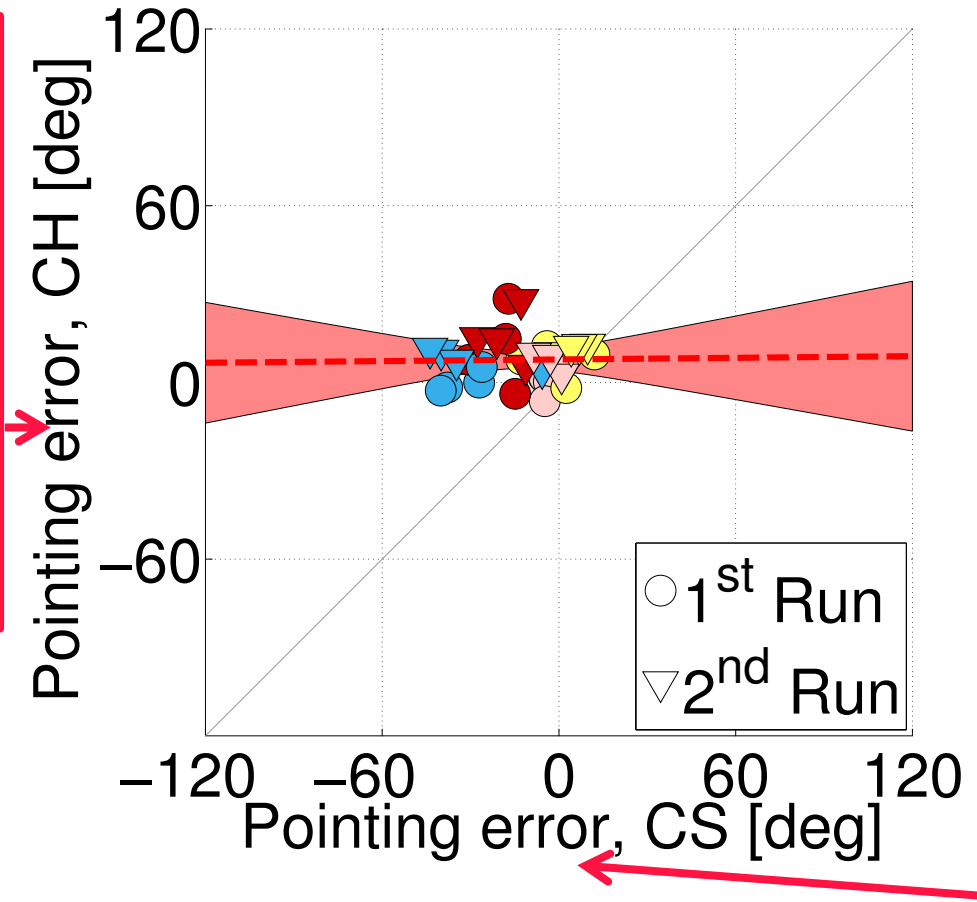
An effect of the wall orientation:



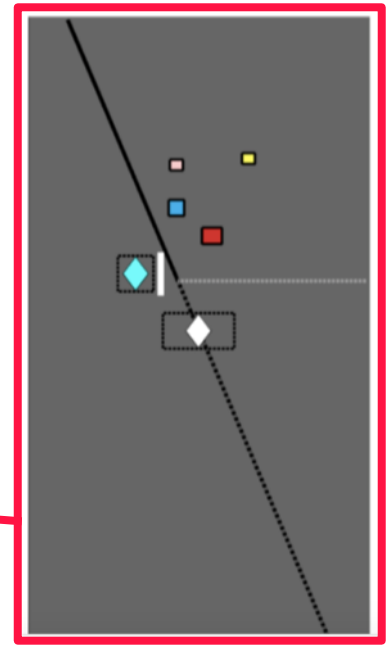
Everything remains the same except the wall orientation

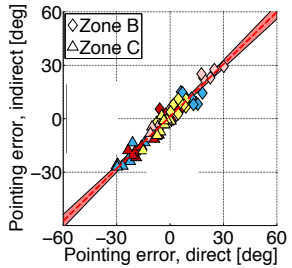
An effect of the wall orientation:



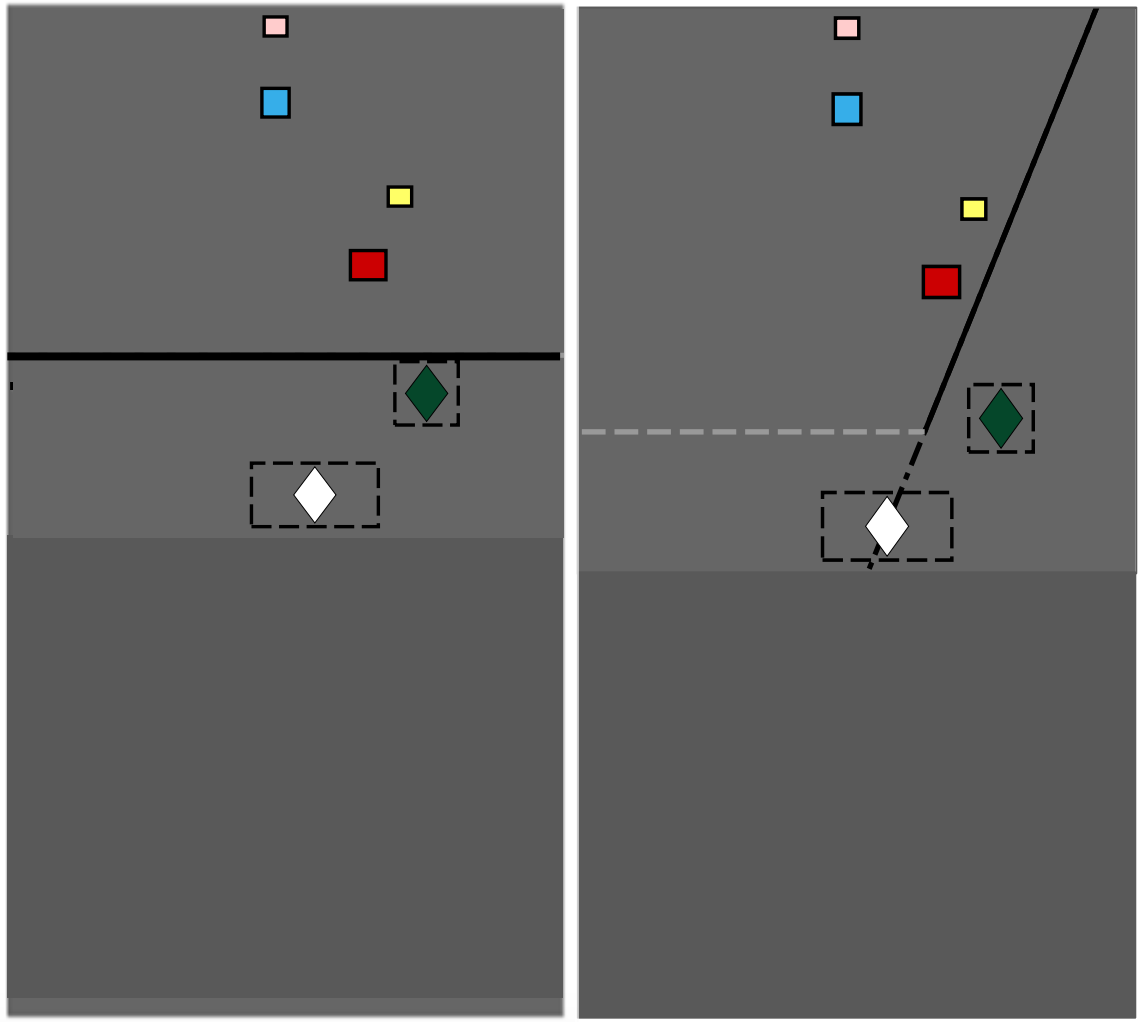


Everything remains the same except the wall orientation



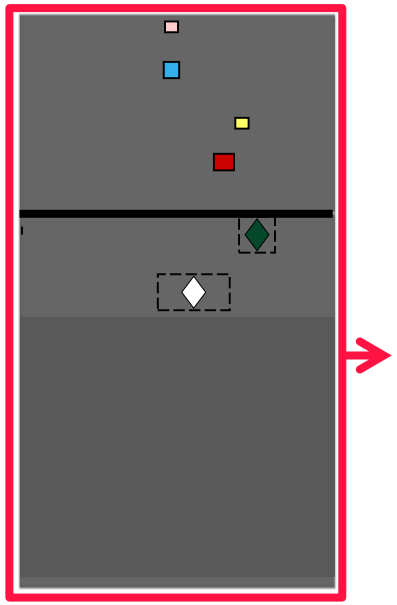


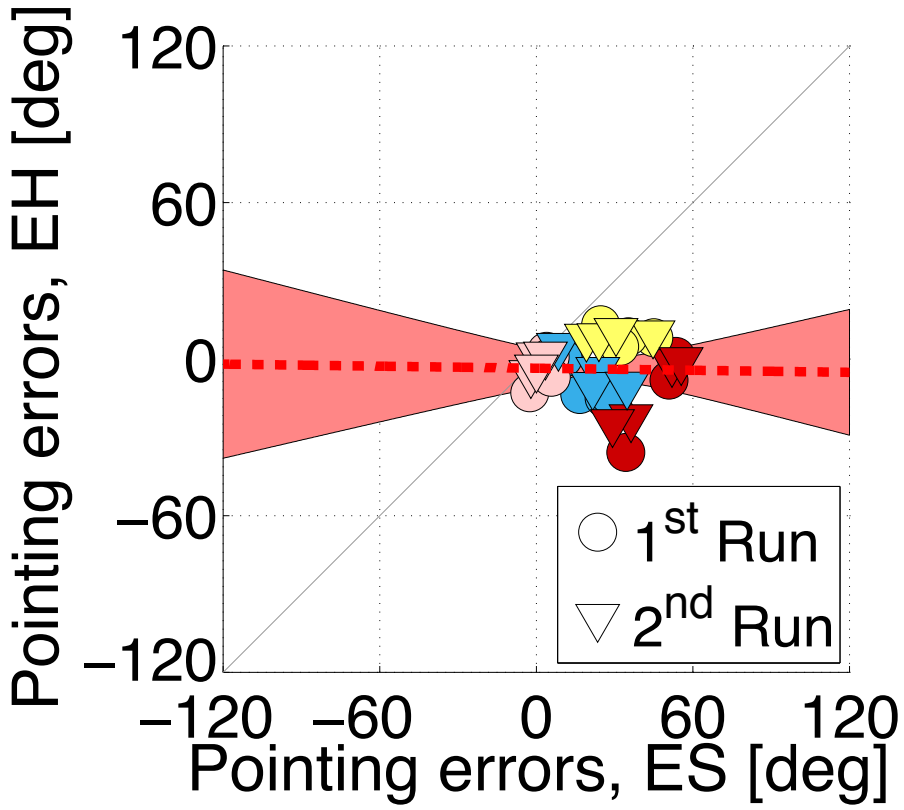
An effect of the wall orientation:

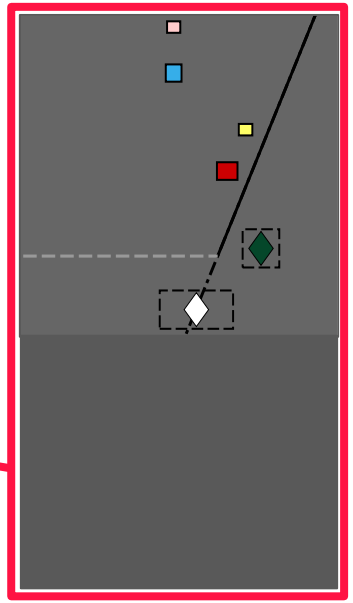


Also tested other viewing zones, other wall orientations

An effect of the wall orientation:

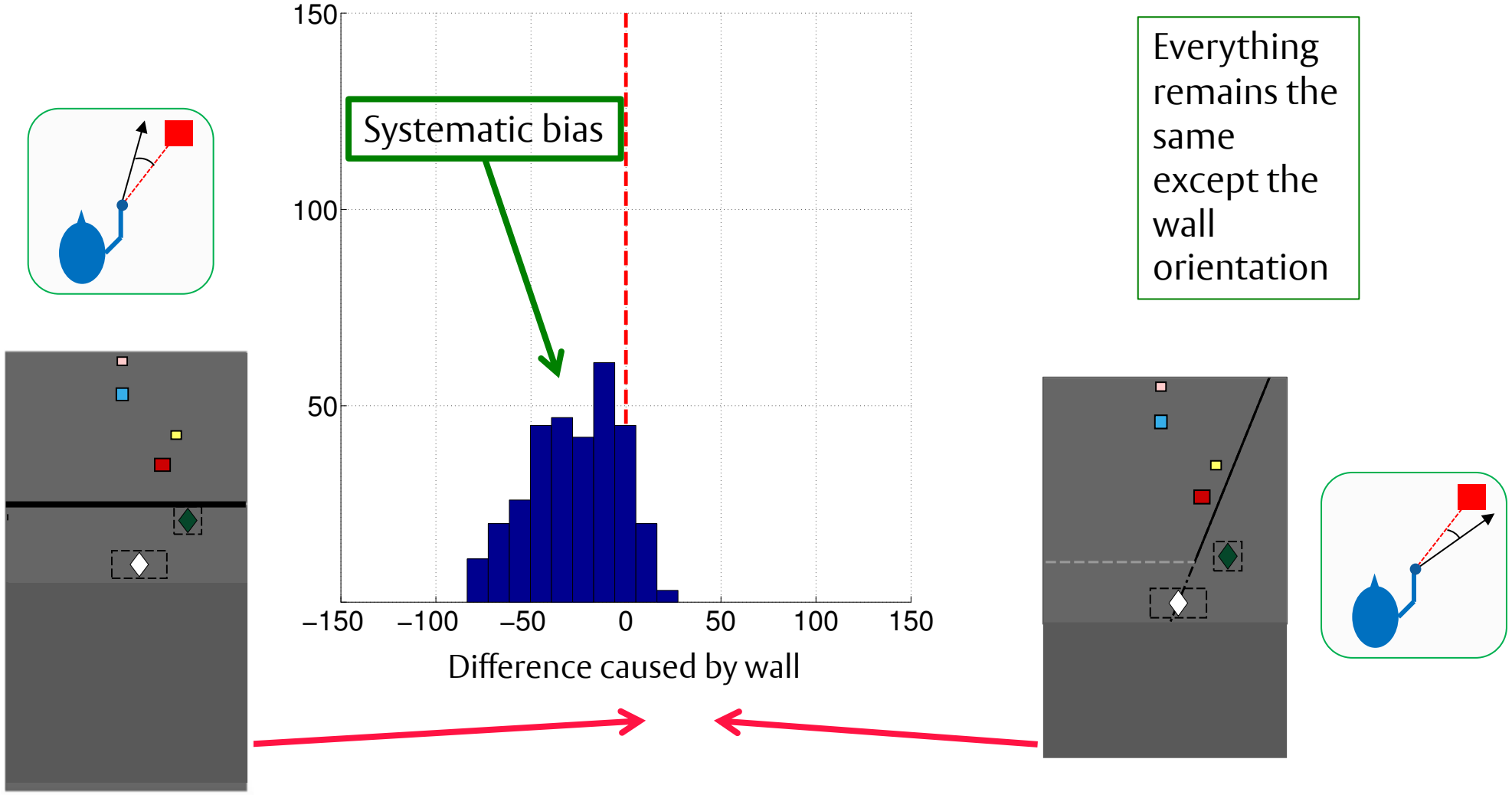




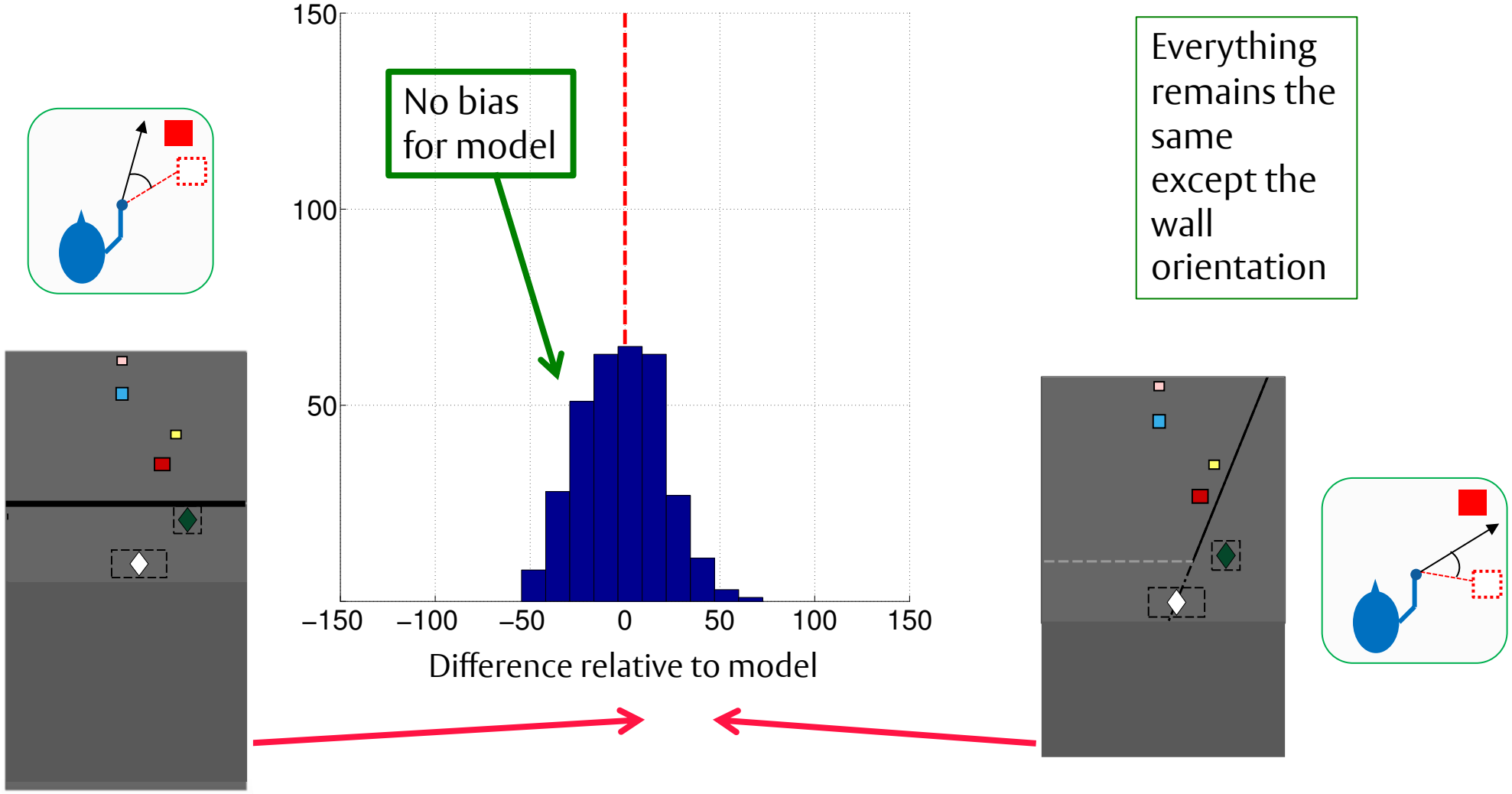


Everything remains the same except the wall orientation

An effect of the wall orientation:



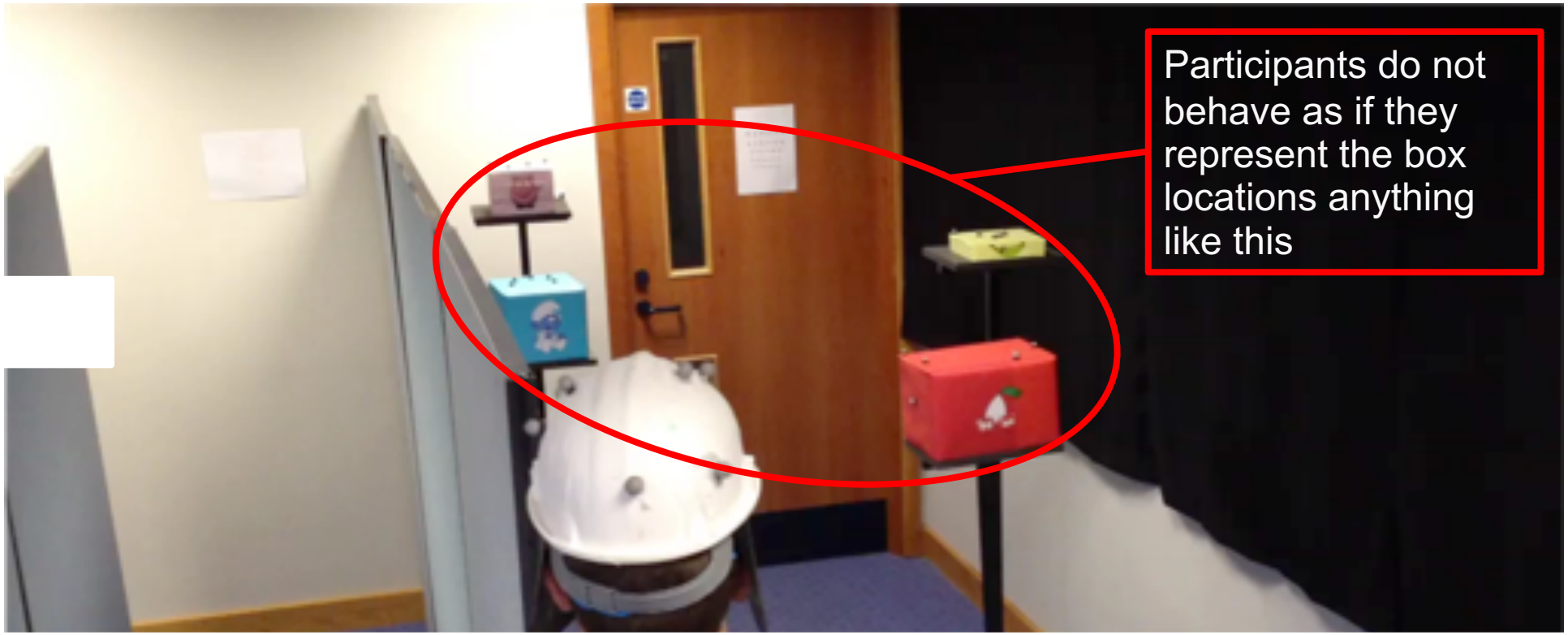
The model accounts for the effects of the wall



Q: Can we update the visual direction of unseen objects as we move?
A: not very well (we have poor heuristics for imagining)

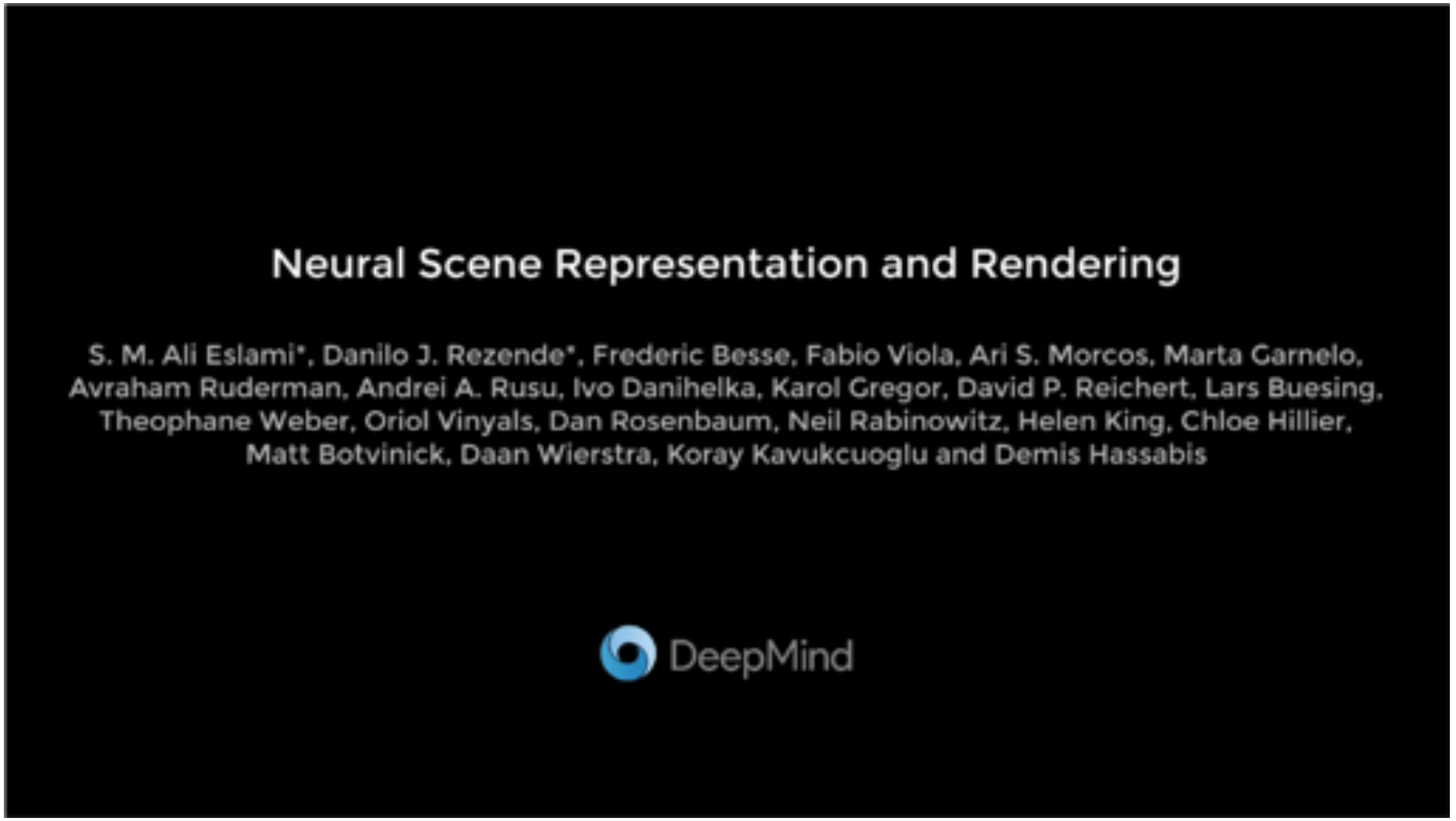
University of Reading

Jenny Vuong



Vuong, Fitzgibbon and Glennerster (bioRxiv, 2018)

‘Neural rendering’ without a 3D reconstruction



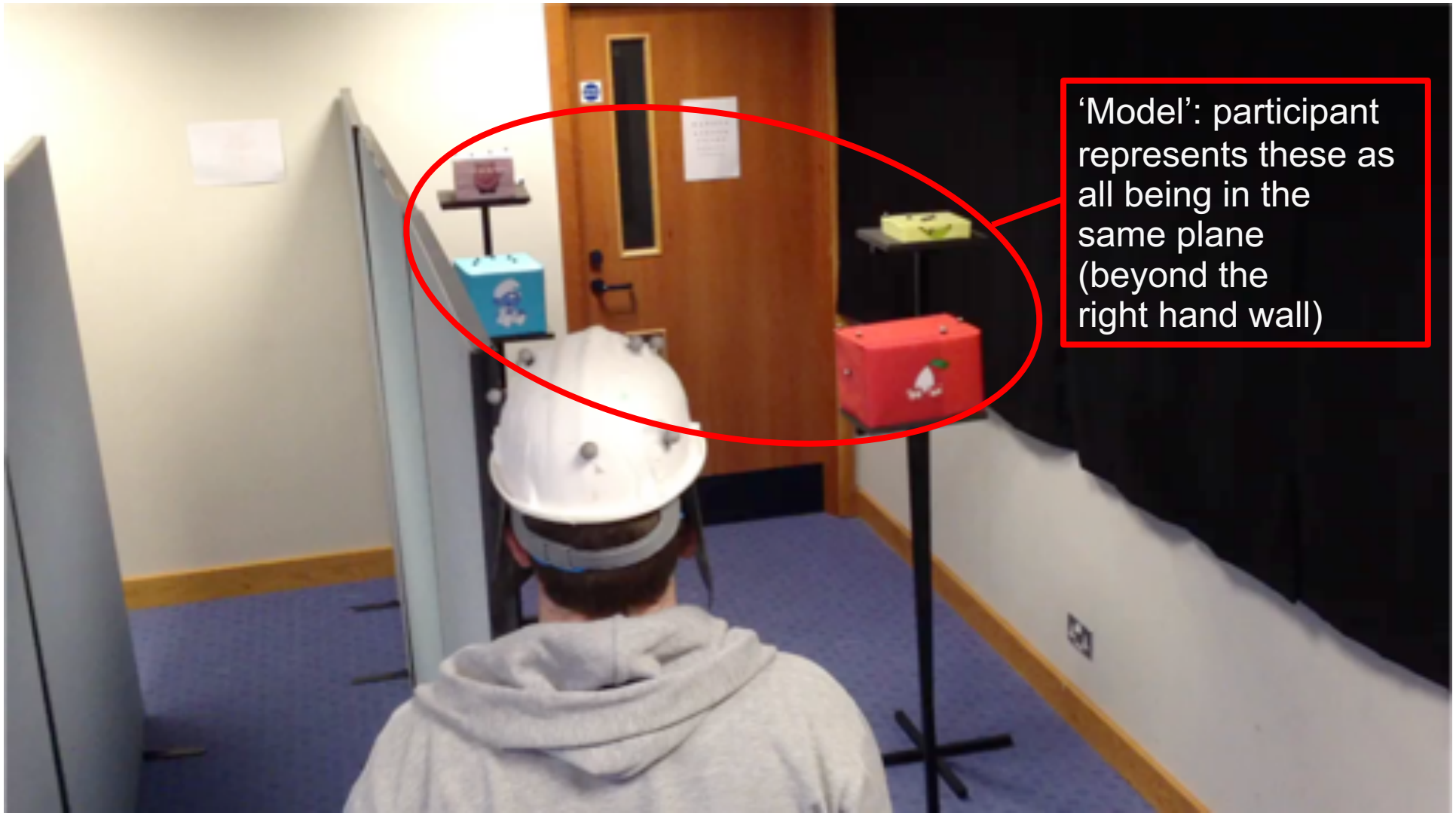
Eslami et al (2018) https://www.youtube.com/watch?v=G-kWNQJ4idw

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

Jenny Vuong

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

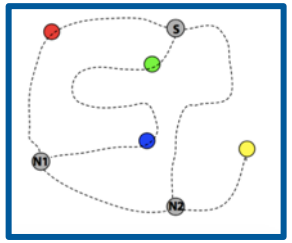
Jenny Vuong



Outline

- Updating visual direction
  - some evidence and a ‘model’

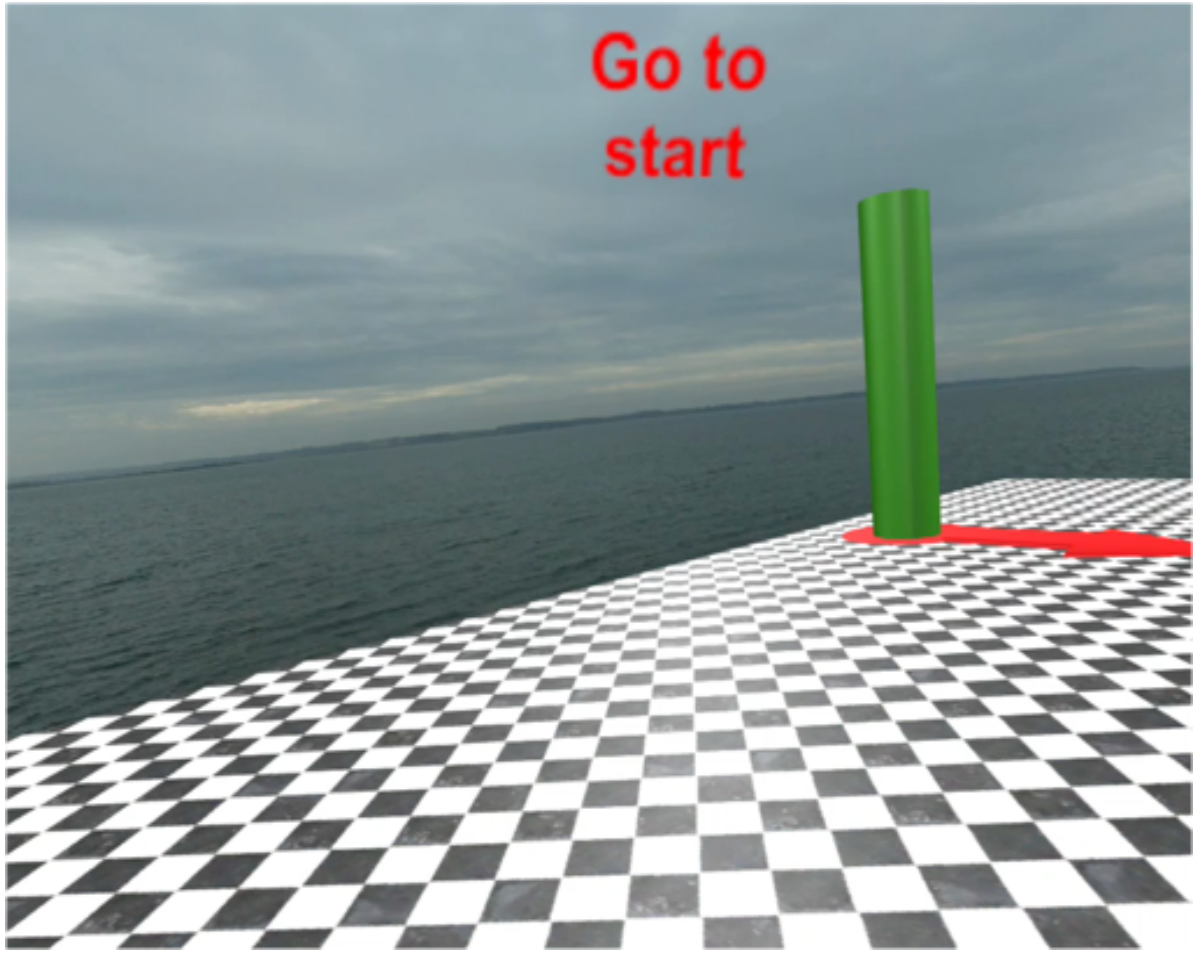
- Navigating through wormholes
  - a 3D model is not the best explanation



Alex Muryy

University of Reading

Learning to point to targets in a maze



Tasks:

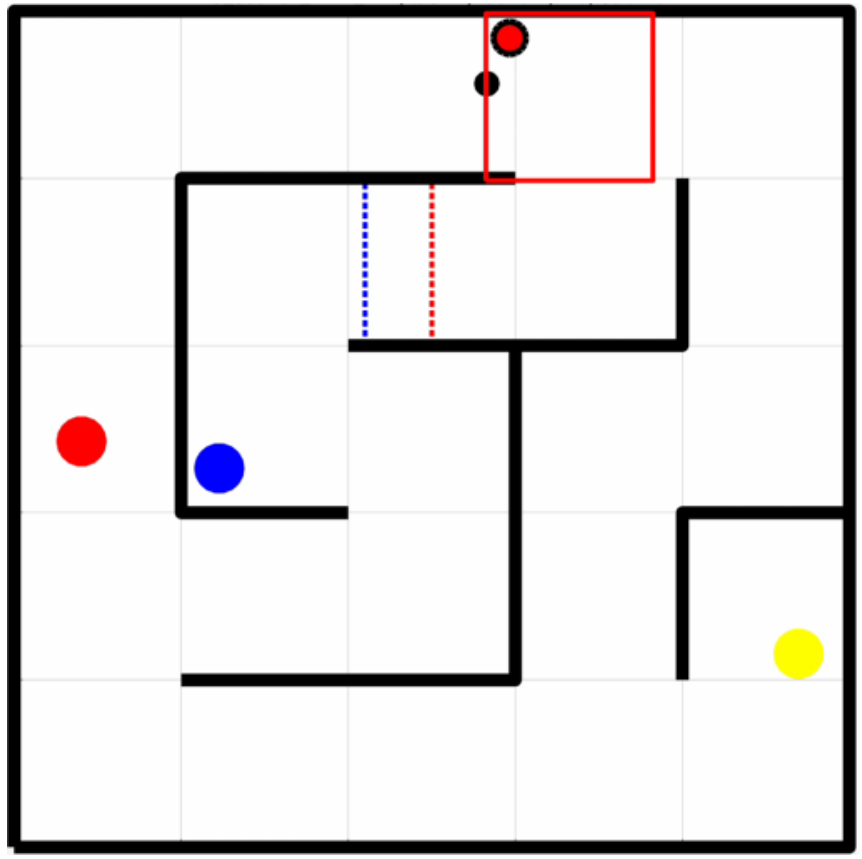
(i) find targets in specified order and

(ii) point to them…

Alex Muryy

University of Reading

Learning to point to targets in a maze



Life gets harder…

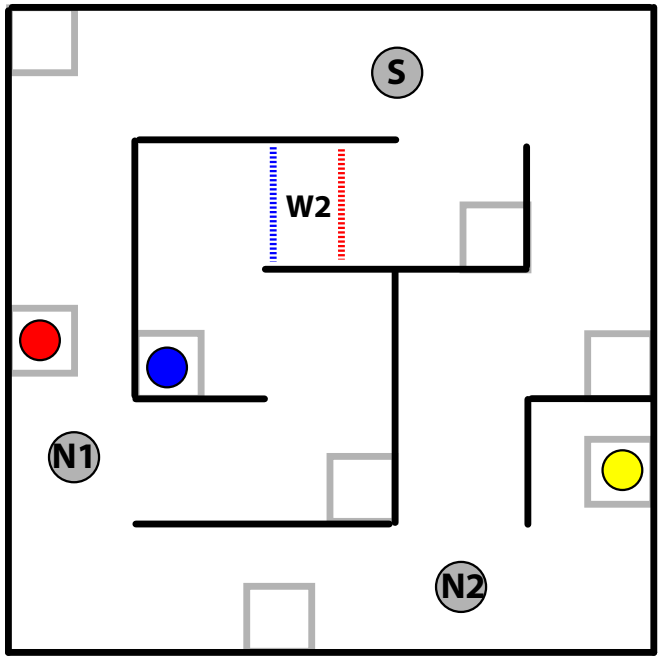
Learning phase (repeat x5):
a) Navigation: go Start-R-G-B-Y
b) Pointing: from Y point to S, R, G, B

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

Alex Muryy

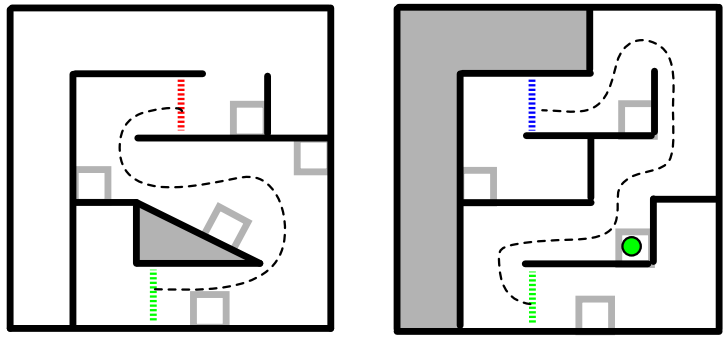
University of Reading

Non-metric scene: 1 wormhole

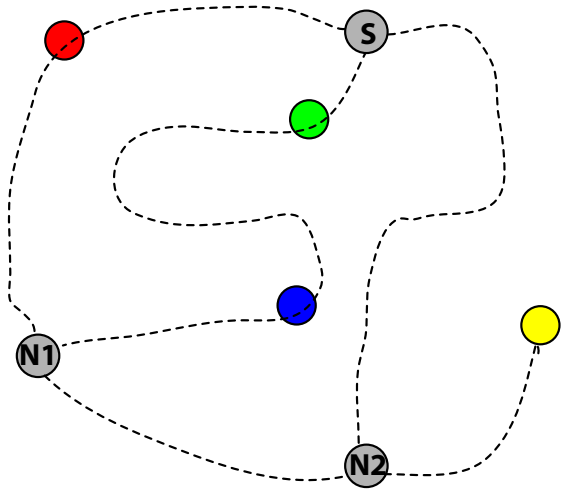


- topological node
- small walls
- wormhole

wormhole



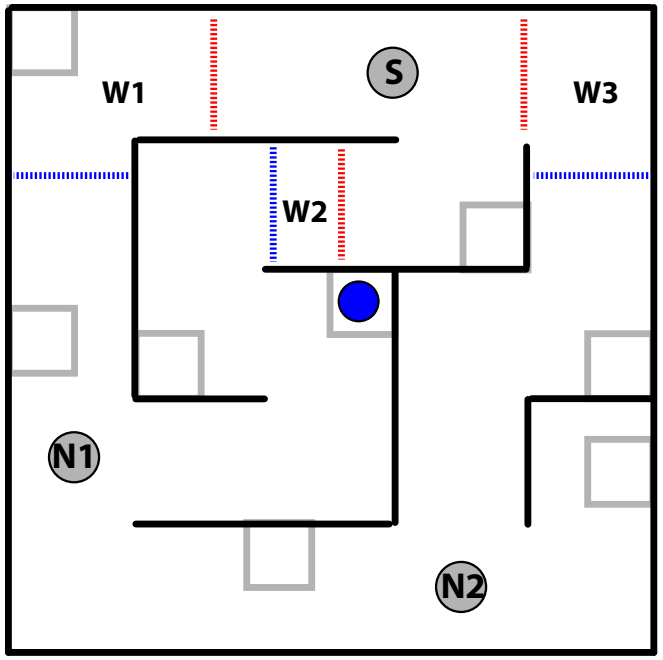
topological graph



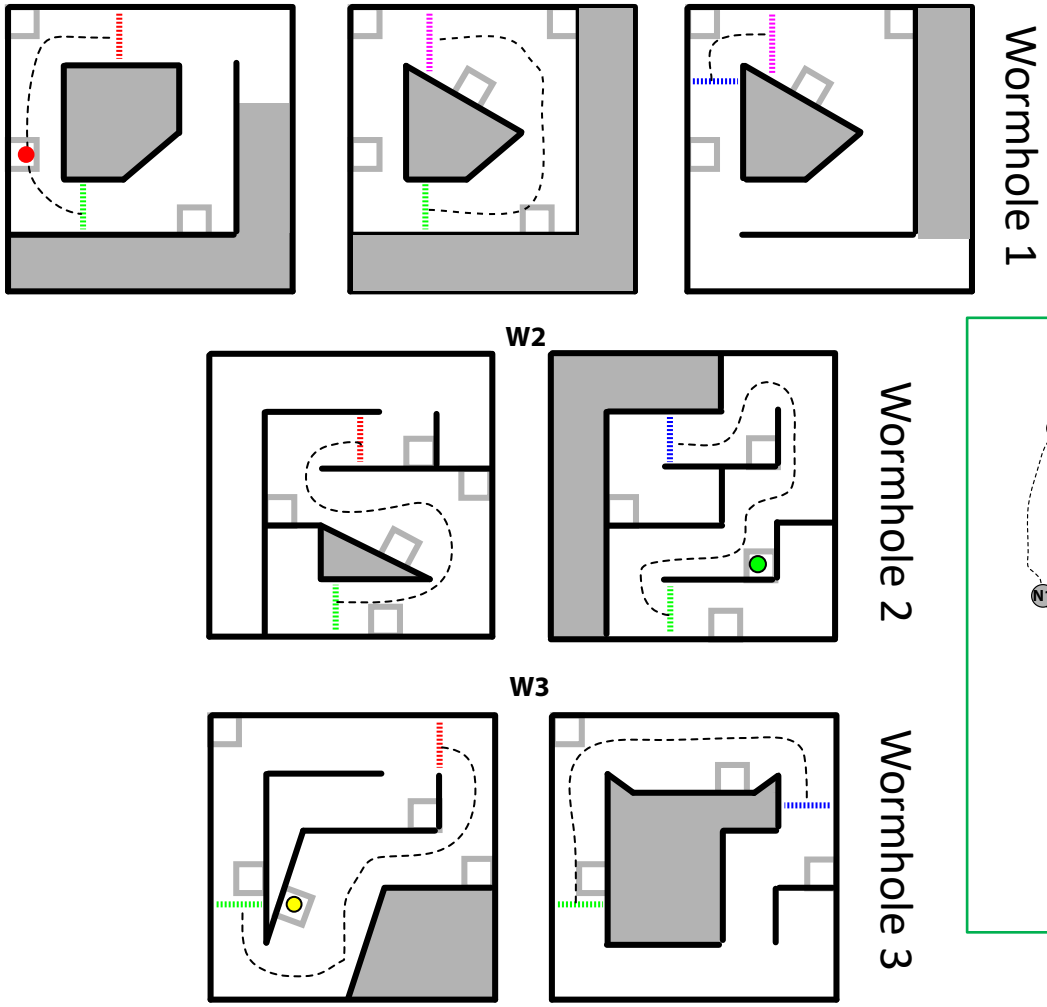
Alex Muryy

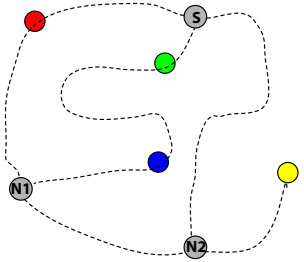
University of Reading

Non-metric scene: 3 wormholes



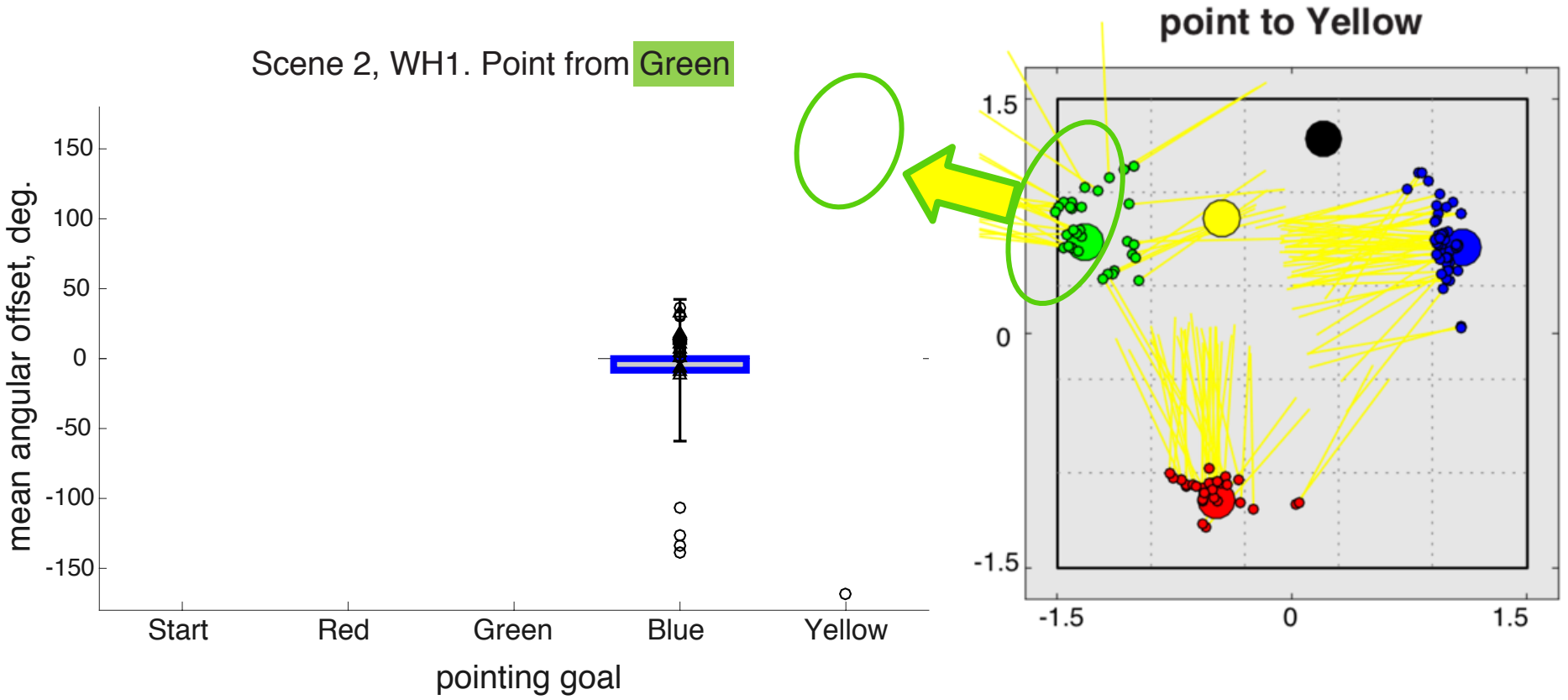
- topological node
- small walls
- wormhole





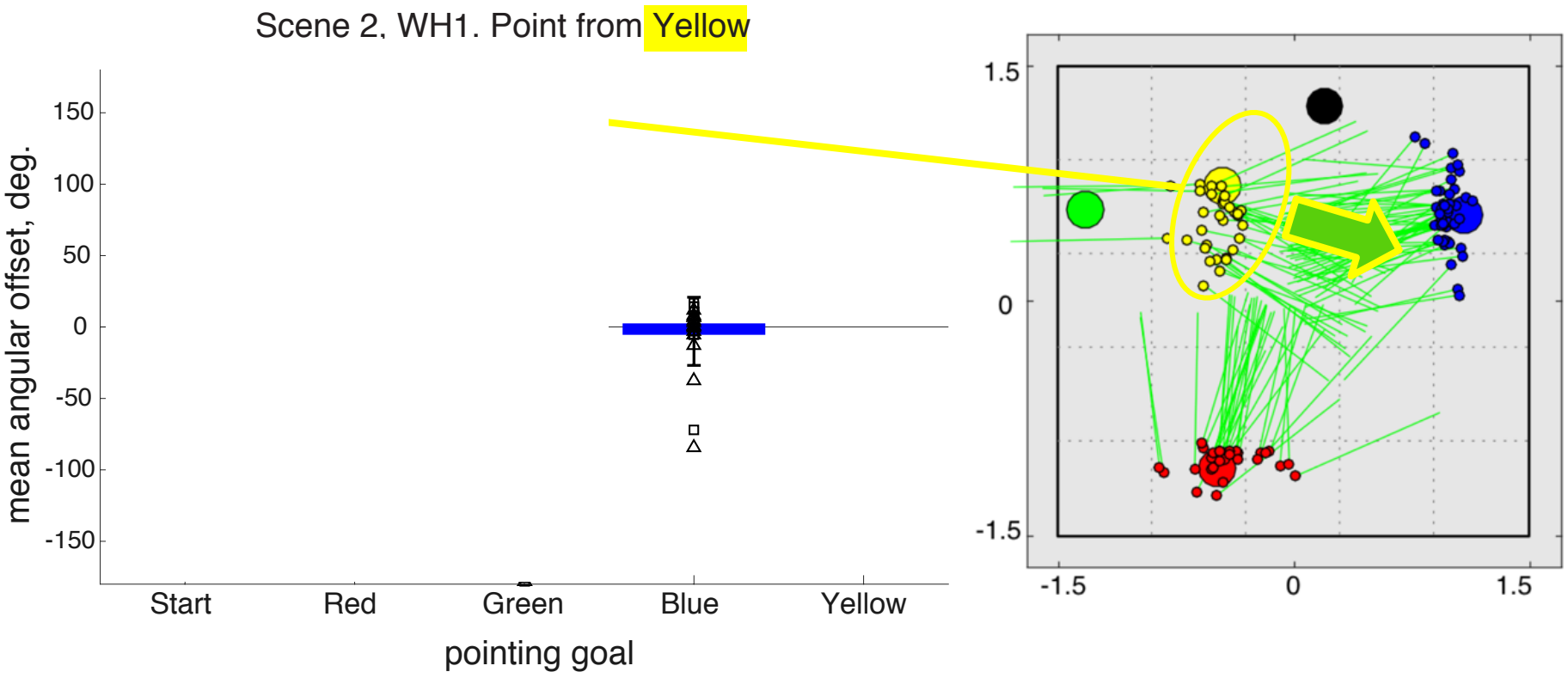
… but same topological structure

Pointing: a 'metric' task



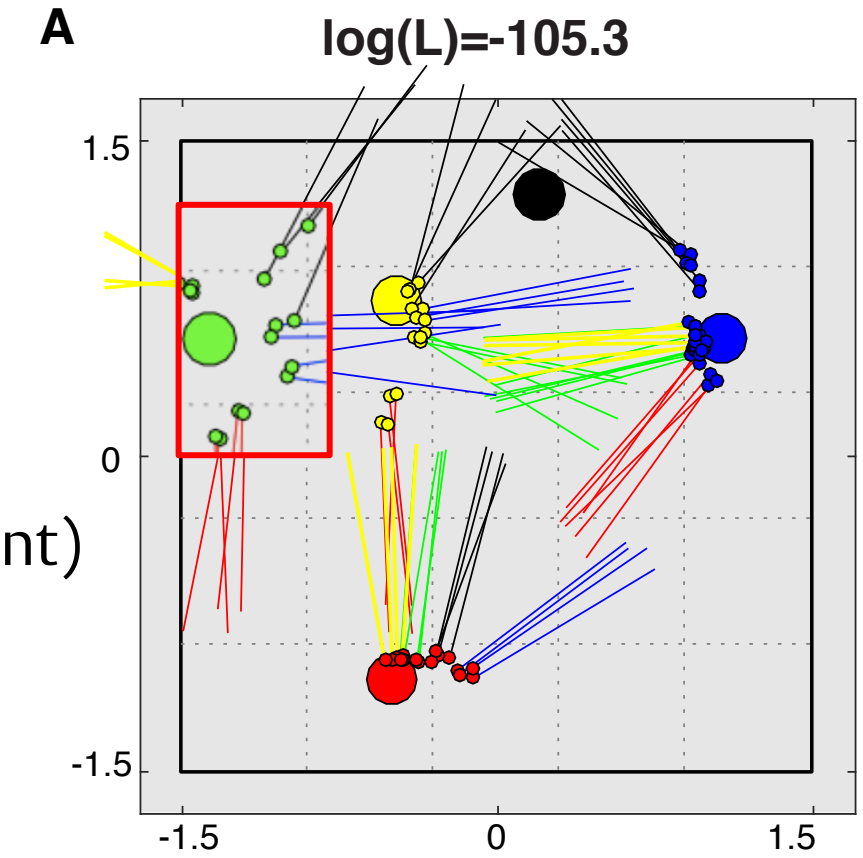
Pointing to some targets leads to very large, systematic errors.

Pointing: a 'metric' task



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

Pointing: a 'metric' task

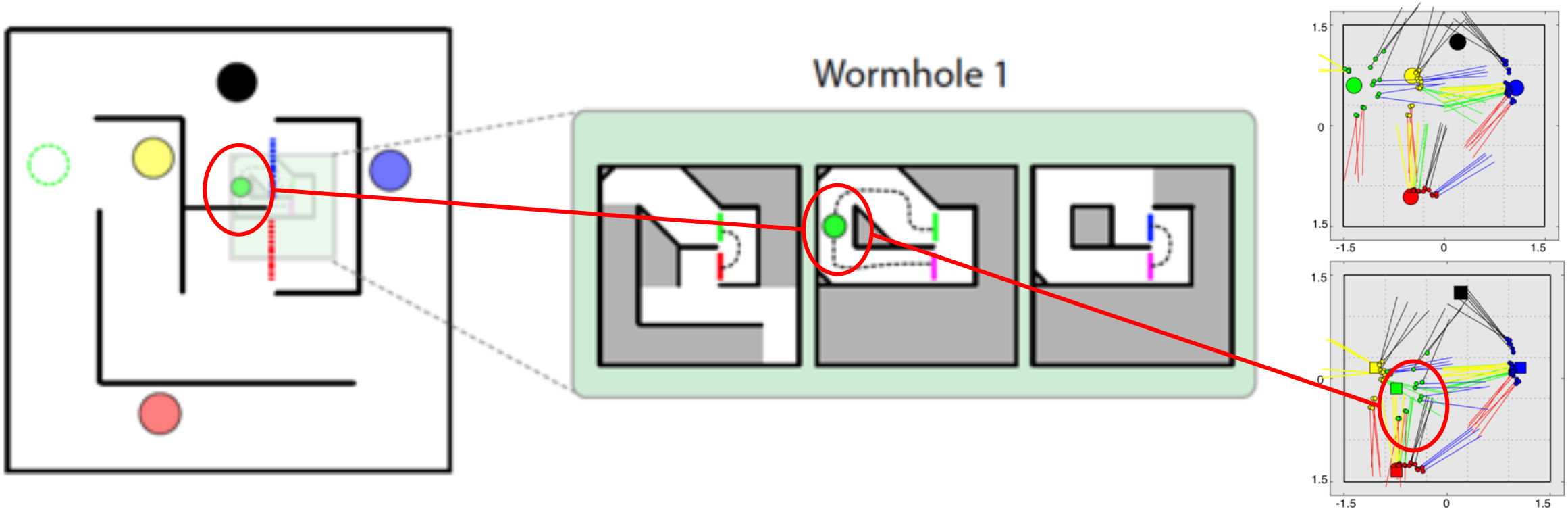


Original (one participant)

Maximum likelihood configuration

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

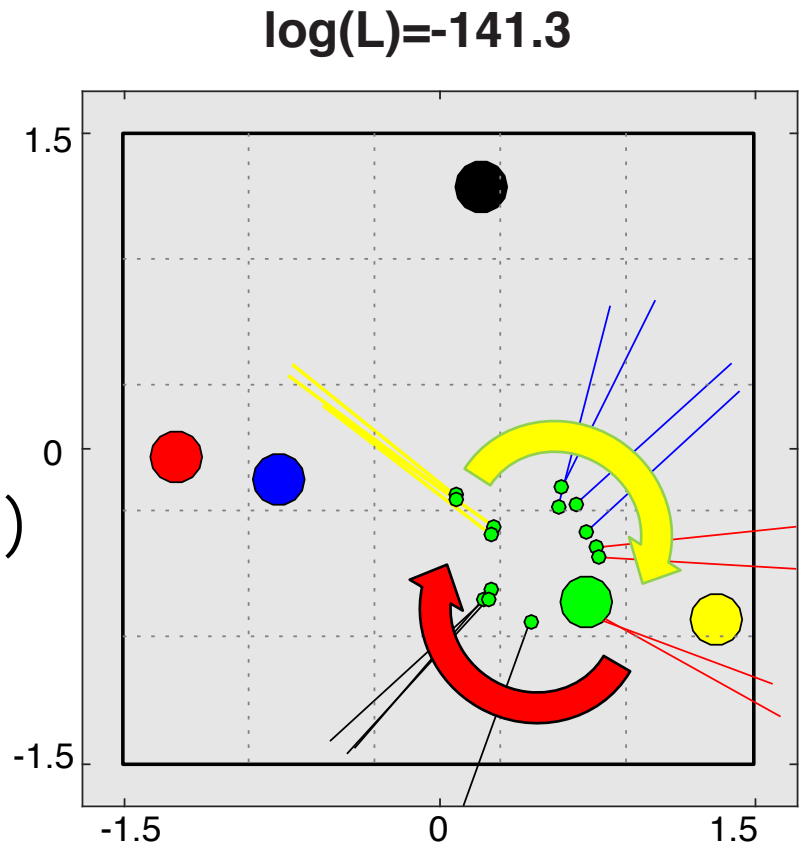
Pointing: a ‘metric’ task



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

Adding in rotation

Original (one participant)



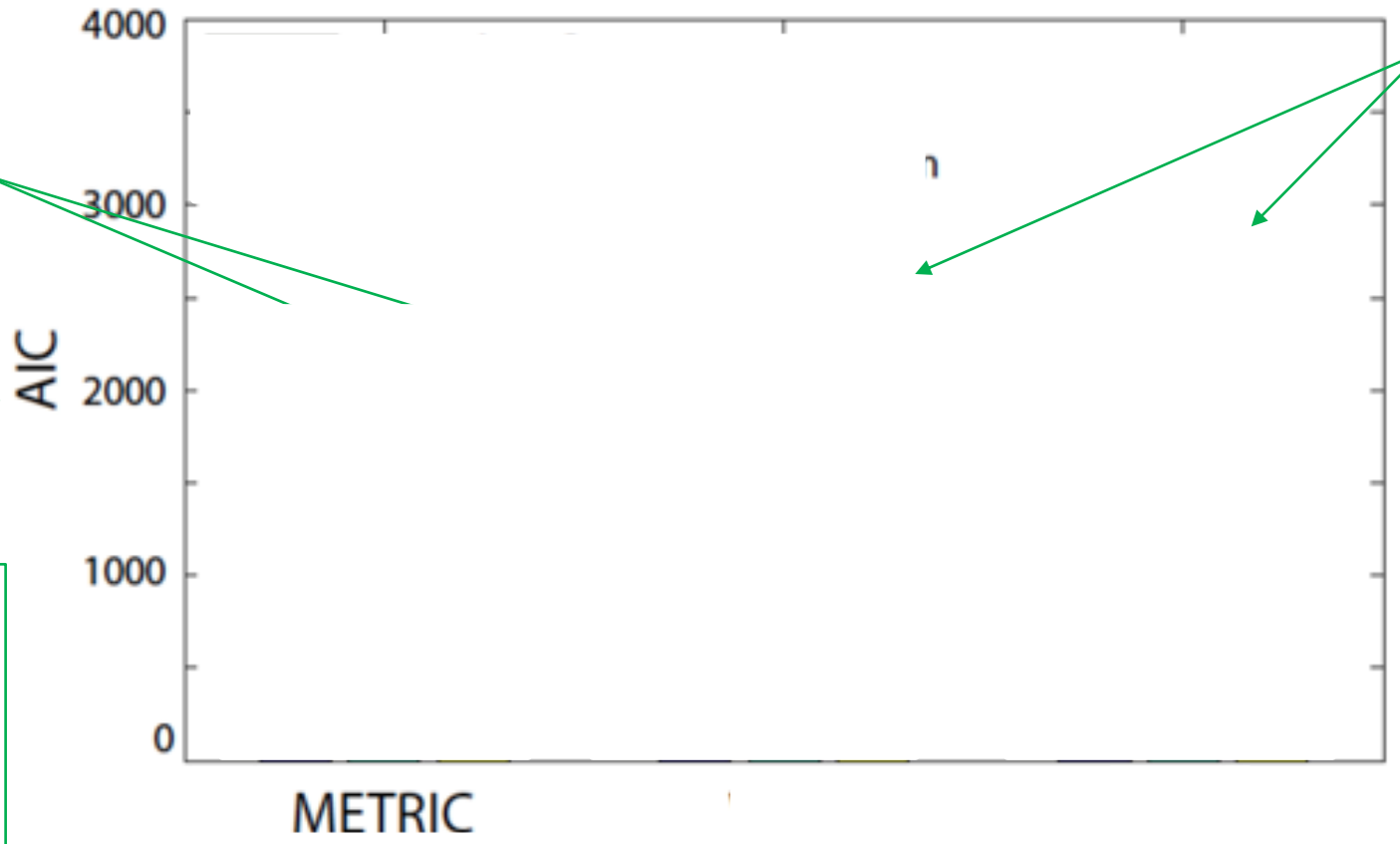
180° rotation of all pointing directions

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)



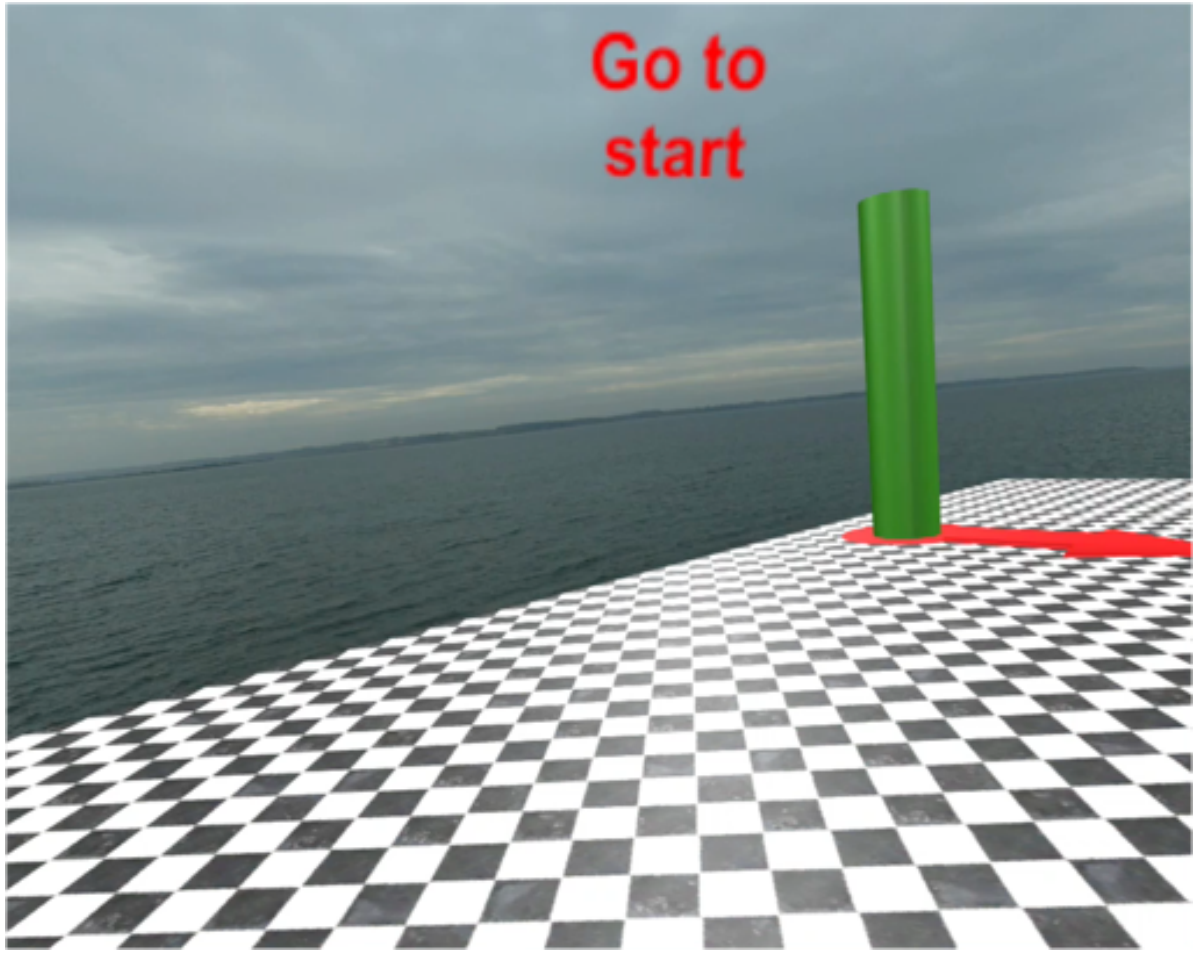
In the WORMHOLE conditions, the best model is one that optimizes the location of the targets (i.e. a distorted world) *and* optimizes the rotation of the observer independently at each pointing zone

This is not a consistent, metric, 3D representation

Alex Muryy

University of Reading

Learning to point to targets in a maze



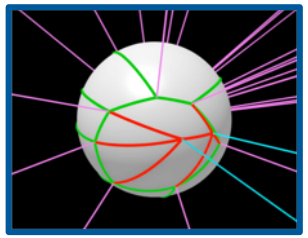
People's ability to point at unseen targets may be built up from an initial topological representation with information about lengths and turns gradually added as they learn about the environment.

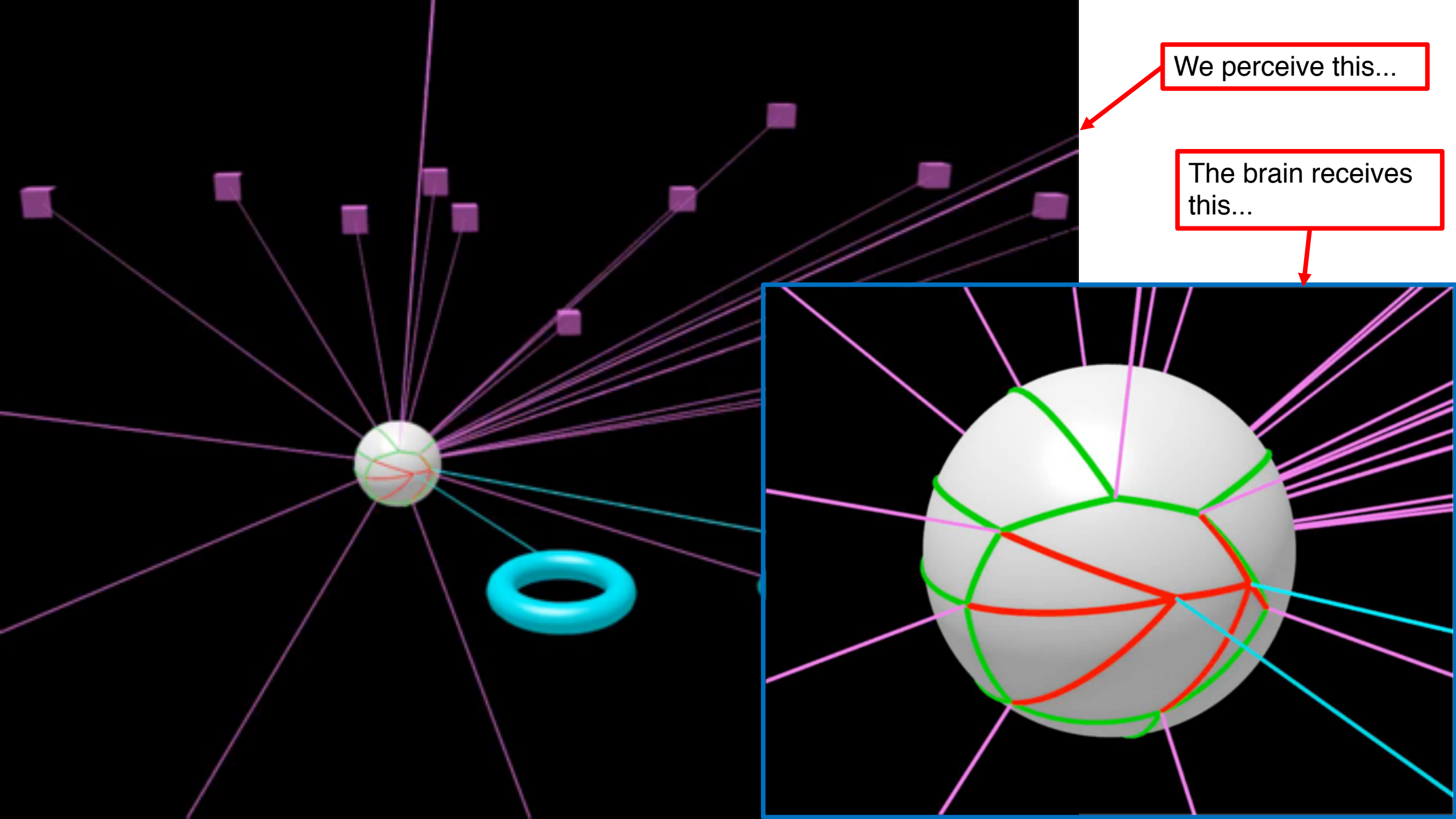
Outline

- Updating visual direction
  - some evidence and a ‘model’

- Navigating through wormholes
  - a 3D model is not the best explanation
  - coarse to fine learning of space

- A sphere of visual directions
  - information about viewing distance
  - A 2½ -D sketch



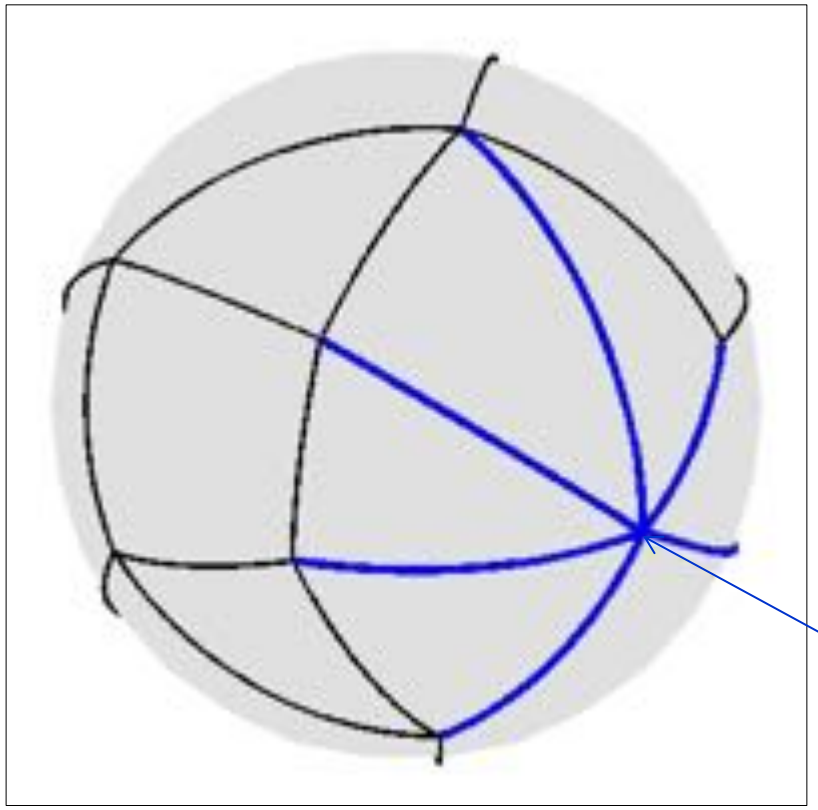

We perceive this...
The brain receives this...

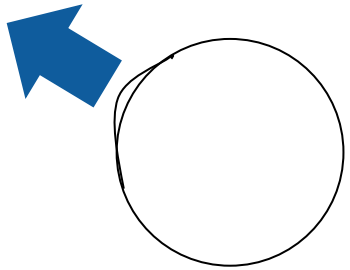
A 6-week old baby

https://www.youtube.com/watch?v=6qFj7Lh0Bbc

- Camera/eye can rotate easily but not much translation of the camera/head
- lots of practice at connecting images that are related by a pure rotation of the camera

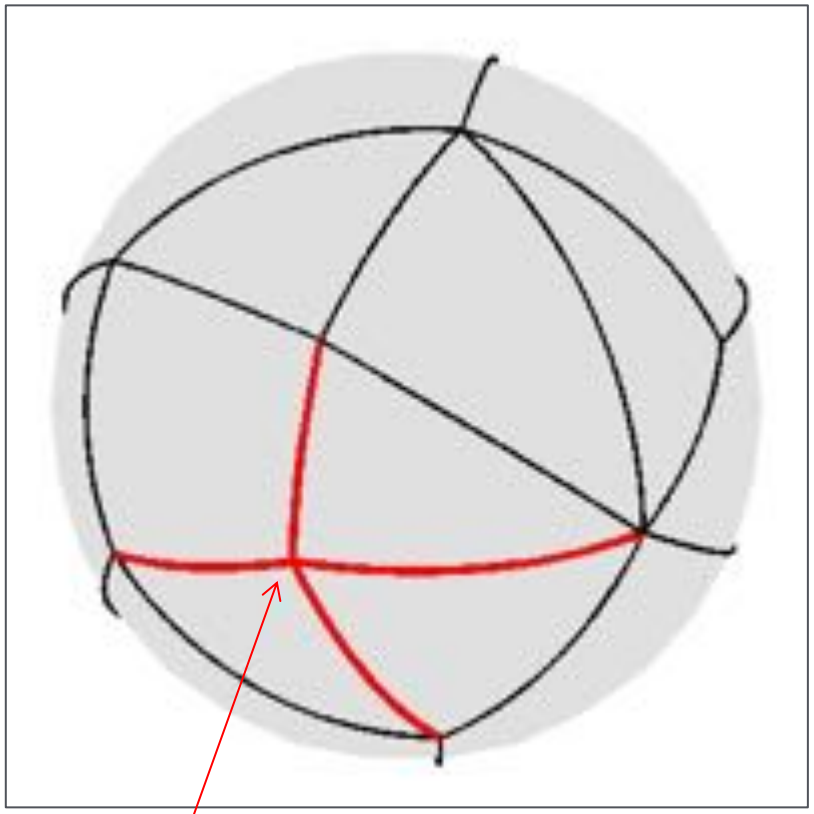
A stable coordinate frame for eye rotation



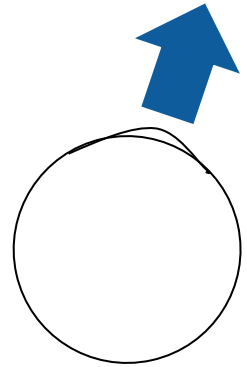


Fixated point here

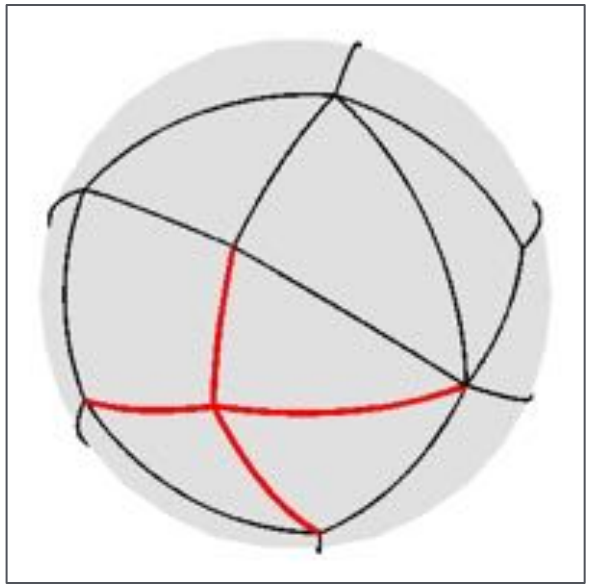
A stable coordinate frame for eye rotation

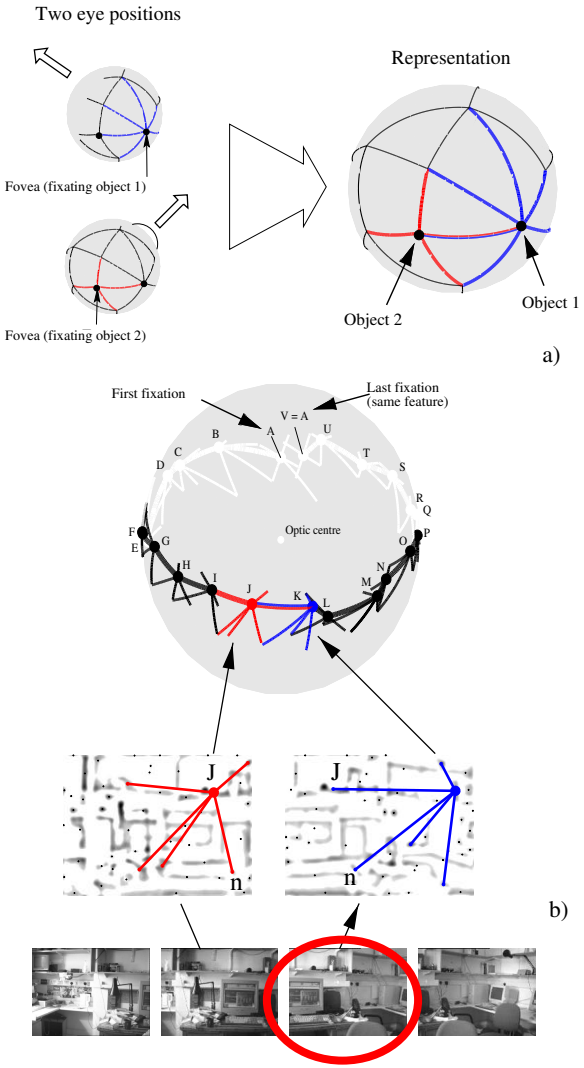


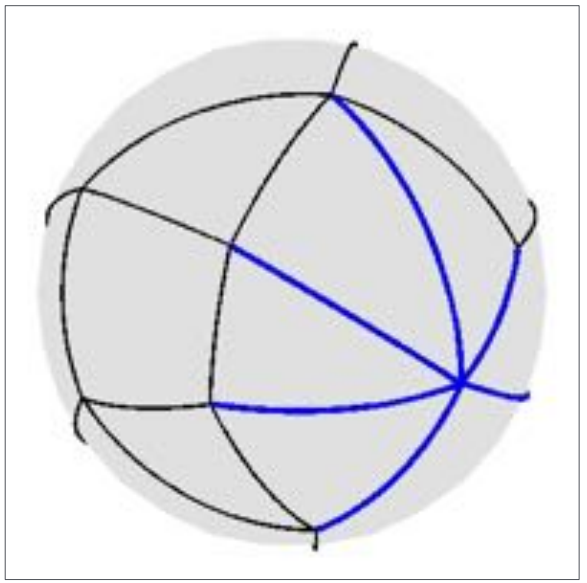
Fixated point here



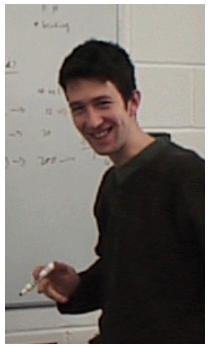
A stable coordinate frame for eye rotation







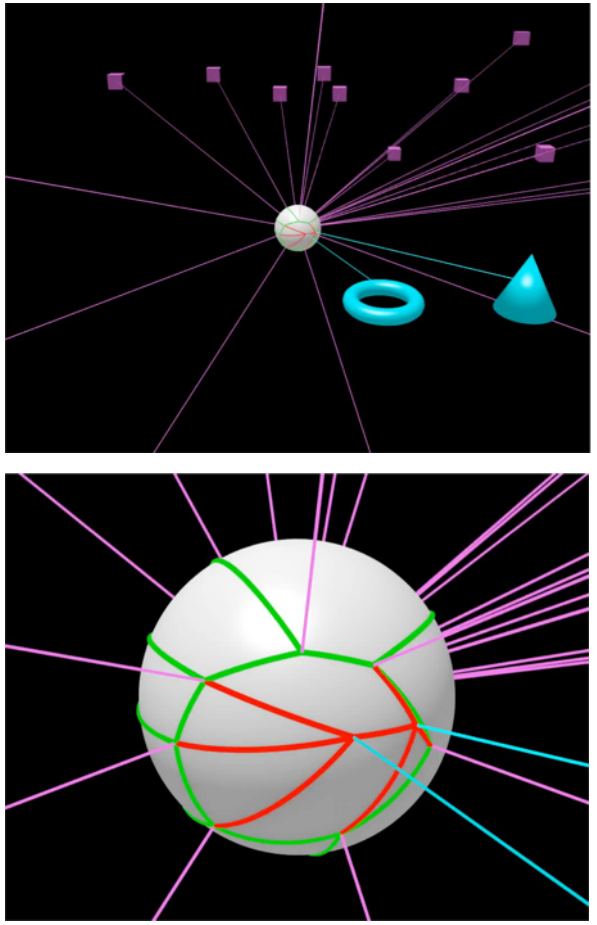
Miles Hansard

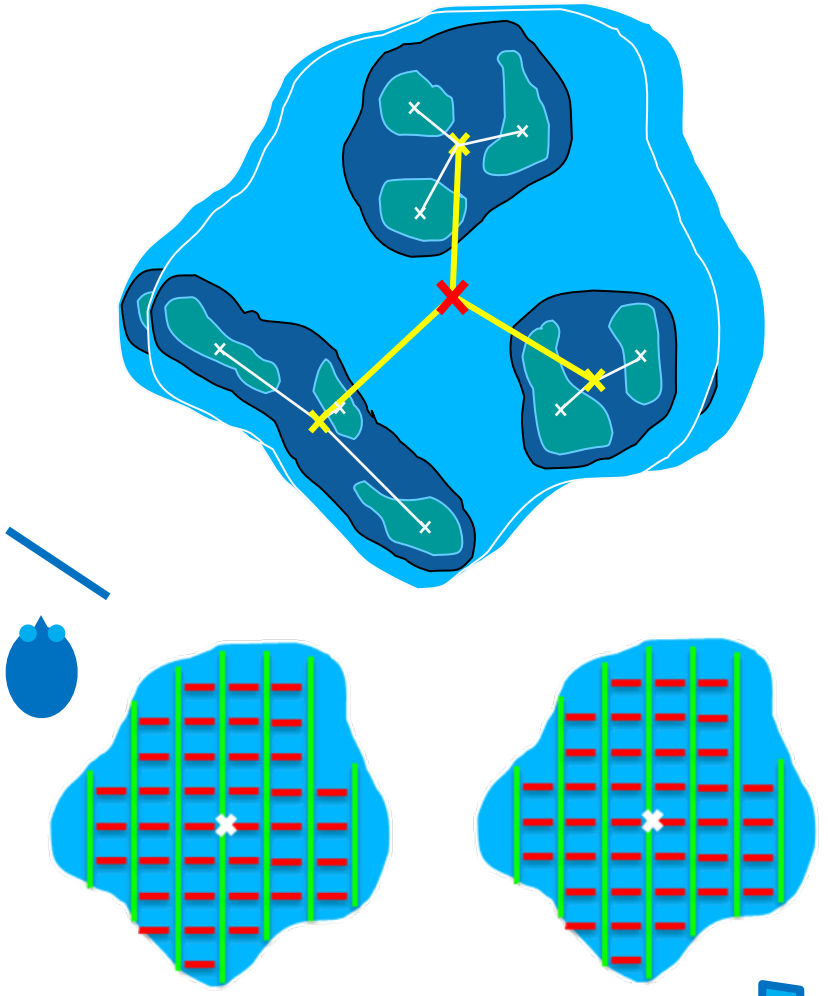


Andrew Fitzgibbon

Glennerster, Hansard and Fitzgibbon (2001)

Elasticity – a property that persists





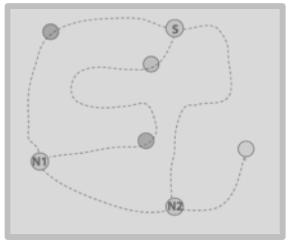
Not just another description of optic flow. Instead, it is a long-lasting useful description with predictive power

Glennerster (2016), Glennerster, Hansard and Fitzgibbon (2001, 2009)

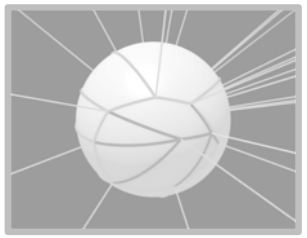
Outline

- Updating visual direction
  - some evidence and a ‘model’

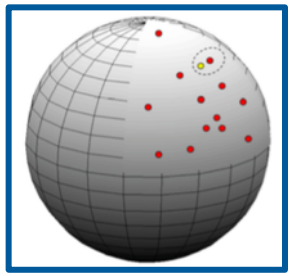
- Navigating through wormholes
  - a 3D model is not the best explanation
  - coarse to fine learning of space



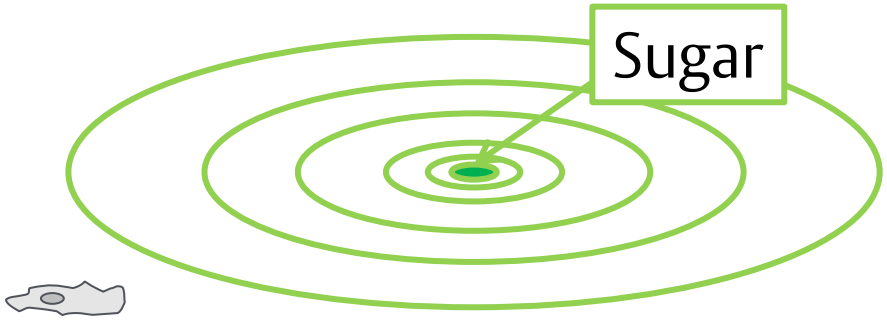
- A sphere of visual directions
  - information about viewing distance
  - A 2½ -D sketch

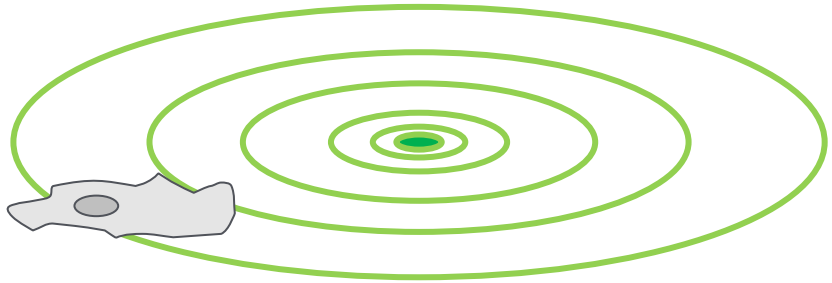


- A sphere of sensory+motivational states
  - a gradual increase in dimensionality

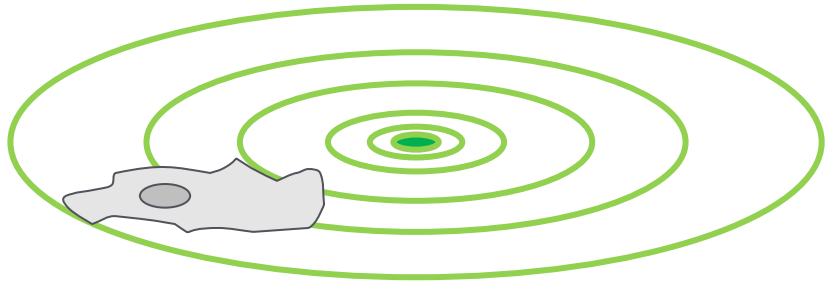


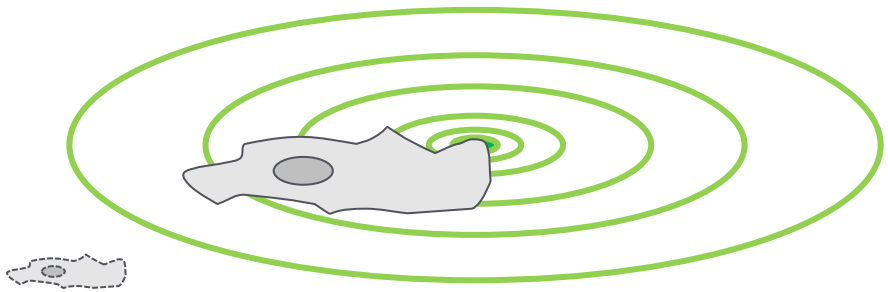
University of Reading

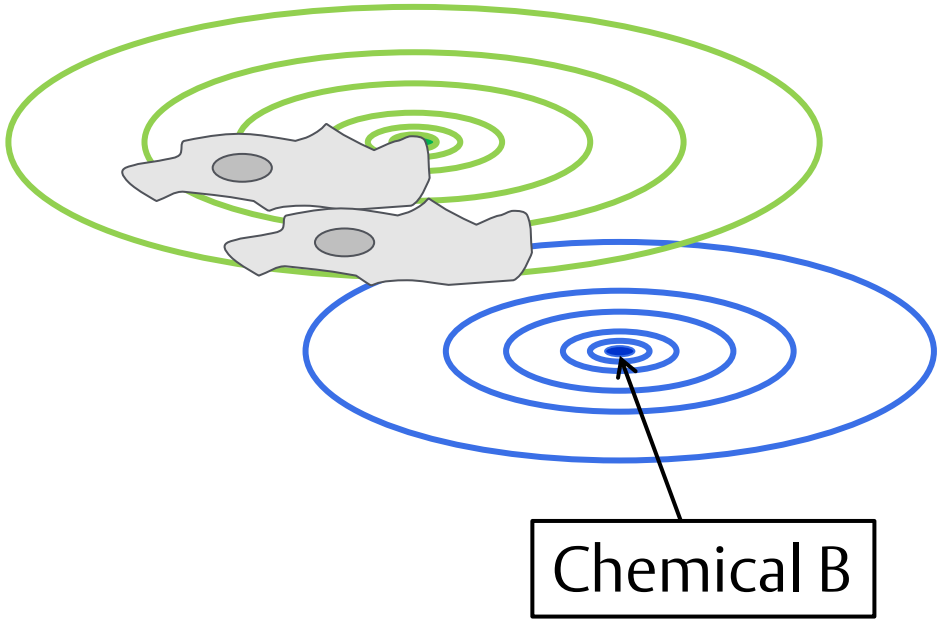

Sugar



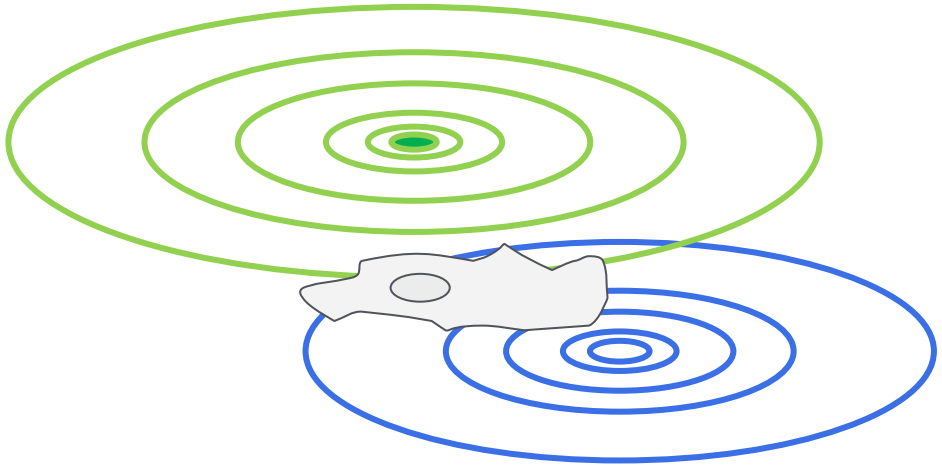
University of Reading



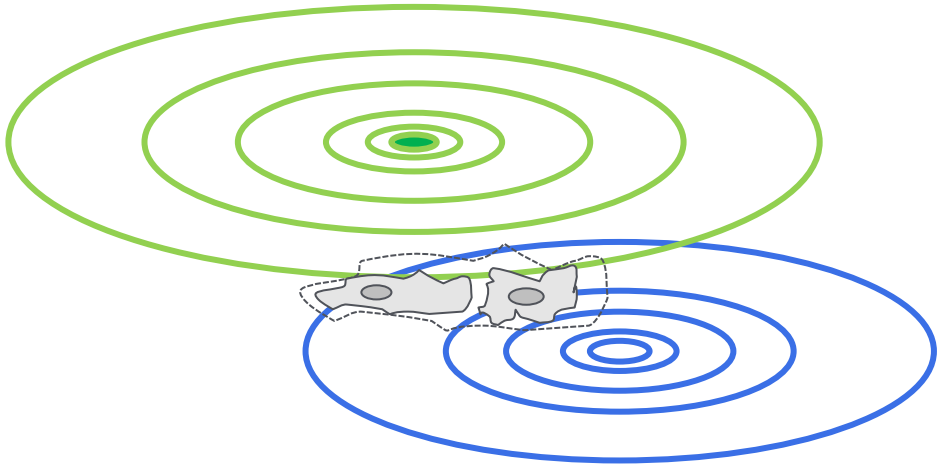


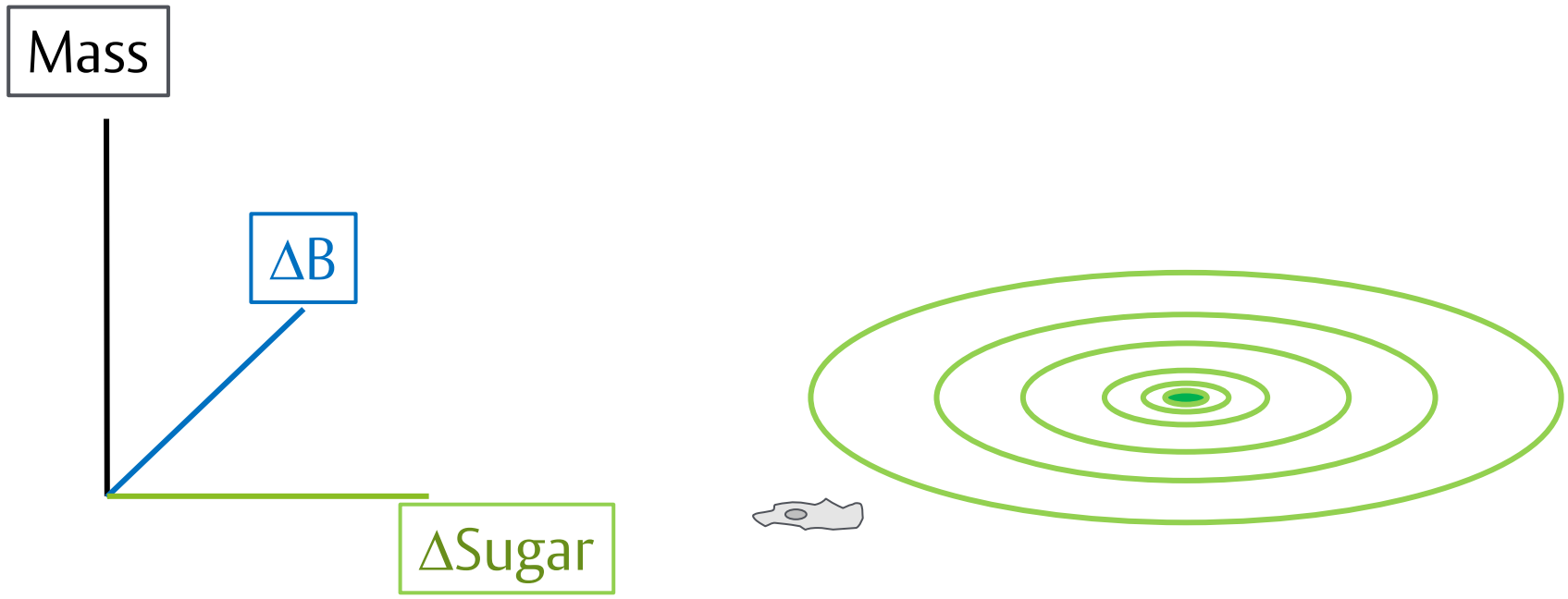

Chemical B

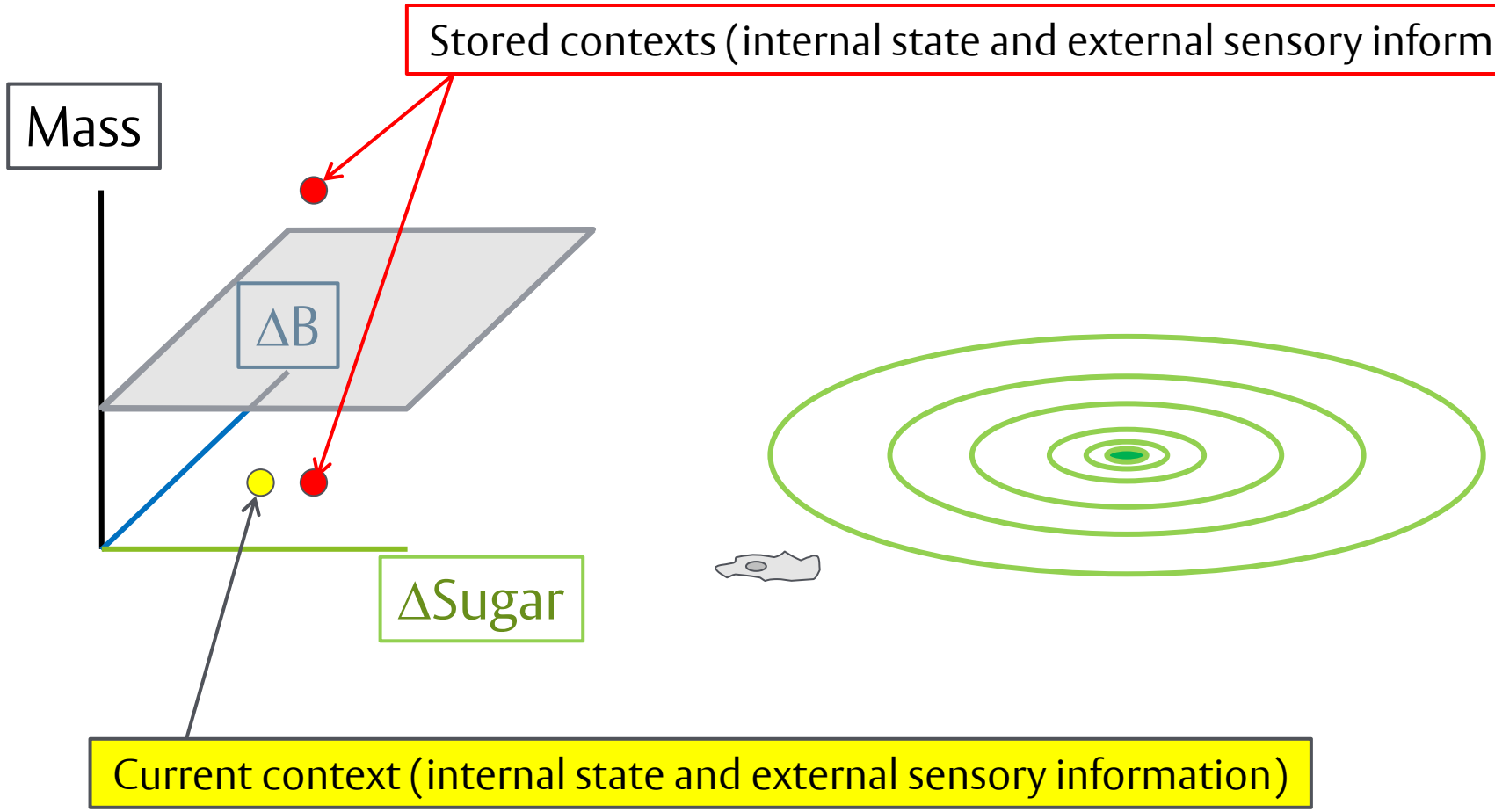
University of Reading

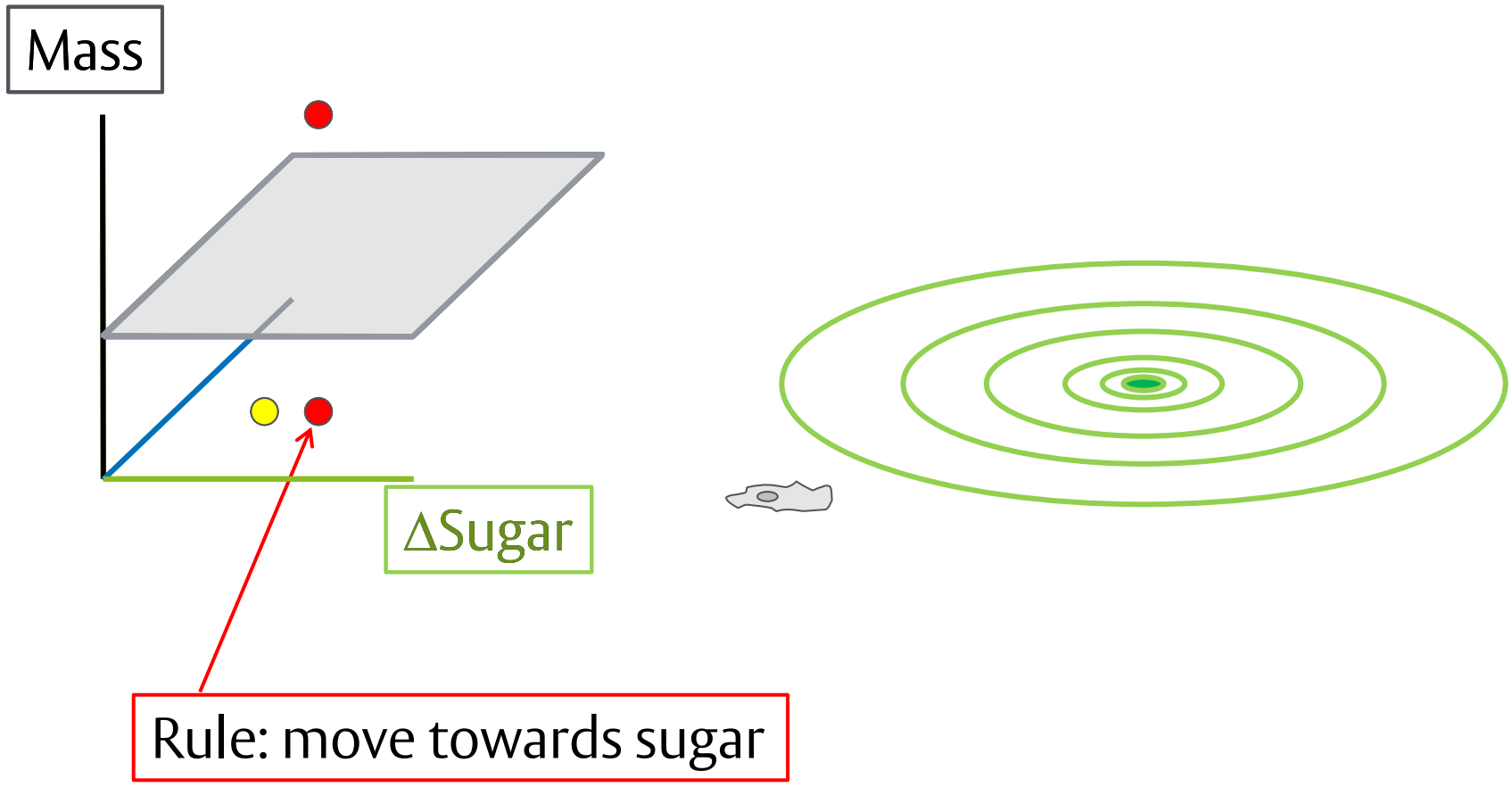


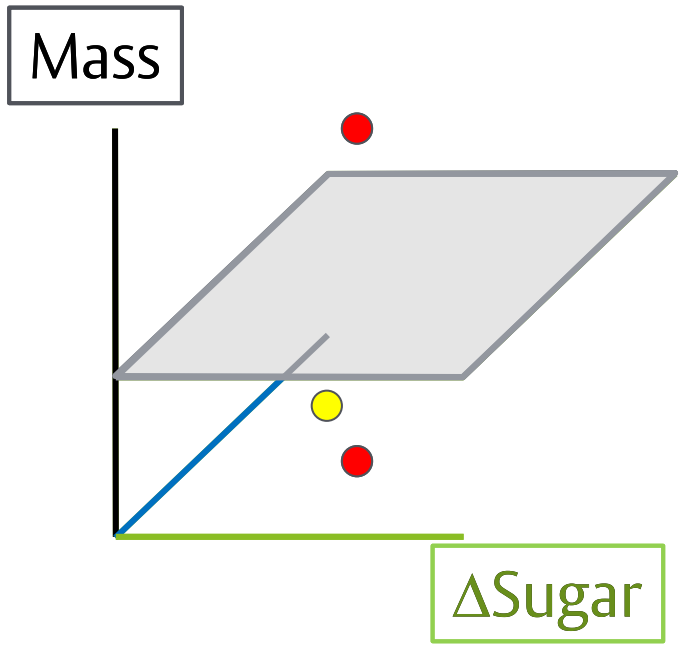
University of Reading

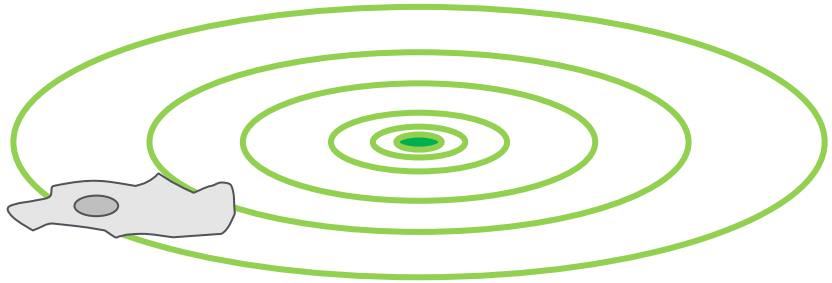


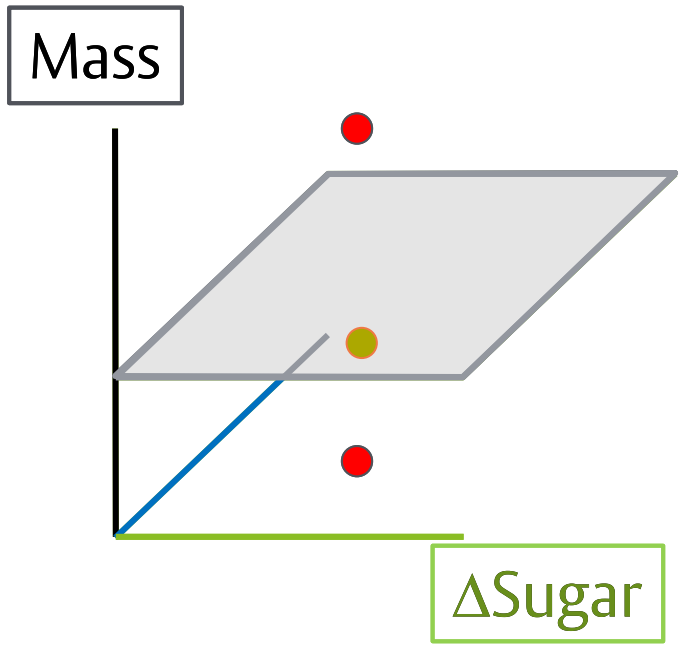

Mass
ΔB
ΔSugar

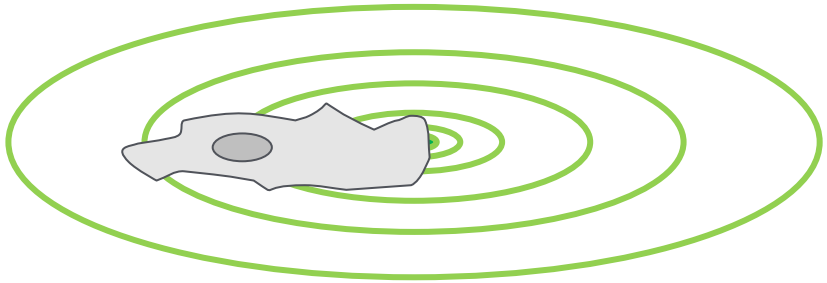

Stored contexts (internal state and external sensory information)
Mass
ΔB
ΔSugar
Current context (internal state and external sensory information)

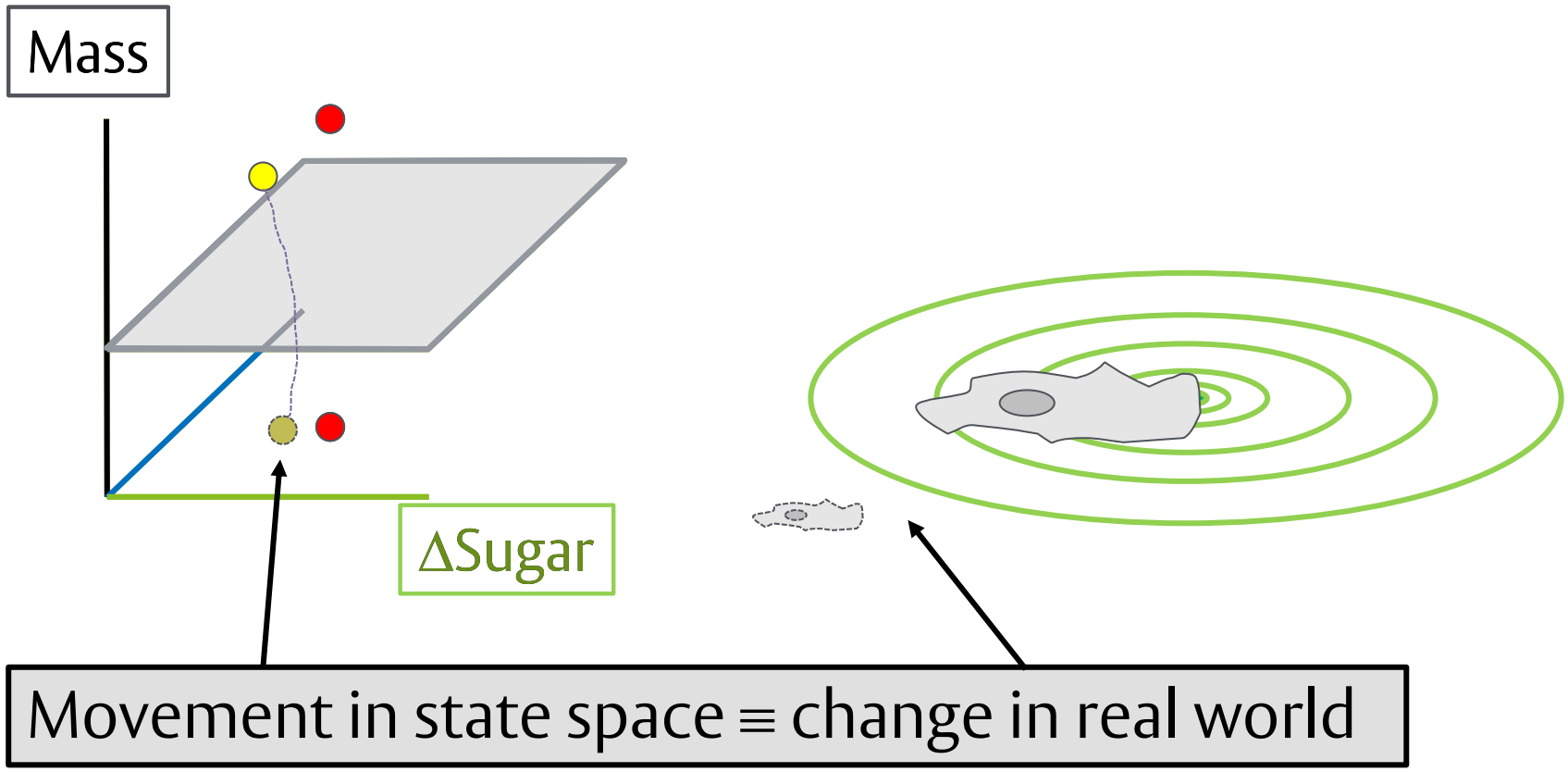

Mass
ΔSugar
Rule: move towards sugar

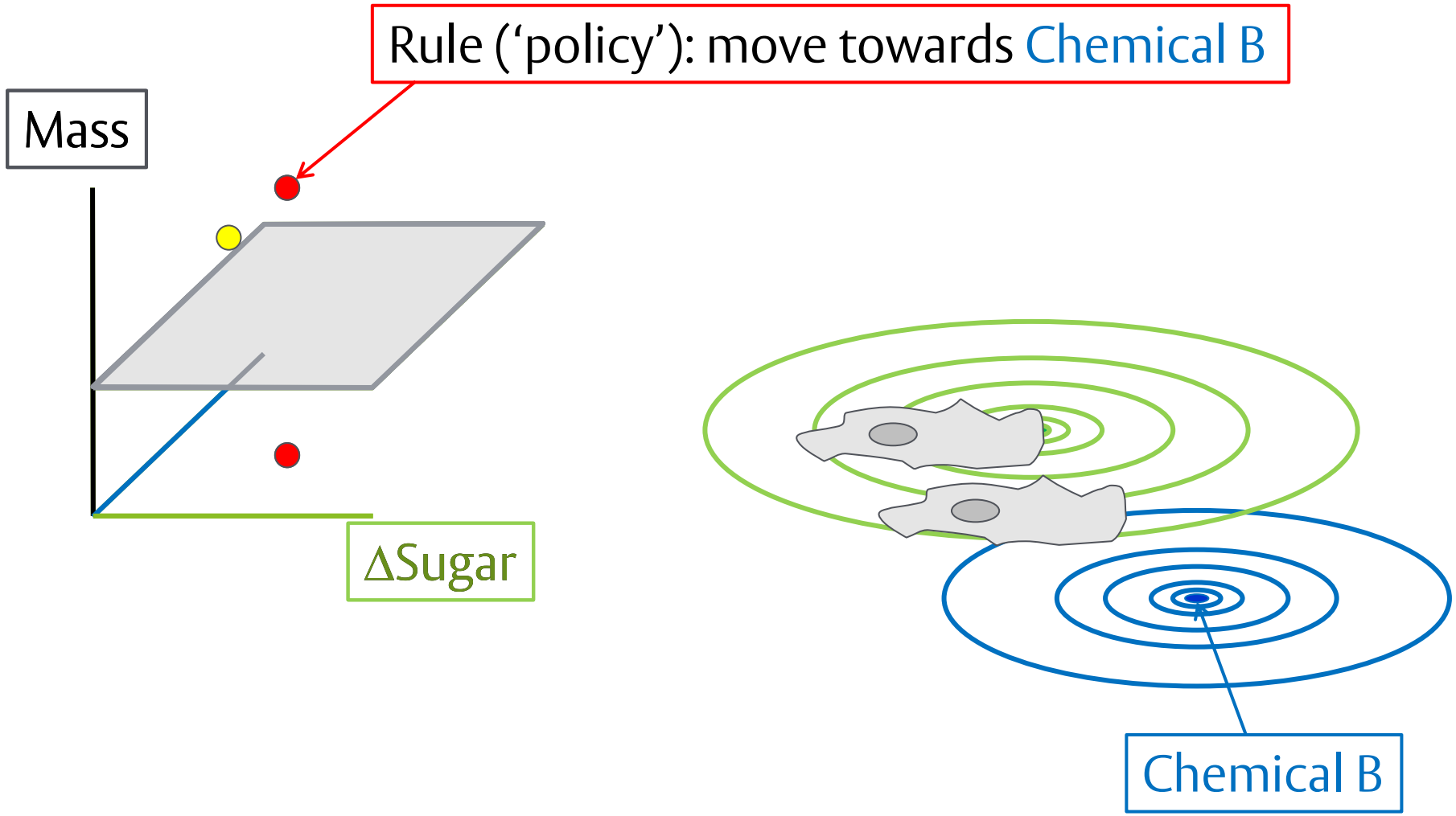

Mass
ΔSugar

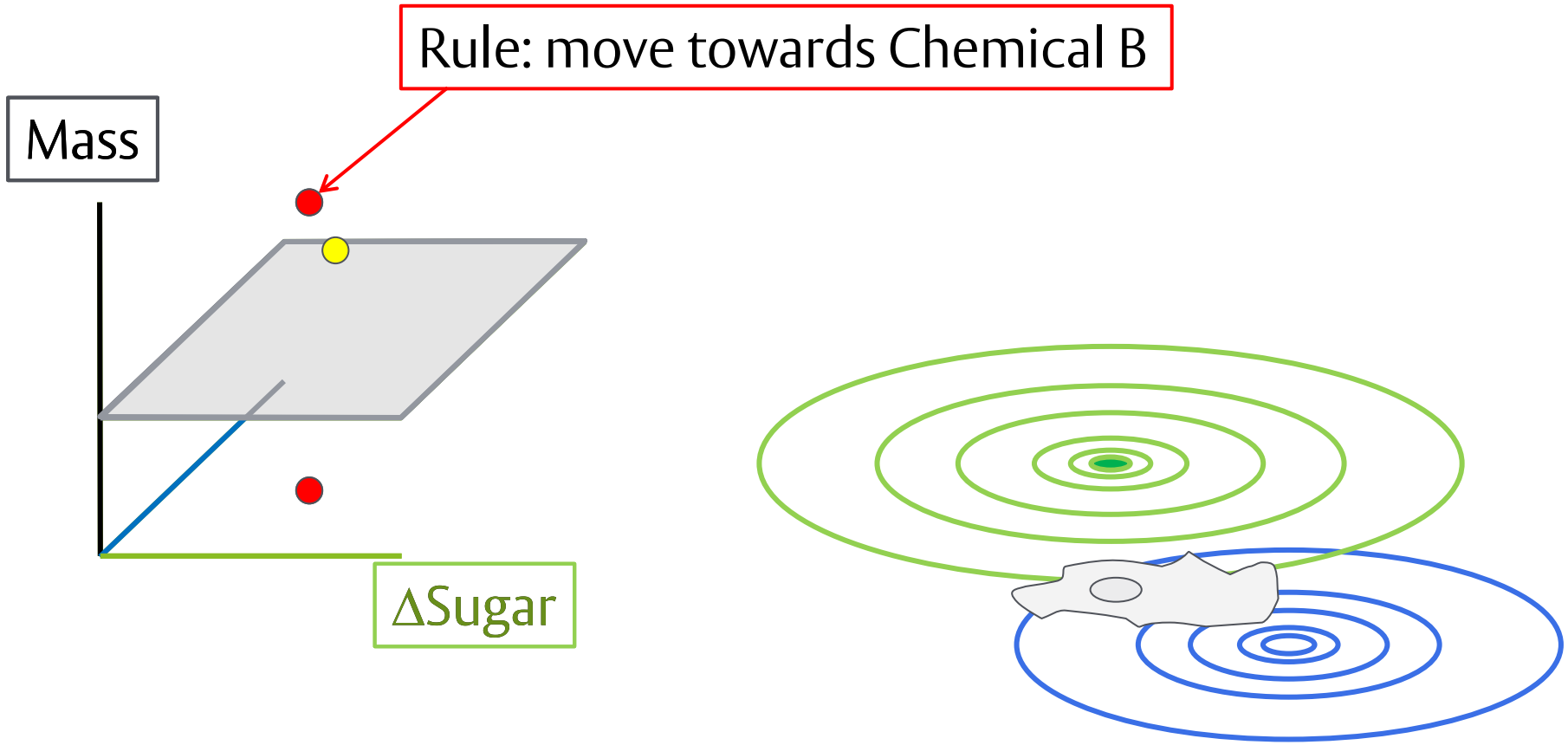


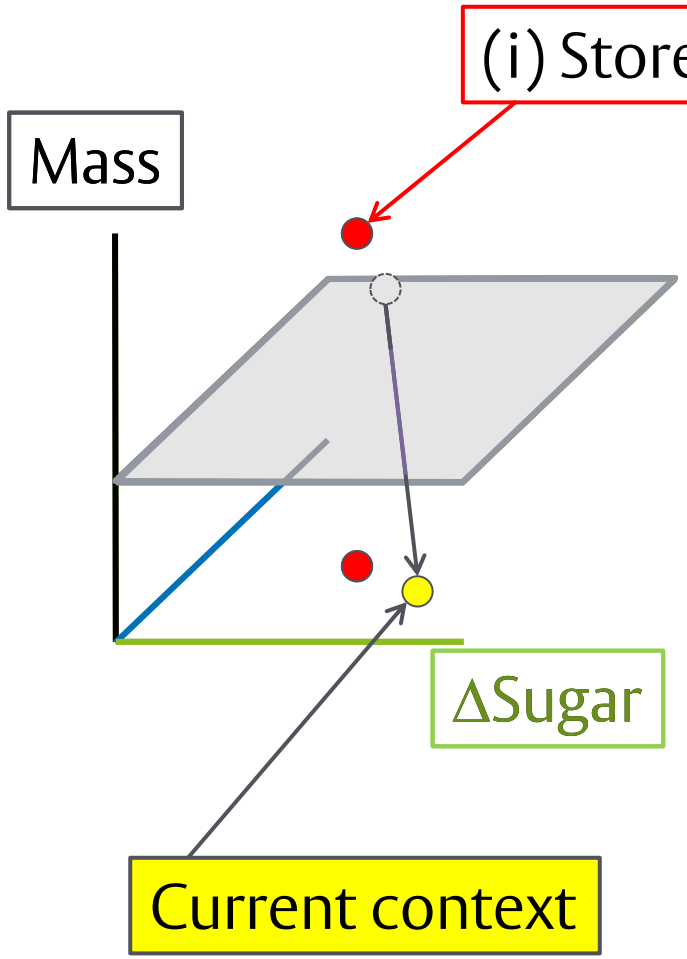

Mass
ΔSugar

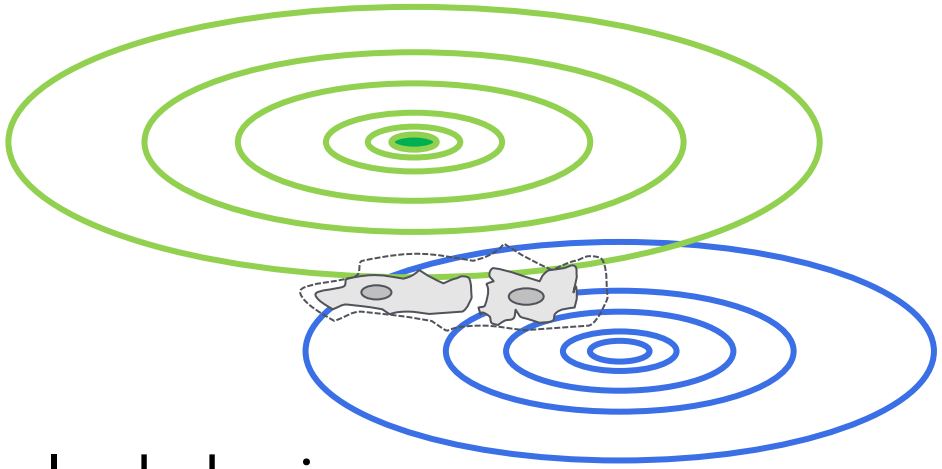



Mass
ΔSugar
Movement in state space ≡ change in real world


Rule ('policy'): move towards Chemical B
Mass
ΔSugar
Chemical B

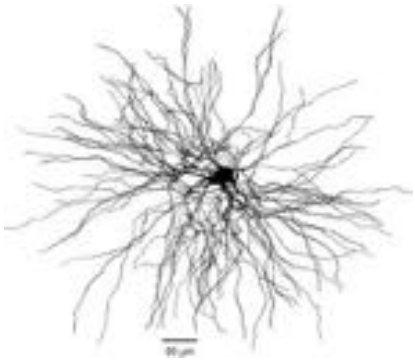

Rule: move towards Chemical B
Mass
ΔSugar



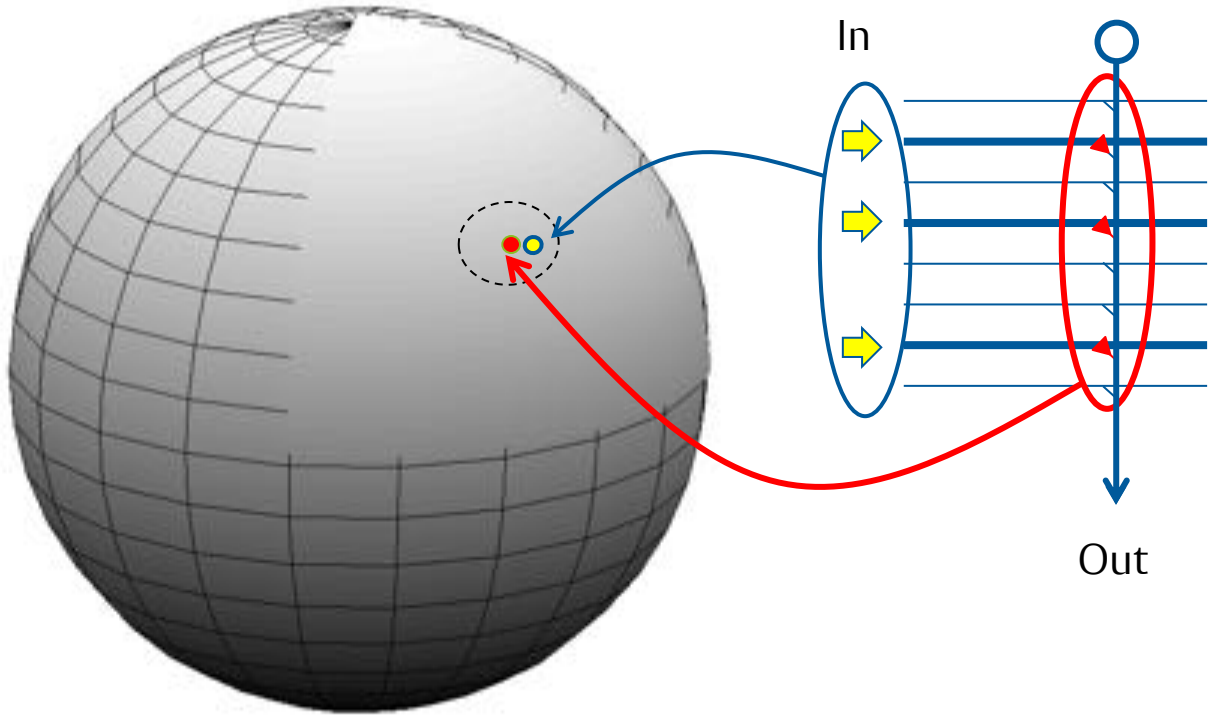


Complex behaviour:
- more dimensions
- more contexts

One neuron



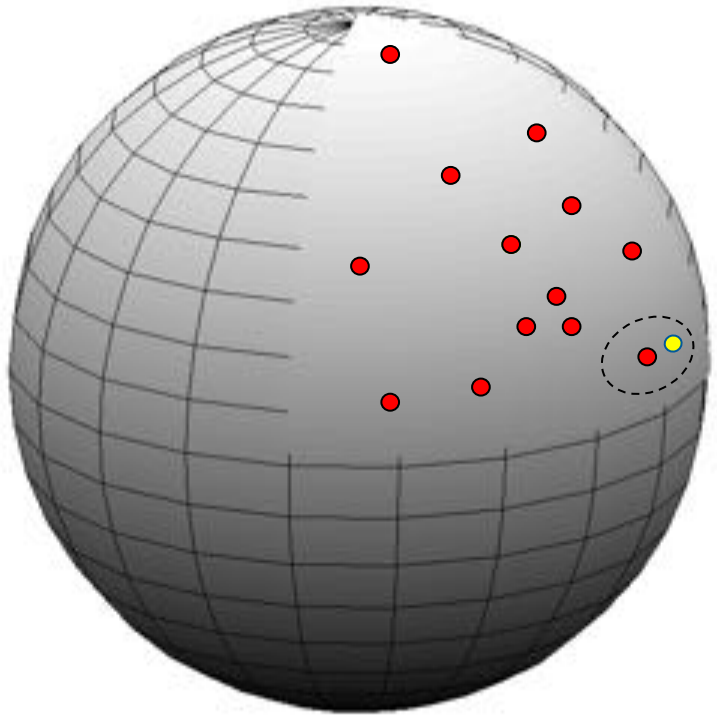
10^4 inputs



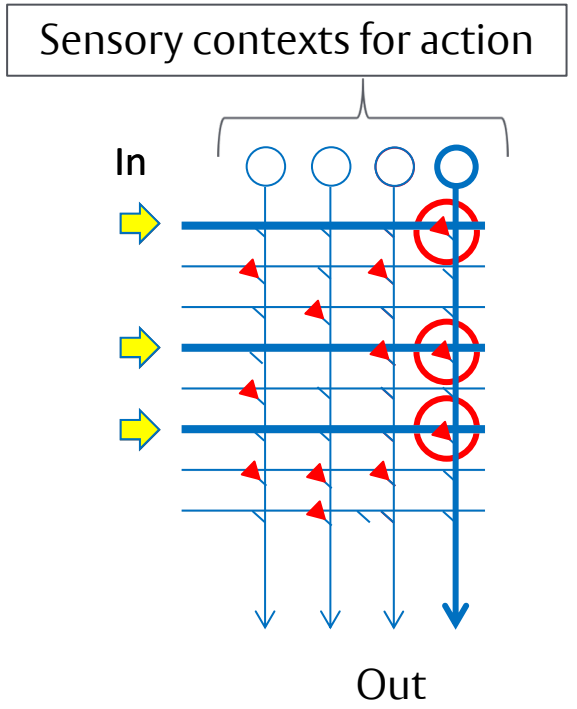
10^4 dimensions

Whole brain

10^{11} inputs

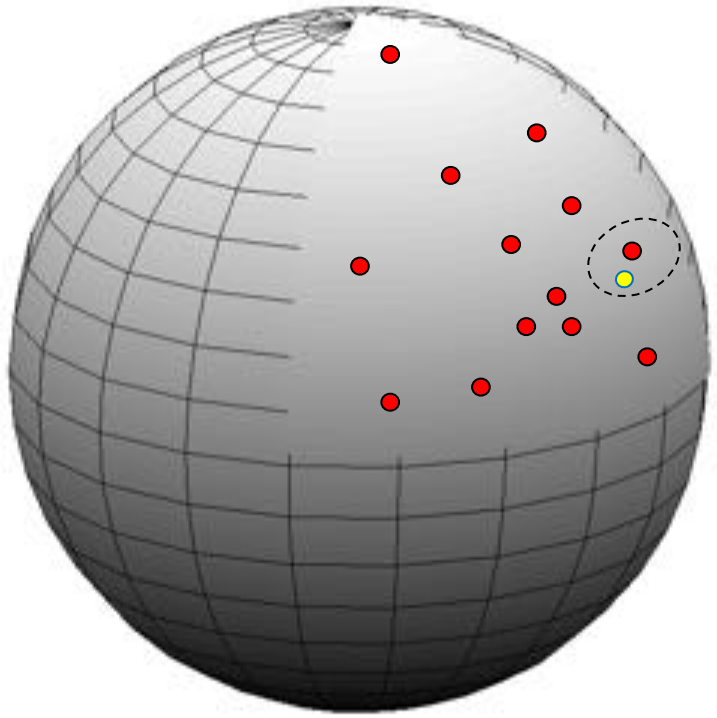


10^{11} dimensions

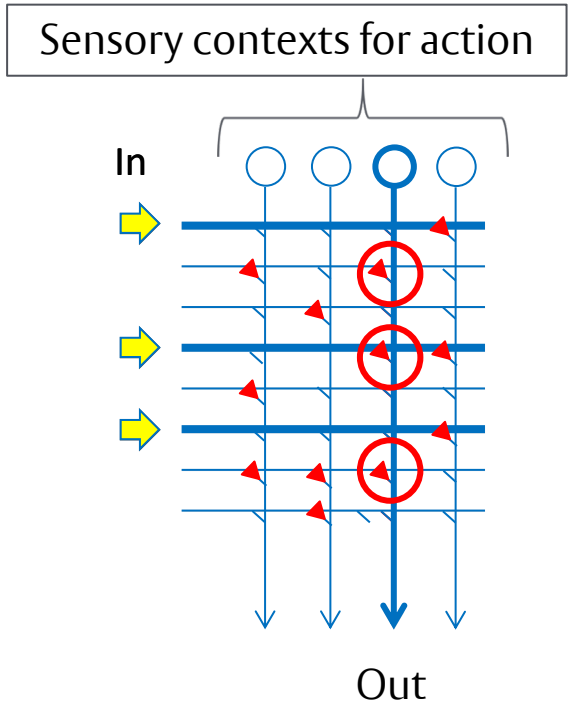


Whole brain

10^{11} inputs



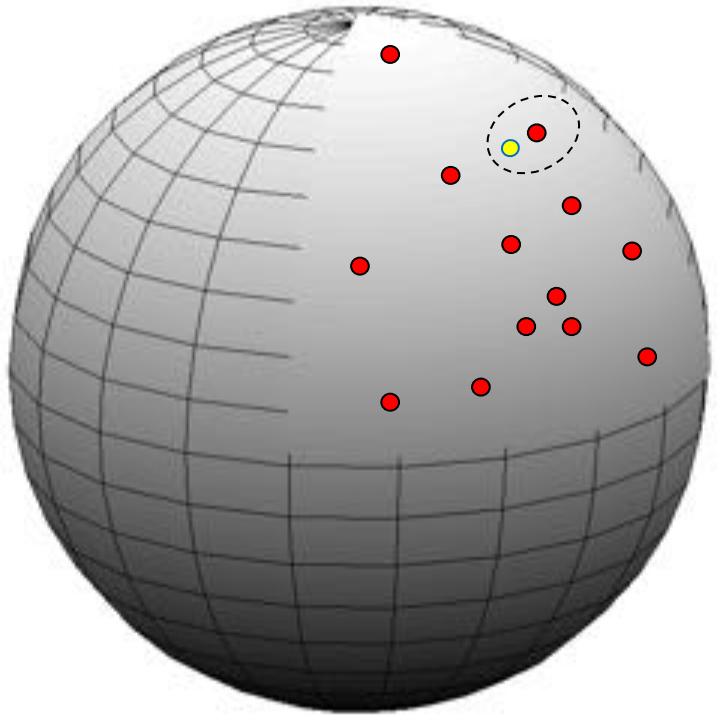
10^{11} dimensions



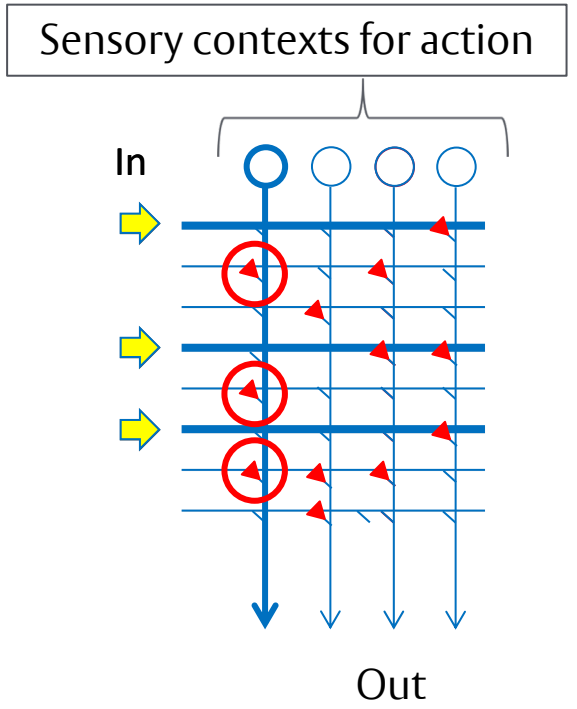
University of Reading

Whole brain

10^{11} inputs



10^{11} dimensions

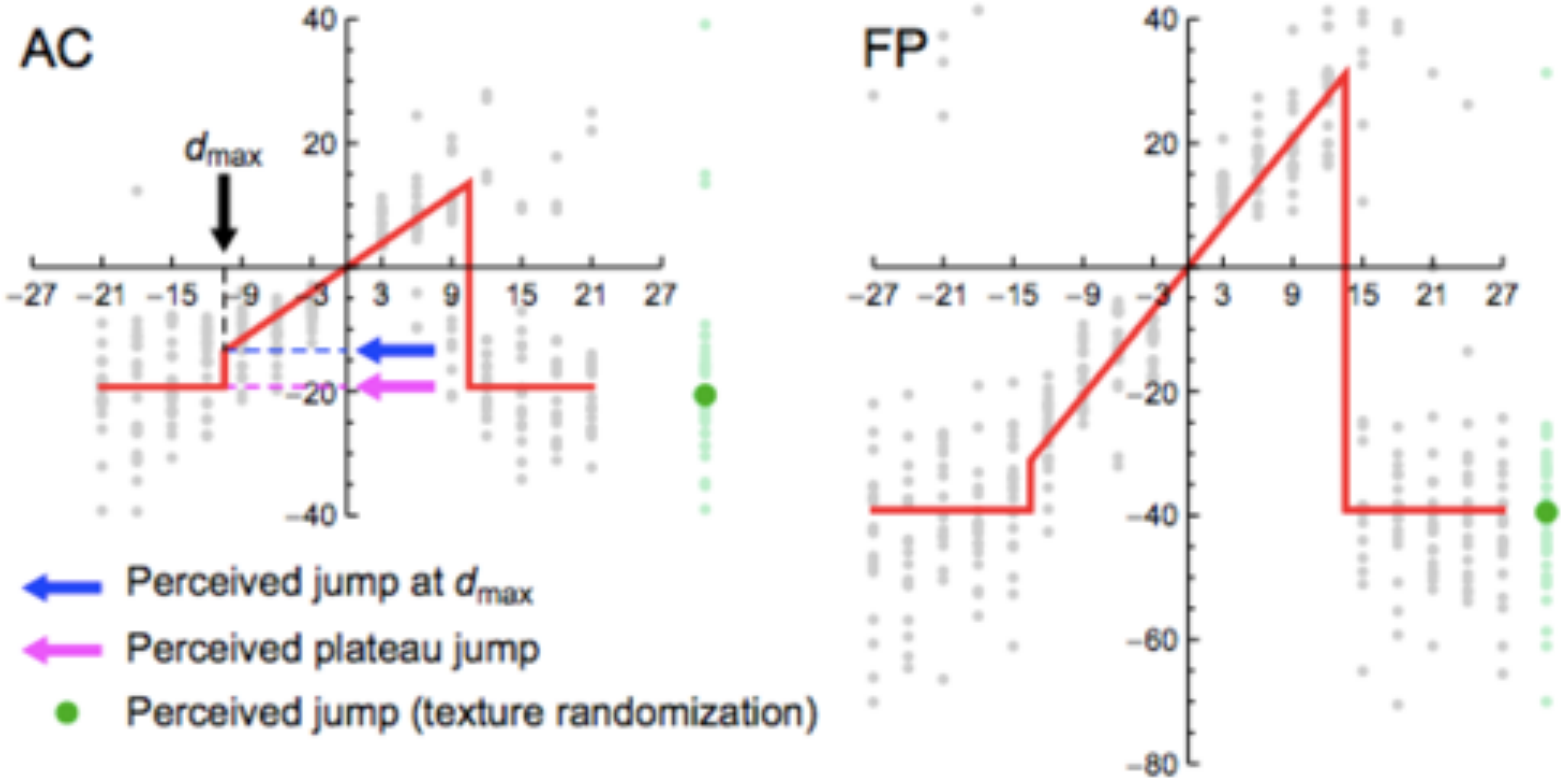


Marr (1969), Albus (1971)

University of Reading

Action

High phi



'Of course, [the Bayesian explanation] is utter BS. Andrew is saying that every study since Hubel and Wiesel has been barking up the wrong tree.' (Stephen Macnick)

Wexler, Glennerster, Cavanagh, Ito and Seno (2013)

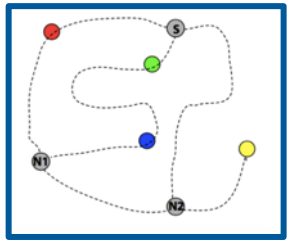
Outline

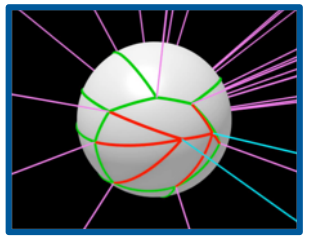
Evidence against a 3D model in the head

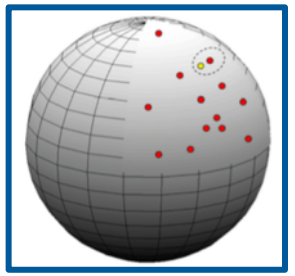
- Updating visual direction
  - some evidence and a 'model'
- Navigating through wormholes
  - a 3D model is not the best explanation

Speculations about an alternative

- A sphere of visual directions
  - information about viewing distance
  - A 2½ -D sketch
- A sphere of sensory+motivational states
  - a gradual increase in dimensionality

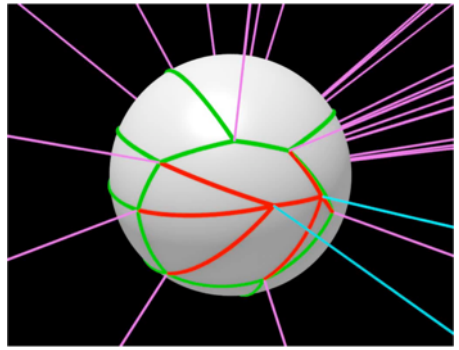






Jenny Vuong

Alex Muryy



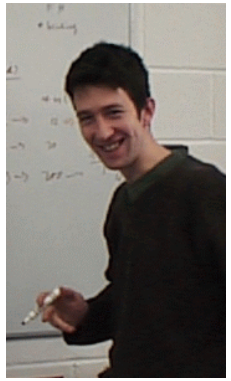
Thanks…

Luise Gootjes-Dreesbach

Peter Scarfe

James Stazicker

Miles Hansard



Andrew Fitzgibbon

University of Reading

Microsoft Research

EPSRC

[dstl]