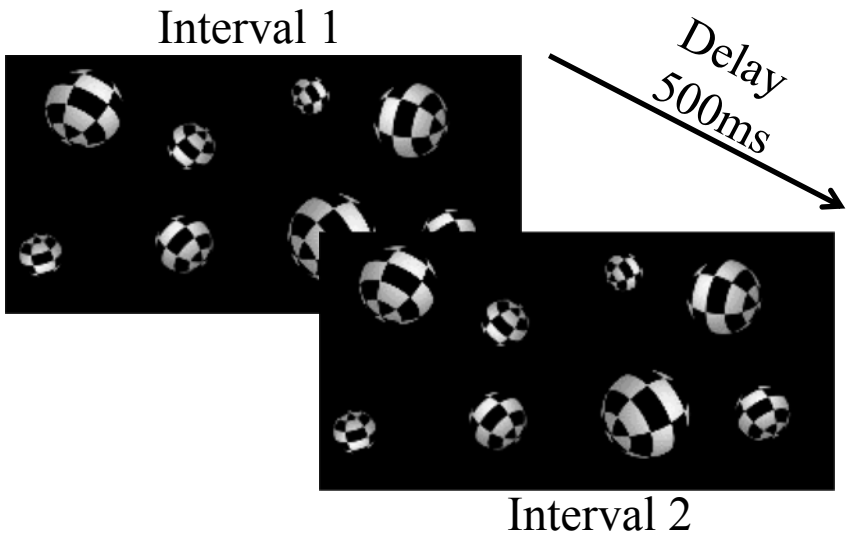
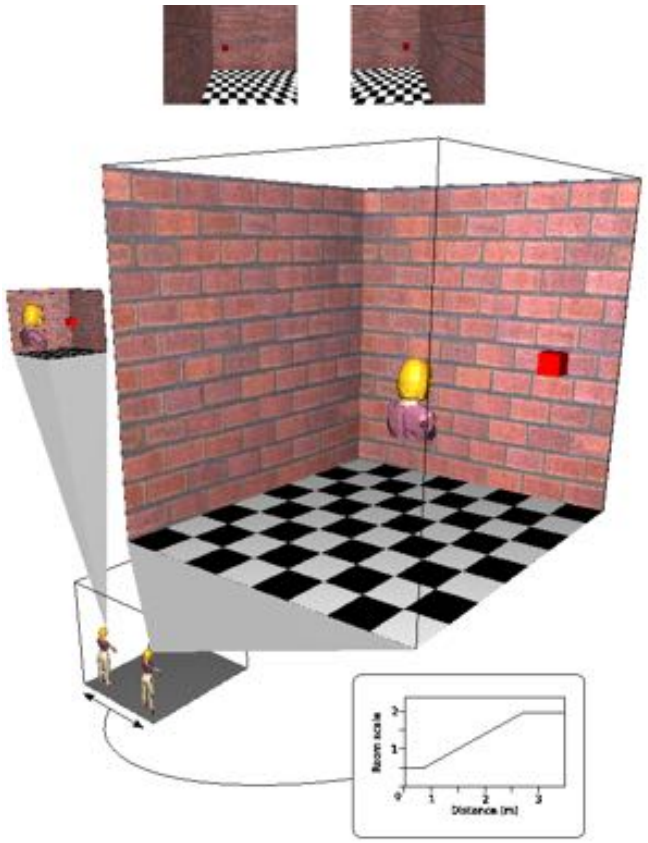


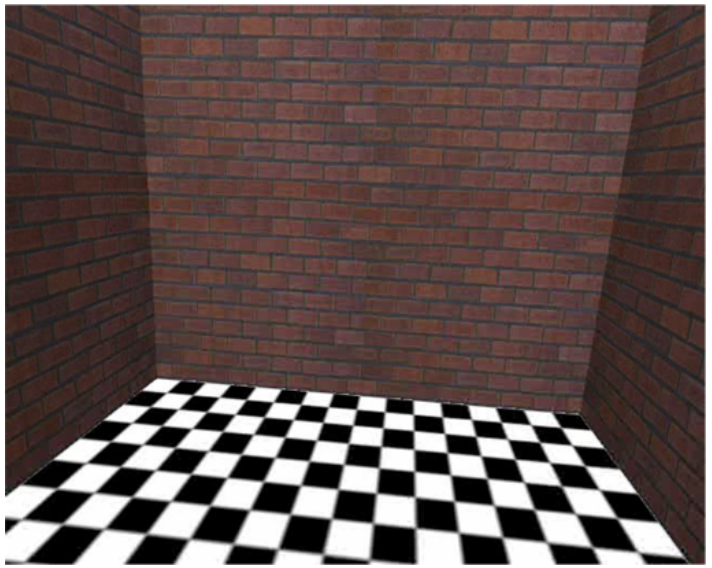
Luise Gootjes-Dreesbach, Peter Scarfe & Andrew Glennerster

RELATIONS IN 3D SCENES

University of Reading, UK

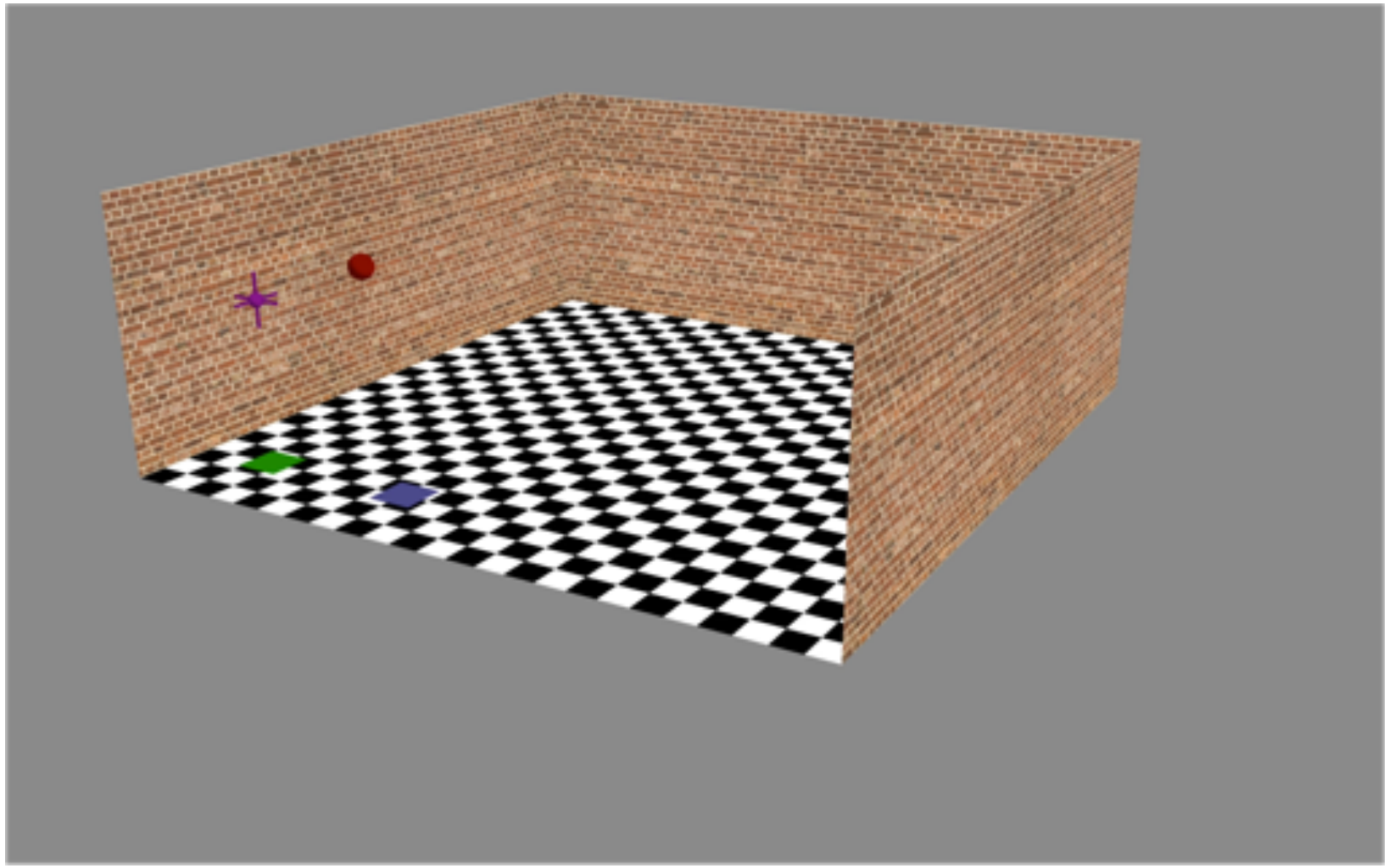
Result





People fail to notice that the room is changing size (eg by a factor of 4).

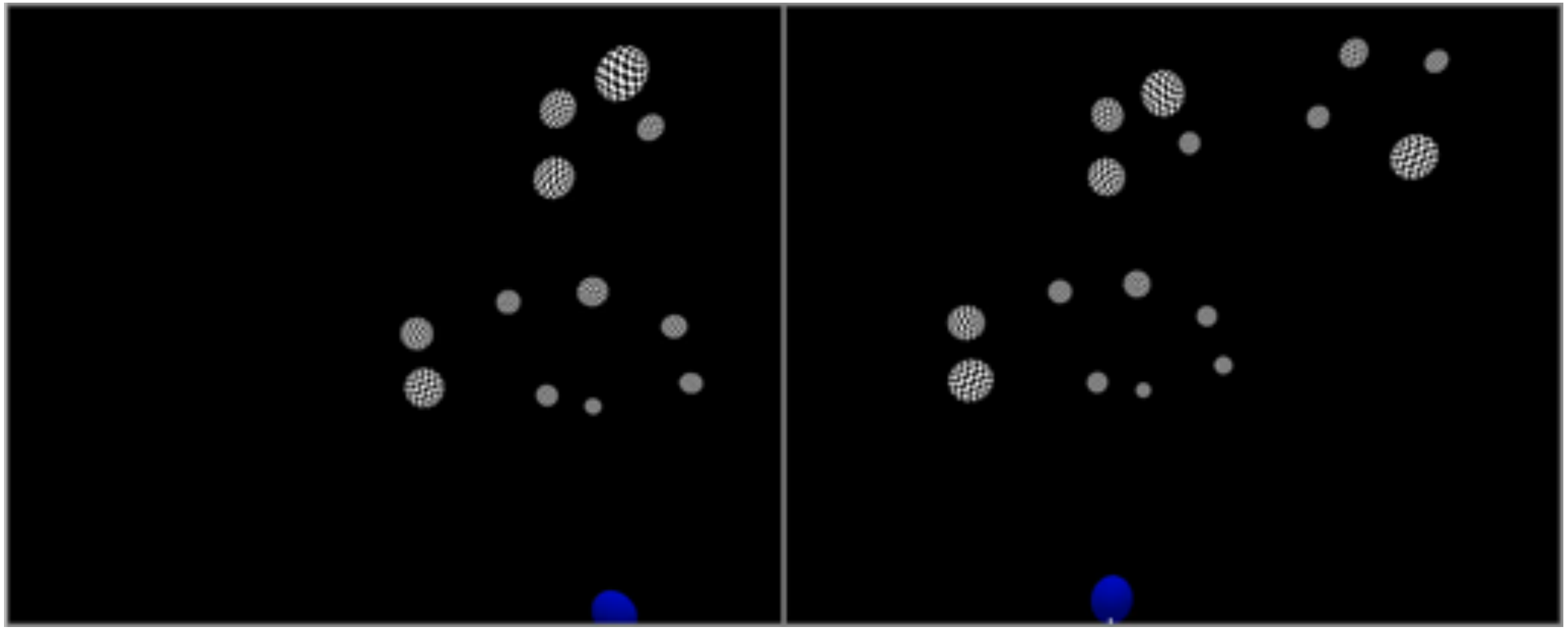
RELATIONS IN 3D SCENES: CHANGE BLINDNESS



Proprioceptive 'global' signal not visual

Participants do not notice this expansion/contraction

RELATIONS IN 3D SCENES: CHANGE BLINDNESS



Change localization task

RELATIONS IN 3D SCENES: CHANGE BLINDNESS

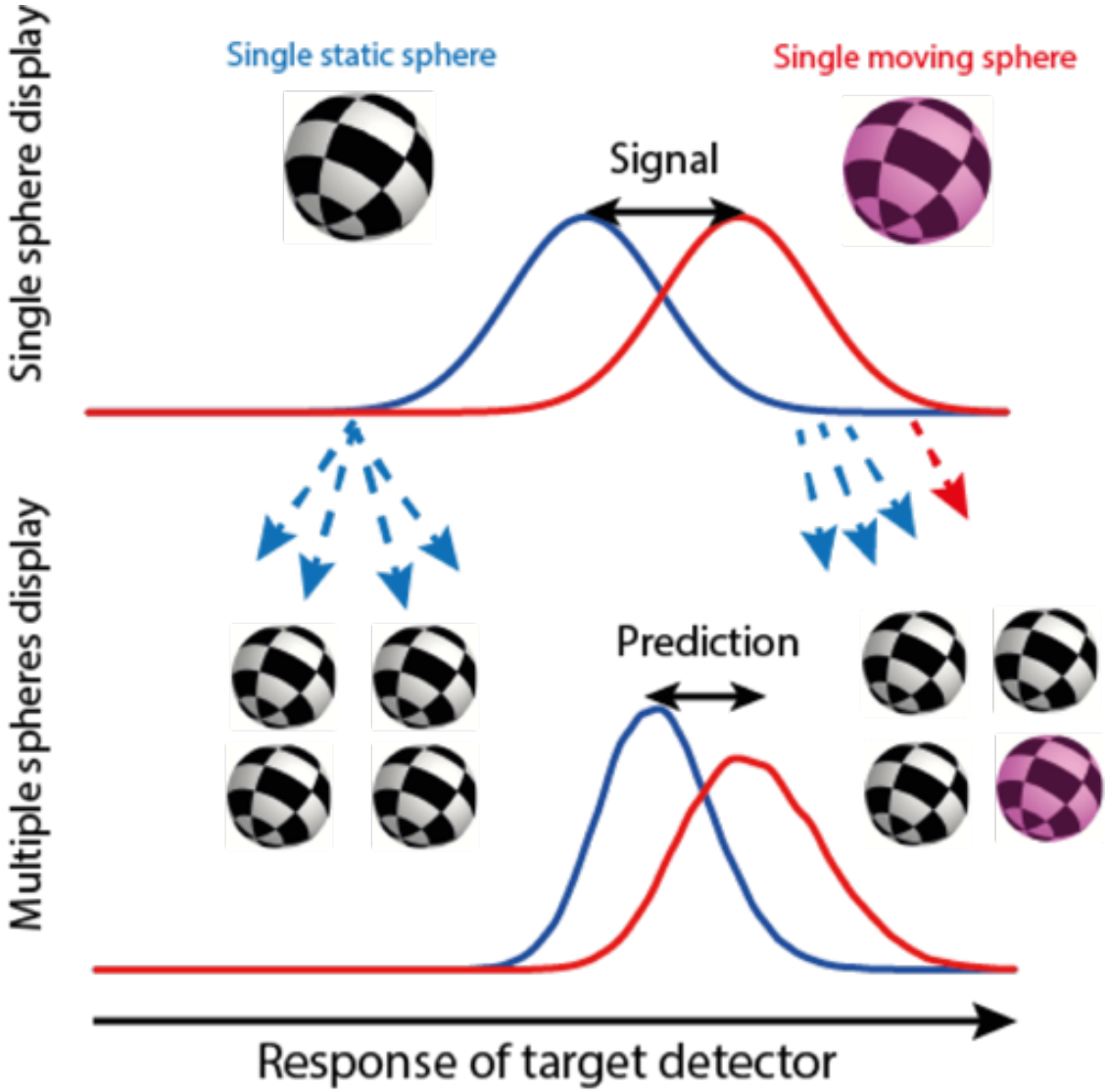
Previously

- Low overall performance in change localisation task -> internal representation can not be easily queried for exact positions
- Task-irrelevant relational information disrupts performance
- Not due to grouping or image change alone (colour grouping)

How important is relative, as opposed to absolute movement?

RELATIONS IN 3D SCENES: MODELLING

Relative importance of "global" vs "relative" signals



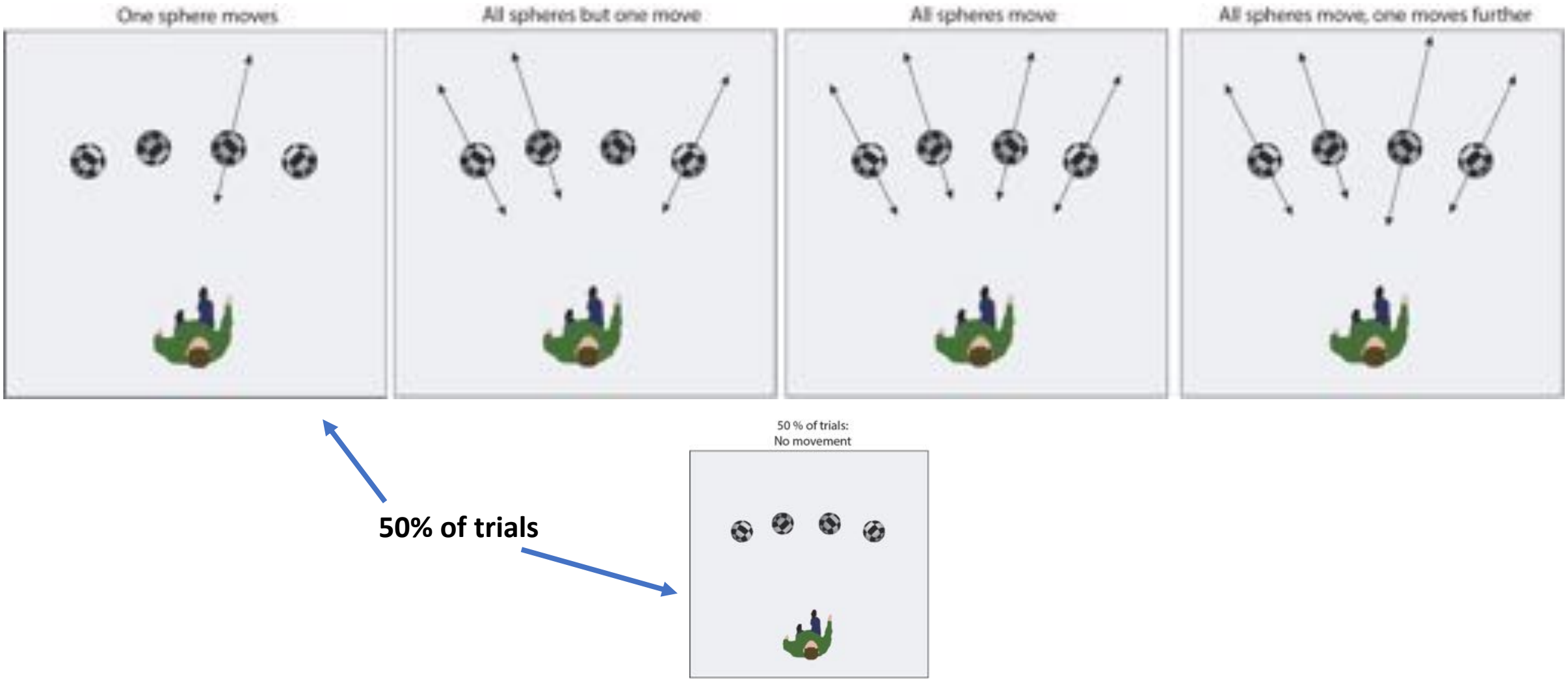
We consider 2 types of underlying signal:

- **Global:** individual sphere movement
- **Relative:** movement relative to other spheres

We are looking for the combination of global and relative signals that give the best model fit.

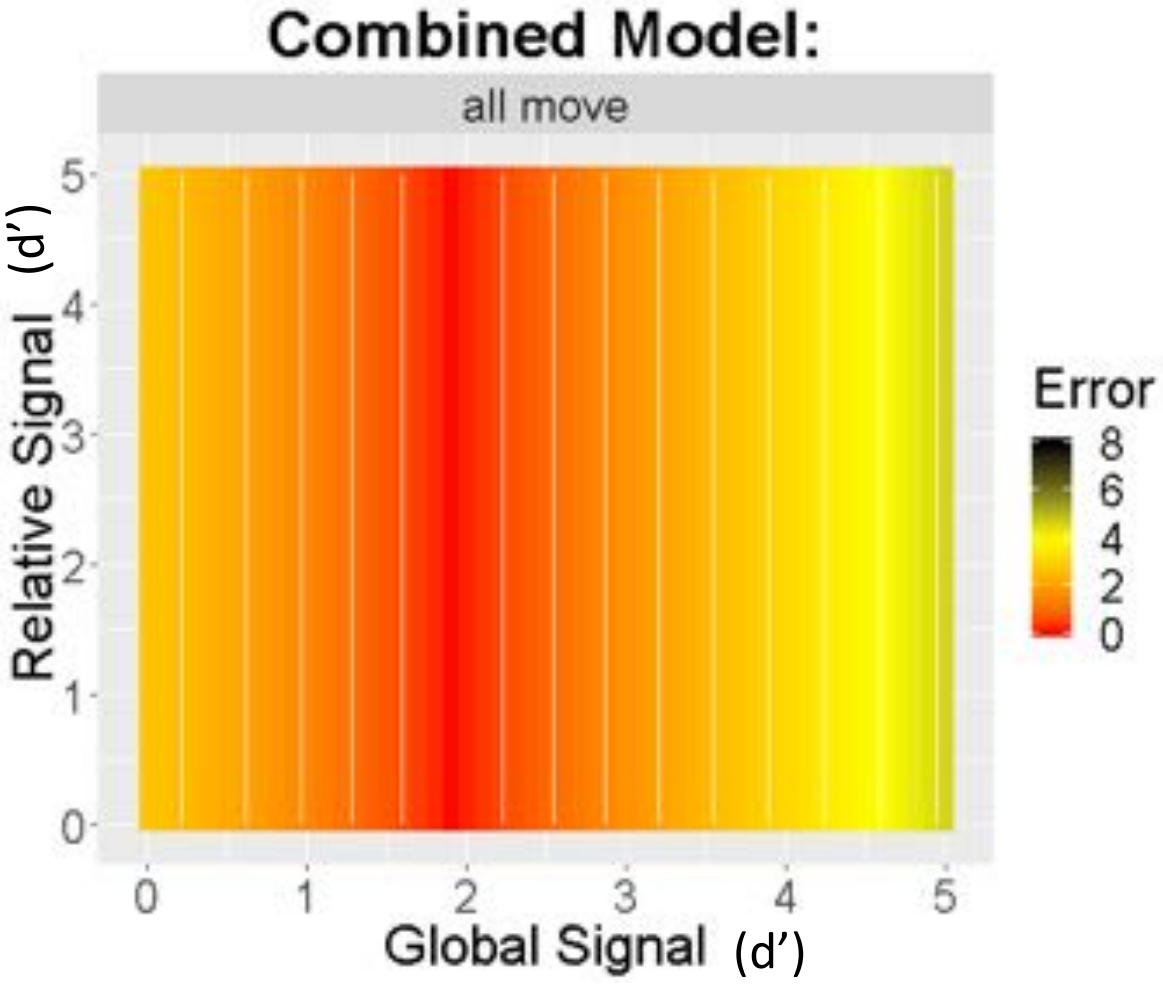
RELATIONS IN 3D SCENES: MODELLING

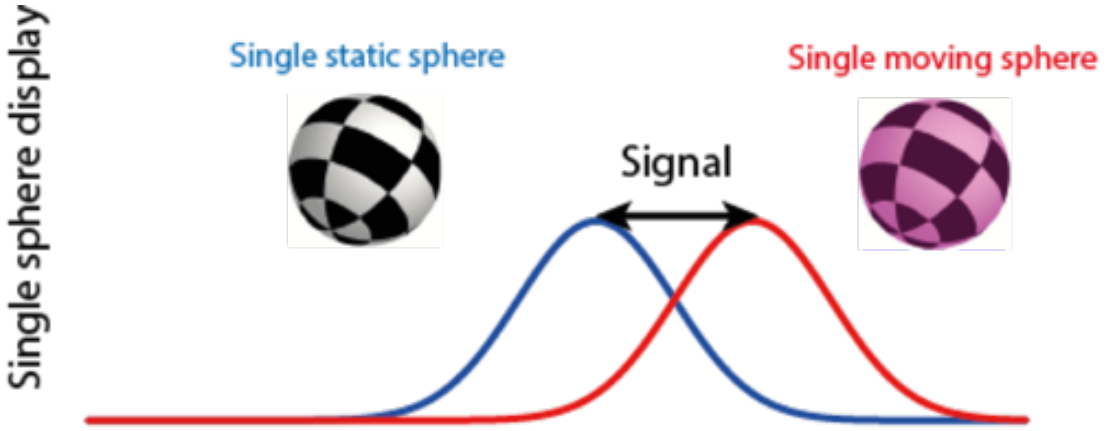
Relative importance of "global" vs "relative" signals



RELATIONS IN 3D SCENES: MODELLING

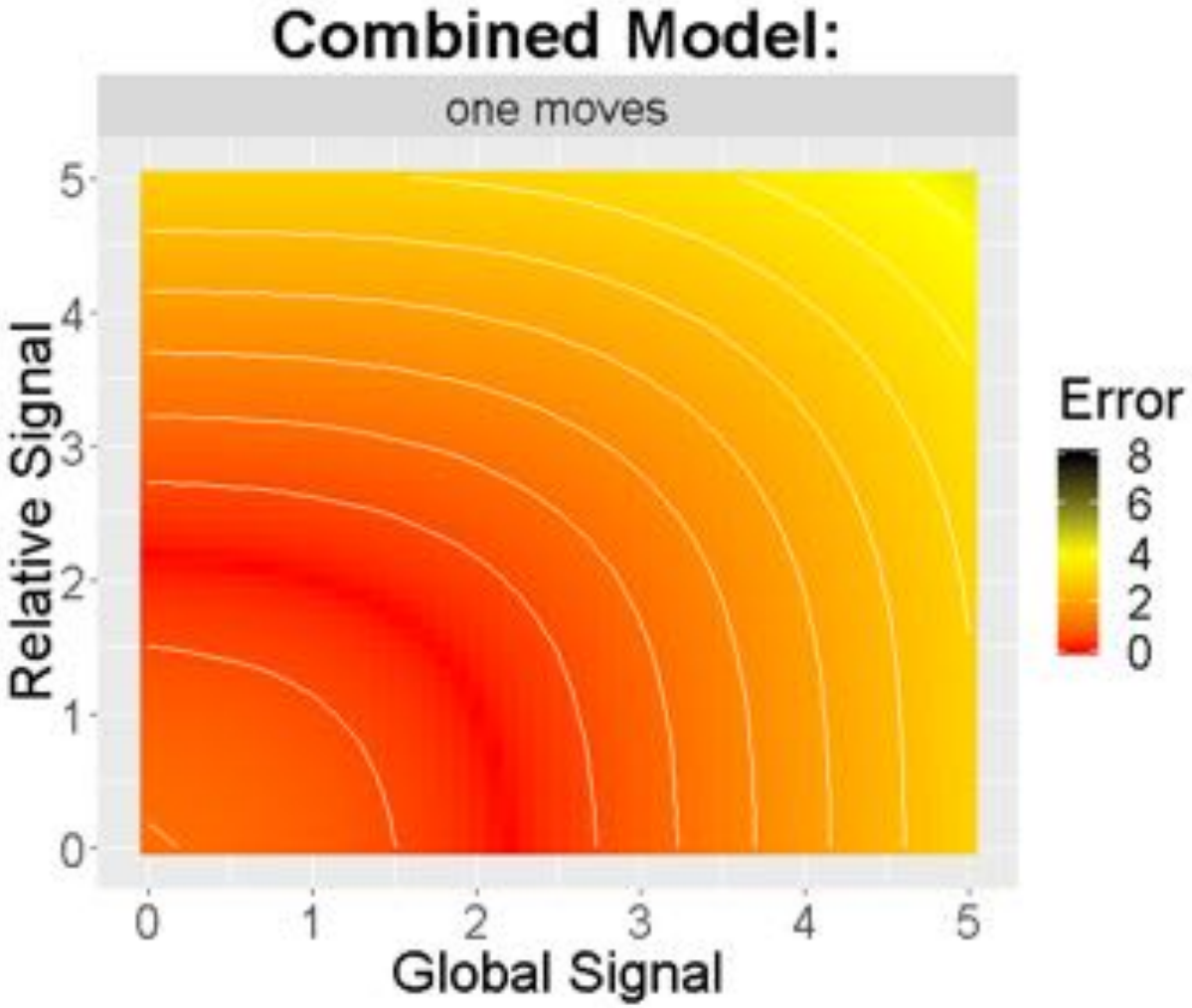
Relative importance of "global" vs "relative" signals

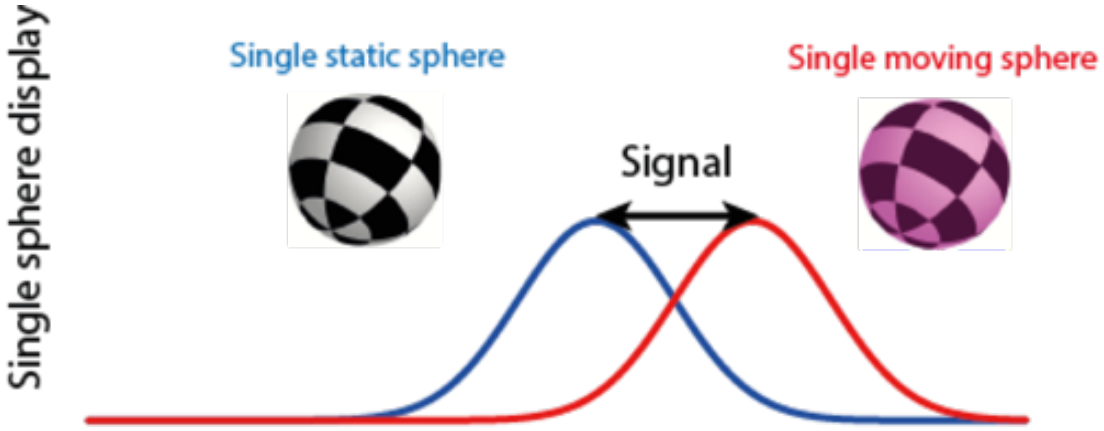


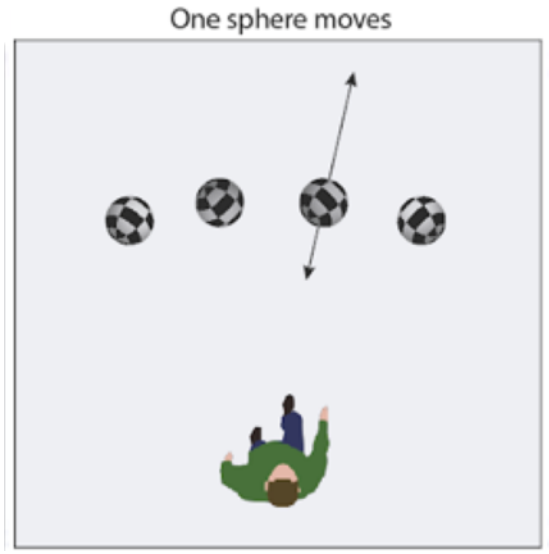


RELATIONS IN 3D SCENES: MODELLING

Relative importance of "global" vs "relative" signals

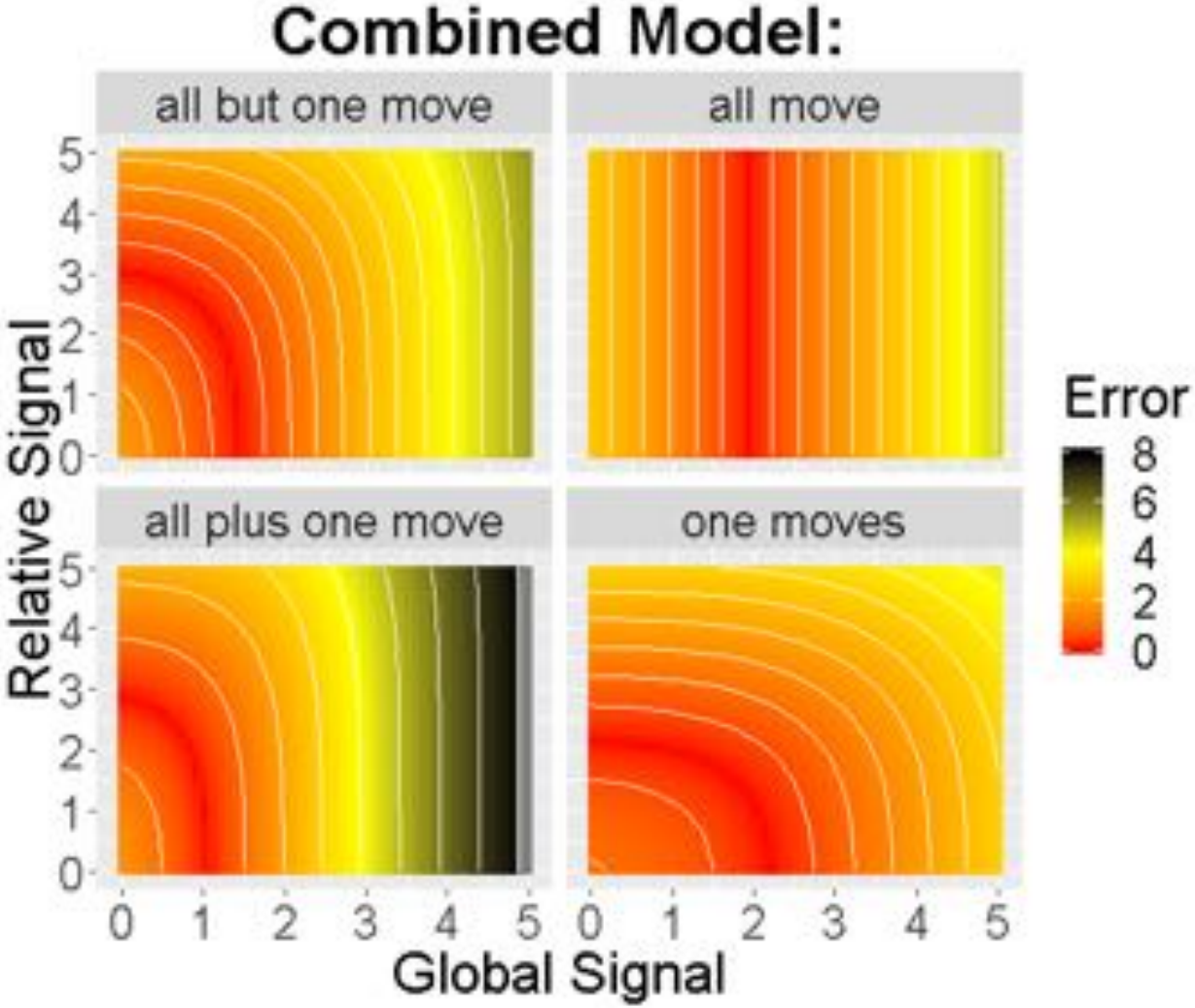


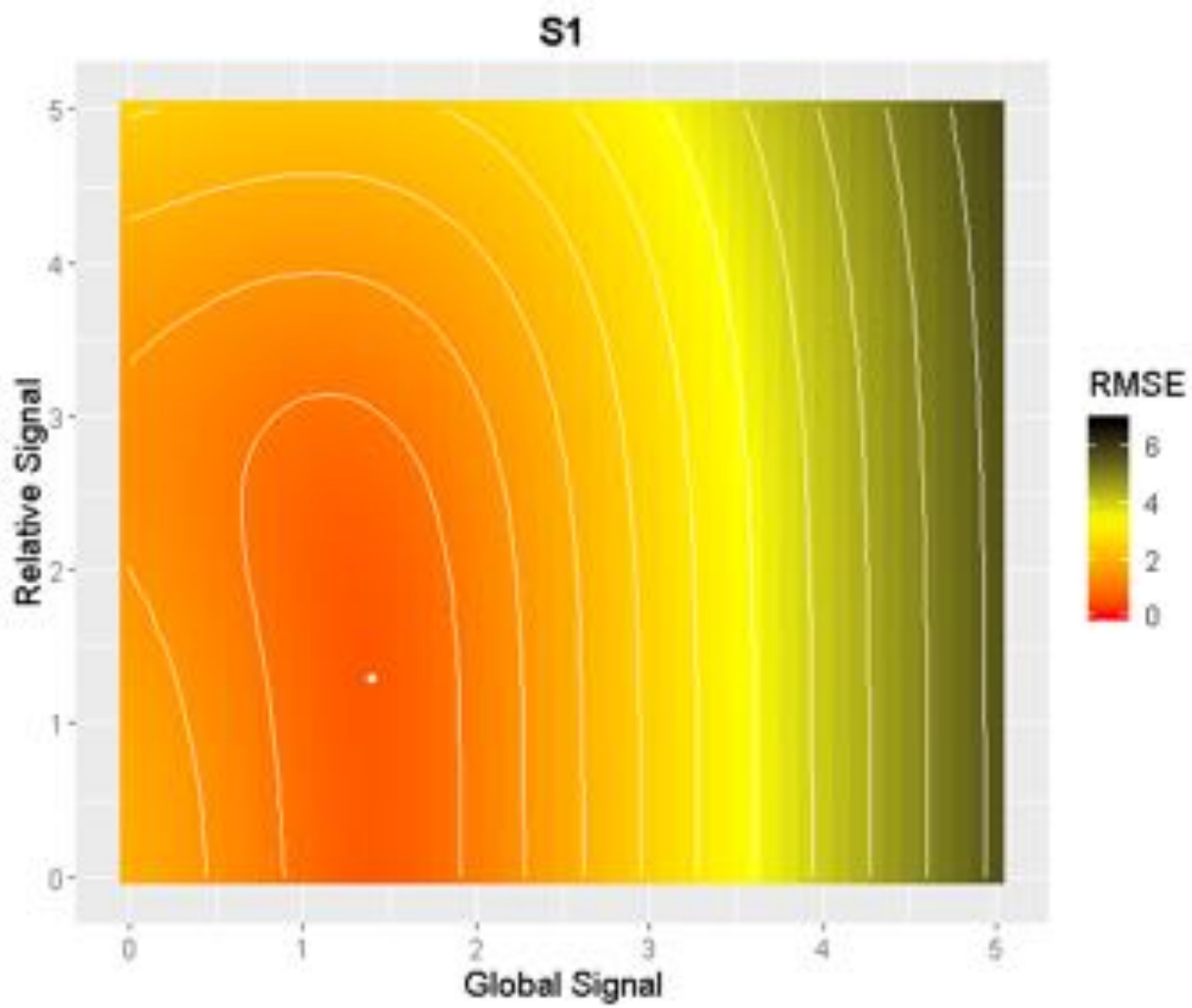




RELATIONS IN 3D SCENES: MODELLING

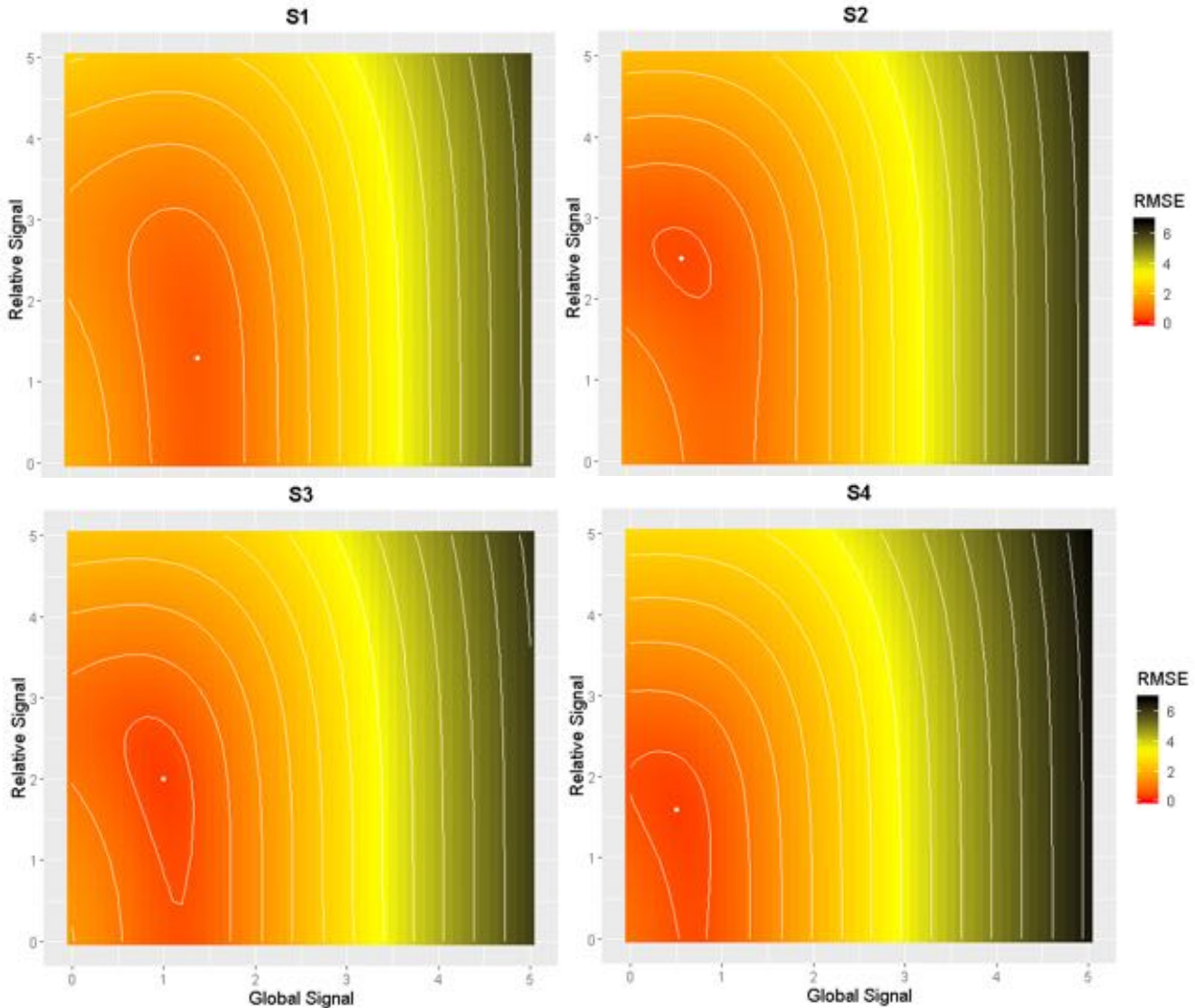
Relative importance of "global" vs "relative" signals





RELATIONS IN 3D SCENES: MODELLING

Relative importance of "global" vs "relative" signals



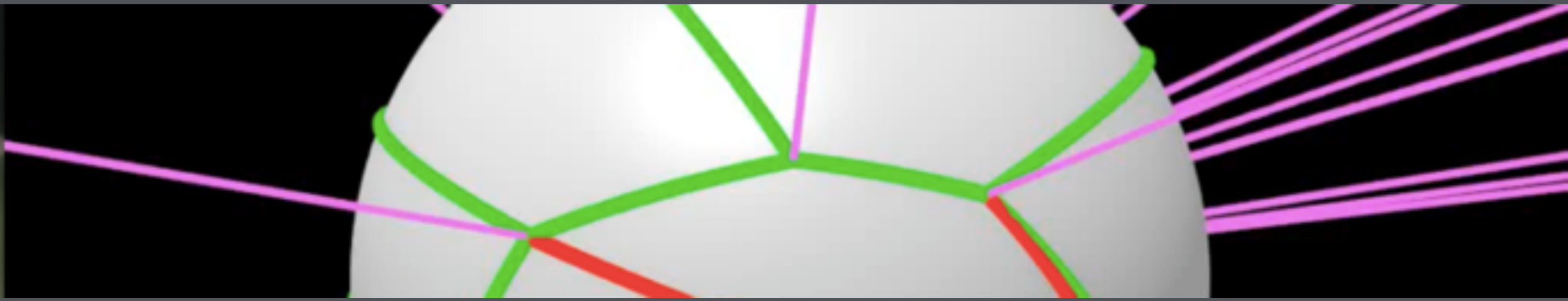
Conclusions

- Relative signal has a stronger influence
- All participants show some sensitivity to the global signal
- Considering these signals separately is different from an approach based on 3D reconstruction

School of Psychology and
Clinical Language Sciences

University of Reading

FISH GILLS AND FACES



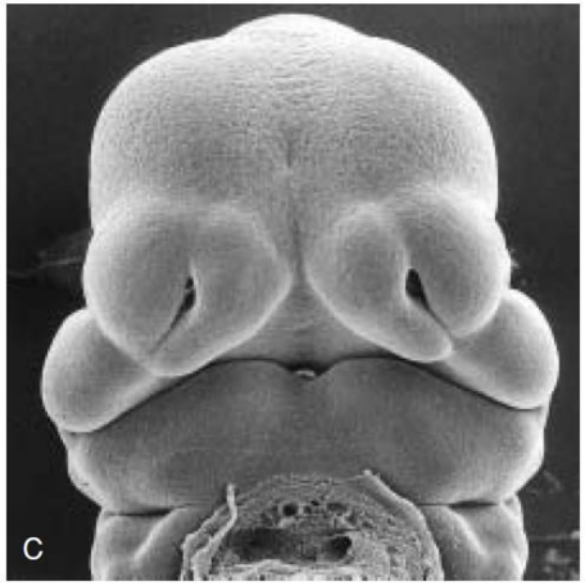
Andrew Glennerster

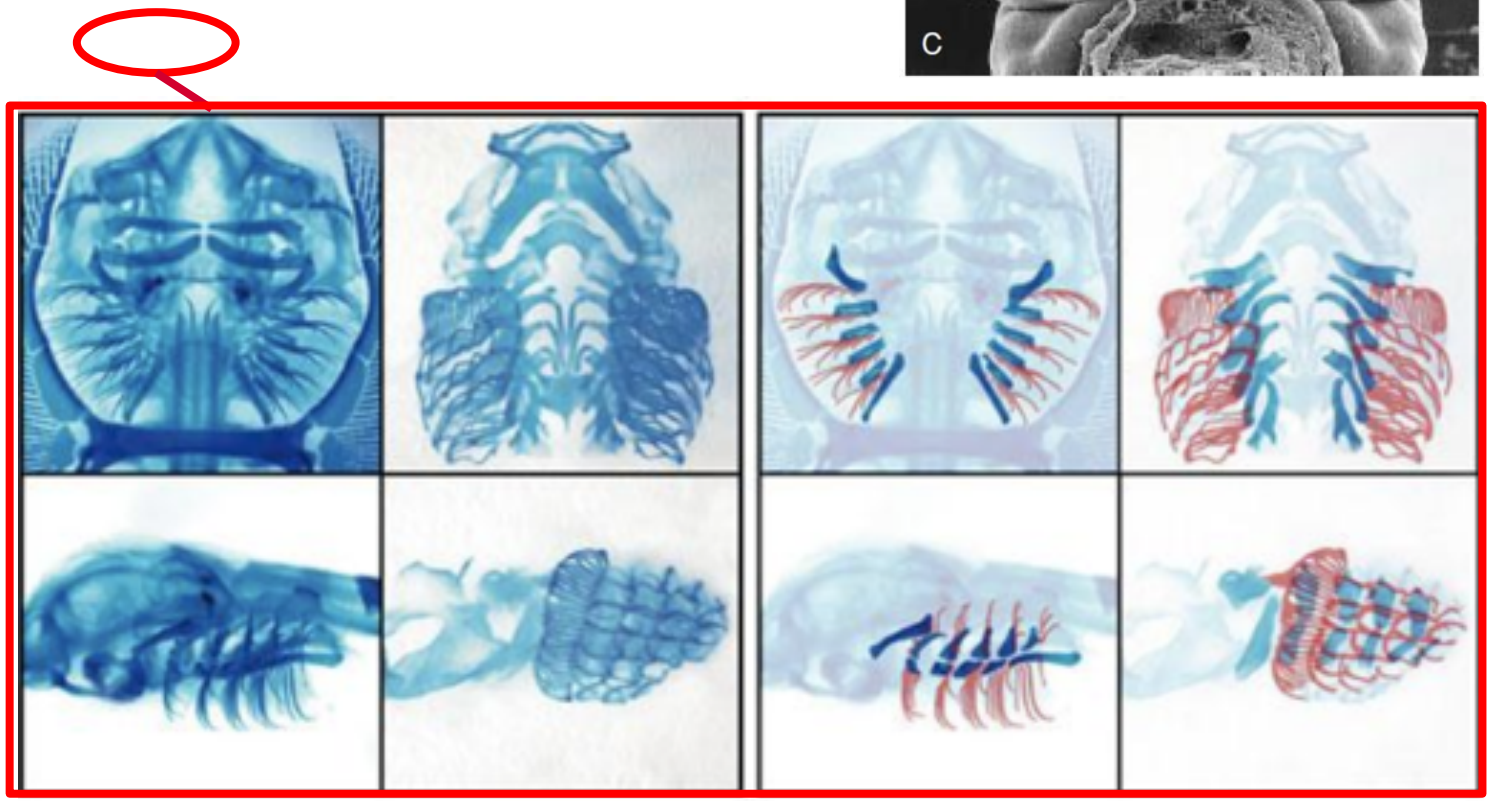
More ‘human-like’ reinforcement learning

Computer vision approaches have focused primarily on recognizing objects and simple actions using small image patches to train classifiers. These approaches do not take advantage of supporting contextual information in the broader scene, or knowledge of the world, and are thus inefficient, brittle and inaccurate for natural images in particular. For humans, however, scene understanding is a simple task. One reason for this is that humans possess what we call visual commonsense, that is, the ability to rapidly understand complex scenes without much deliberation using their knowledge and contextual reasoning consistent with the available information.

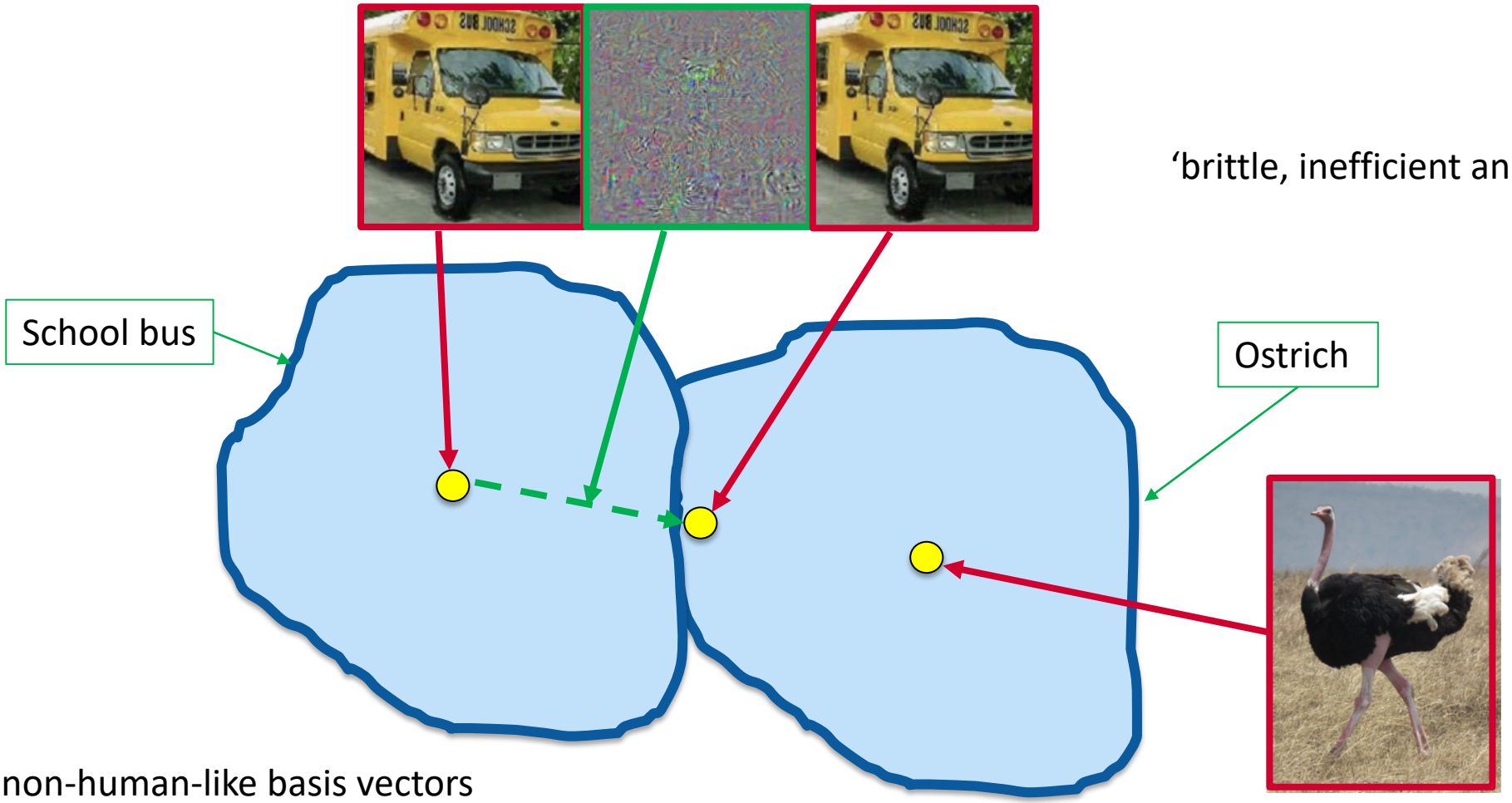
Today

- Question: what RL would have to do to be more ‘human-like’?
- Answer: build up tasks gradually (hierarchically) and add dimensions as the system learns new tasks


C

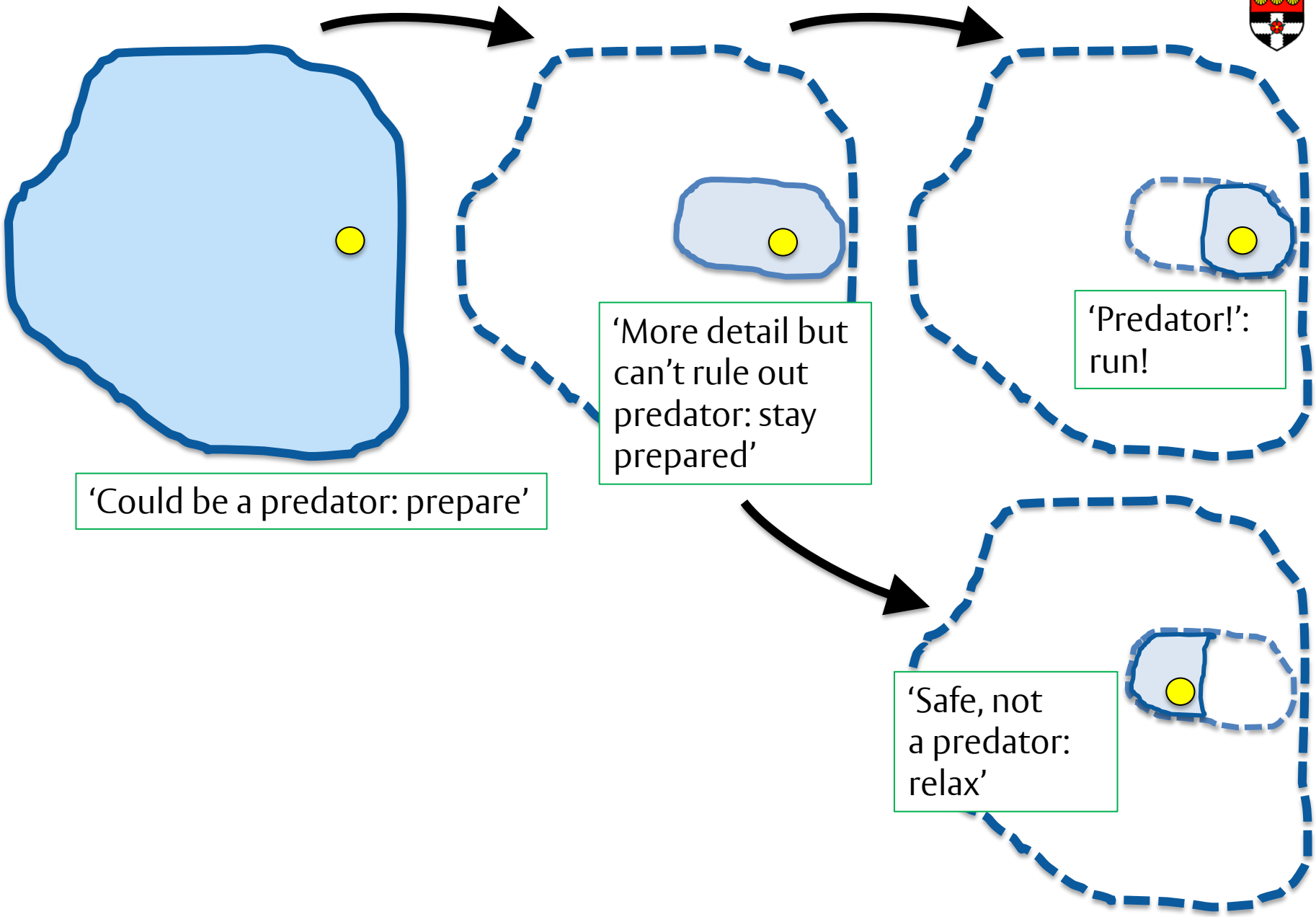


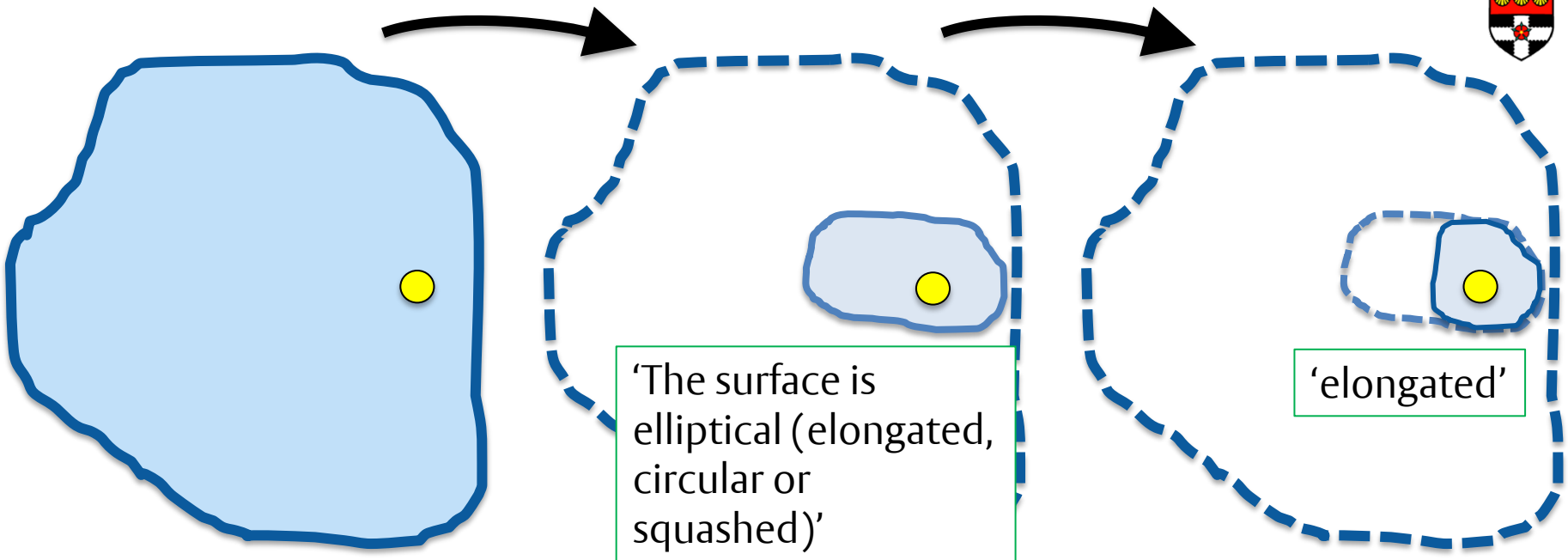
More ‘human-like’ reinforcement learning



‘brittle, inefficient and inaccurate’

- completely non-human-like basis vectors
- how should RL learn human-like basis vectors?


'Could be a predator: prepare'
'More detail but can't rule out predator: stay prepared'
'Predator!': run!
'Safe, not a predator: relax'

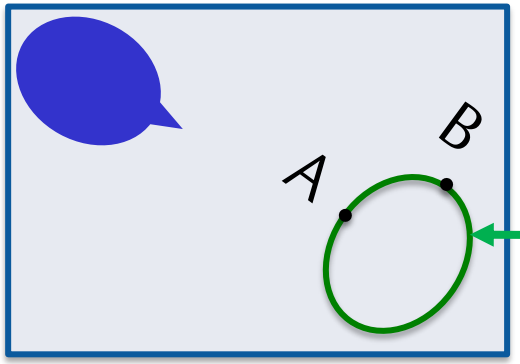


‘A is in front of B’

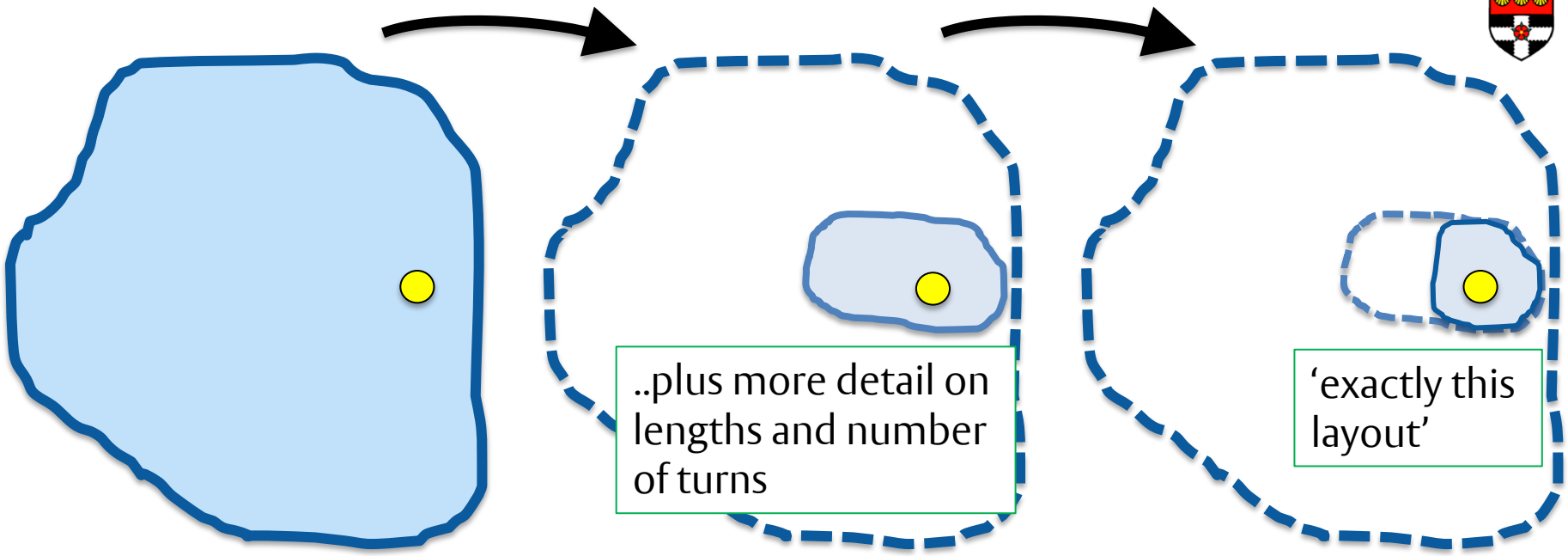
‘The surface is elliptical (elongated, circular or squashed)’

‘elongated’

‘squashed’



i.e. perception as well as action

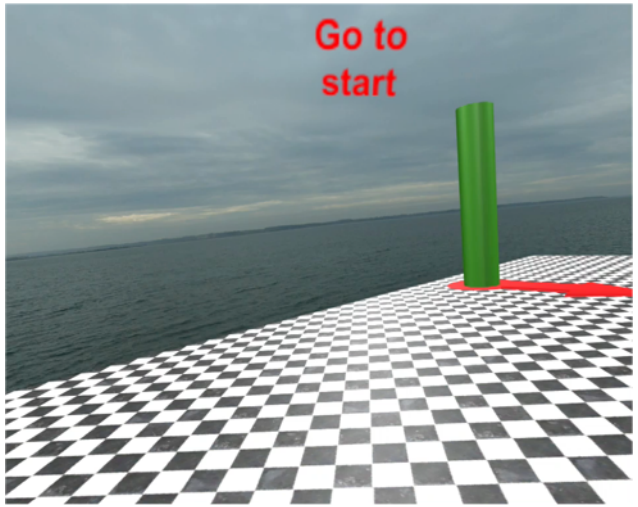


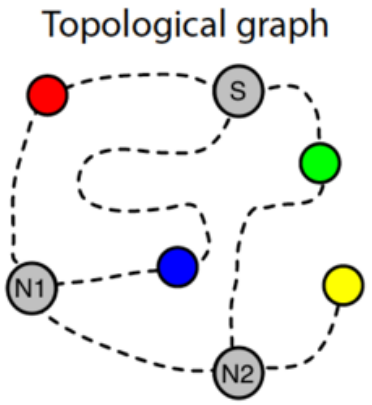
Topological layout

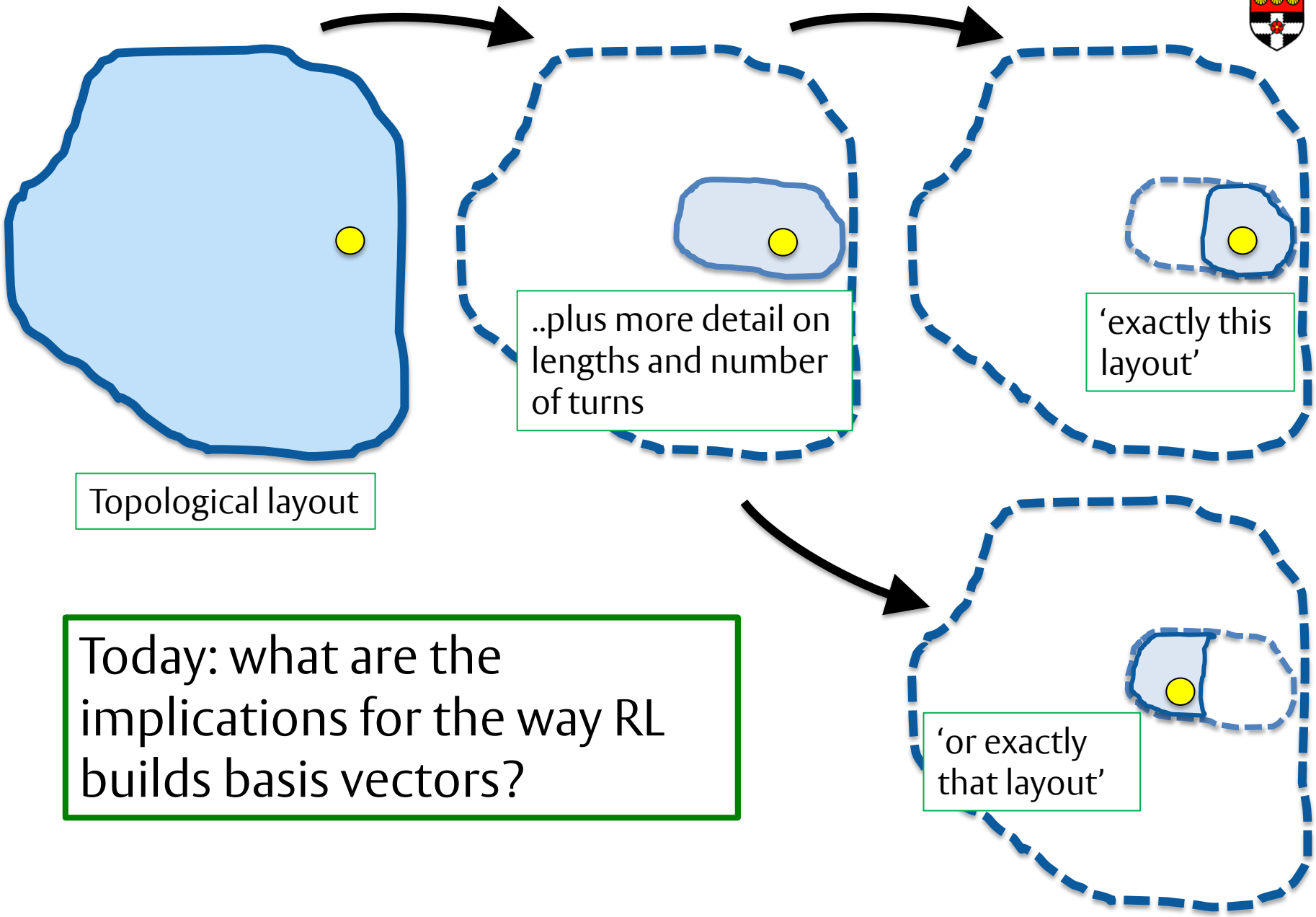
..plus more detail on lengths and number of turns

'exactly this layout'

'or exactly that layout'

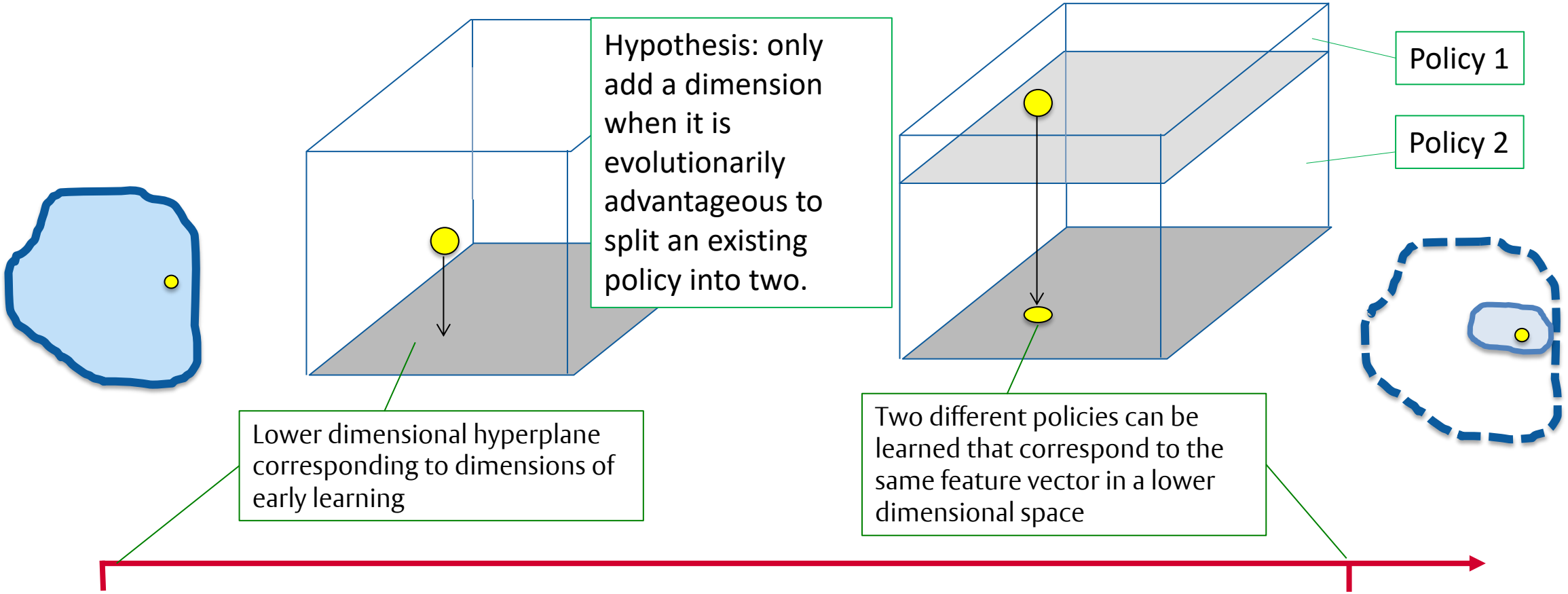






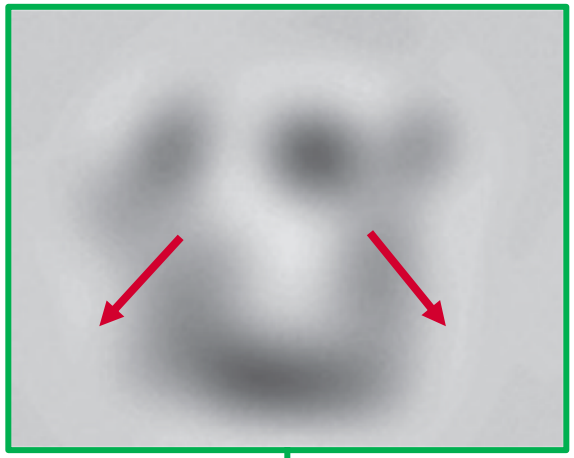
Today: what are the implications for the way RL builds basis vectors?

Adding dimensions as tasks evolve



Contexts for action/policies over the first 200ms of seeing a face

(Phillipe spoke yesterday…)



Action:
orient towards possible face

Action:
definite face; identity?

Action:
identify individual; shift gaze to help with next question

It is tempting to speculate that perception recapitulates evolution – reinforcement learning does not seem to do that (eg ostrich/schoolbus)

Early in learning OR
early in perception, eg 10ms exposure

Late in learning OR
late in perception, eg 200ms

Learning appropriate basis vectors: categorization

Action:
"An old pot"

Action:
"A delftware pharmaceutical pot"

Action:
"… and it is 18th Century and English"

Early in learning OR
early in categorization process

Late in learning OR
late in categorization process

Learning appropriate basis vectors: evolution

Action:
chemotaxis;
phototaxis

Early in learning (low dimensional: higher dimensions not yet invented)

Late in learning (higher dimensions used to split categories hierarchically)

Learning appropriate basis vectors: evolution

Action:
chemotaxis;
phototaxis

Action:
lots;
still no brain

Early in learning (low dimensional: higher dimensions not yet invented)

Late in learning (higher dimensions used to split categories hierarchically)

Learning appropriate basis vectors: evolution

Action:
chemotaxis;
phototaxis

Action:
lots;
still no brain

Action:
need not be overt
movements

Early in learning (low dimensional space: higher dimensions not yet invented)

Late in learning (higher dimensions used to split categories hierarchically)

Learning appropriate basis vectors: action

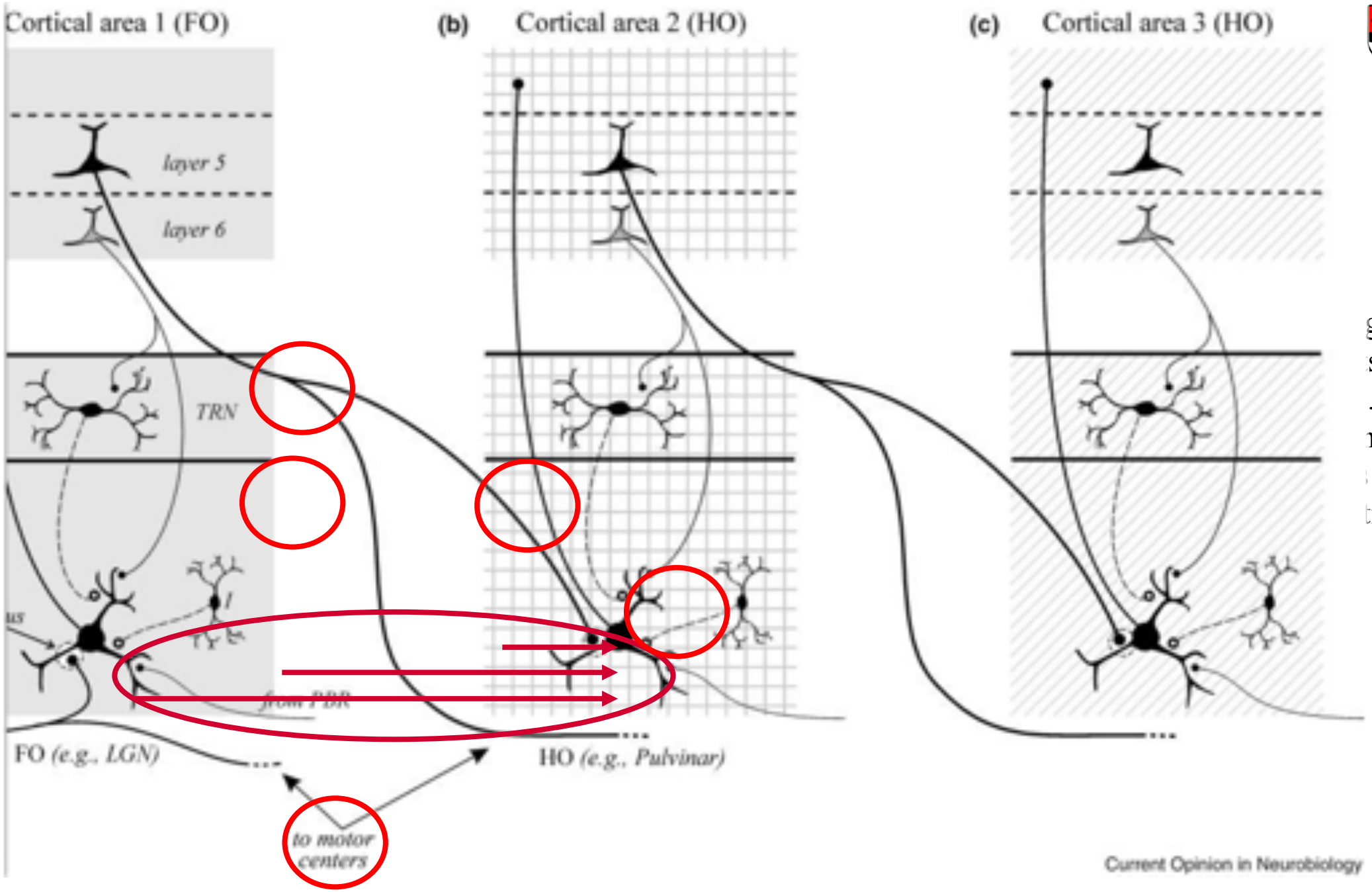
Action:
move proximal muscles

Action:
extend limb

Action:
guide hand

Early in learning OR
early in the action

Late in learning OR
late in the action



University of Reading

gh thalamus to cortex
s that our perceptual
… Whatever the
ıy or all inputs
that also target motor
tion."

What is left to do?

Mark Edmonds

University of Reading

	Easy (or Given)	Hard
Physics Engine	Learning representation of environment Transfer learning	Computing representation of current environment Task inclusion
RL	Computing representation of current environment Task inclusion	Learning representation of environment Transfer learning

Similar to known capacities of brains

You guys keep on doing more of this

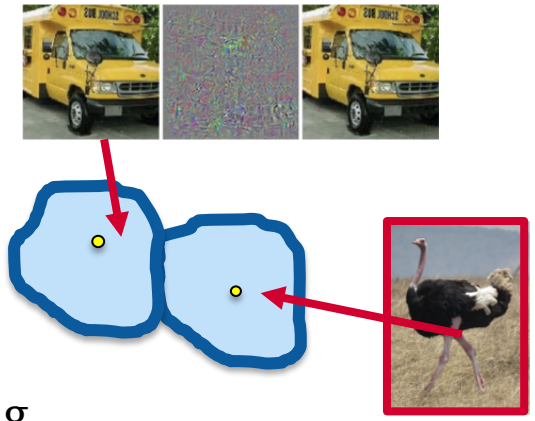
More ‘human-like’ reinforcement learning

Yesterday:

- how human 3D vision might be based on a set of policies
- some examples of hierarchical learning, eg navigating a maze

Today

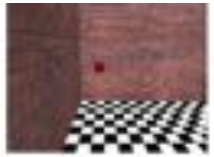
- Question: what RL would have to do to be more ‘human-like’?
- Answer: build up tasks gradually (hierarchically) and add dimensions as the system learns new tasks (to get more ‘human-like’ basis vectors)



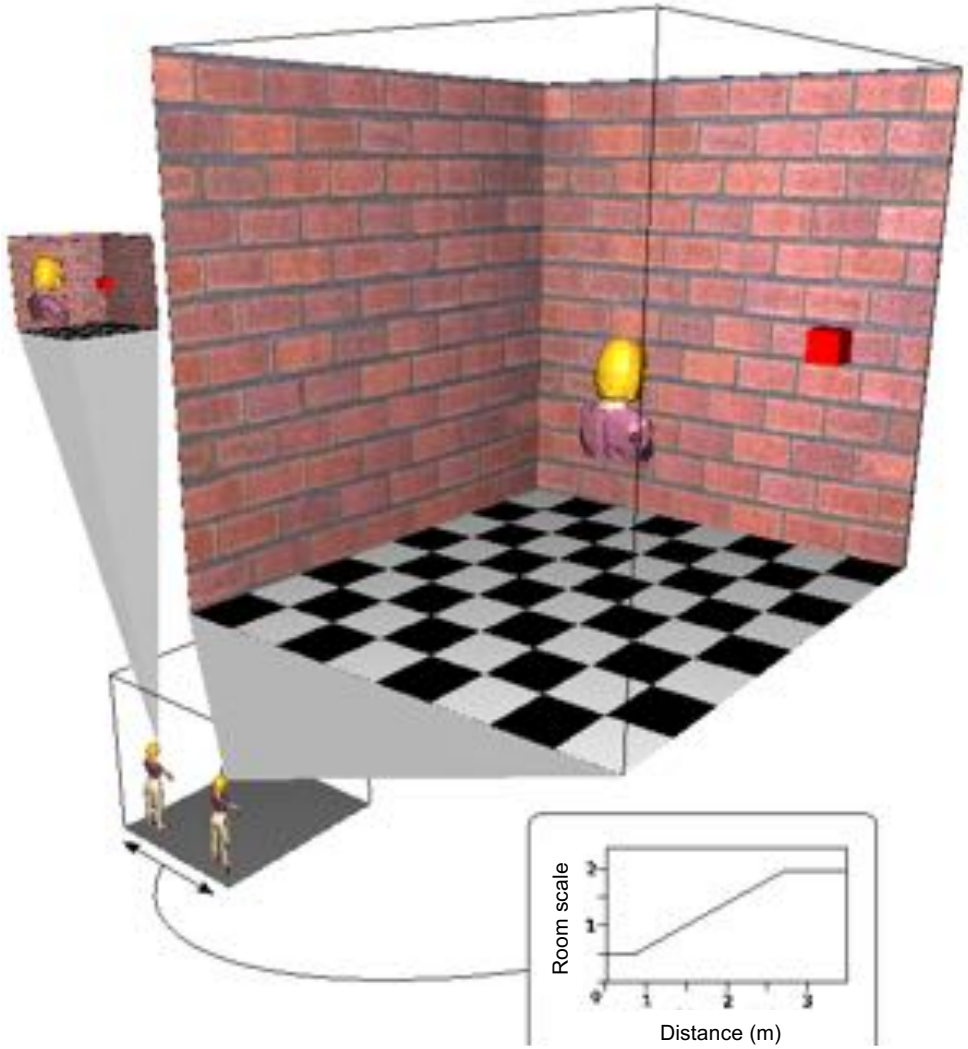
- I am not qualified to discuss the literature but incremental learning is relevant, e.g.
  - Chaudry, Dokania, Ajanthan, Torr (2018) arXiv:1801.10112
  - Wang, Y. X., Ramanan, D., & Hebert, M. Growing a brain: Fine-tuning by increasing model capacity. CVPR 2017.

University of Reading

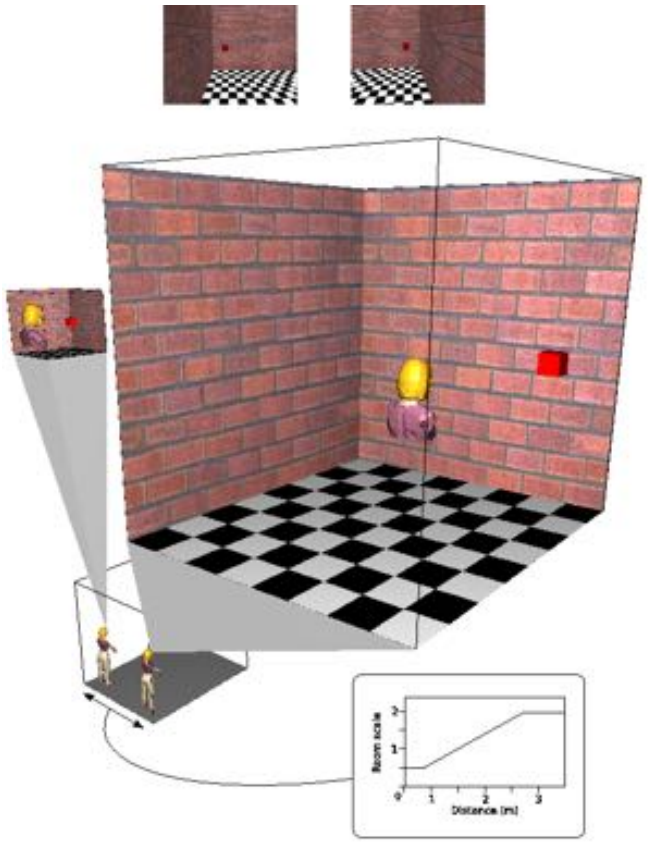
View in small room

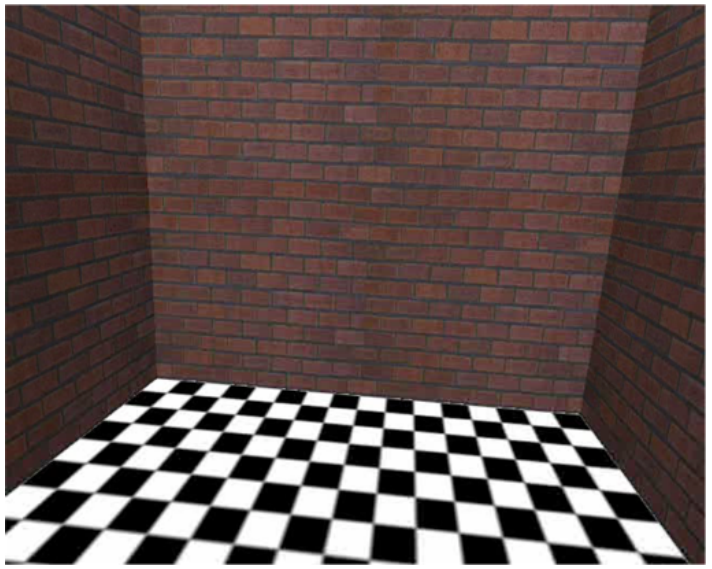


View in large room



Result

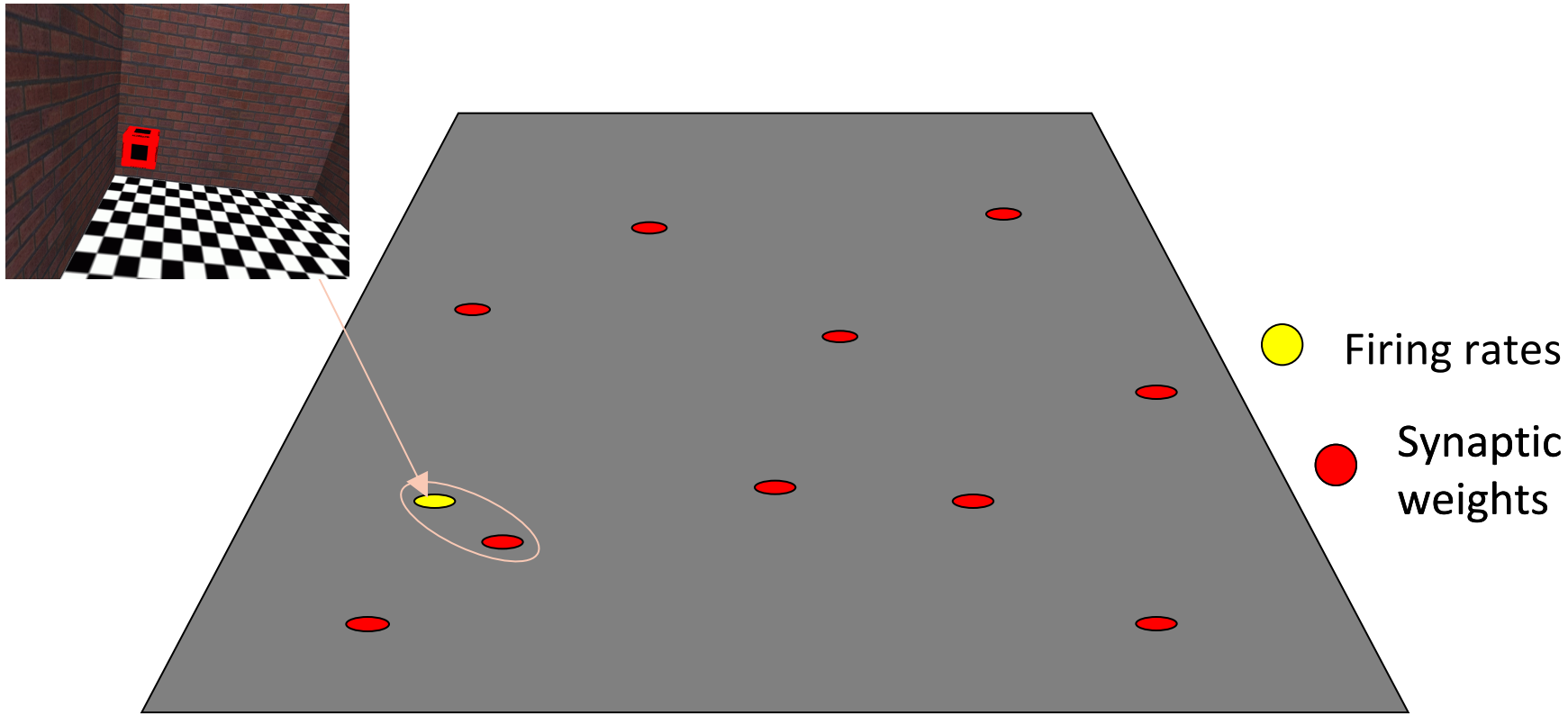




People fail to notice that the room is changing size (eg by a factor of 4).

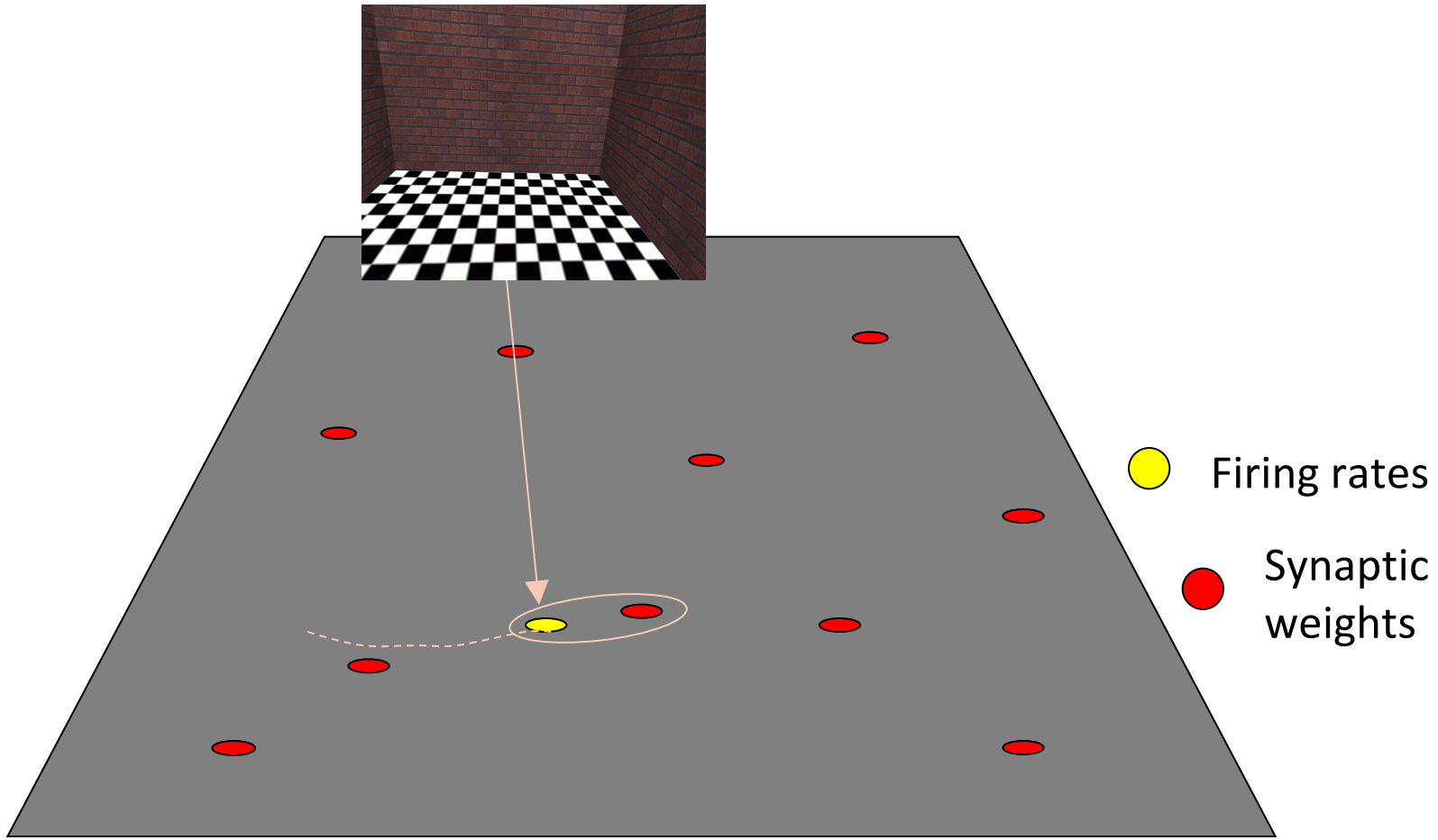
An expected path across sensory space

University of Reading



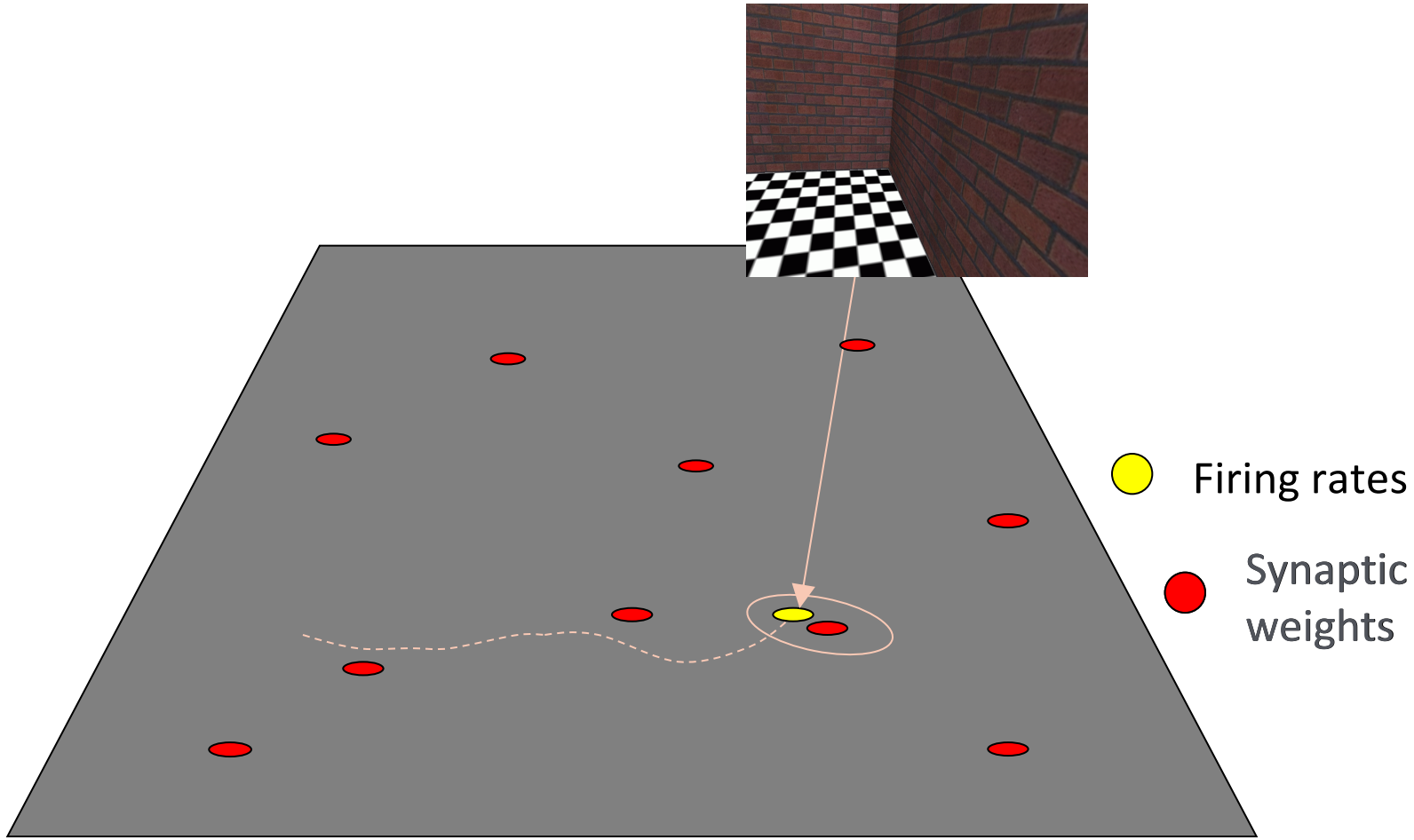
Marr (1969); Albus (1971)

An expected path across sensory space



Marr (1969); Albus (1971)

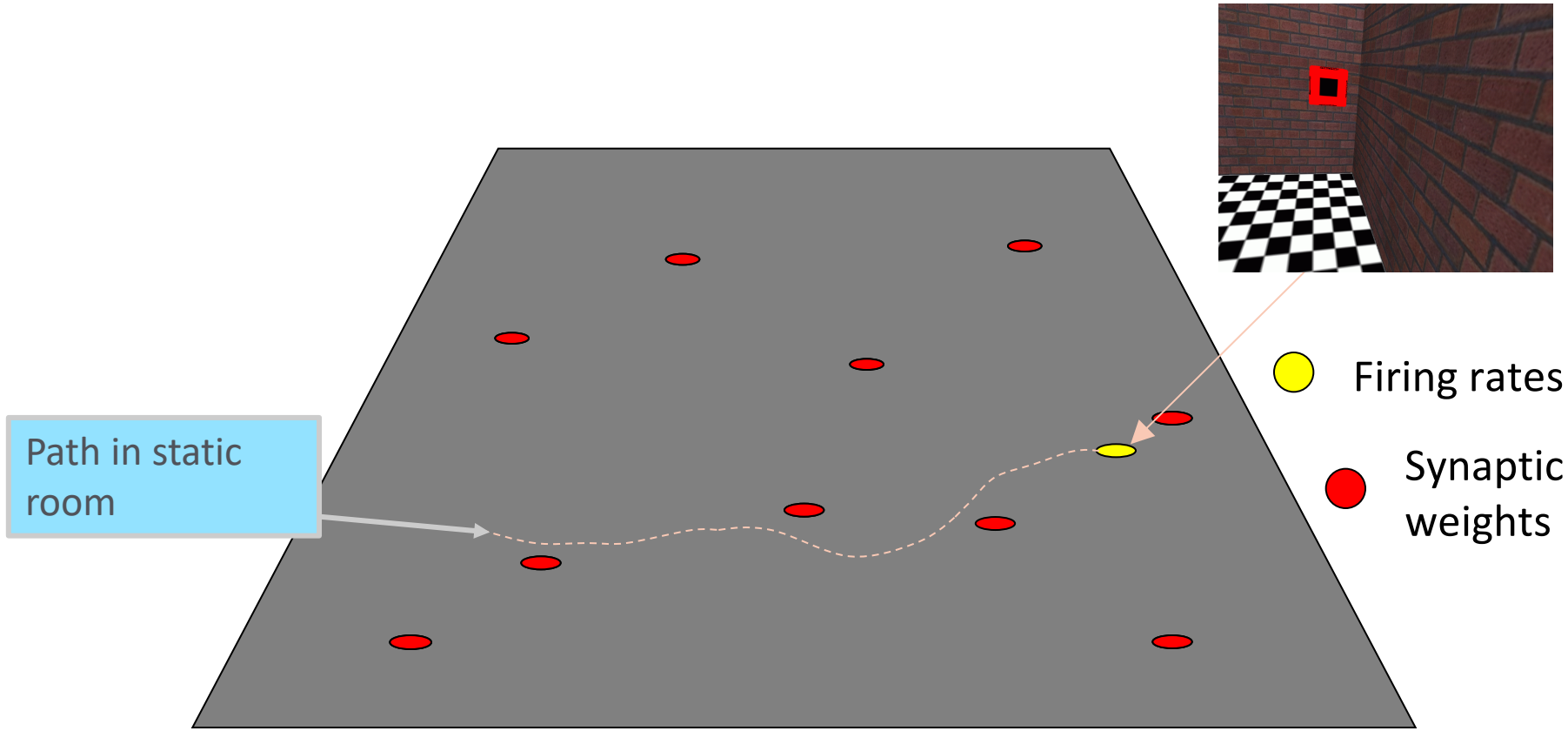
An expected path across sensory space



Marr (1969); Albus (1971)

An expected path across sensory space

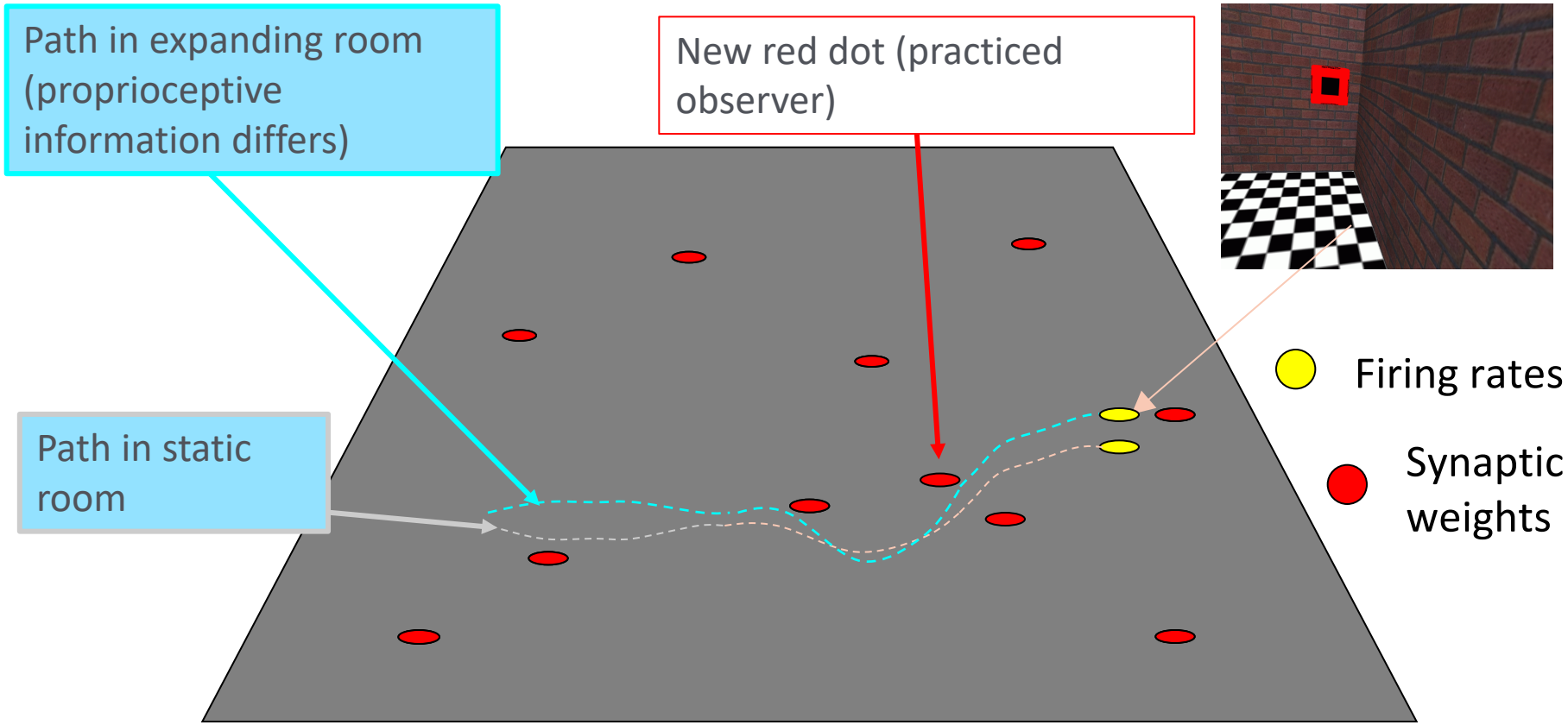
University of Reading



Marr (1969); Albus (1971)

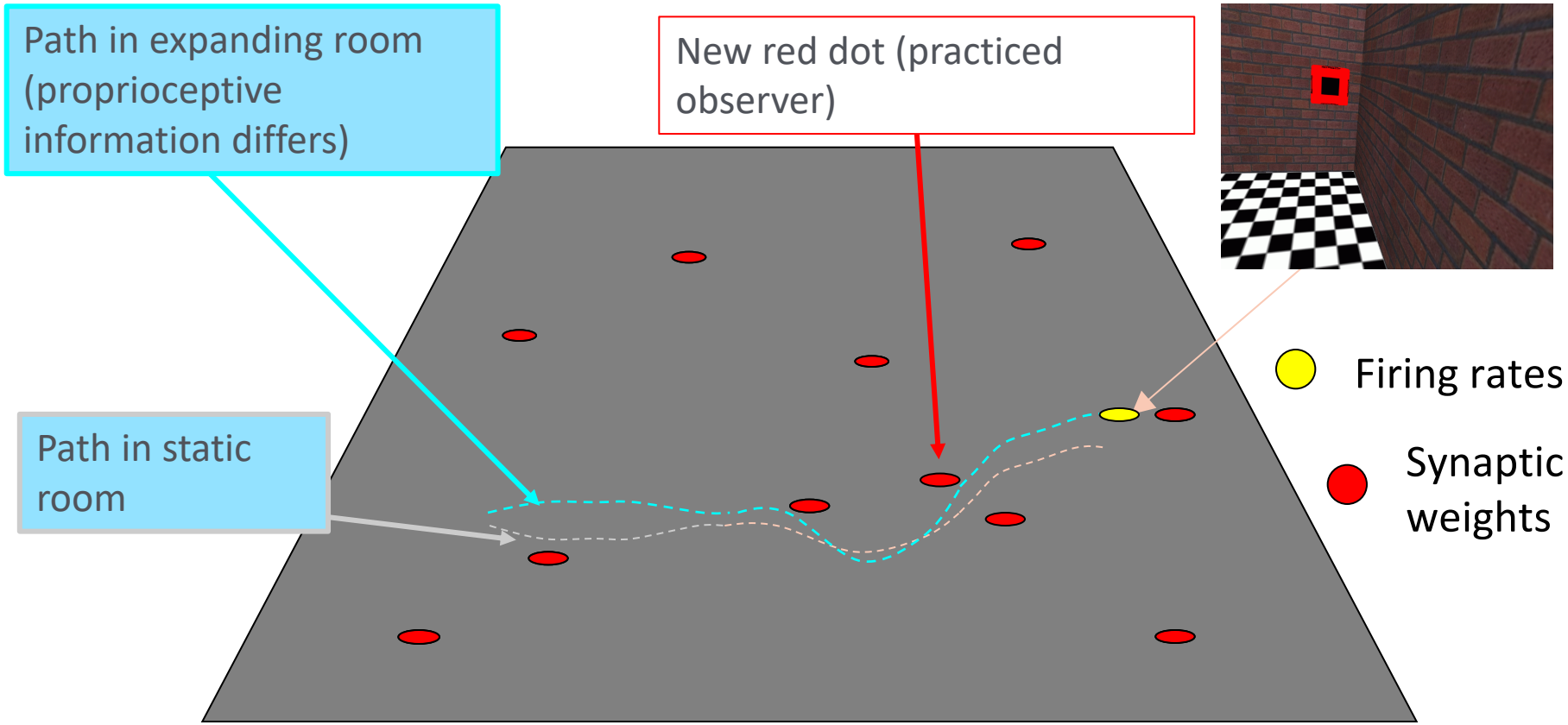
An expected path across sensory space

University of Reading



Marr (1969); Albus (1971)

Learning



Marr (1969); Albus (1971)