Multivariate classification shows associative learning reduces working memory load

William X. Q. Ngiam

School of Psychology, Adelaide University

Symposium: A model collaboration

AUSTRALASIAN COGNITIVE NEUROSCIENCE SOCIETY

MELBOURNE – 26 NOV 2025

A quick introduction to me

I study **attention and working memory** – how information is selected and held in mind for ongoing perception and cognition.

A quick introduction to me

I study **attention and working memory** – how information is selected and held in mind for ongoing perception and cognition.

One key feature of this system is that it is **capacity-limited.** Measuring the capacity limits of visual attention and working memory requires understanding how information is **represented** in working memory.

The point of this talk

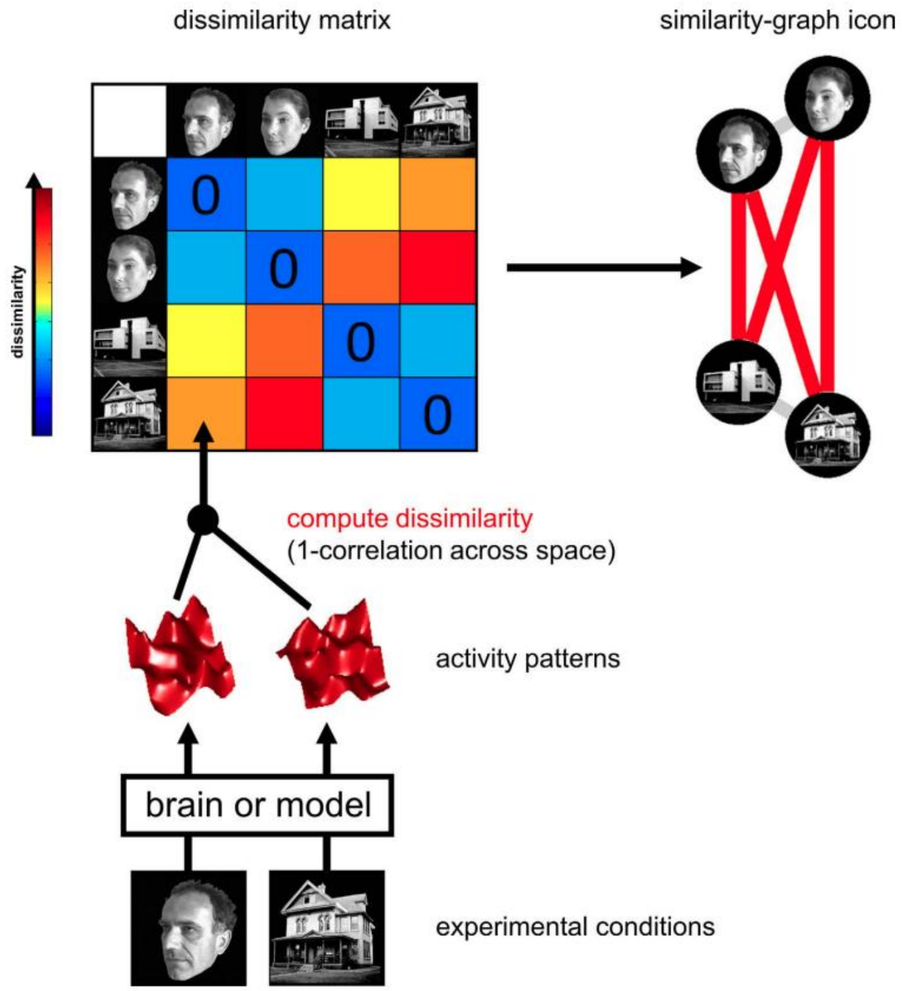
Cognitive theories ought to make predictions about neural signals and decoding results, such as those from **representational similarity analysis**.

A formal modeling approach that **incorporates representation as part of the cognitive model** might help make theory-driven predictions about neural representations.

Representations to neuroscientists

Other neuroscientists are using training neural network models to achieve human-like performance or using machine learning to decode neural representations.

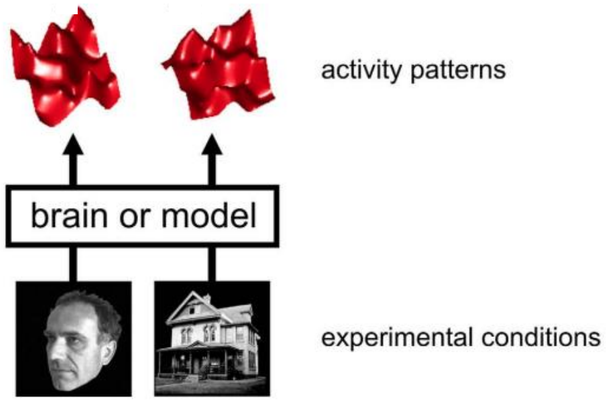
One such method is **representational similarity analysis:**



Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

Representational similarity analysis

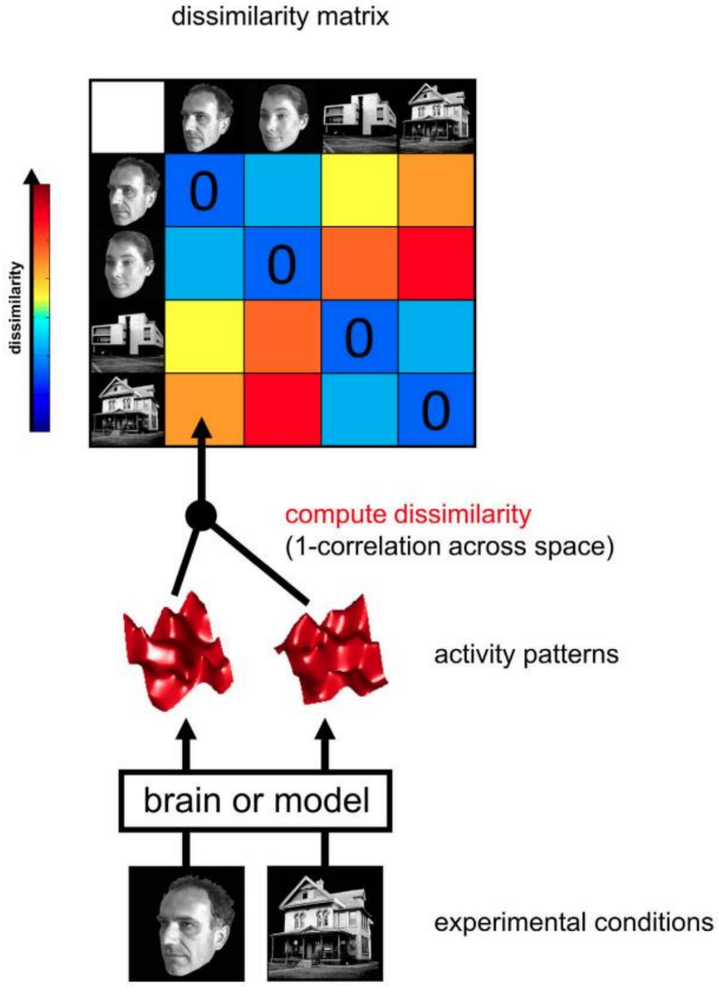
The researcher collects neuroimaging data under experimental conditions



Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

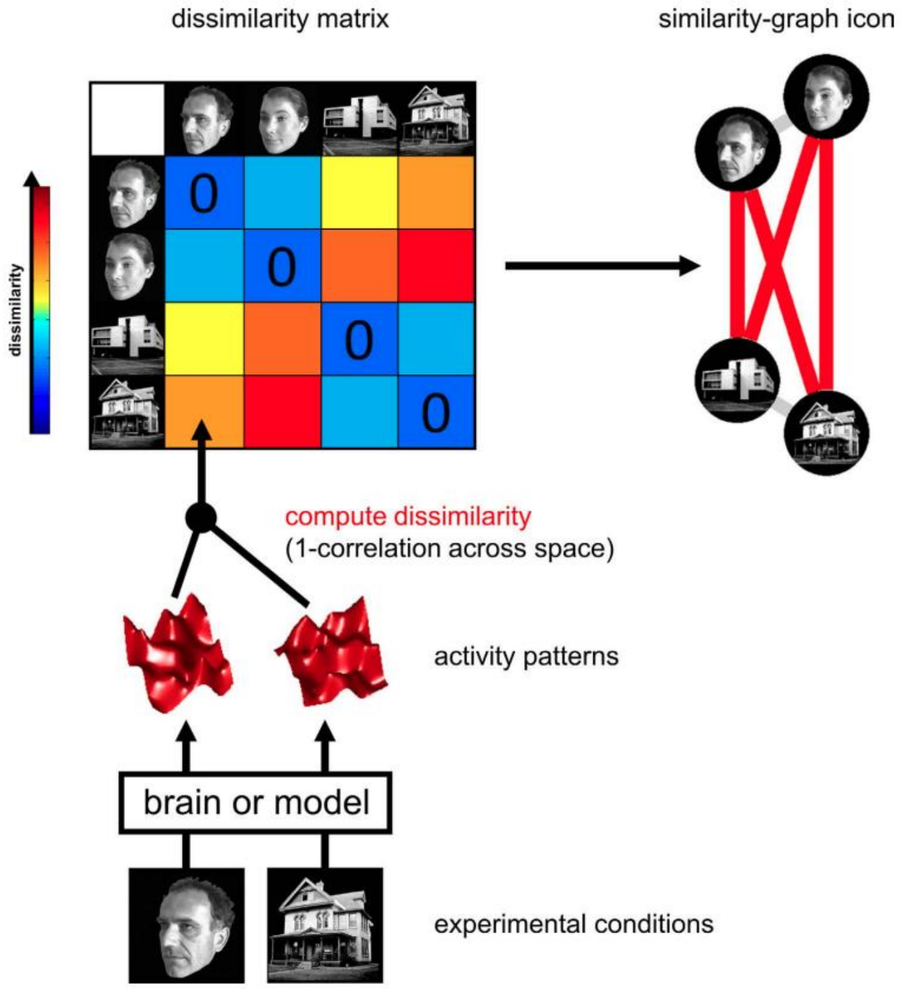
Representational similarity analysis

The researcher uses machine learning to decode the conditions and computes the distance between conditions



Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

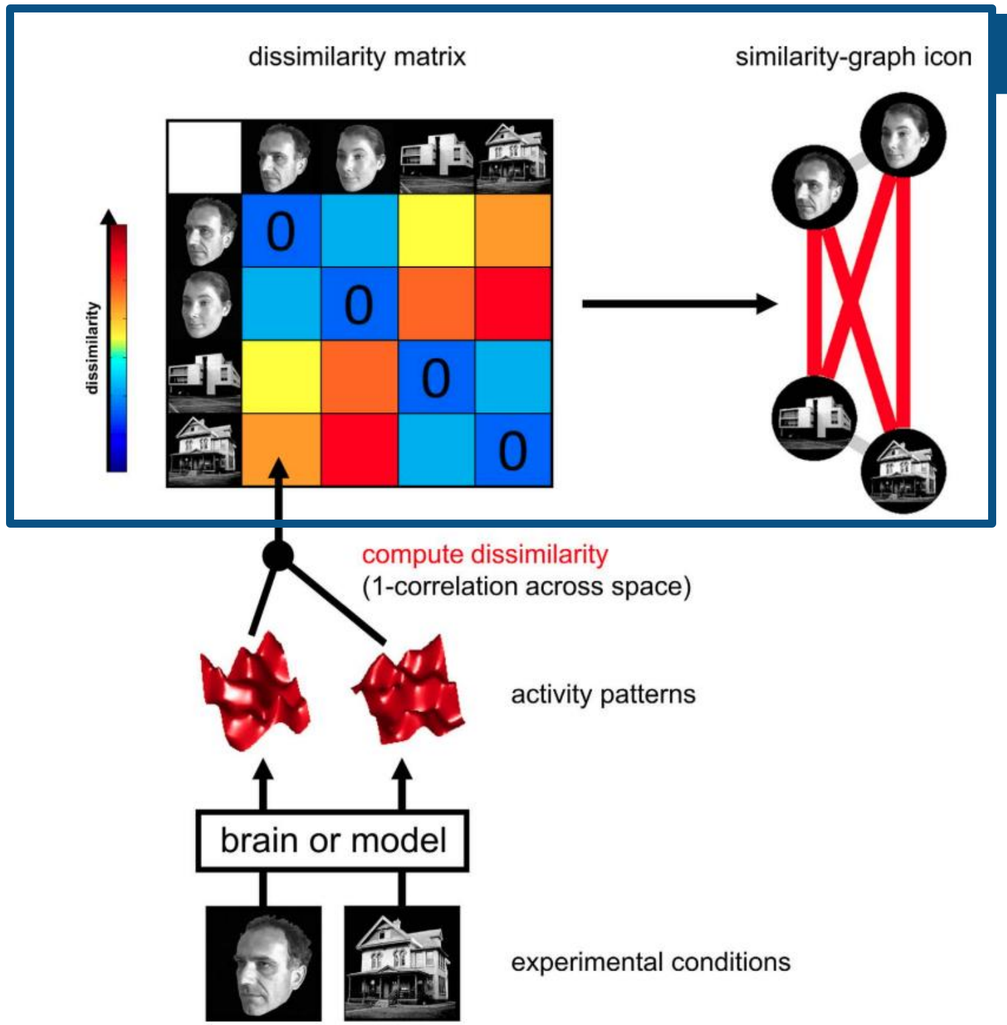
Representational similarity analysis



The dissimilarity can be visualised using multidimensional scaling or other methods.

Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

Representational similarity analysis



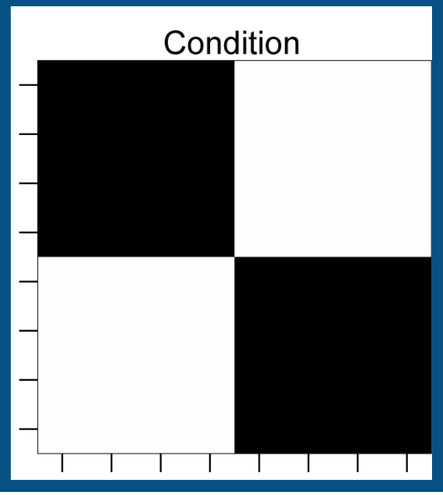
Do these results provide a veridical basis for models of cognition?

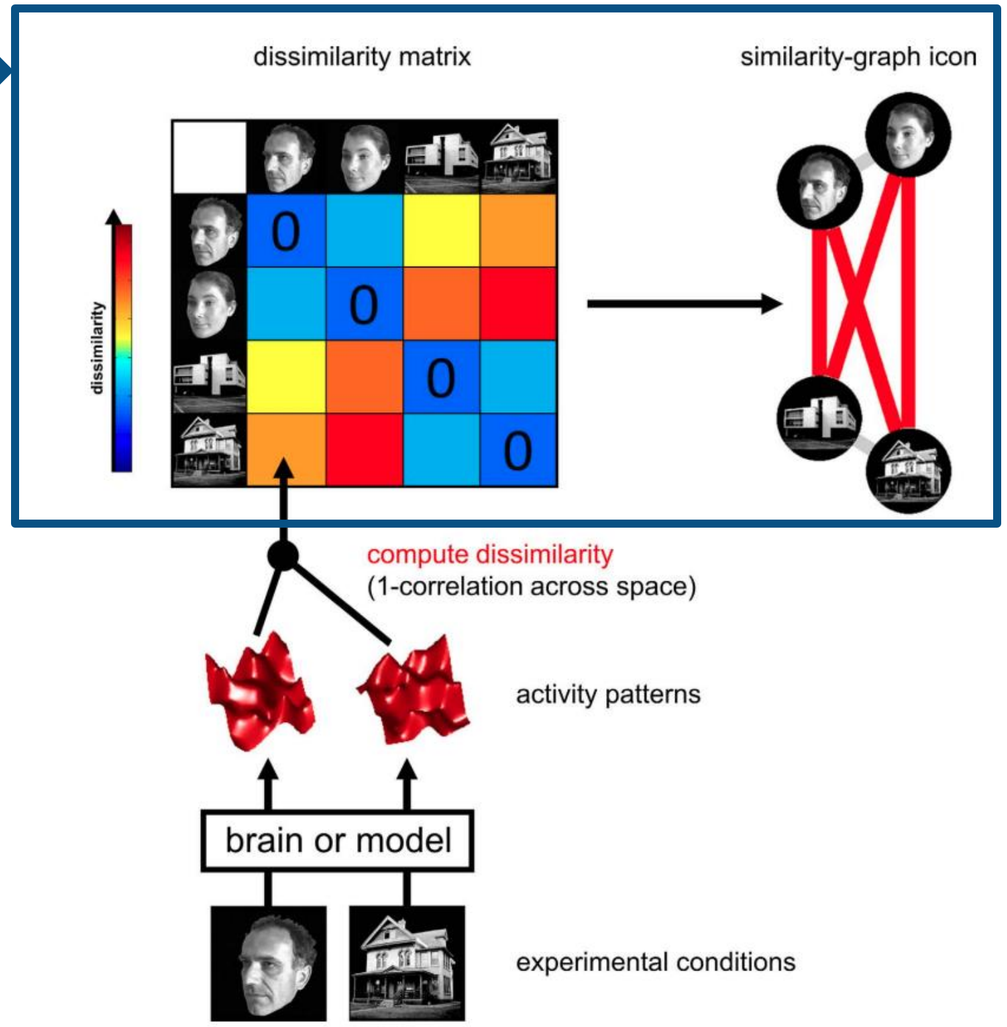
We might be **HARKing** and/or drawing ad-hoc conclusions.

Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

Representational similarity analysis

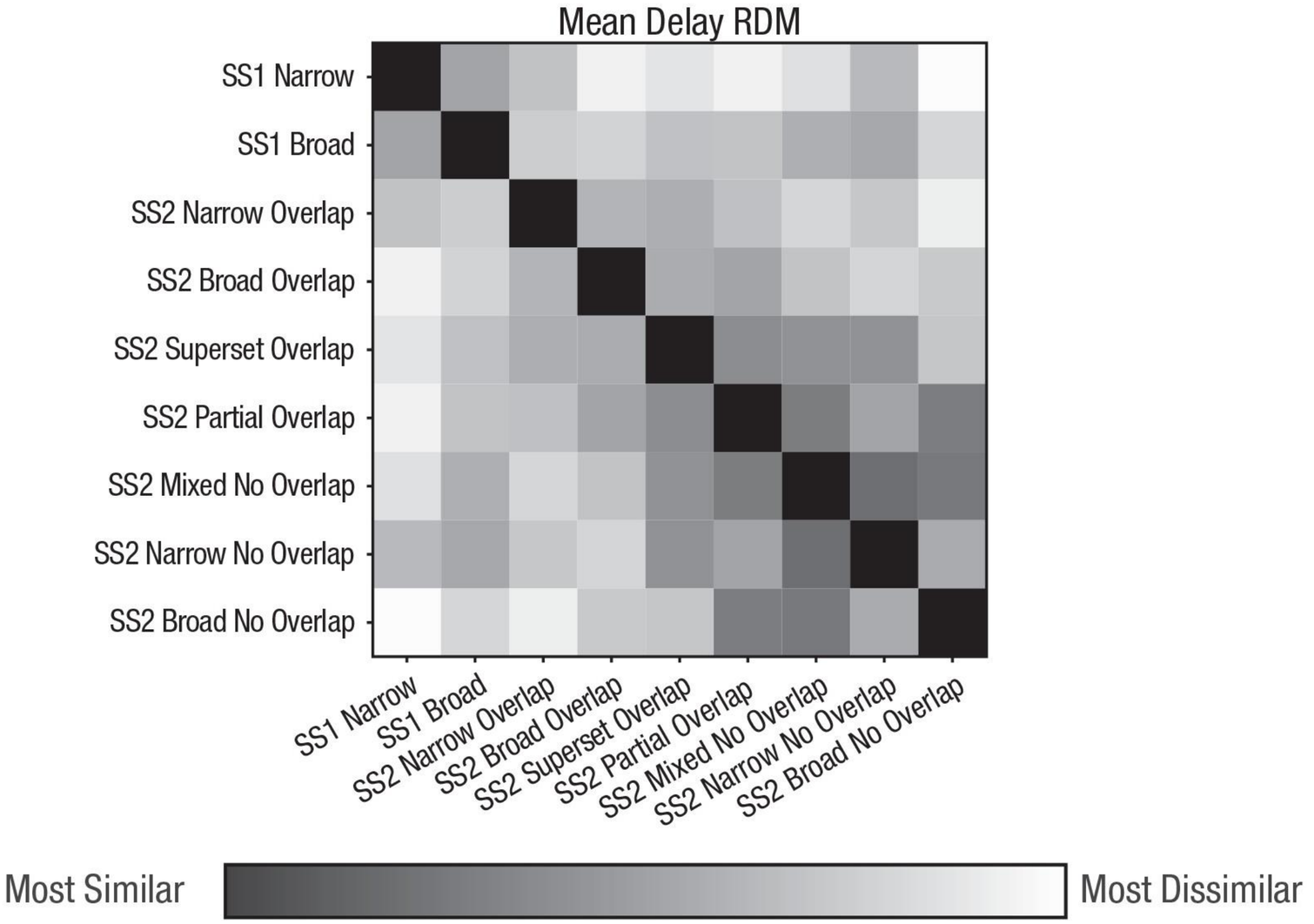
We might create models to predict the observed representational (dis)similarity





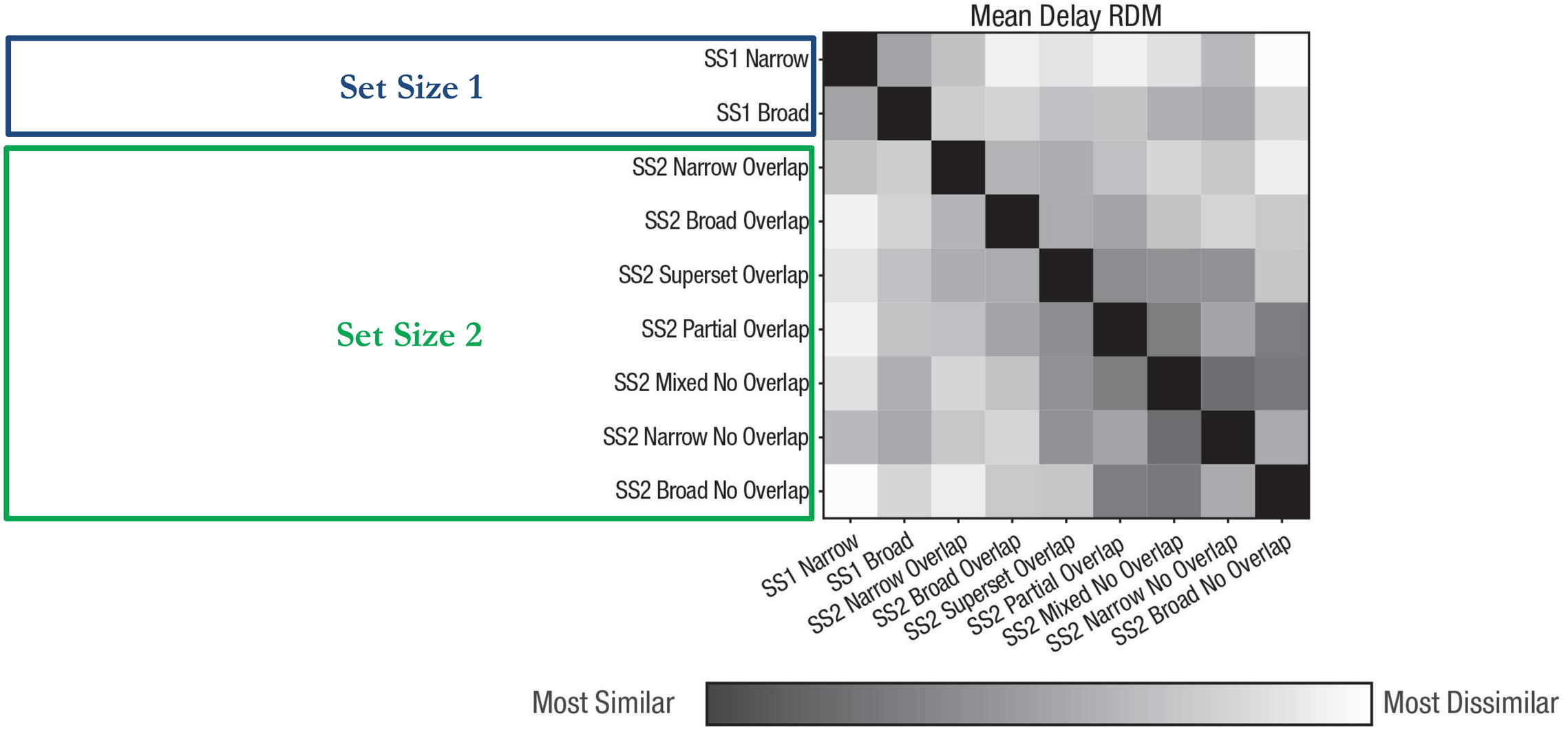
Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

Decoding load in working memory



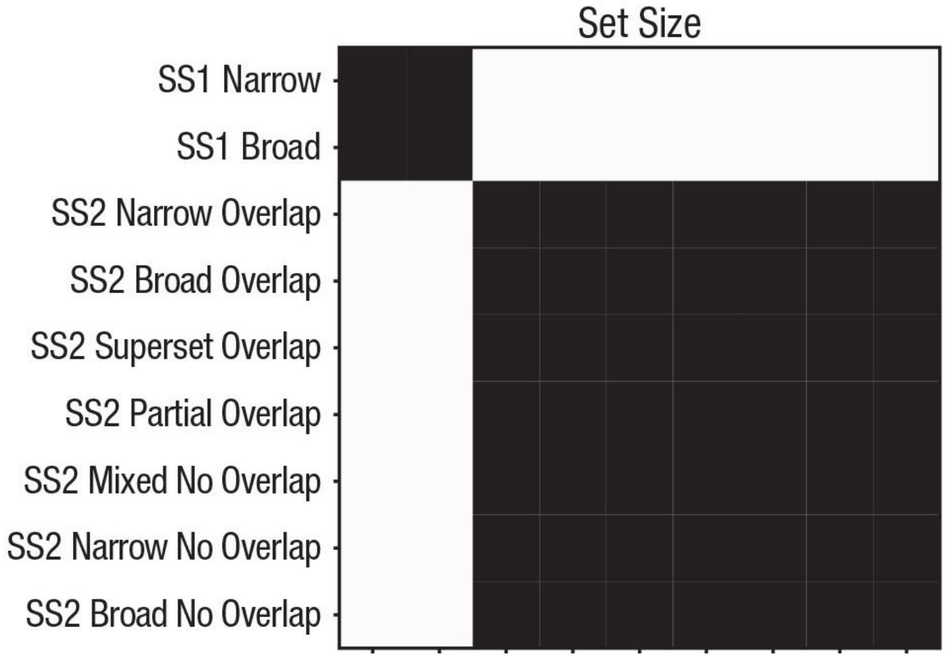
Jones, H. M., Diaz, G. K., Ngiam, W. X. Q., & Awh, E. (2024). Psychological Science, 35(10), 1108-1138.

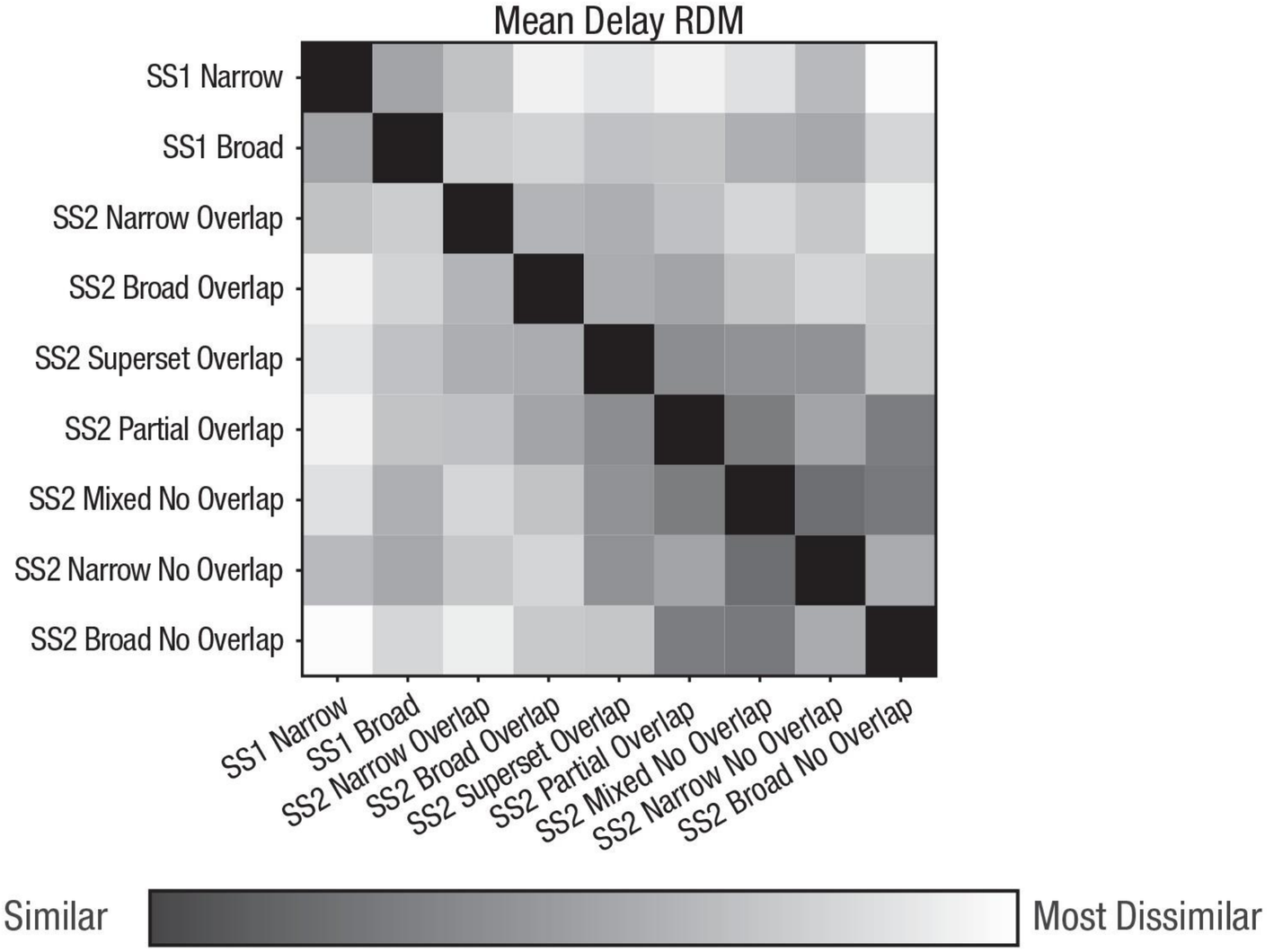
Decoding load in working memory



Jones, H. M., Diaz, G. K., Ngiam, W. X. Q., & Awh, E. (2024). Psychological Science, 35(10), 1108-1138.

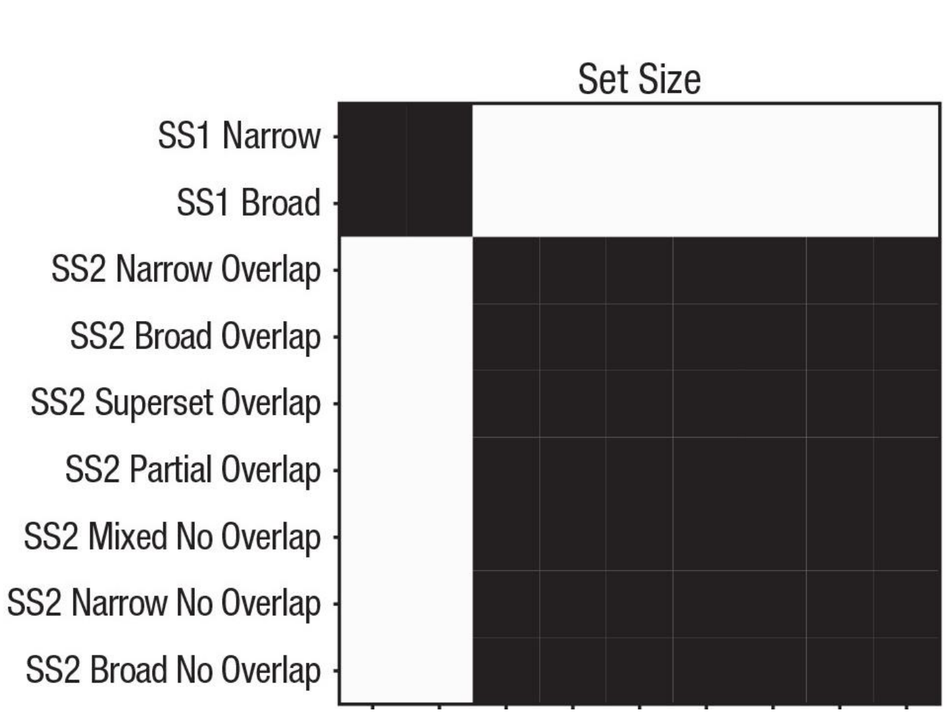
Decoding load in working memory



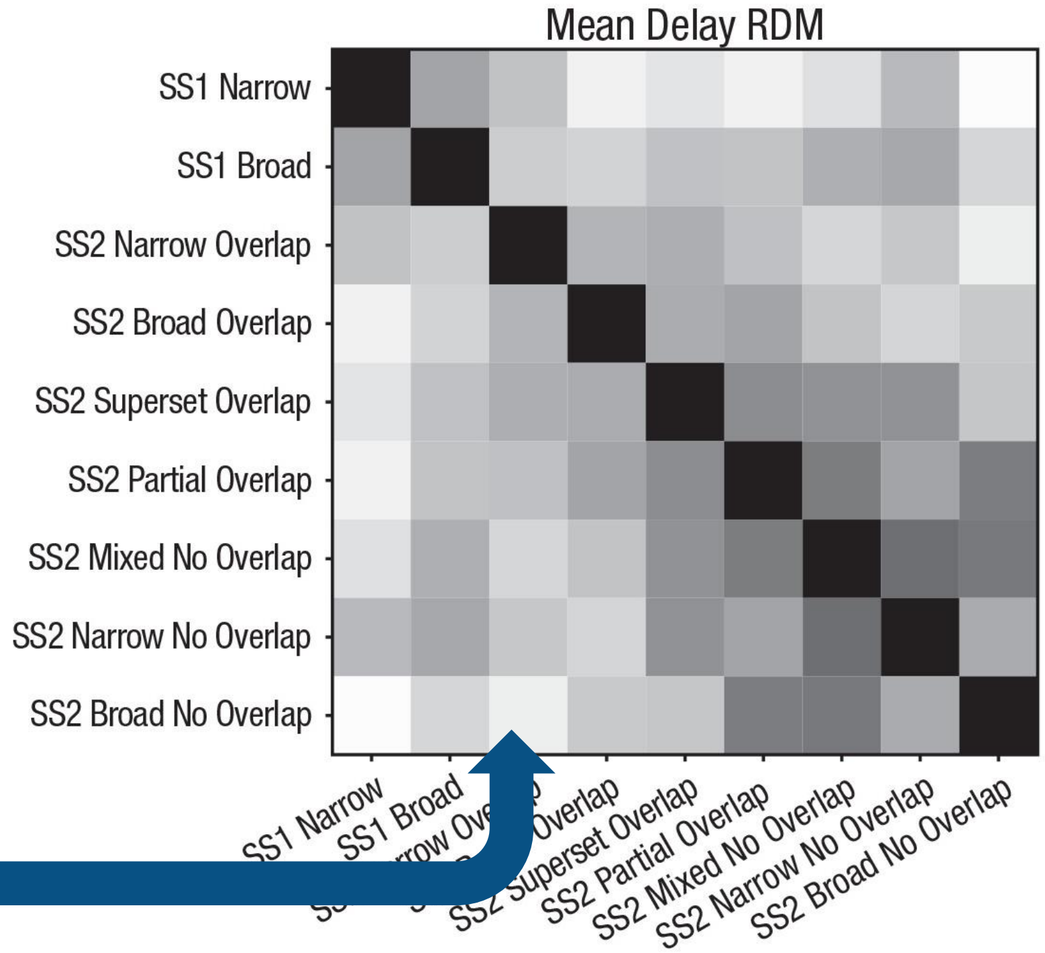


Jones, H. M., Diaz, G. K., Ngiam, W. X. Q., & Awh, E. (2024). Psychological Science, 35(10), 1108-1138.

Decoding load in working memory



A "model" to predict neural representational (dis)similarity



Jones, H. M., Diaz, G. K., Ngiam, W. X. Q., & Awh, E. (2024). Psychological Science, 35(10), 1108-1138.

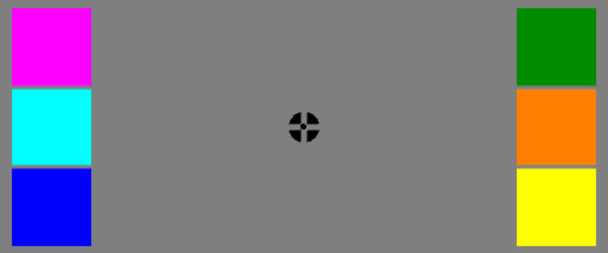
Decoding load in working memory

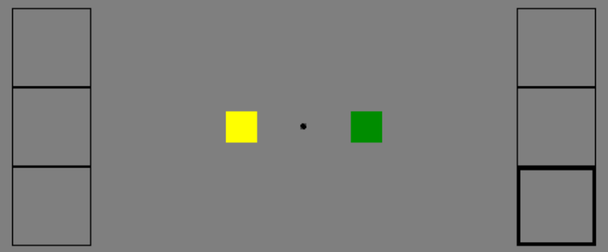
We were interested in how associative learning influences working memory operations.

One proposed operation is that associative learning leads to "chunking" processes – representations where separate items are bound into a "chunk".

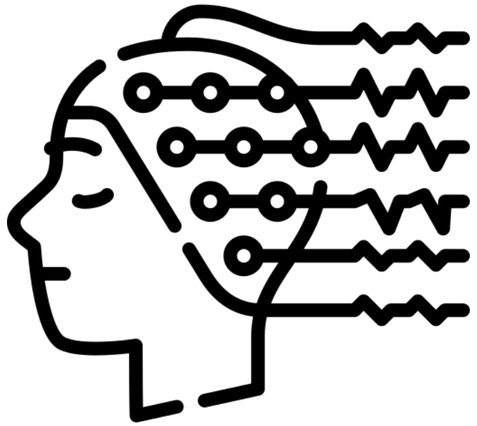
Training

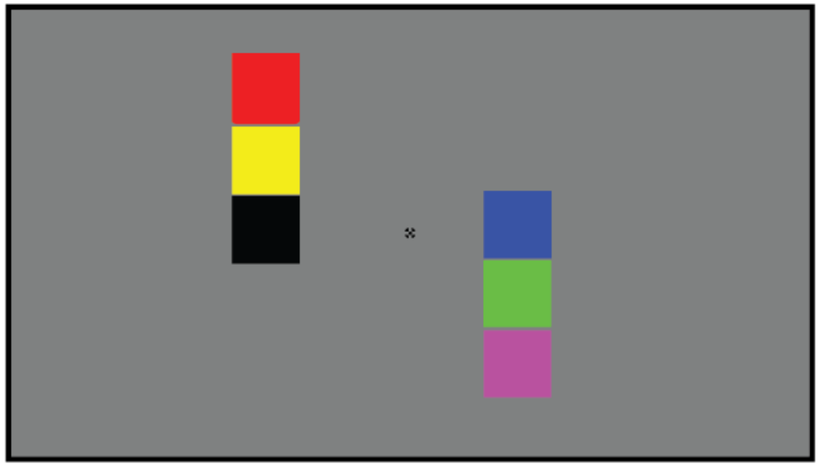
Trained subjects to learn three color triplets

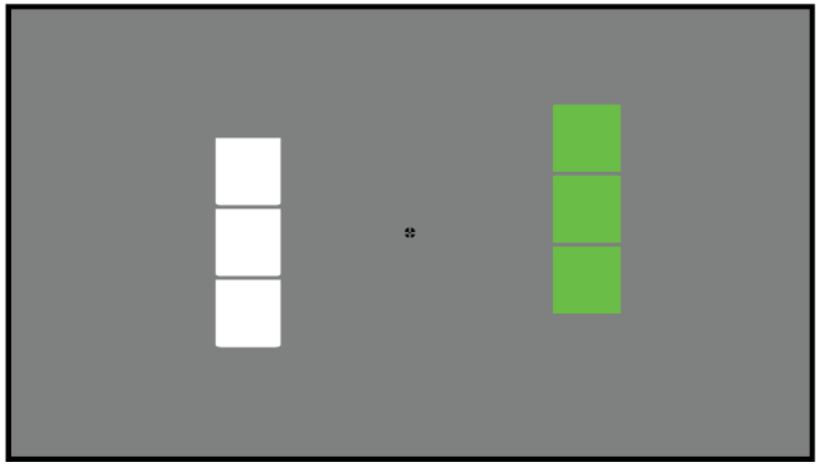




EEG







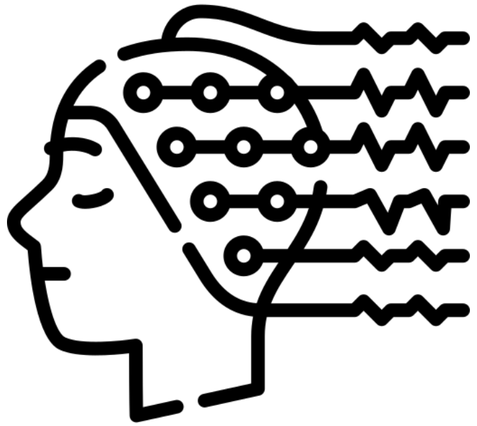
Perceptually equivalent

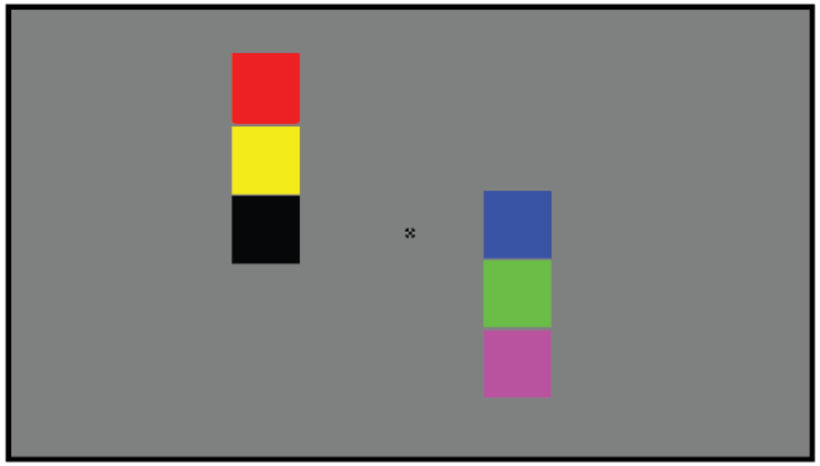
Six random

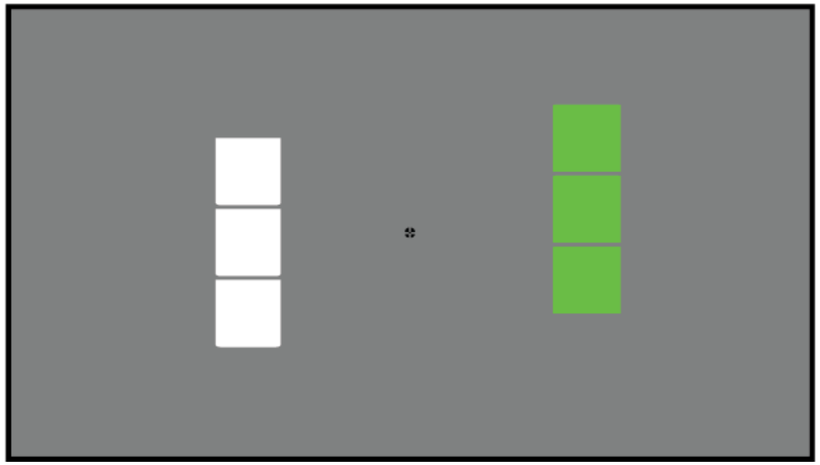
Six chunked

Two random

EEG







Perceptually equivalent

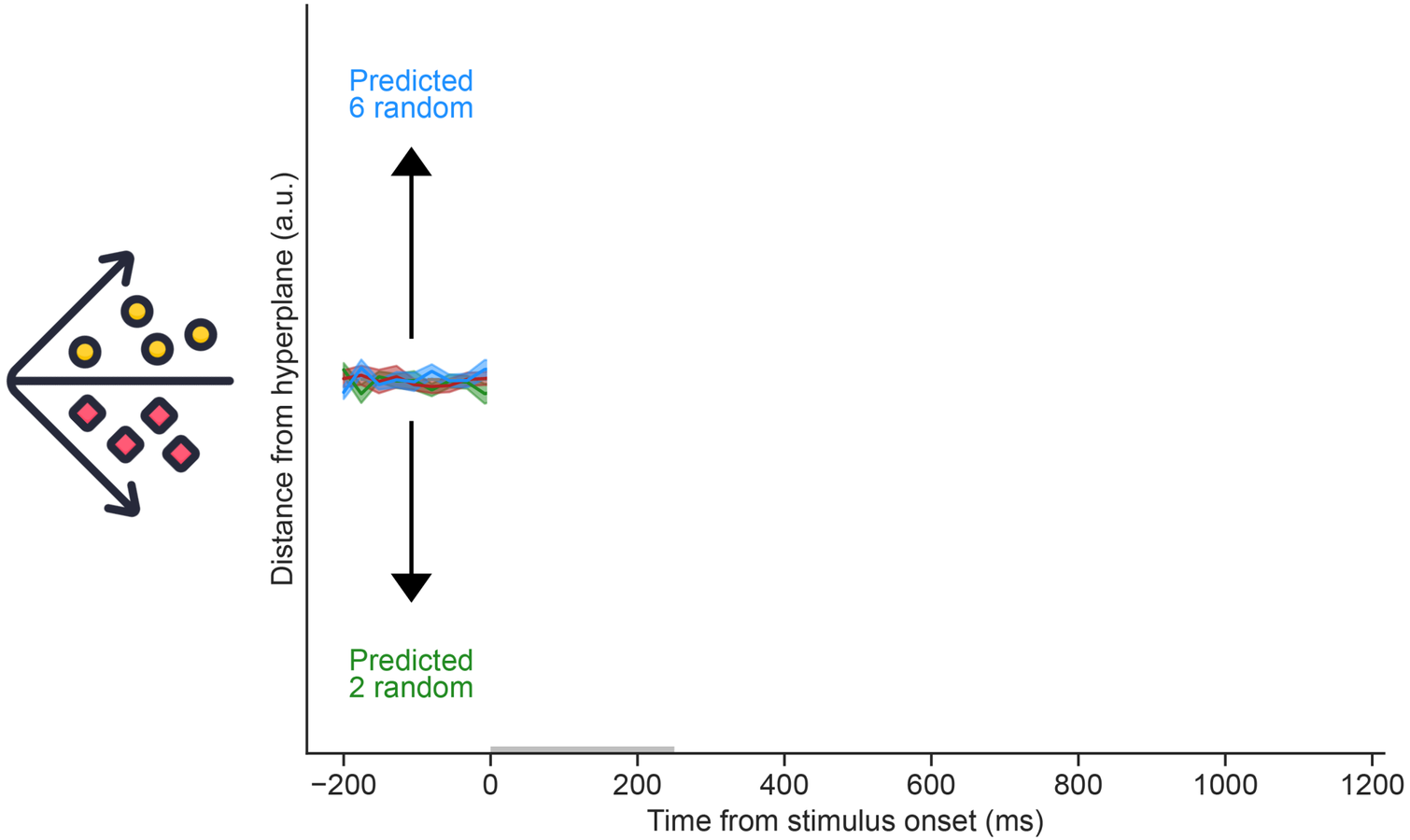
Six random

Six chunked

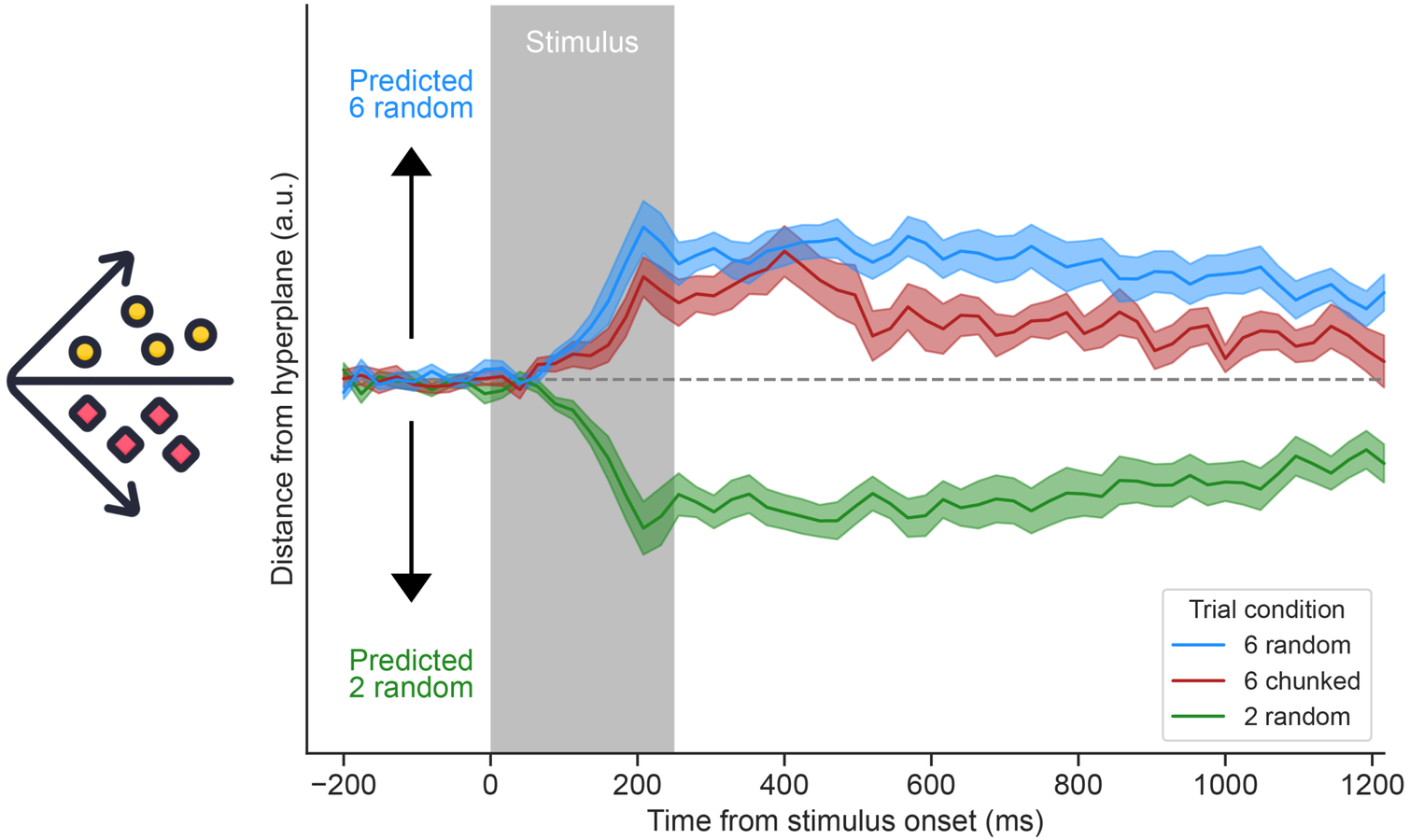
Two random

Expectation: "chunking" results in a reduction of item-based load that should be reflected in neural representations

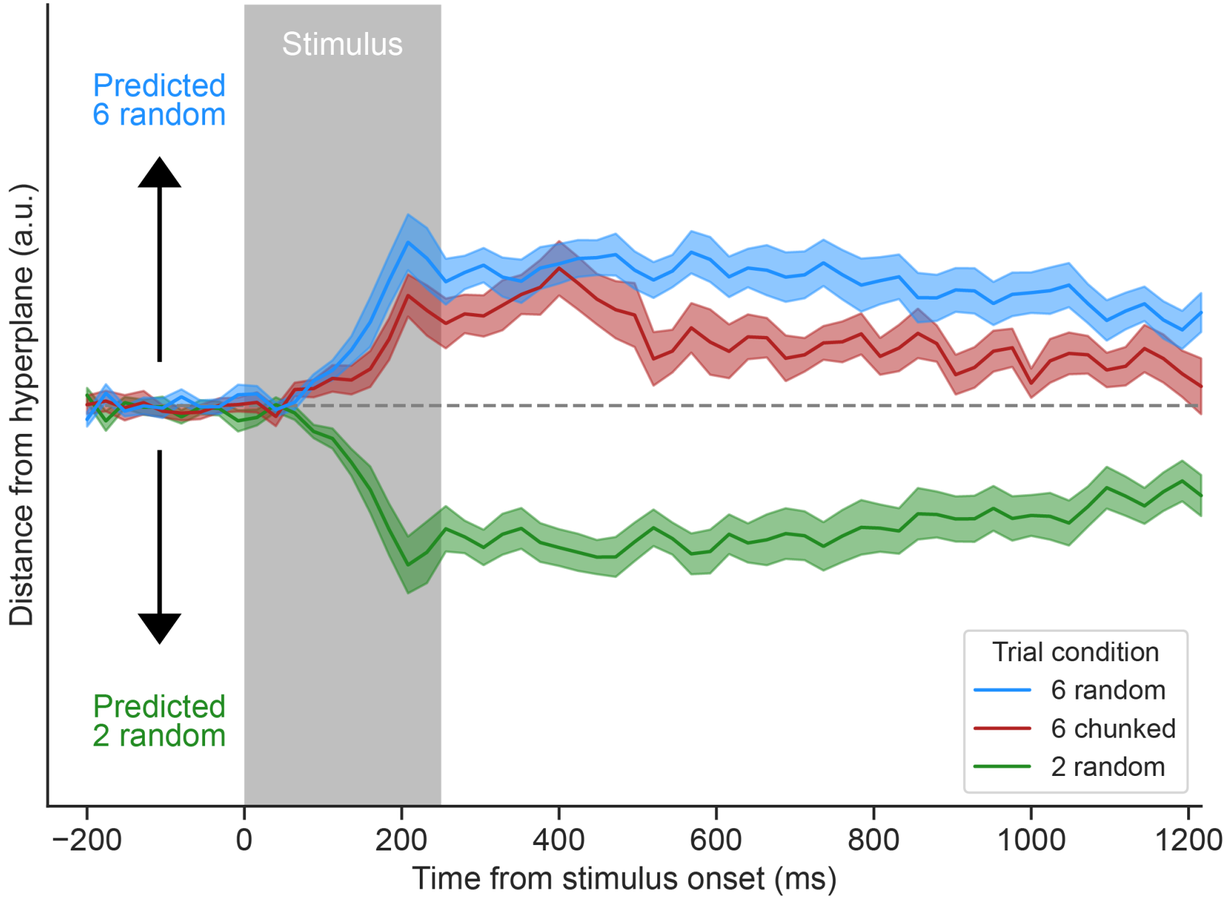
Train 6 random versus 2 random, test 6 chunked



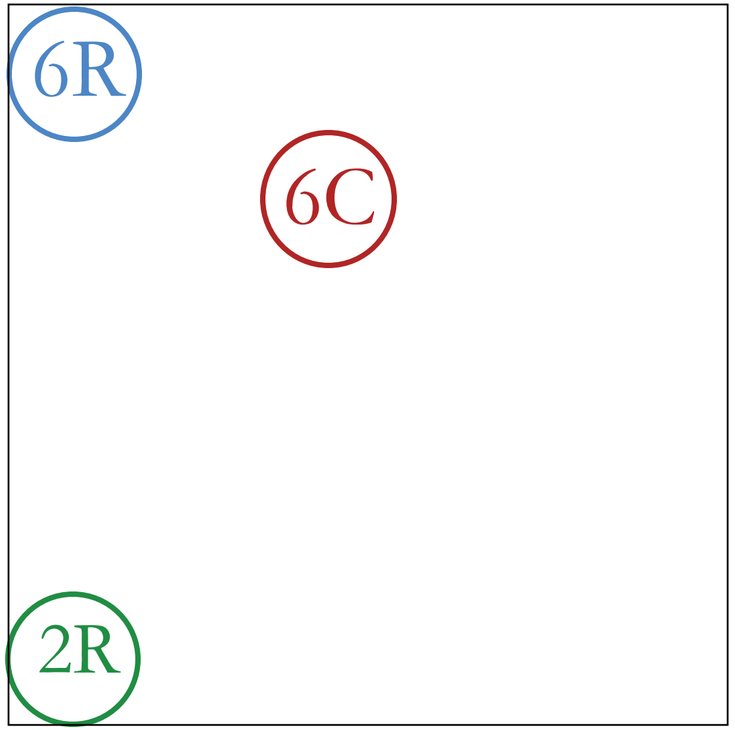
Train 6 random versus 2 random, test 6 chunked



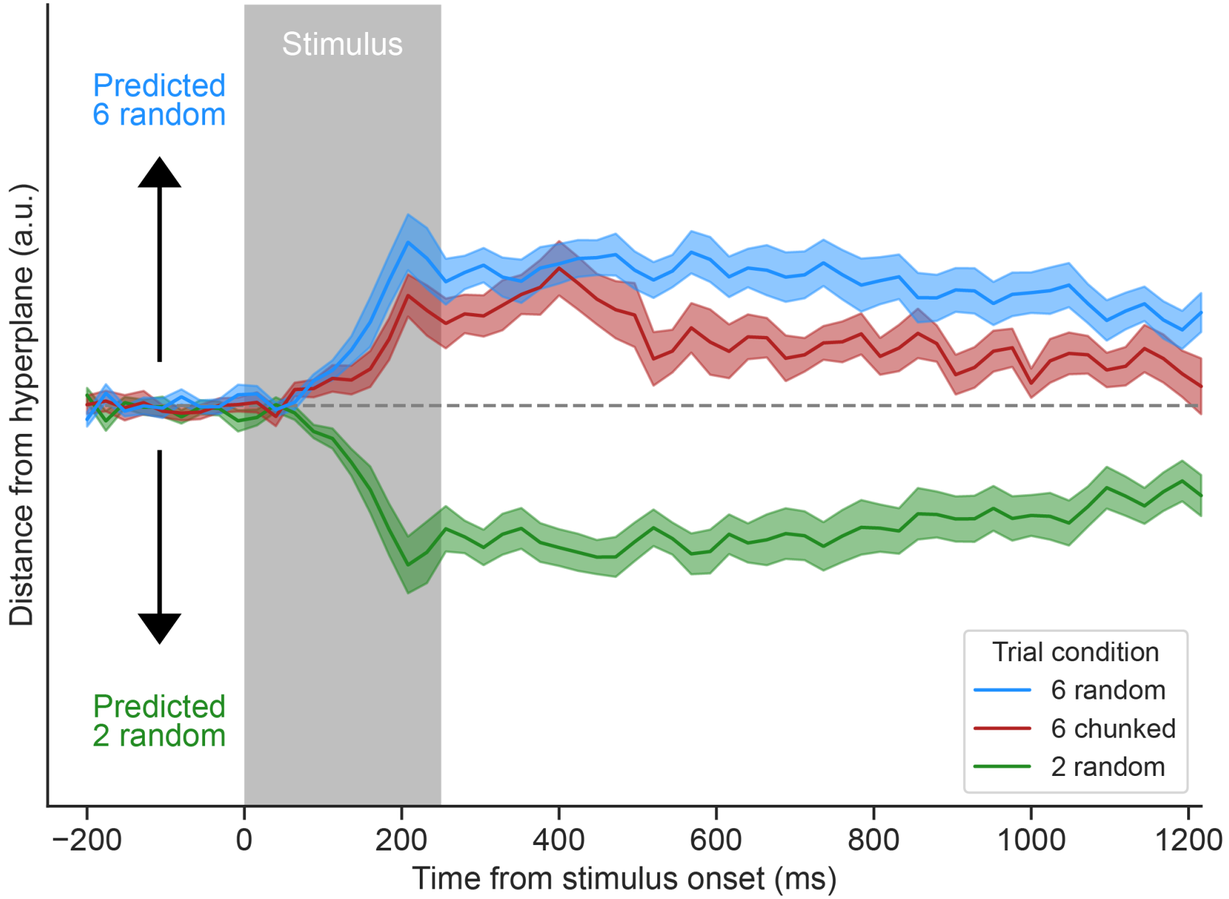
Train 6 random versus 2 random, test 6 chunked



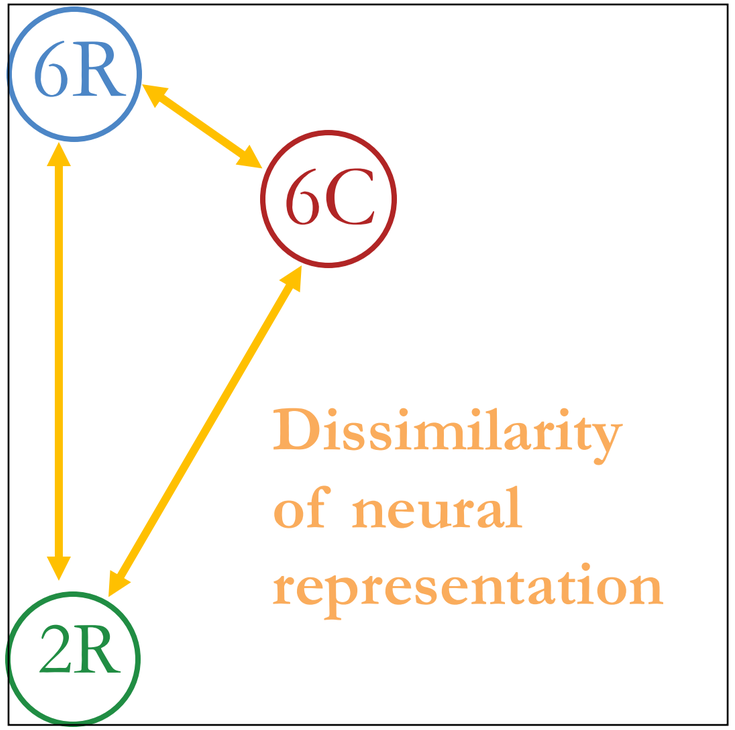
Multidimensional scaling



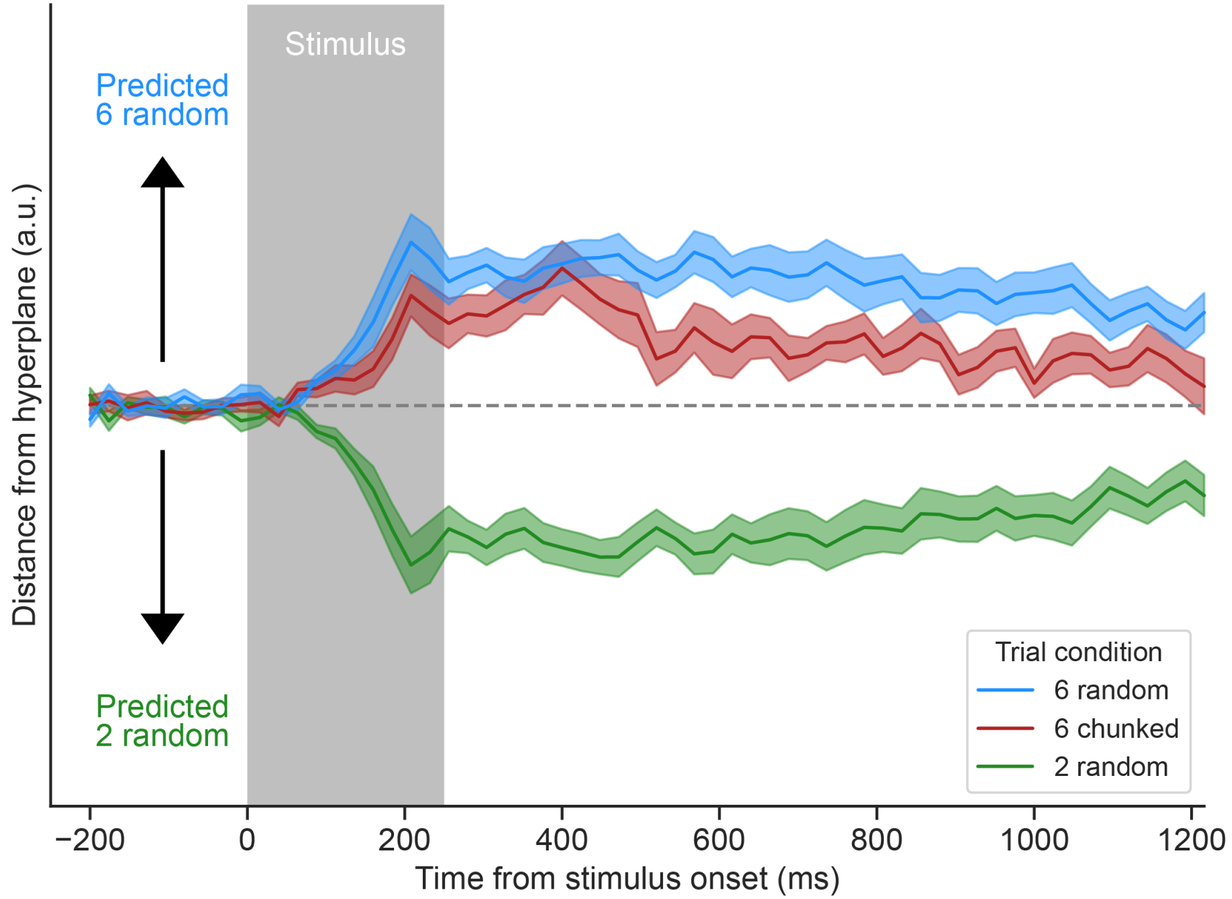
Train 6 random versus 2 random, test 6 chunked



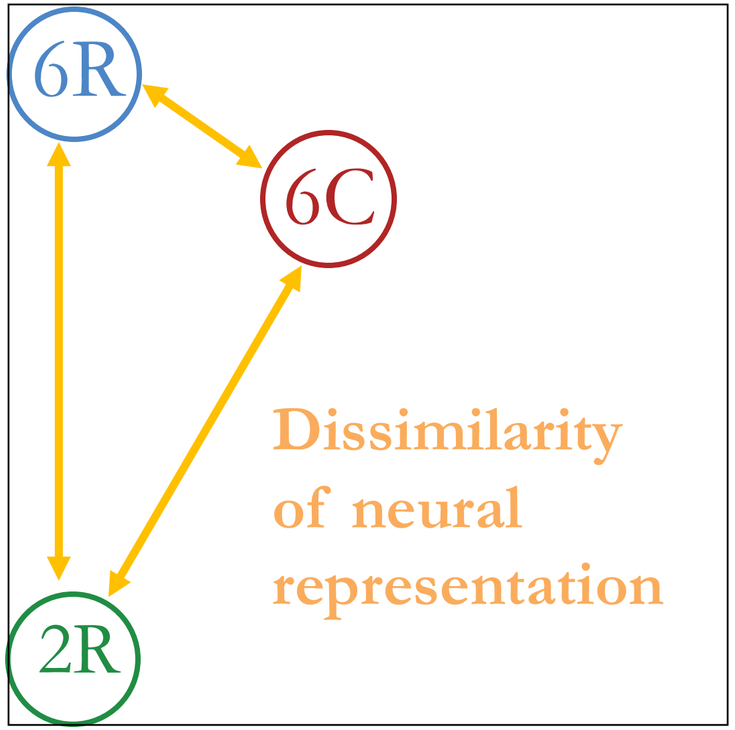
Multidimensional scaling



Train 6 random versus 2 random, test 6 chunked

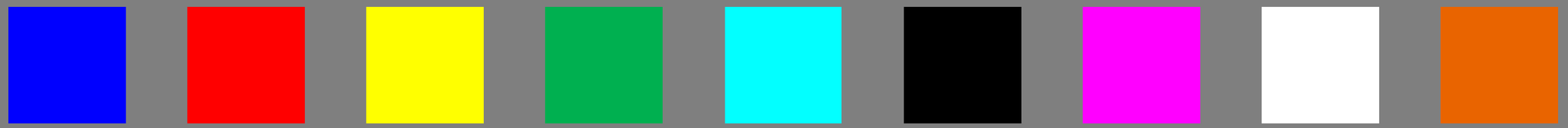


Multidimensional scaling



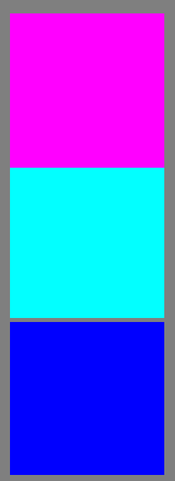
What can we infer from these results?
What cognitive model would predict or could explain this pattern of representational similarity?

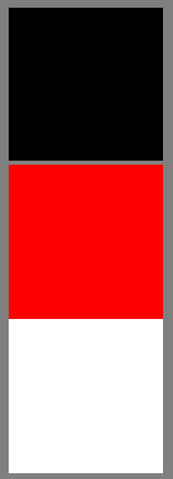
Awareness Test



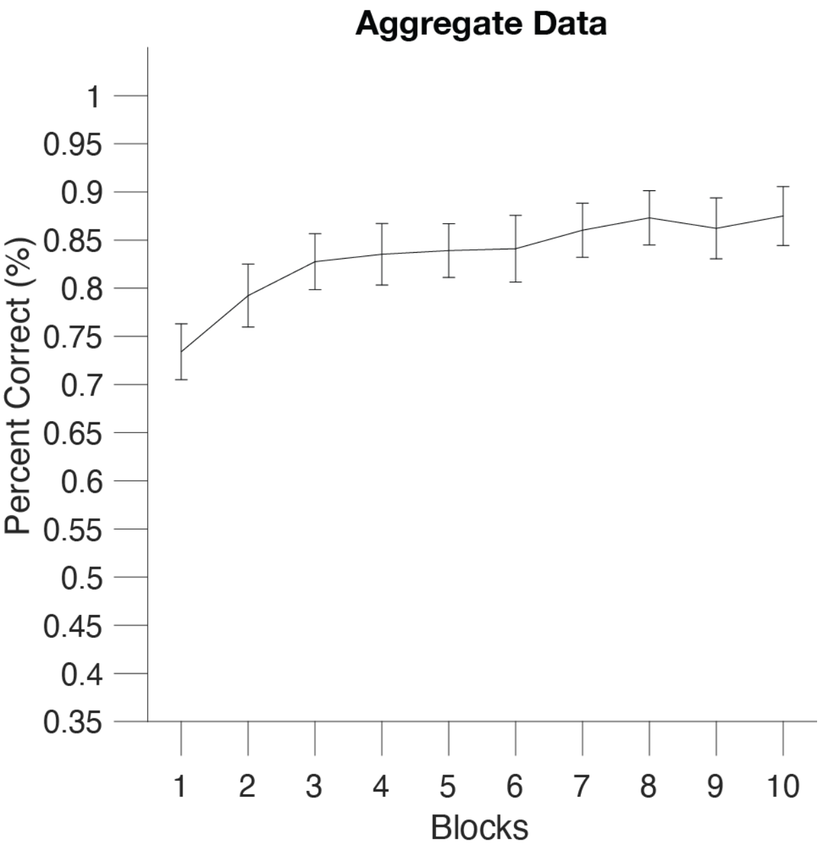
Awareness Test

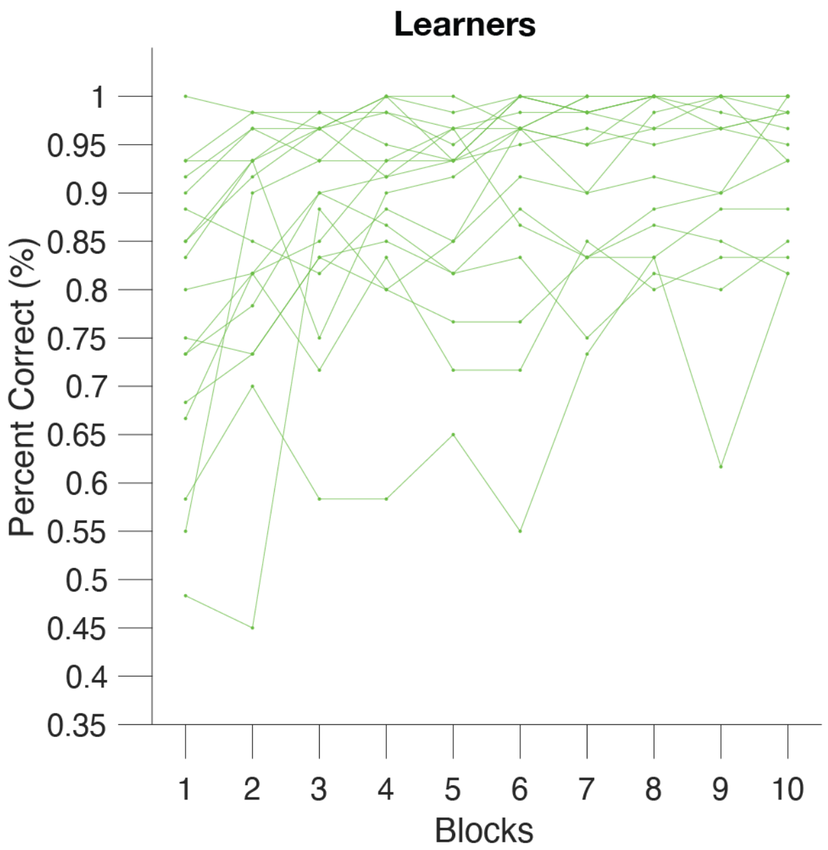
Only subjects that recreated **all** triplets were considered “aware”

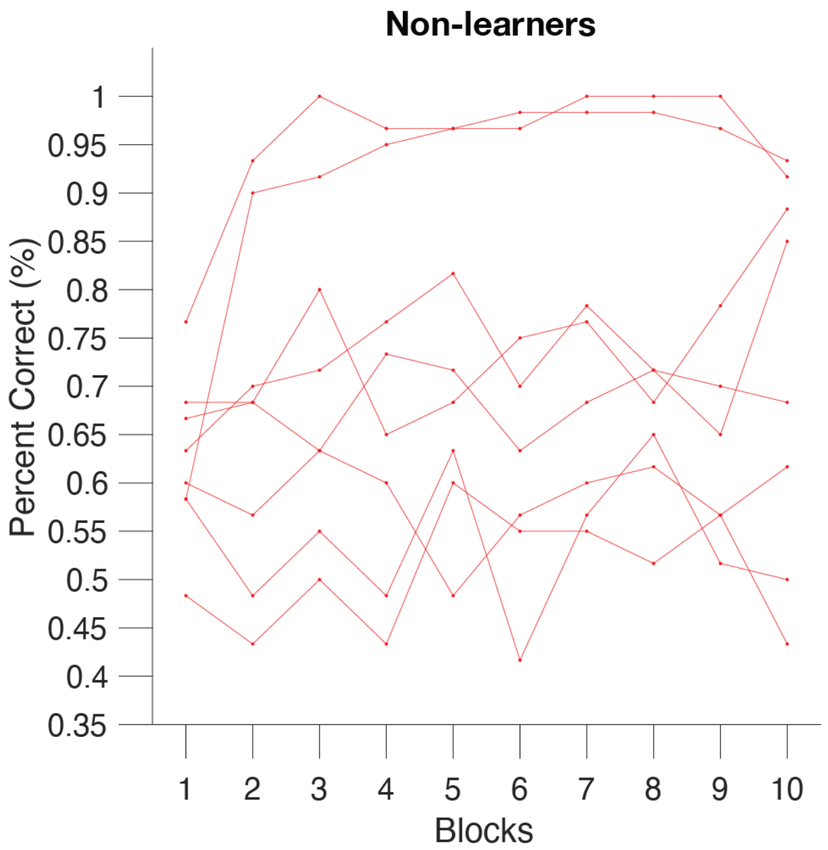




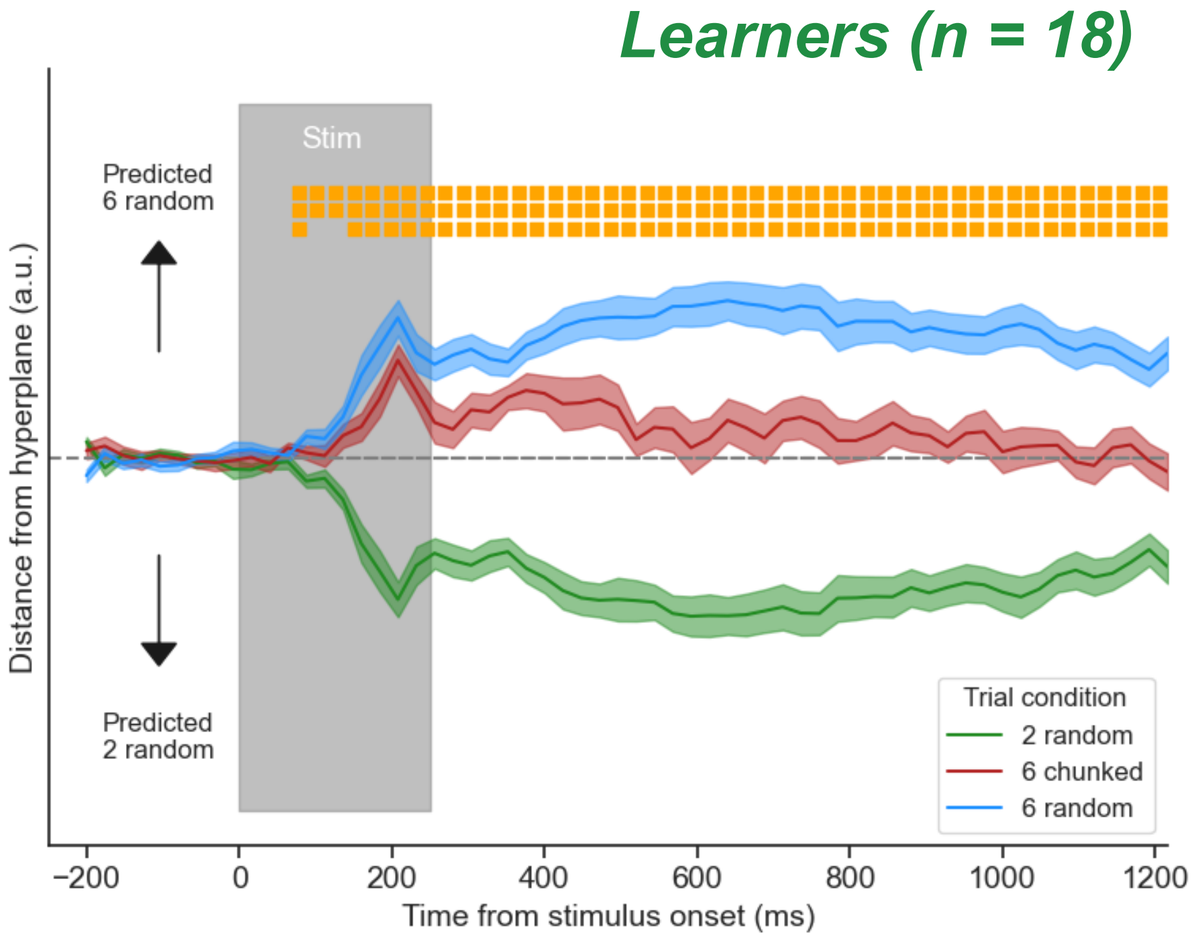
Training Results

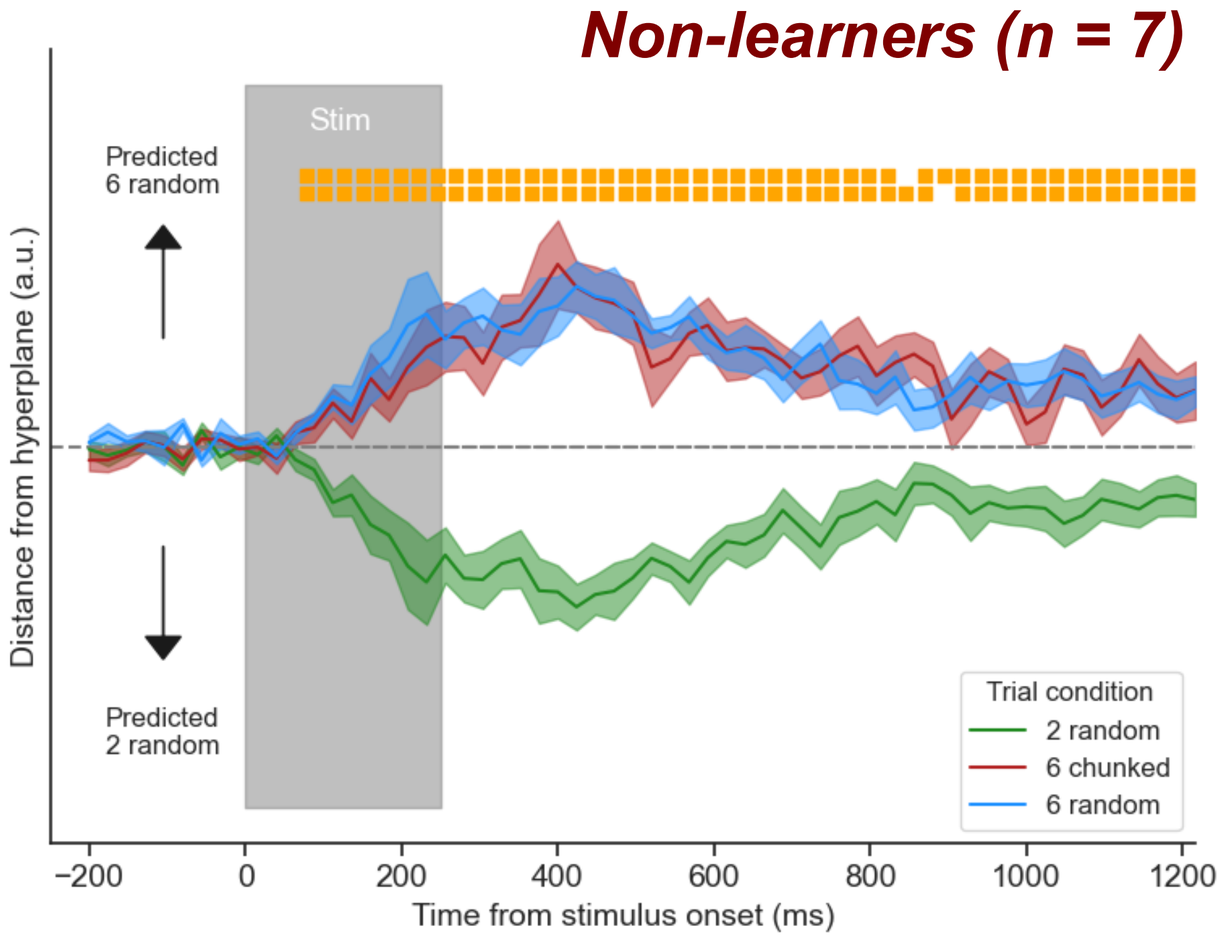




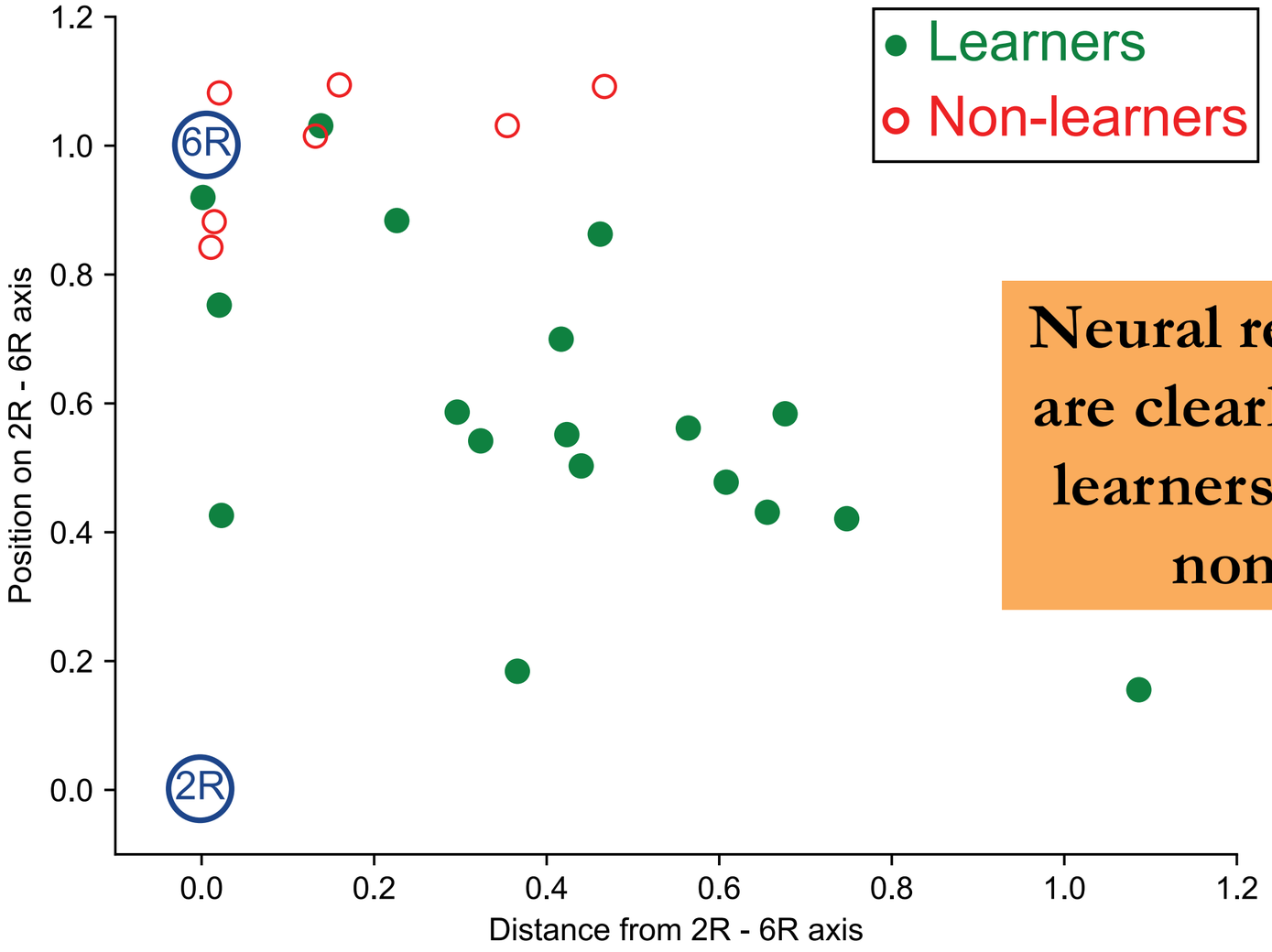


Learners vs non-learners





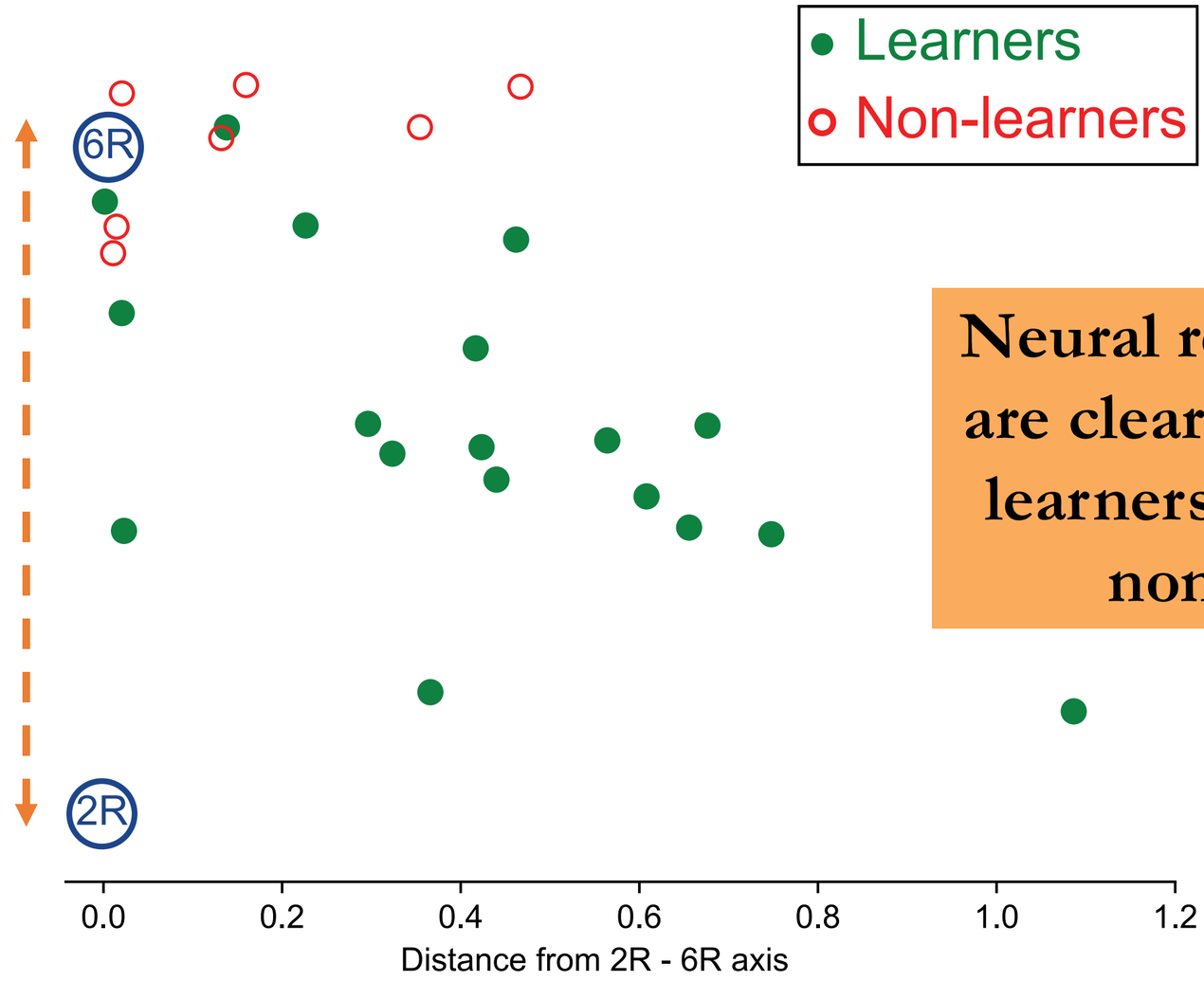
Multidimensional scaling on each subject



Neural representations are clearly different for learners compared to non-learners

Multidimensional scaling on each subject

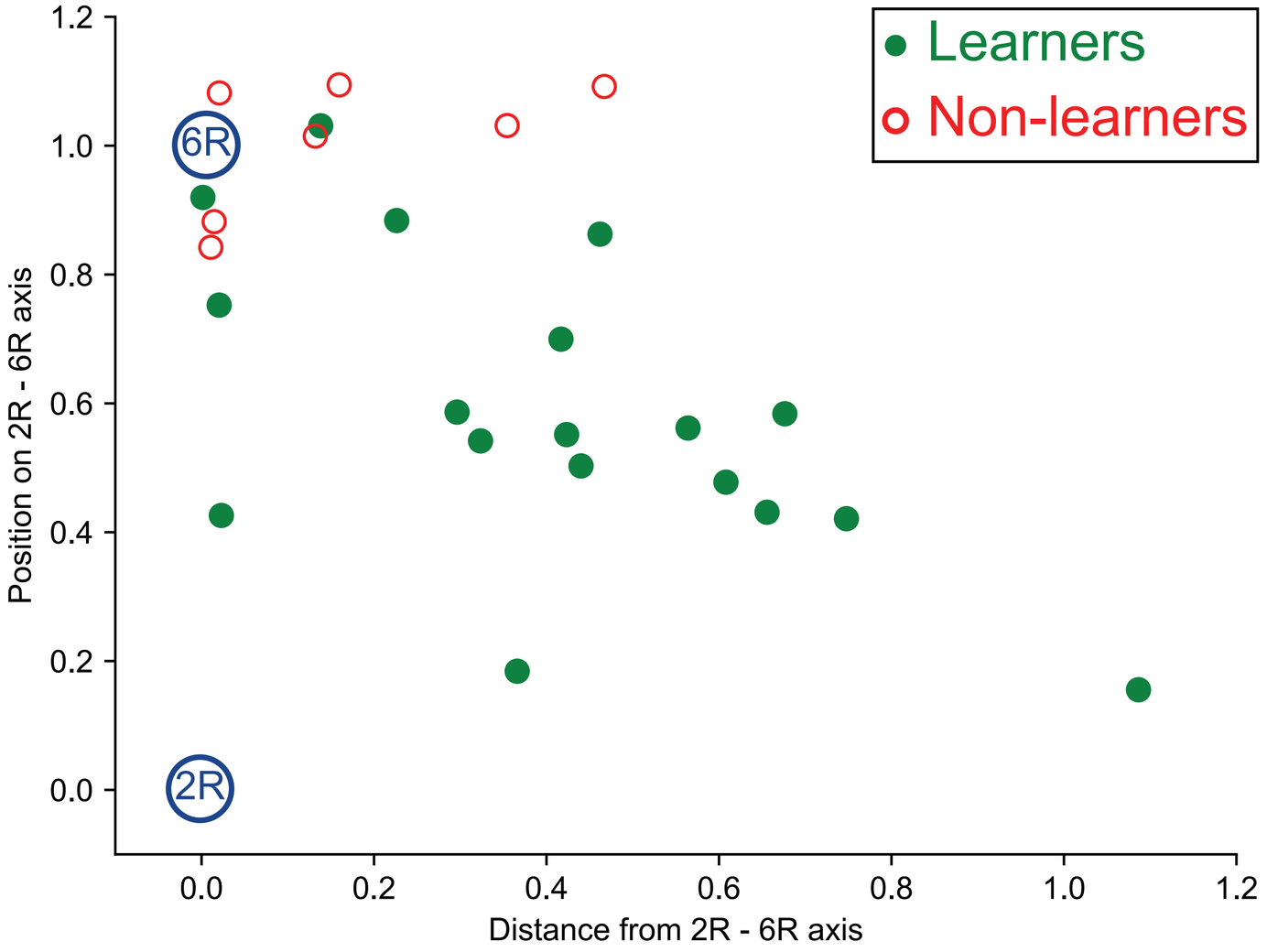
We assume increased neural similarity towards 2R (and dissimilarity to 6R) suggests a reduction in working memory load



Neural representations are clearly different for learners compared to non-learners

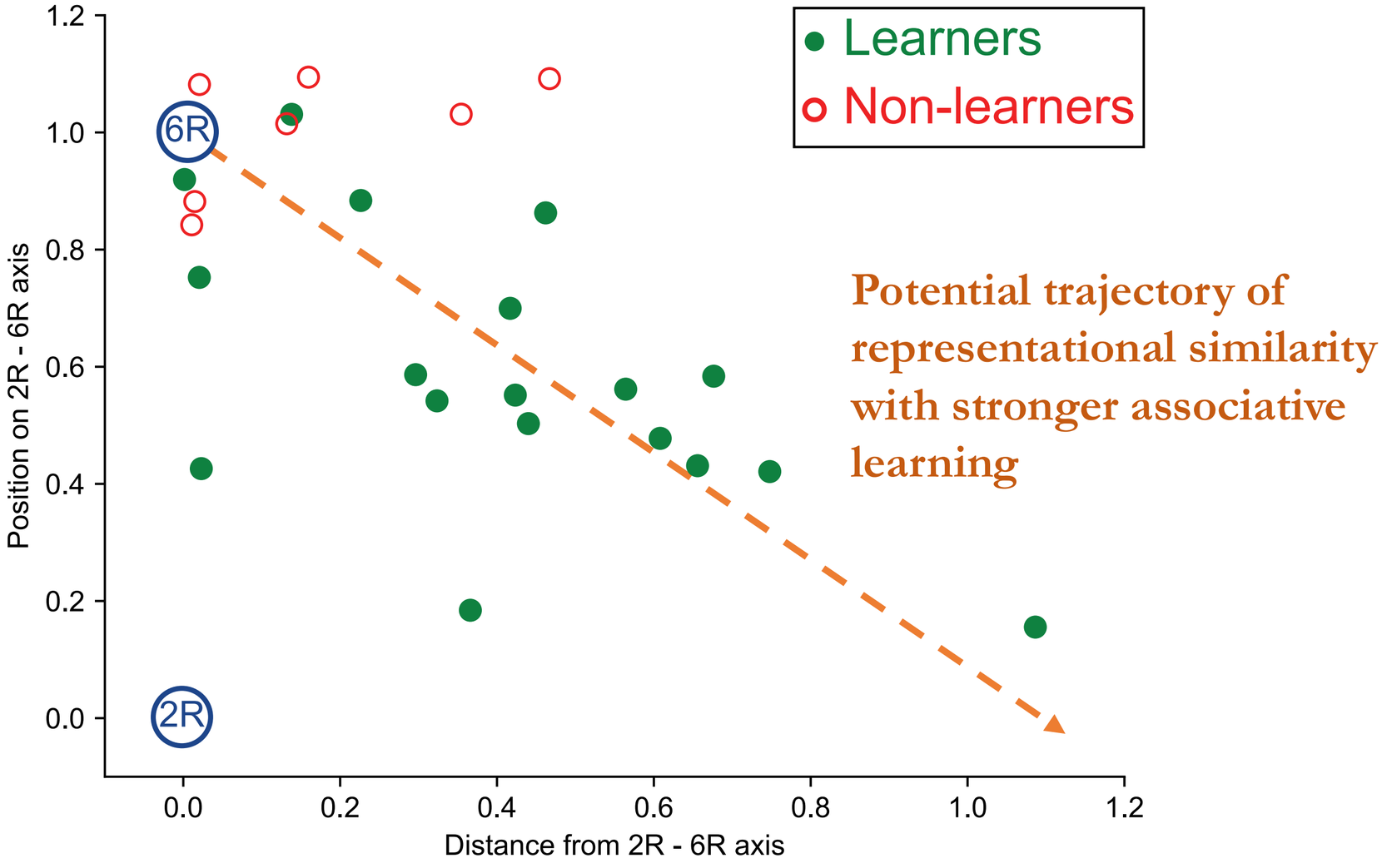
Multidimensional scaling on each subject

We ought to have a cognitive model that **explains** the neural representation and **make specific predictions** of the representational similarity so that it is **falsifiable.**



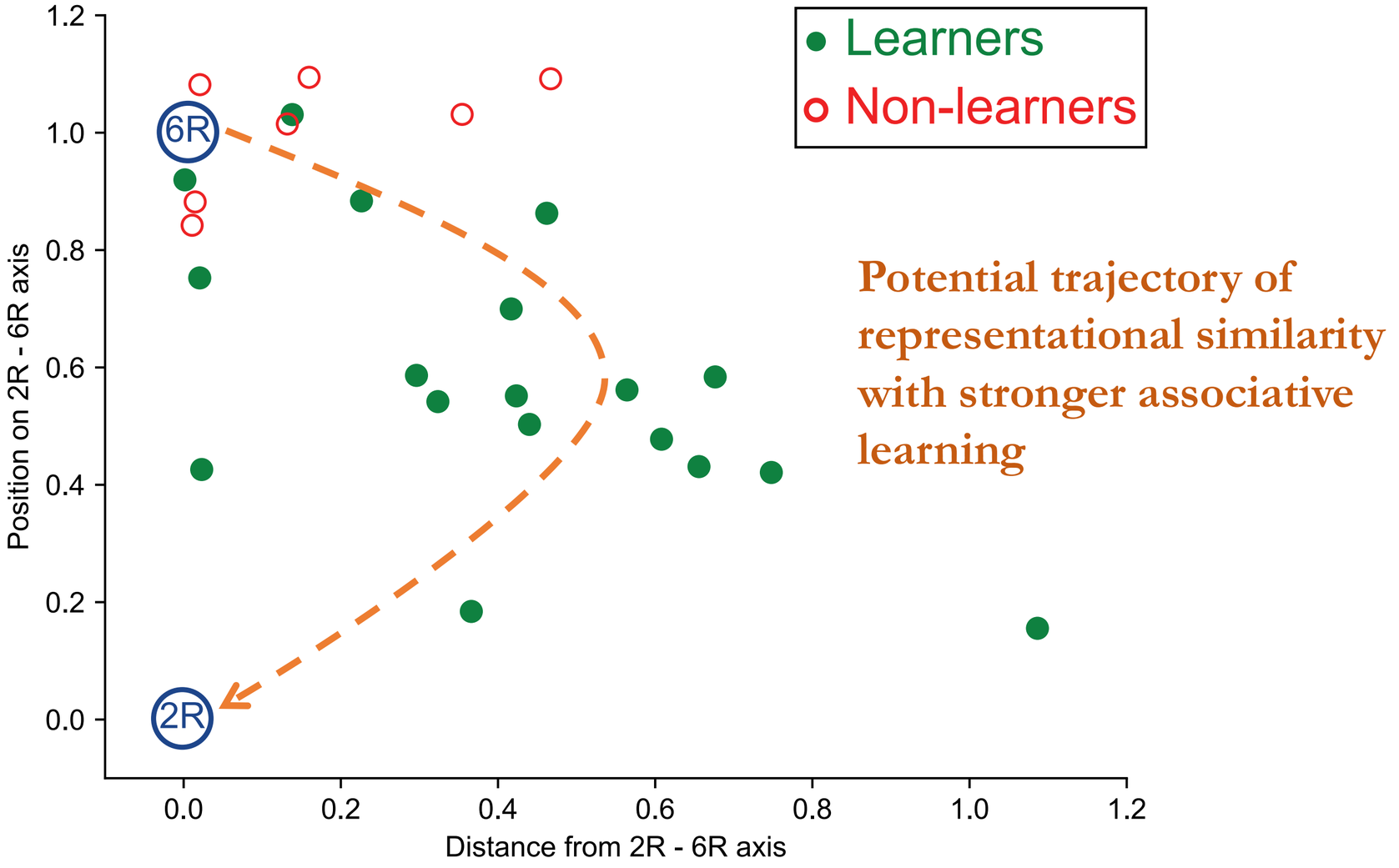
Multidimensional scaling on each subject

We ought to have a cognitive model that **explains** the neural representation and **make specific predictions** of the representational similarity so that it is **falsifiable.**



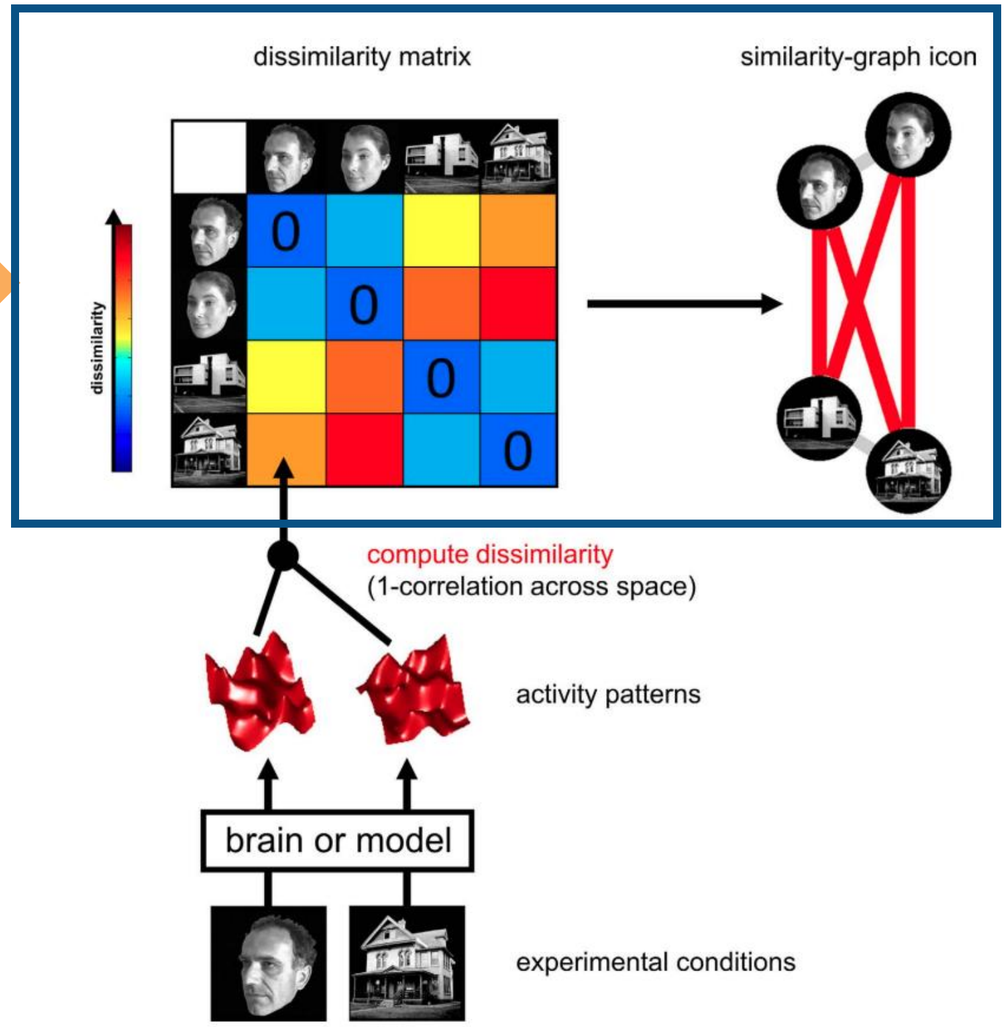
Multidimensional scaling on each subject

We ought to have a cognitive model that **explains** the neural representation and **make specific predictions** of the representational similarity so that it is **falsifiable.**



Representational similarity analysis

We need **cognitive models** to predict the observed representational (dis)similarity



Kriegeskorte, Mur, & Bandettini (2008). https://doi.org/10.3389/neuro.06.004.2008

Do similarity ratings predict neural similarity?

Representations in formal cognitive models

Classic formal models of cognition would derive the psychological representation using multidimensional scaling (MDS) of **similarity judgments**

- Collect similarity ratings for pairs of items in the stimulus set
- Reduce those ratings into a representation that best preserves the distance between the items; the closer the items, the more similar.

Representations in formal cognitive models

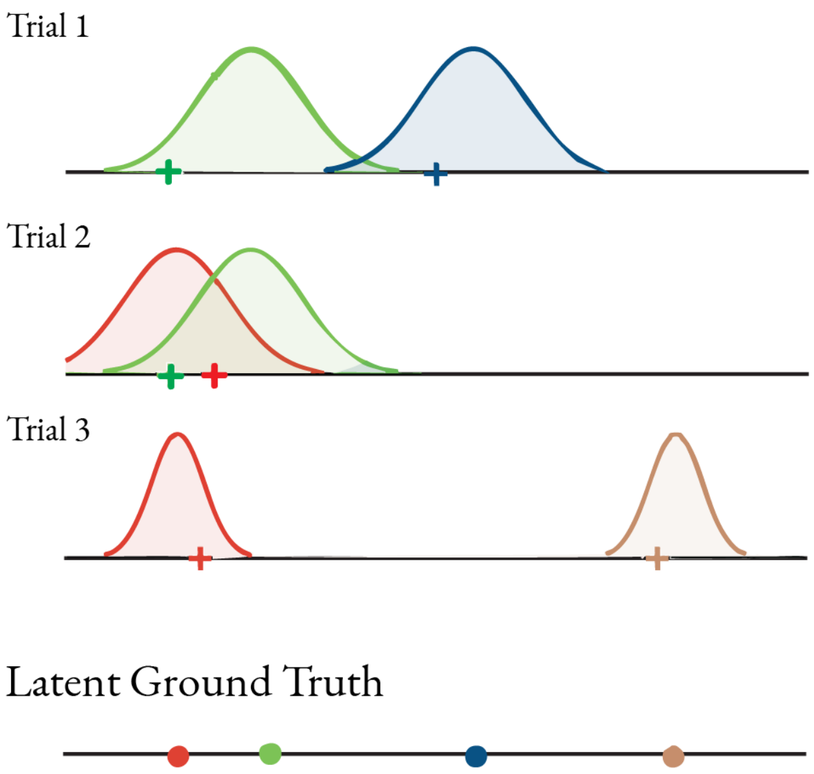
For historical reasons, the similarity-based MDS-representation has been considered **the psychological representation** underlying cognition.

But **similarity judgments and cognition may not share the same mental representation.**

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual review of Psychology*, *43*(1), 25-53.
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317-1323.

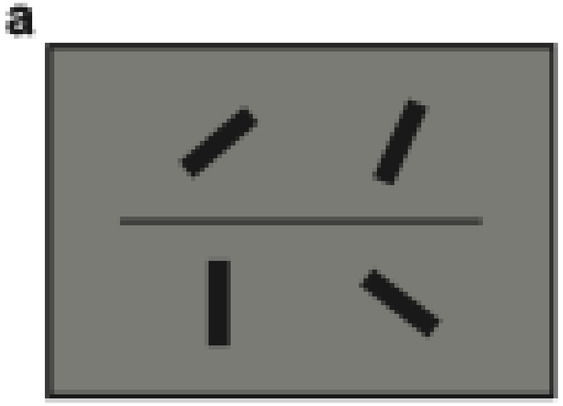
Our modeling approach

In brief, we used Bayesian MCMC methods to recover the **latent representation of oriented lines** used in three cognitive tasks.

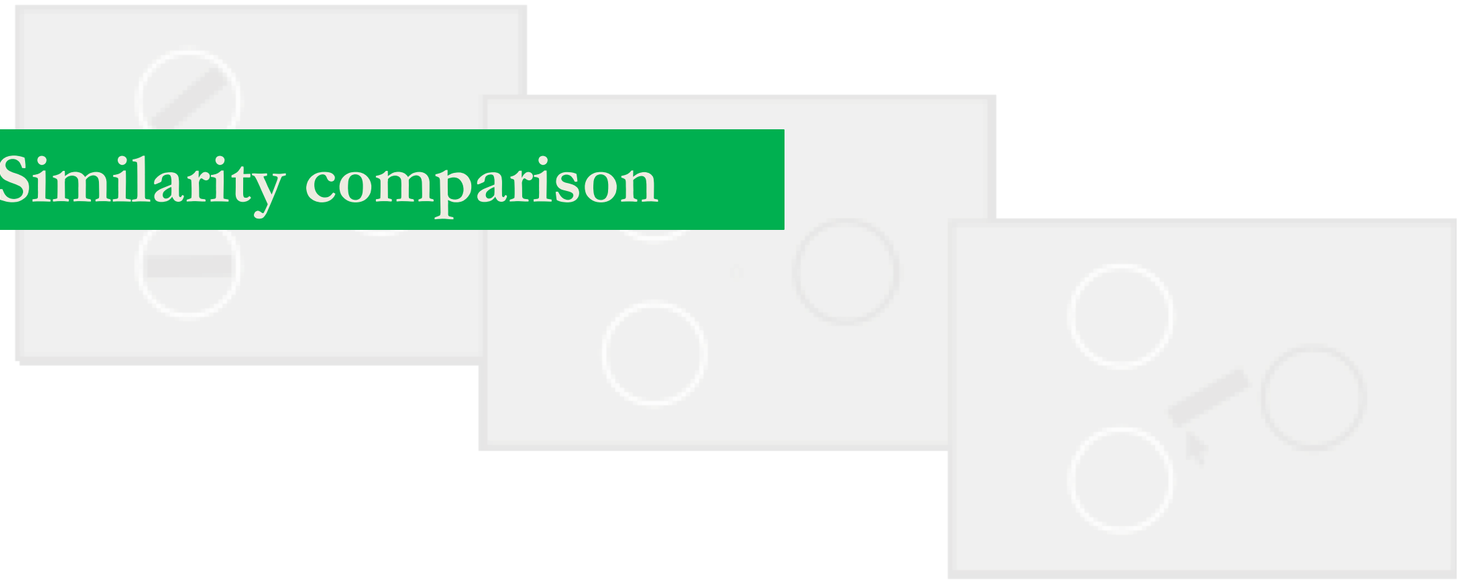


We use the observed similarity ratings to infer the latent representation of those items.

The cognitive tasks

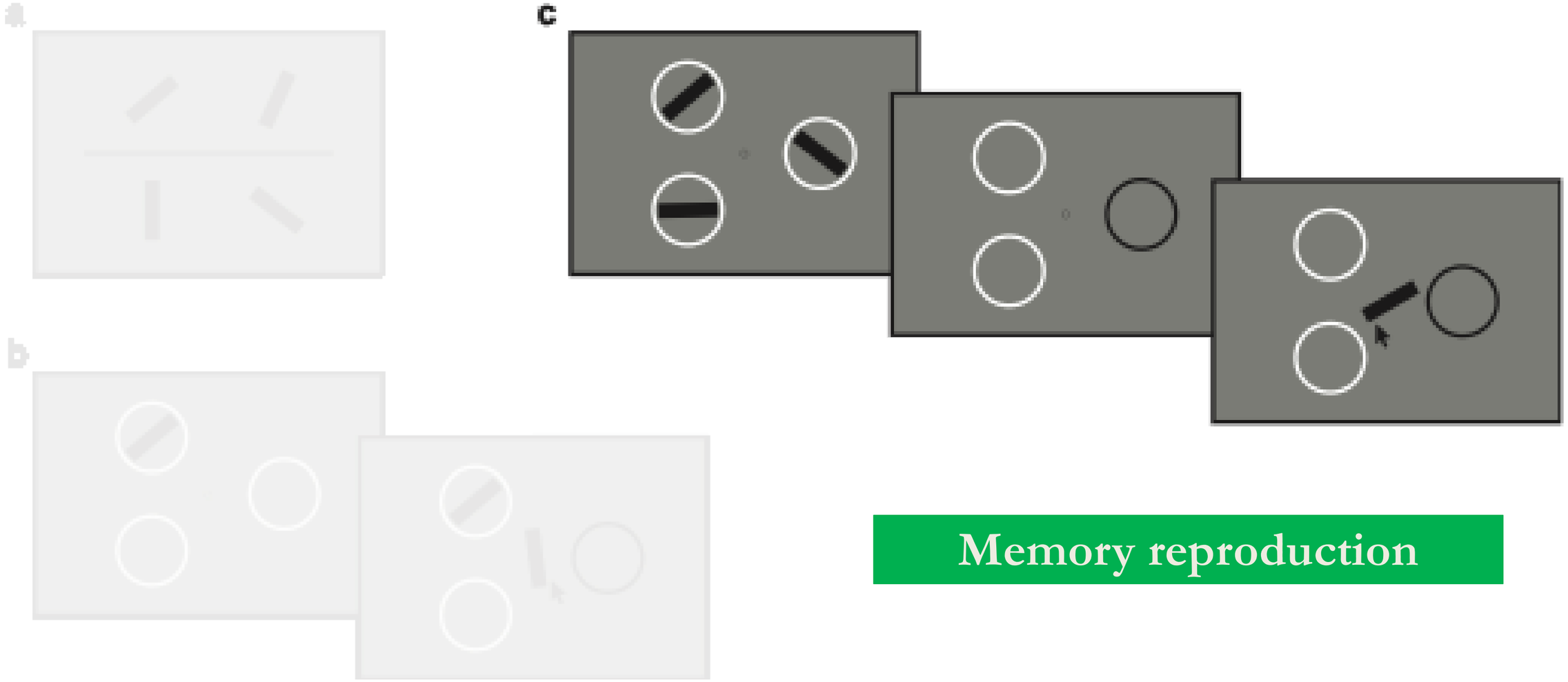


Similarity comparison



Tomić, I., & Bays, P. M. (2024). Perceptual similarity judgments do not predict the distribution of errors in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 50(4), 535.

The cognitive tasks

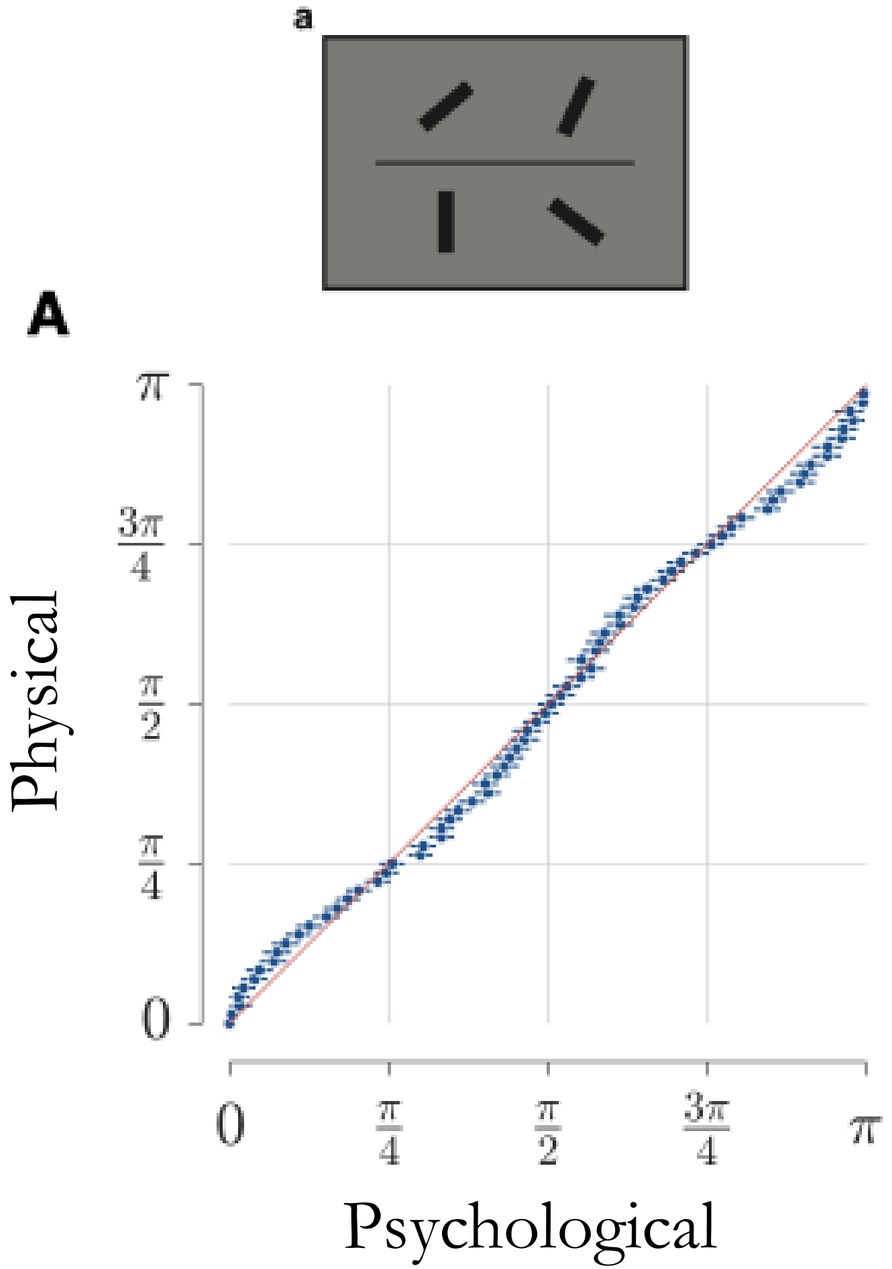


Tomić, I., & Bays, P. M. (2024). Perceptual similarity judgments do not predict the distribution of errors in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 50(4), 535.

Similarity comparison

The representation does not match the physical stimulus space – it is not exactly a diagonal line.

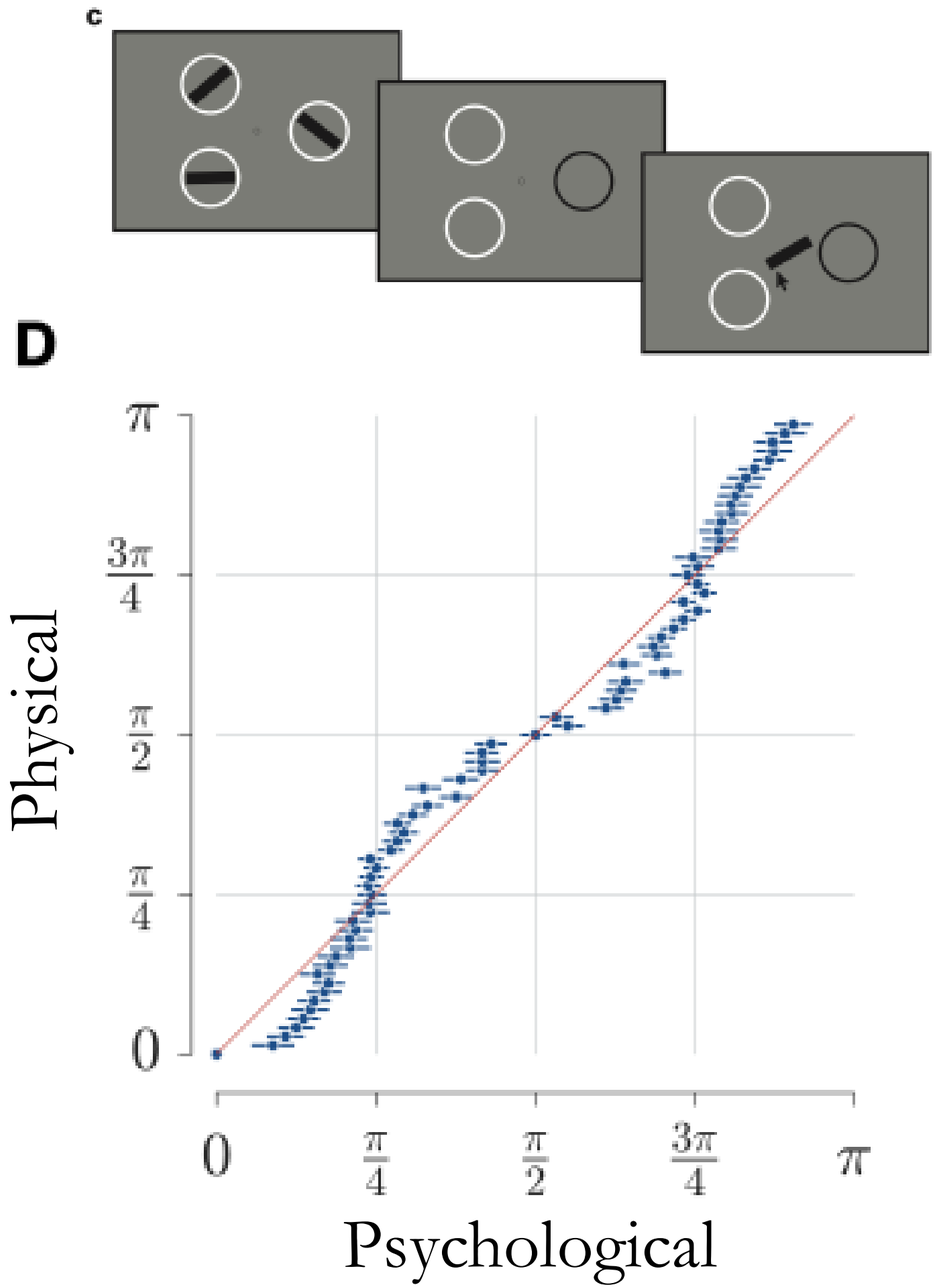
Clear deviations where close to vertical lines appear more vertical, and close to horizontal lines appear more horizontal.



Memory reproduction

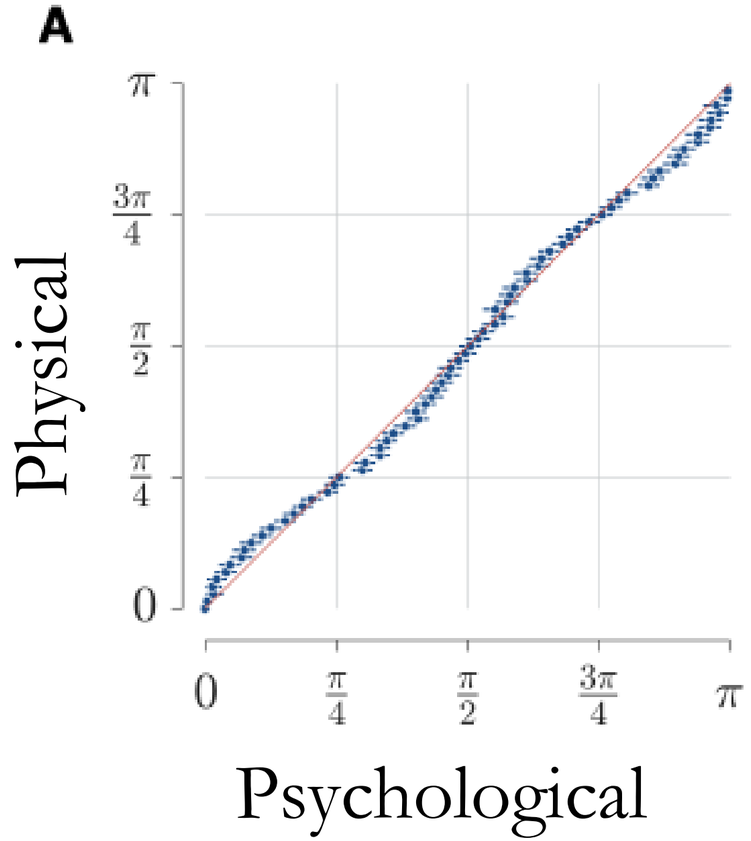
The representation for oriented lines do not exactly match the physically presented stimulus.

In working memory, the oriented lines are represented more towards the oblique directions than they actually are.

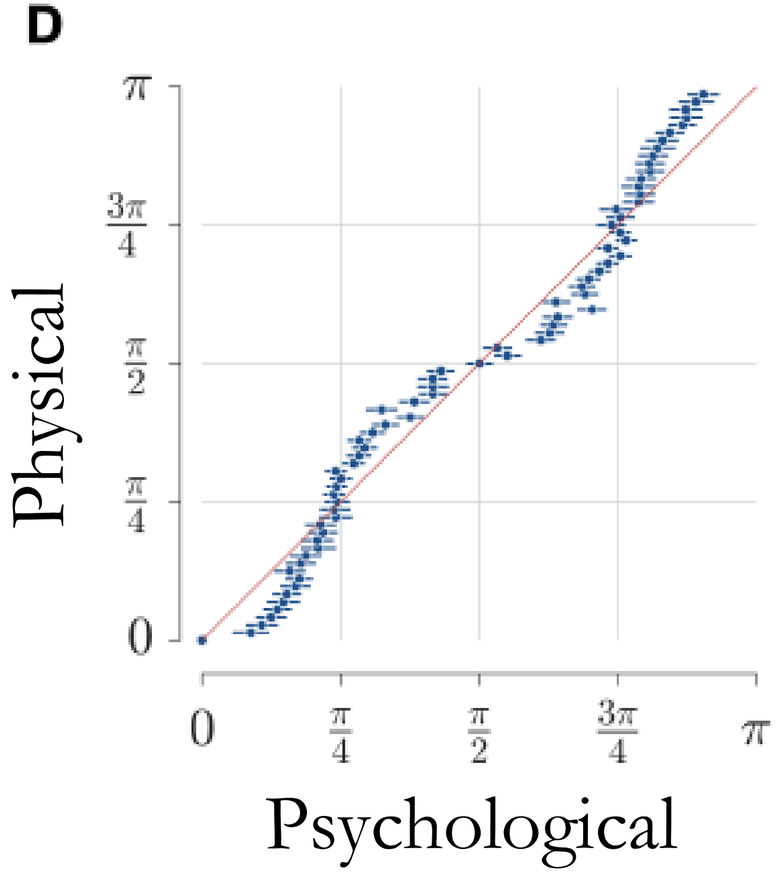


Similarity comparisons and both reproduction tasks **do not share the same cognitive representation.**

Similarity comparison



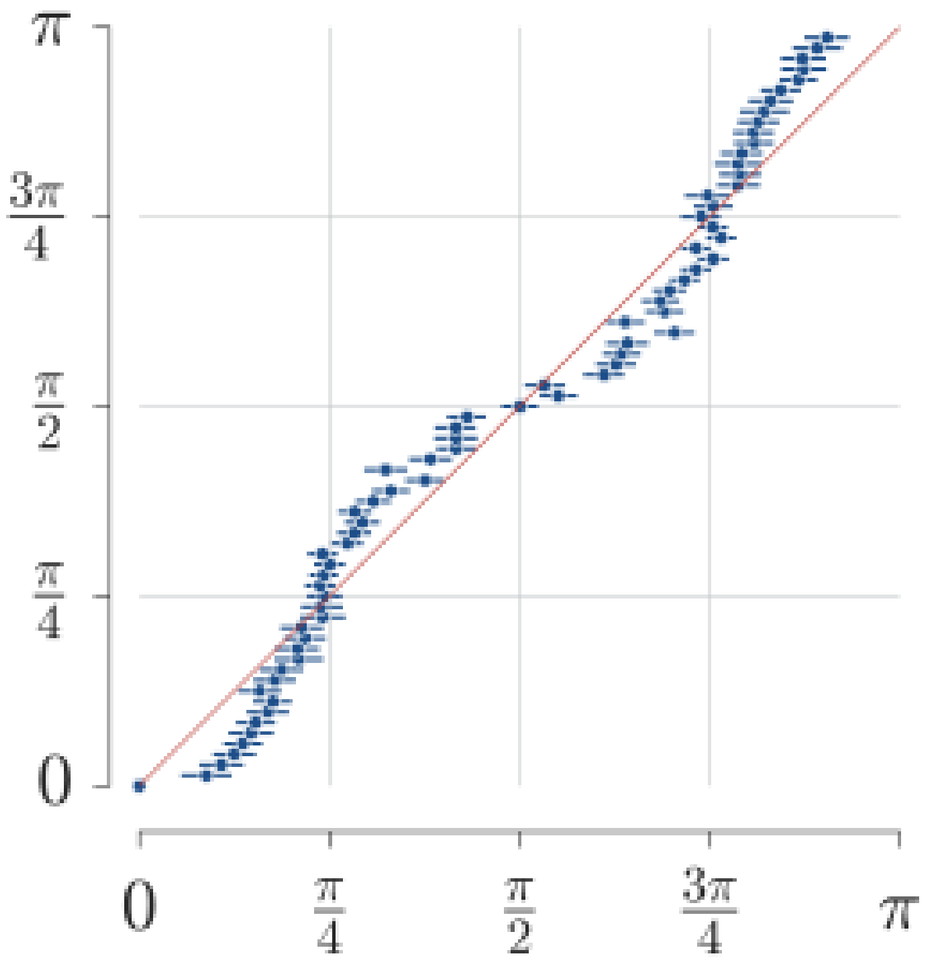
Memory reproduction



Does psychological similarity predict neural similarity?

Psychological similarity cannot be assumed to be the basis for cognition (and **possibly not the basis for similarity of neural representations**).

D



Takeaways

We ought to have formal cognitive models that predict and explain the similarity of neural representations.

We cannot assume that psychological similarity will predict neural similarity as the cognitive representation may not be the same.

How do we integrate formal models of cognition with empirical neuroscience? Come to **MathPsych** next year to find out…

Thank you!

william.ngiam@adelaide.edu.au

https://palm-lab.github.io

@williamngiam.github.io

Adelaide University

Preprint