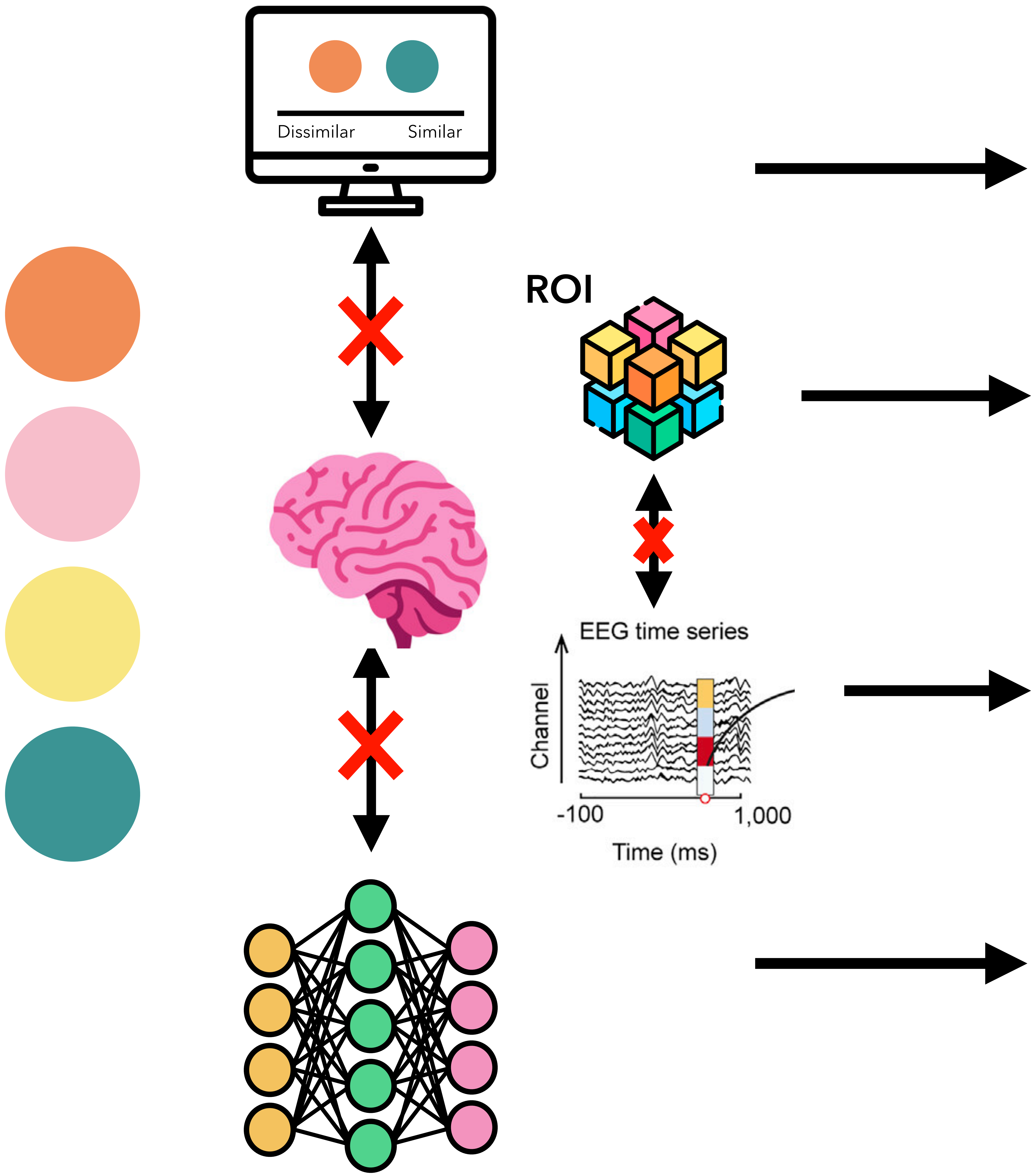
REPRESENTATIONAL SIMILARITY ANALYSIS

FOR THE COMMET (DREAM) TEAM

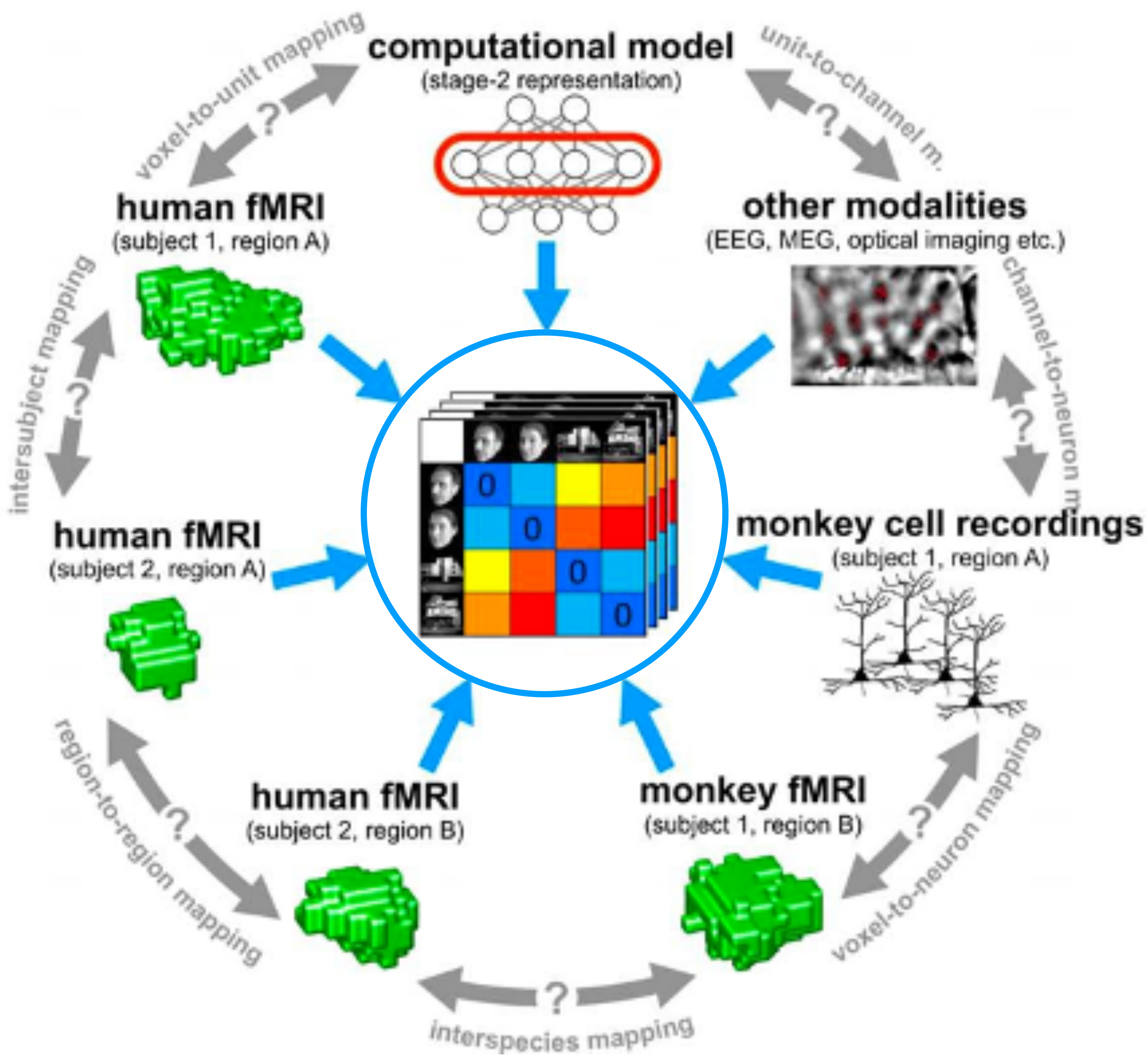
CAMILLE GRASSO - 22/05/26

Representational Similarity Analysis



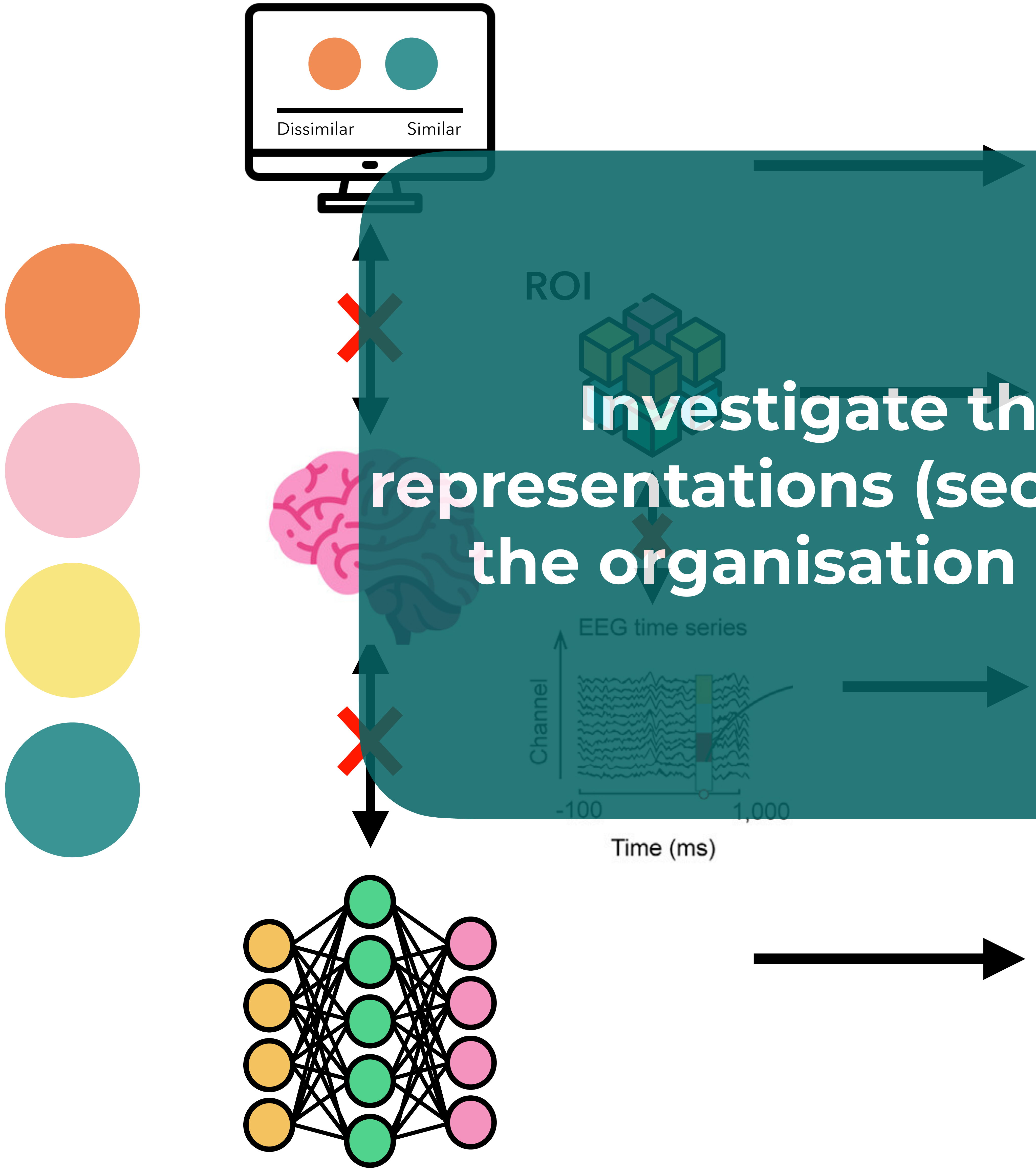
Representational dissimilarity matrix (RDM)

Common representational space

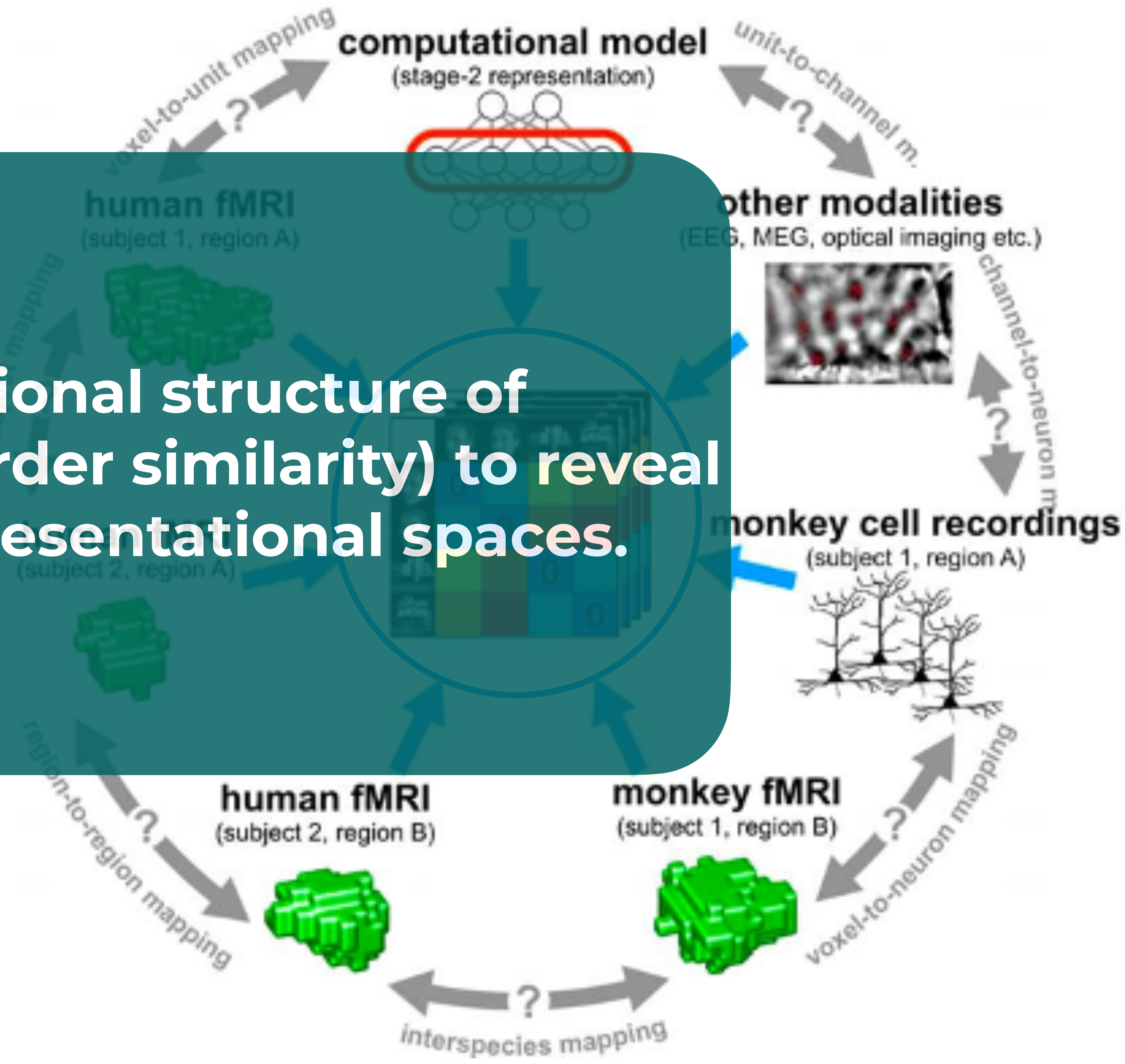


Representational Similarity Analysis

Common representational space



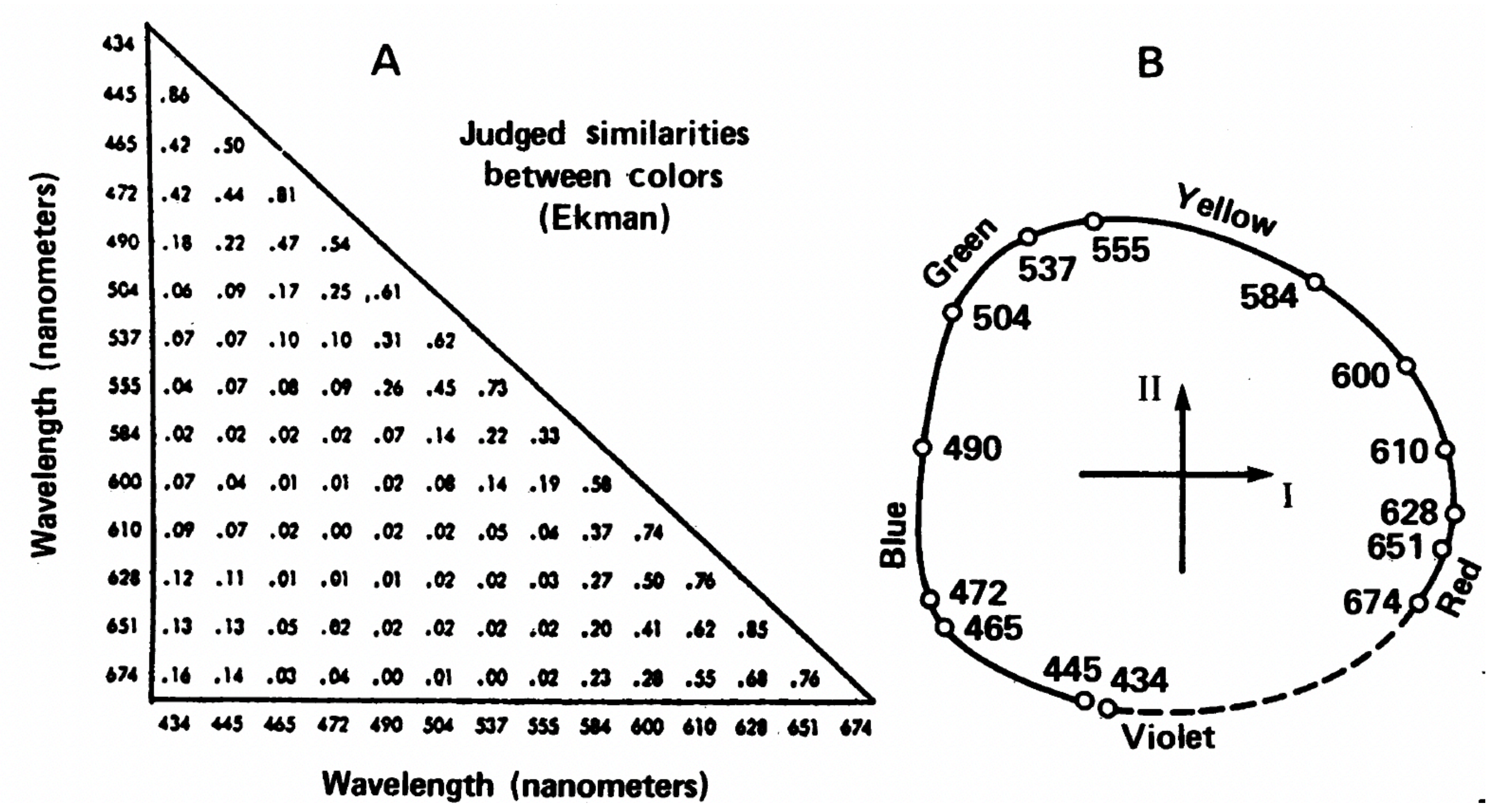
Representational dissimilarity matrix (RDM)

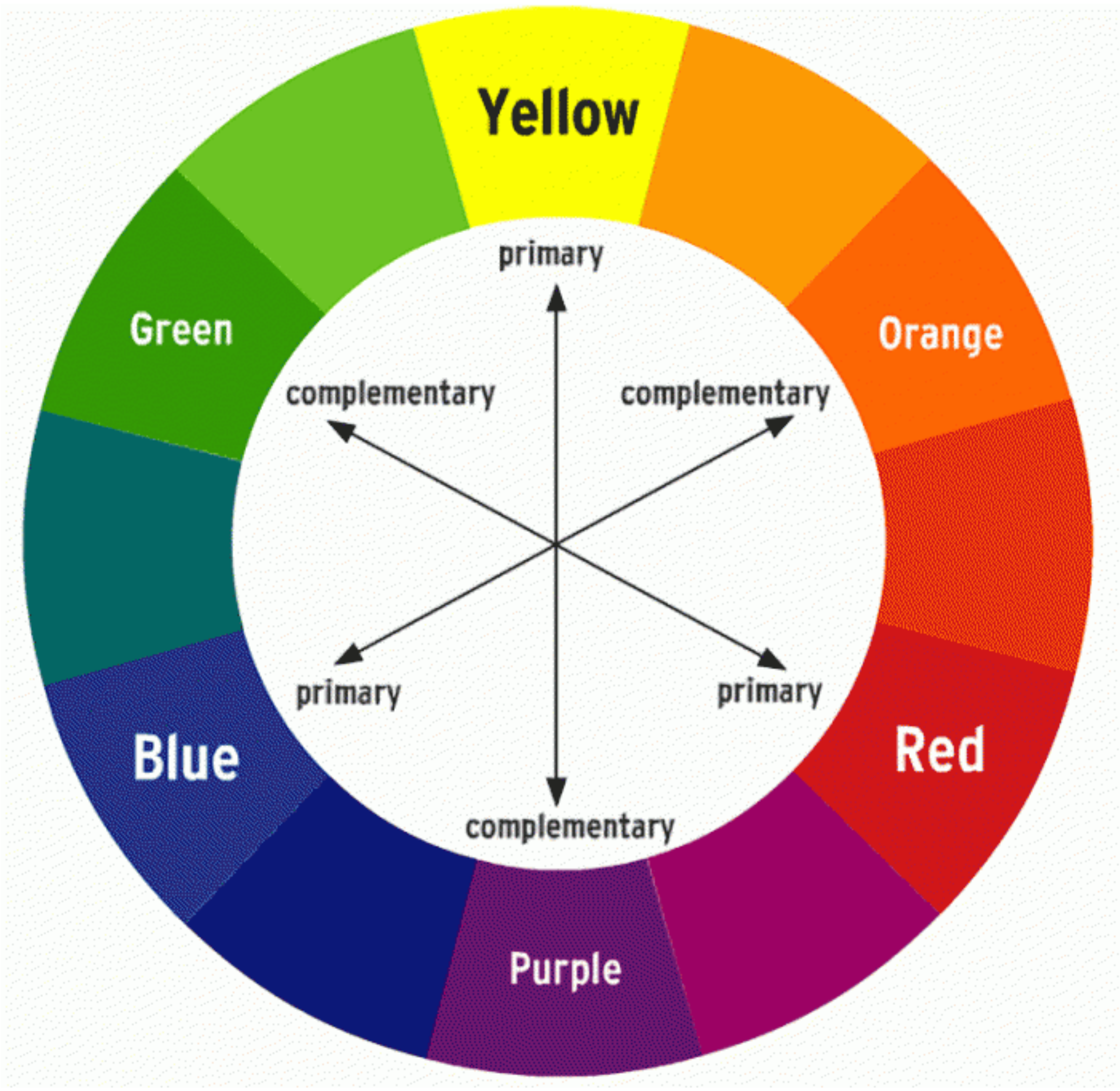


Investigate the relational structure of representations (second-order similarity) to reveal the organisation of representational spaces.

Investigating the (geometrical) structure of representations

Investigating the relational structure of subjective experiences (second-order similarity) to reveal the organisation of conceptual spaces.





Edelman (1995; 1998); Edelman & Duvdevani-Bar (1997); Shepard (1980); Shepard & Chipman (1970); Shepard et al. (1975).

Investigating the (geometrical) structure of representations

Inspired by work from Shepard (1982) and others. The pitch helix "explains" why tones separated by an octave sound similar despite the pitch distance (height vs. chroma).

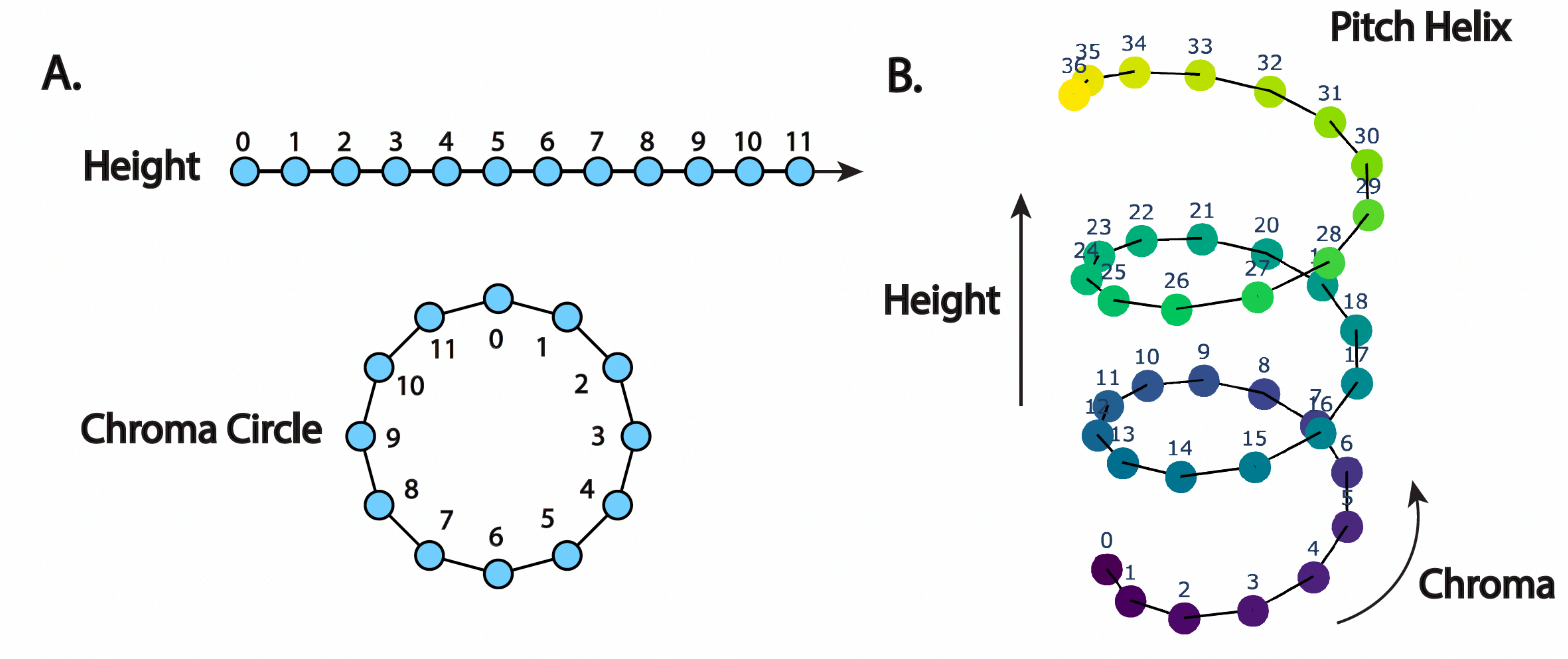


Fig. 1 The pitch helix representation and its underlying components, namely, the pitch height line and the chroma circle.

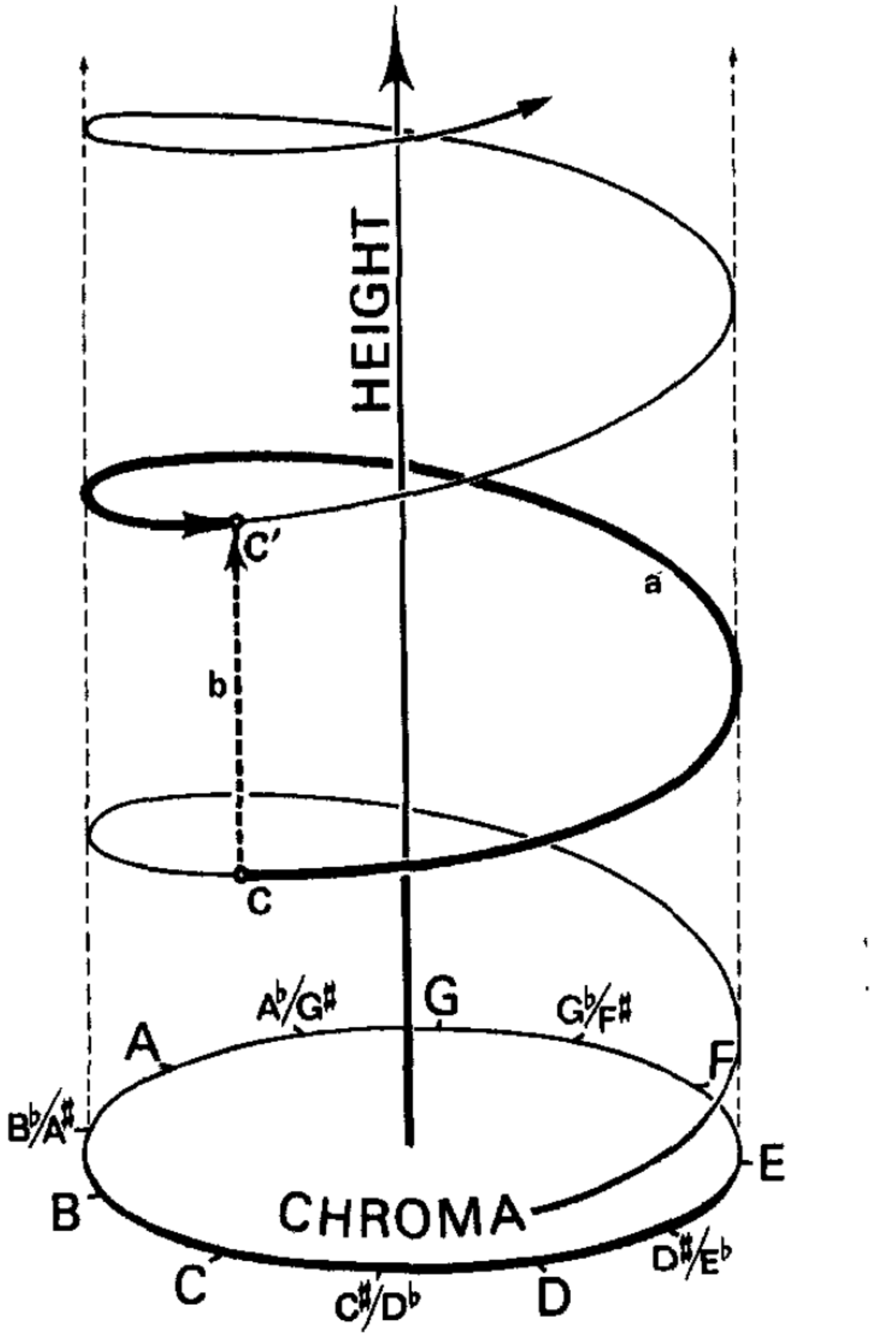


Figure 1. A simple regular helix to account for the increased similarity between tones separated by an octave. (From "Approximation to Uniform Gradients of Generalization by Monotone Transformations of Scale" by Roger N. Shepard. In D. I. Mostofsky (Ed.), *Stimulus Generalization.* Stanford, Calif.: Stanford University Press, 1965, p. 105. Copyright 1965 by Stanford University Press. Reprinted by permission.)

Marjieh et al., (2024); Kriegerkorte et al. (2008).

RDMs as a common representational space

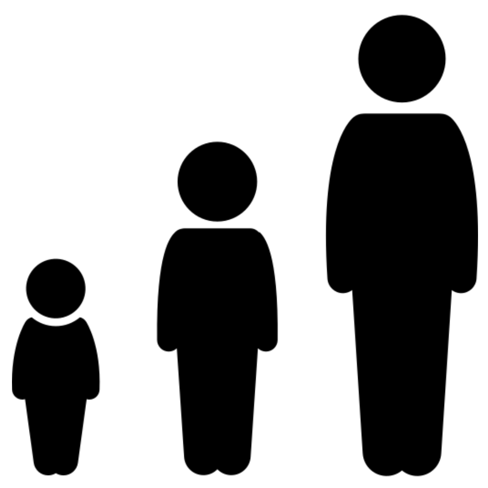
Neural data: s/EEG, MEG, fMRI, single units
Behavior: (explicit) similarity judgments, (implicit) reaction times
Models: RNNs, LLMs
Individuals: typical vs. atypical, children vs adults, species, humans vs. models

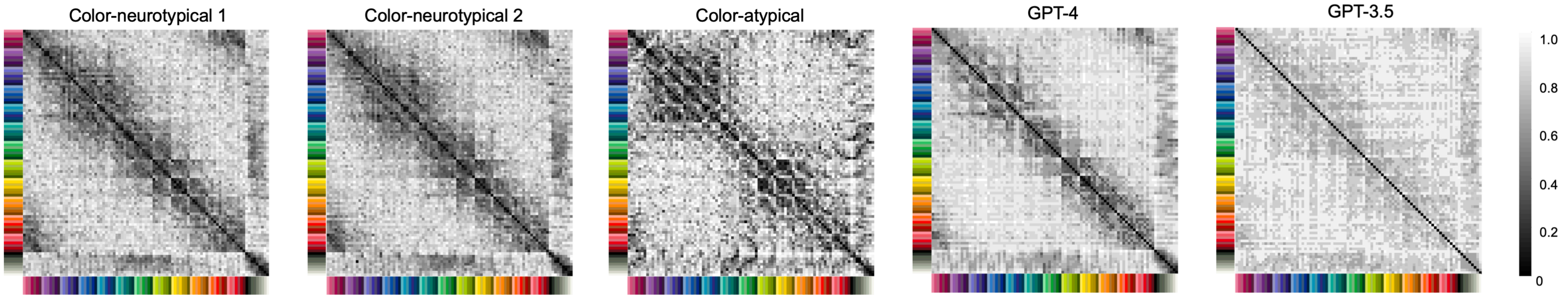
PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES OPEN ACCESS

Spontaneous eye movements reflect the representational geometries of conceptual spaces

Simone Viganò[a,b,1], Rena Bayramova[a,c], Christian F. Doeller[a,d,2], and Roberto Bottini[b,1,2]

Edited by Michael Goldberg, Columbia University, New York, NY; received February 27, 2024; accepted March 13, 2024



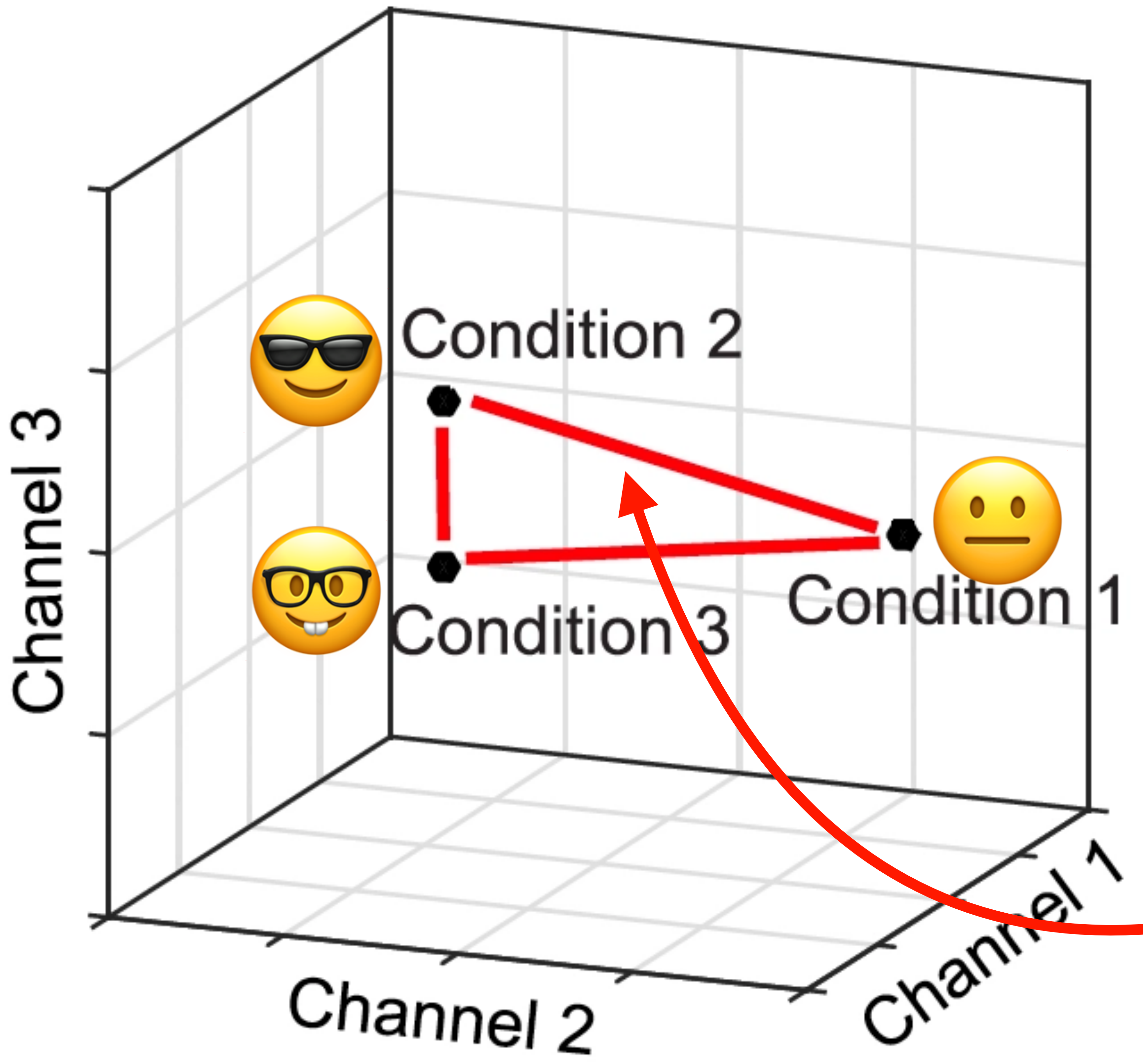


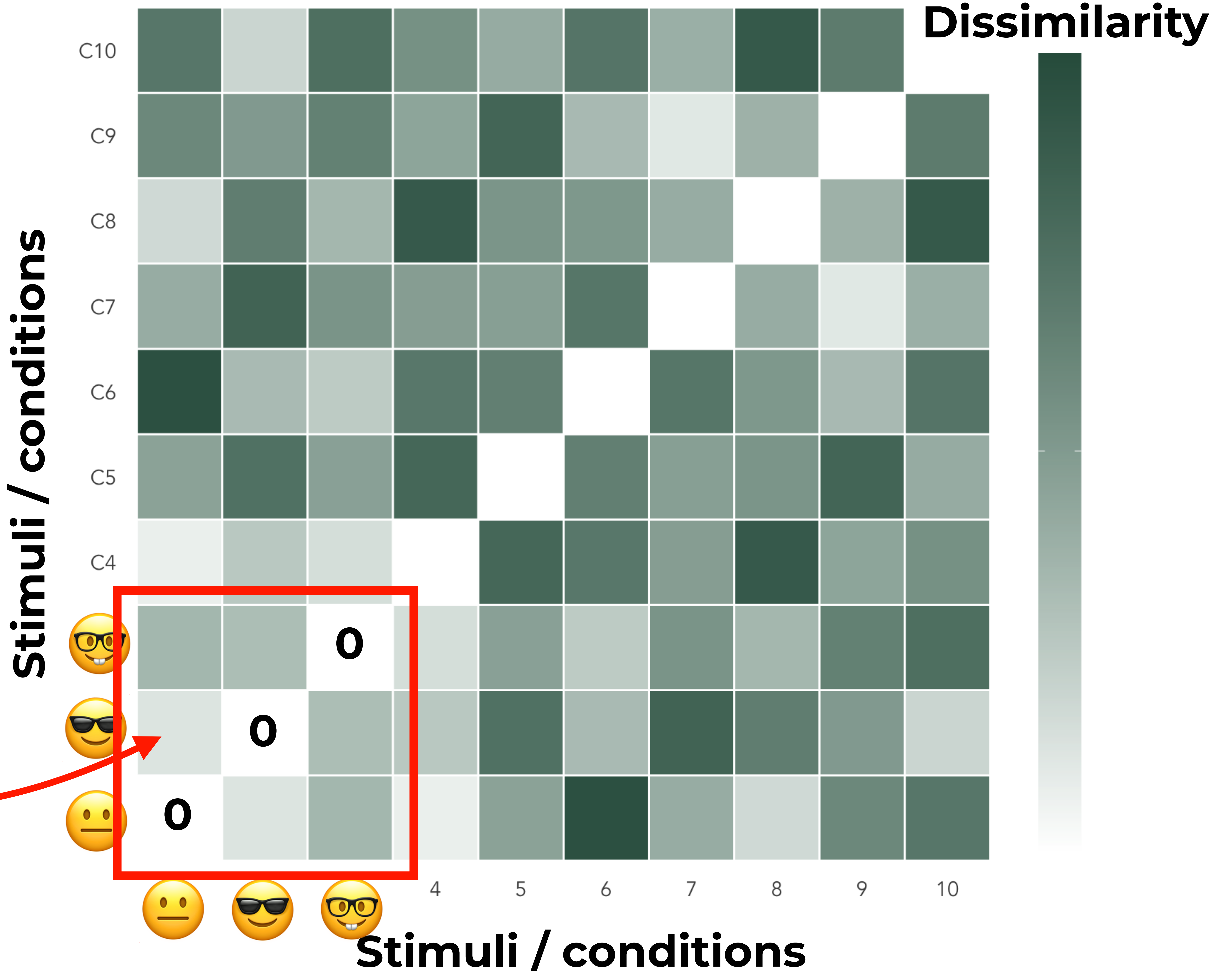
We need : same conditions + meaningful pairwise dissimilarities.

Vigano et al., 2024; Kawakita et al., 2023a, 2023b; Xie et al., 2022;

RDMs as a common representational space

- Square matrix (n conditions x n conditions)
- Diagonal = 0 (condition compared to hitself)
- Each cell = 1 dissimilarity value
- Usually symmetric



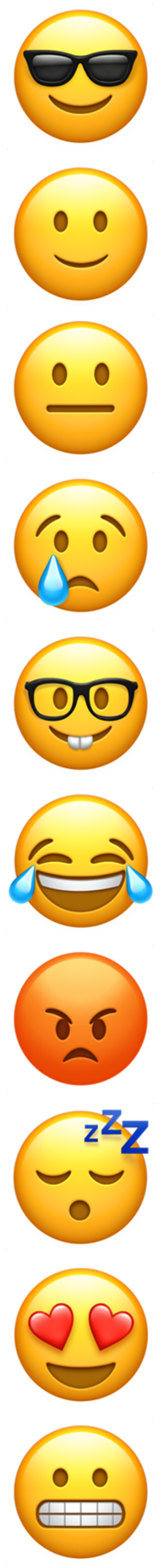


STEPS OF RSA

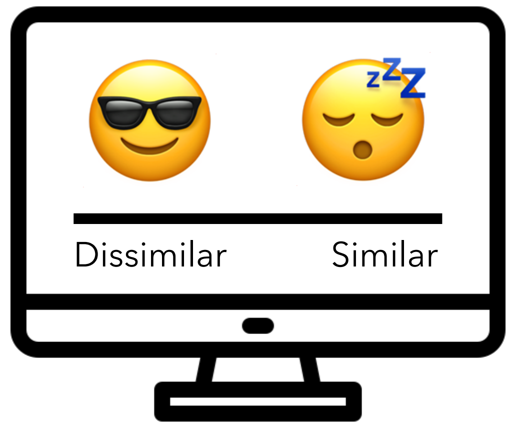
BUILD RDMs

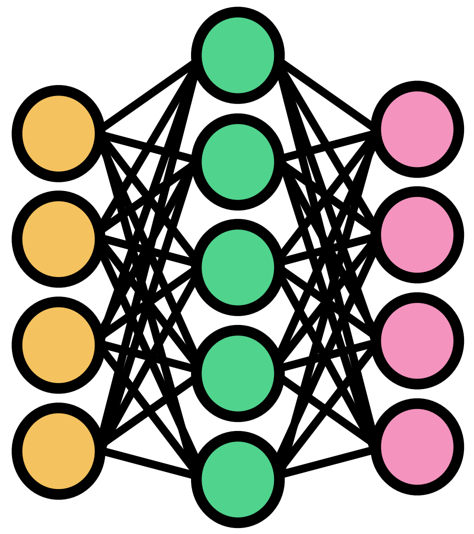
COMPARE RDMs

Stimuli

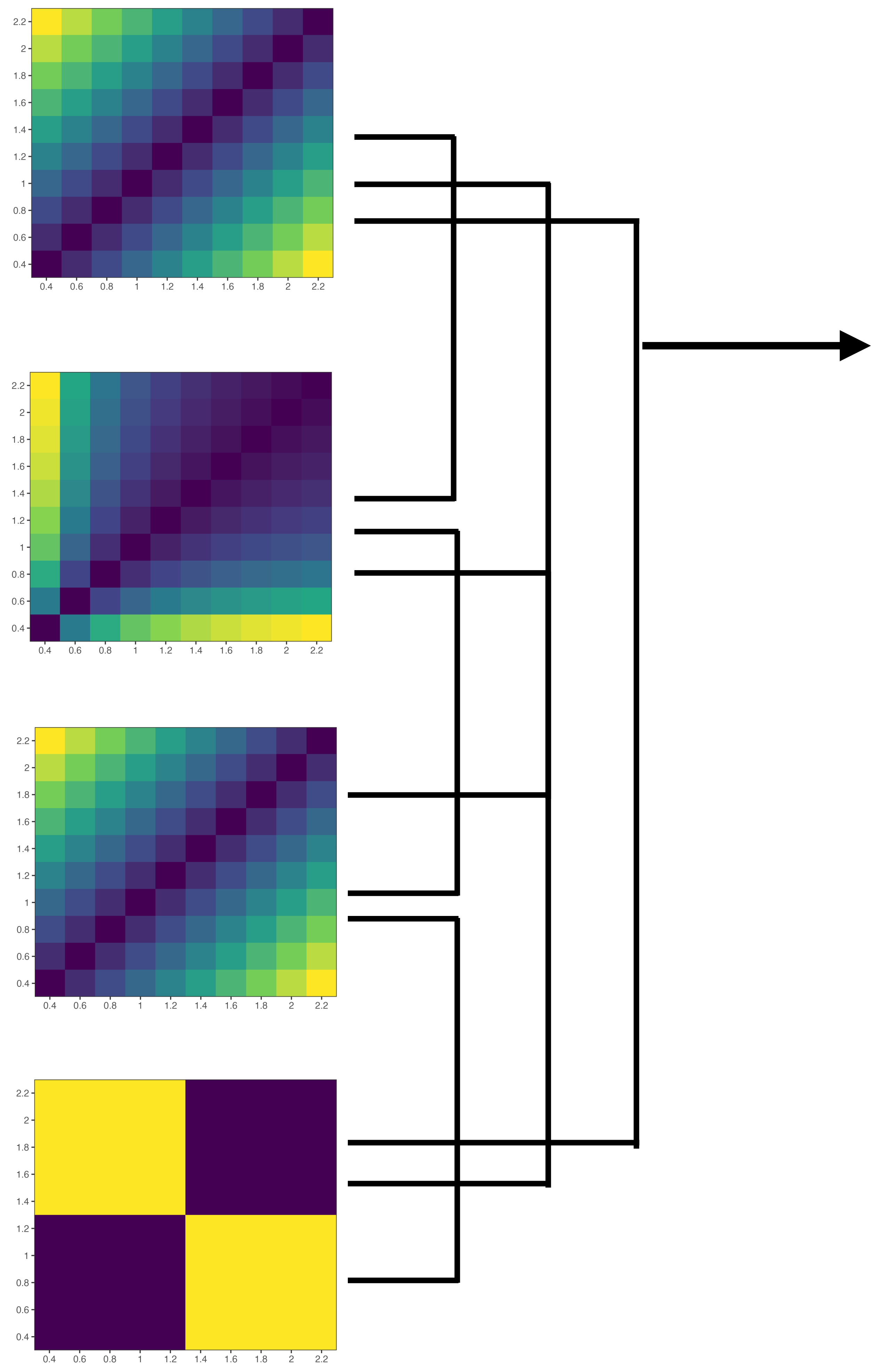


Empirical RDMs





Theoretical RDMs

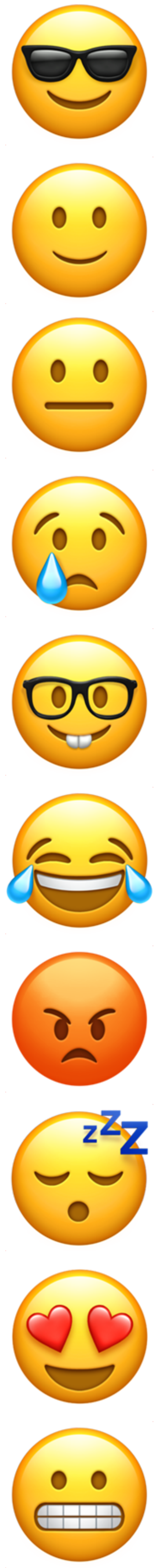


Spearman’s ρ
Kendall’s τ
Pearson r
Regression / GLM
Cross-validated prediction

Is the match between RDMs reliable?

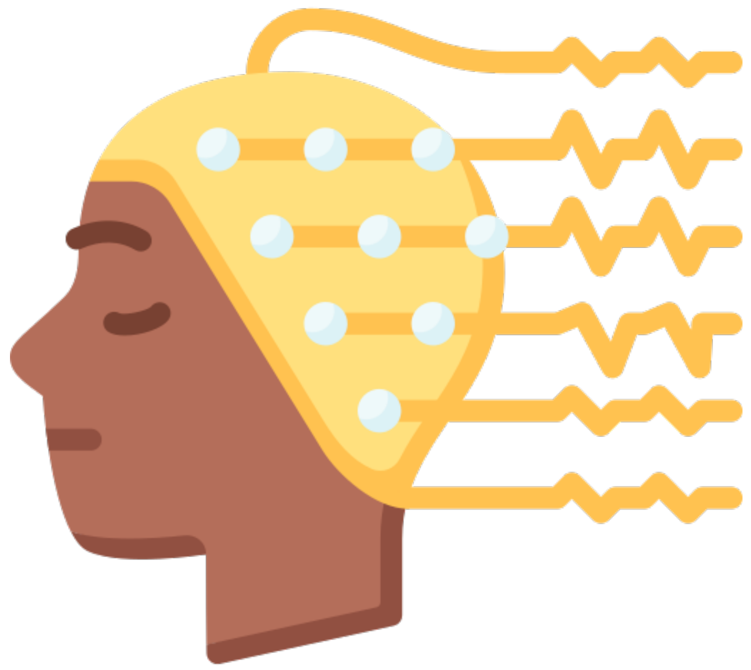
Which structure/model best explains the empirical geometry?

Stim

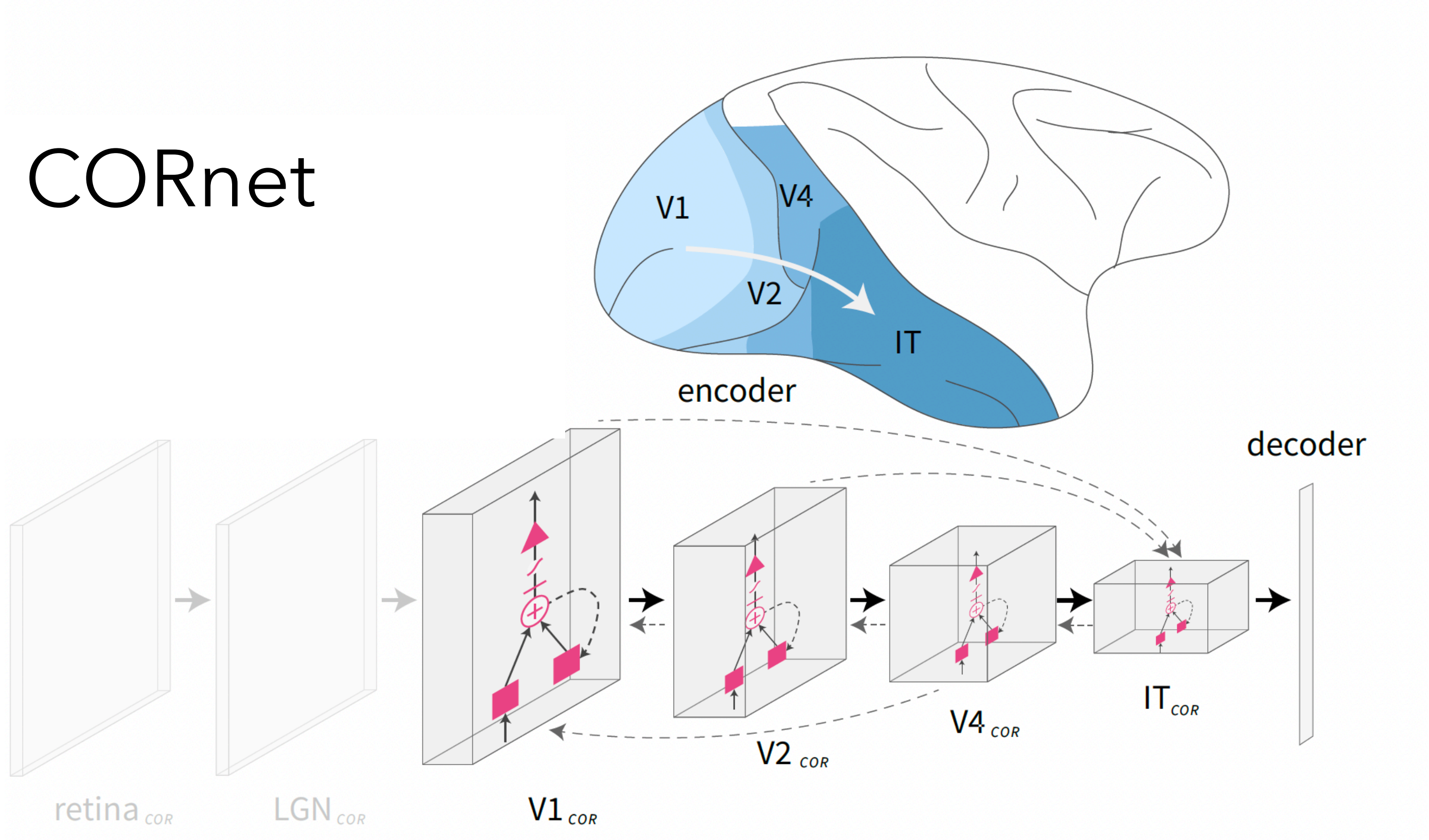


fMRI

EEG



CORnet



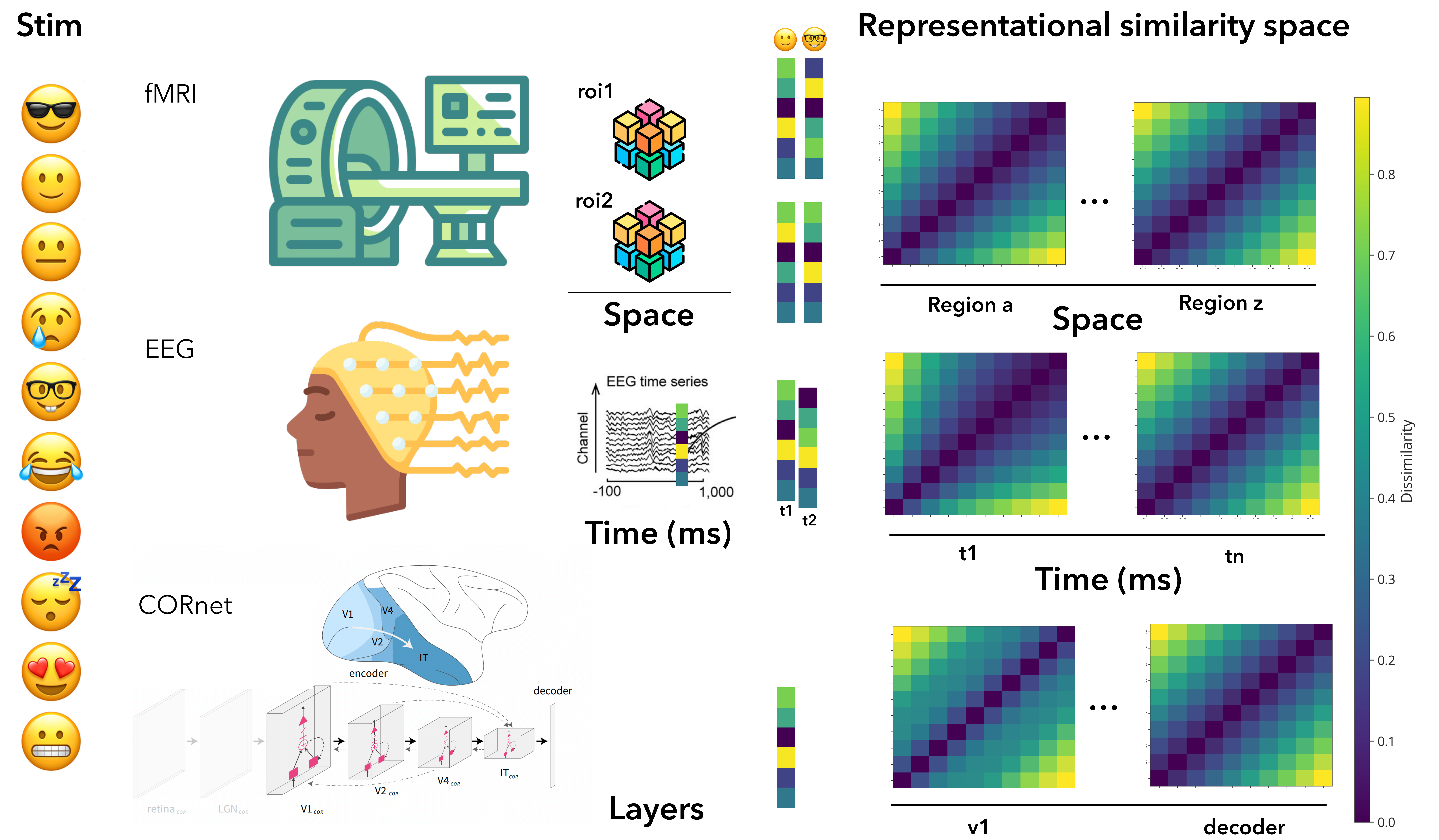
Where **is the geometry of emojis representations?**

When **does it emerge?**

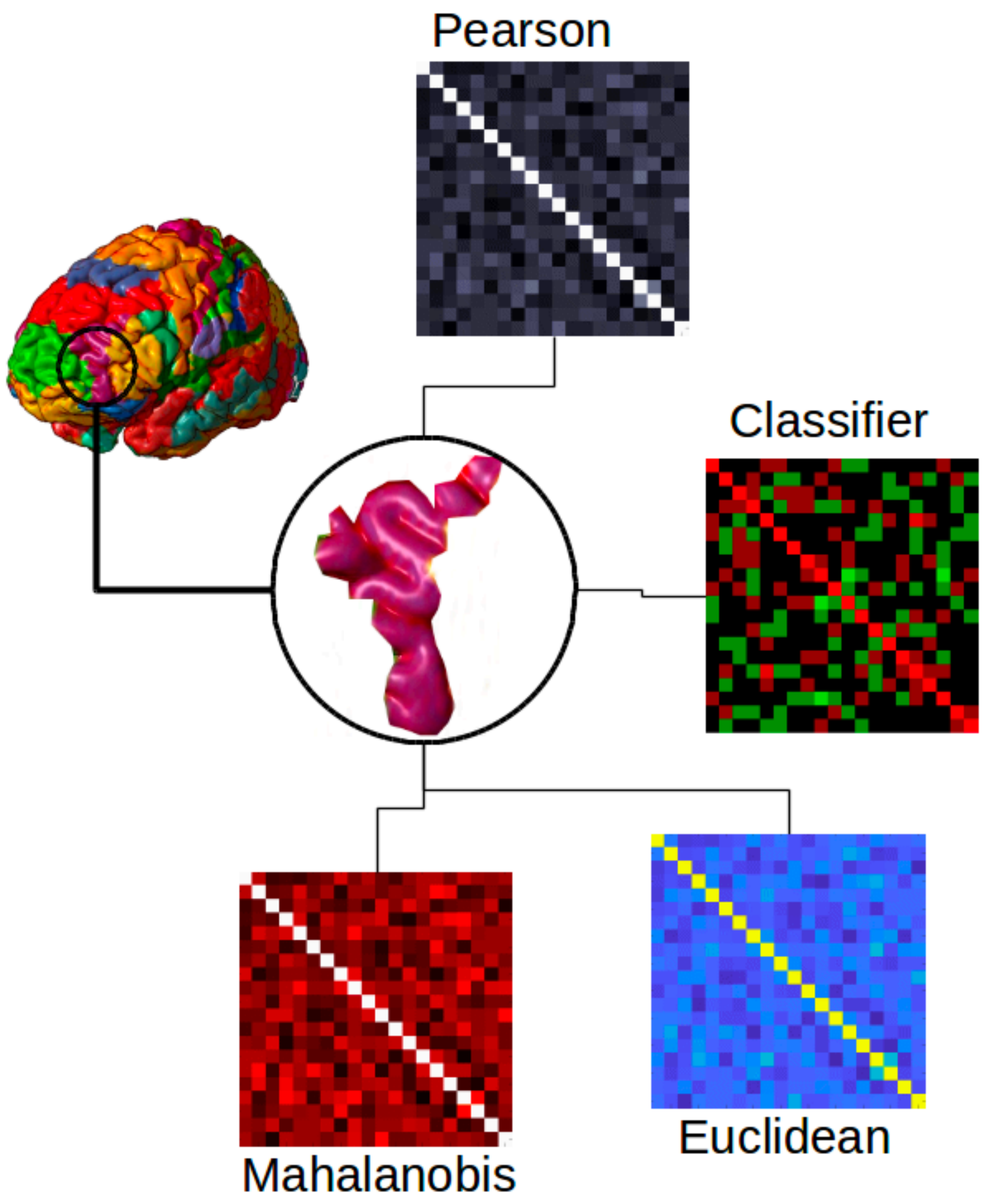
Which computational stage **does it resemble?**

Cichy & Oliva, 2020; Hebart, 2018

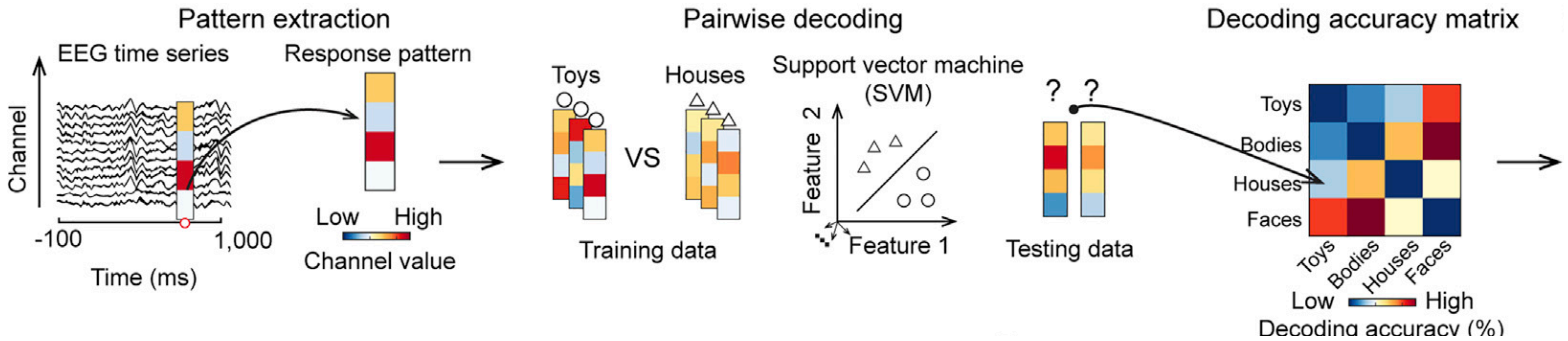

One
Eternity
Later


Stim
fMRI
roi1
roi2
Space
EEG
EEG time series
Channel
-100
1,000
Time (ms)
CORnet
V1
V4
V2
IT
encoder
decoder
retina
LGN
V1
V2
V4
IT
Layers
t1
t2
Representational similarity space
Region a
Space
Region z
t1
Time (ms)
tn
v1
decoder
Dissimilarity
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0

Which dissimilarity measure?



Correlation distances : 1-correlation (Pearson, Spearman, Kendall): look at the shape and ignore the global activation level. Weak with noisy data.
Euclidean distance : assumes that all dimensions are equally reliable/ independent.
Mahalanobis distance : neural-data-friendly Euclidean distance, take into account the covariance structure of the noise. Requires to estimate the noise covariance well.
Cross-validated dissimilarity (e.g., crossnobis = Mahalanobis cross-validated) : reduce the influence of noise. Zero a meaningful interpretation.
Decoding accuracy

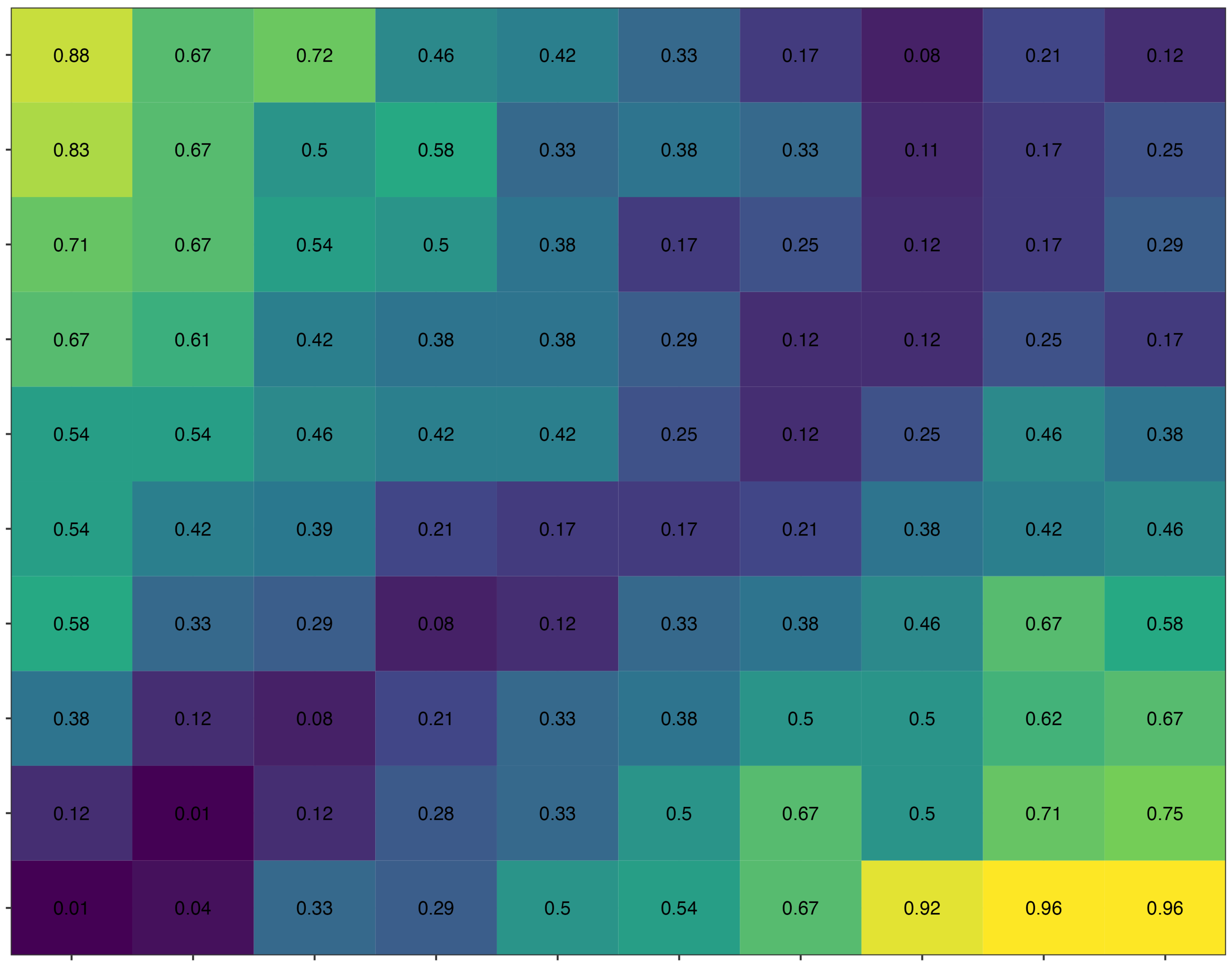


RSA best practice: **Crossnobis** or similar cross-validated measures are now often preferred for their superior reliability and interpretability.

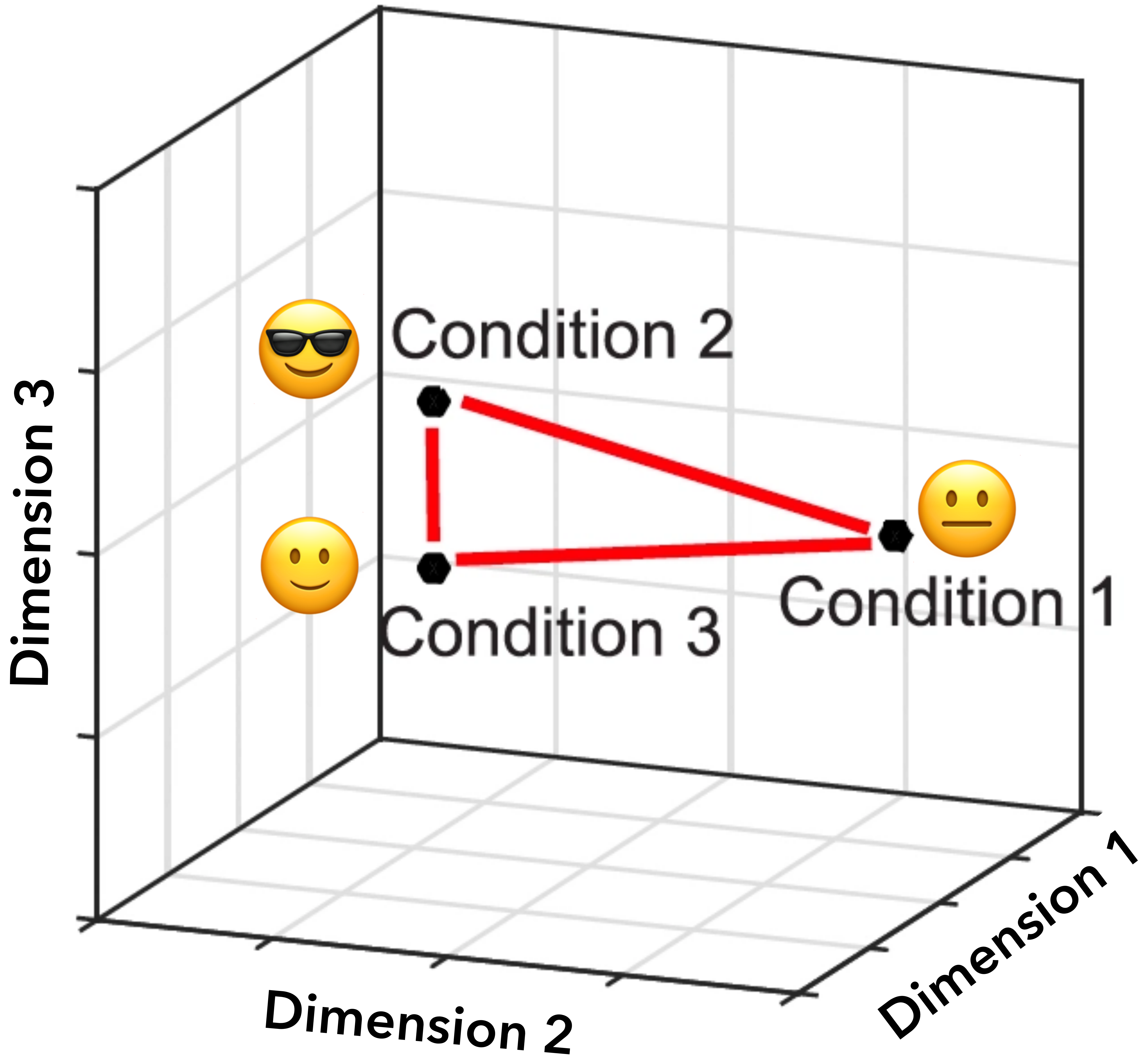
Bodadilla-Suarez et al., 2018; Xie et al. 2022; https://nikokriegeskorte.org/2019/01/09/whats-the-best-measure-of-representational-dissimilarity/

Building empirical RDMs

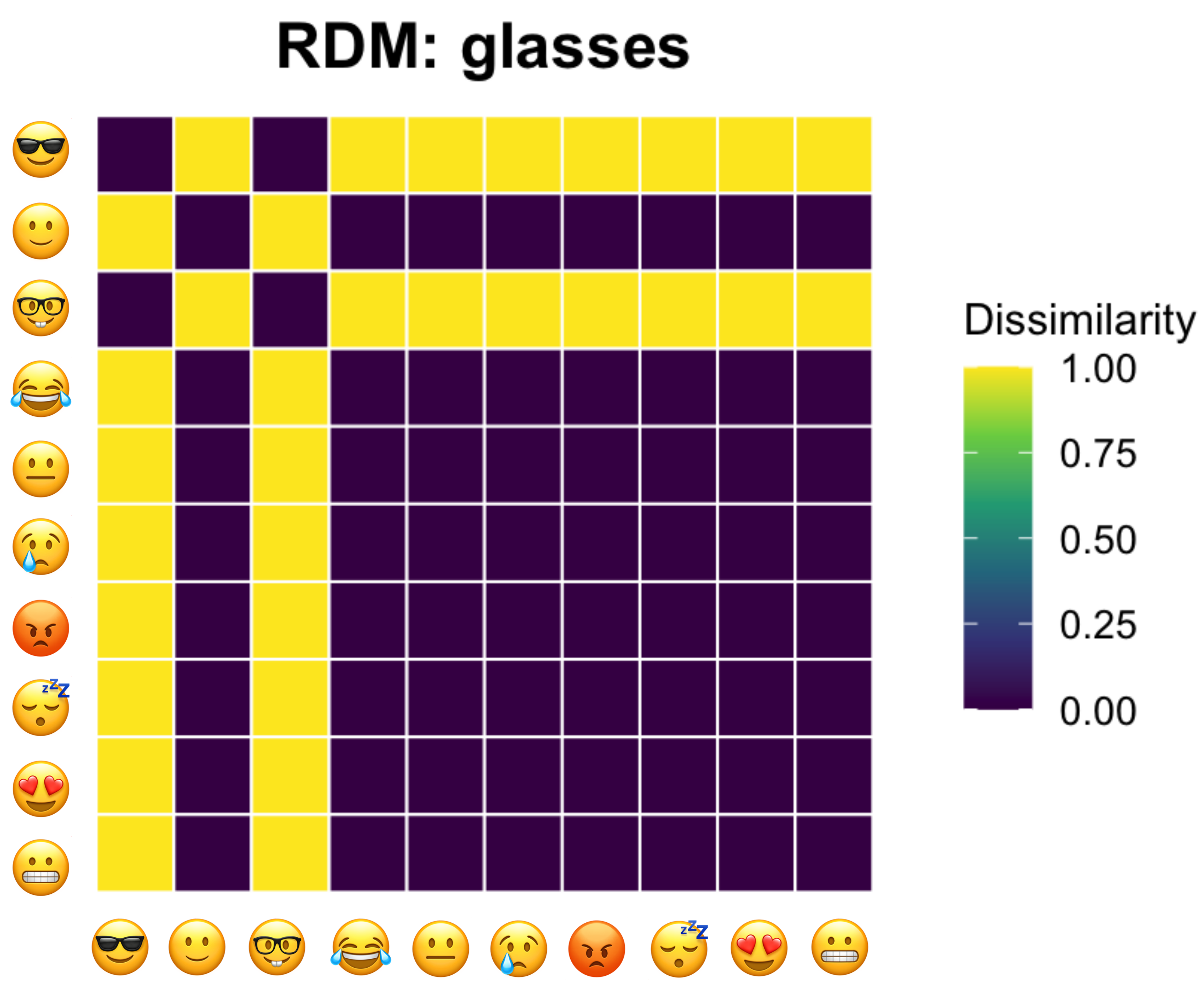
Visualizing the representational geometry

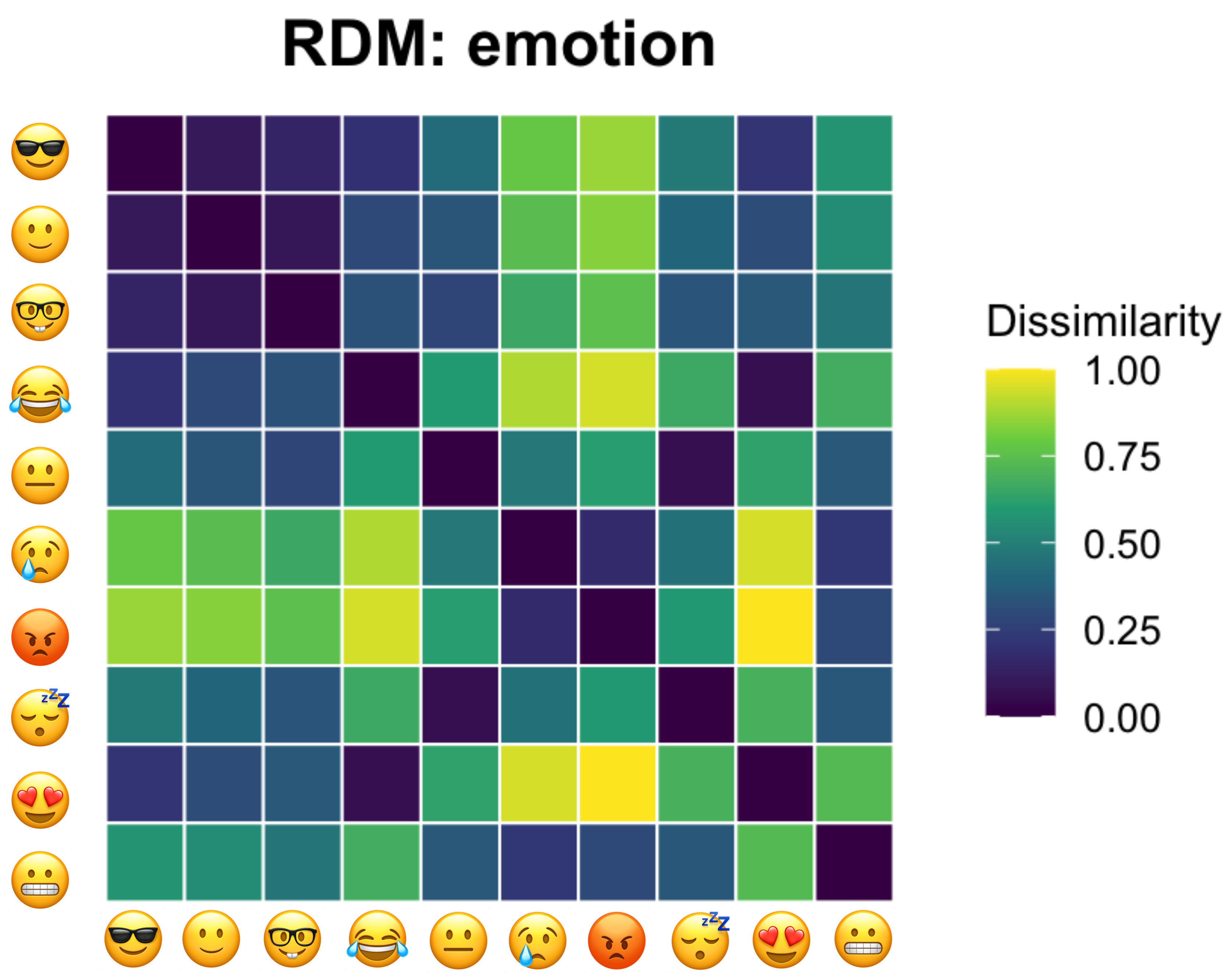


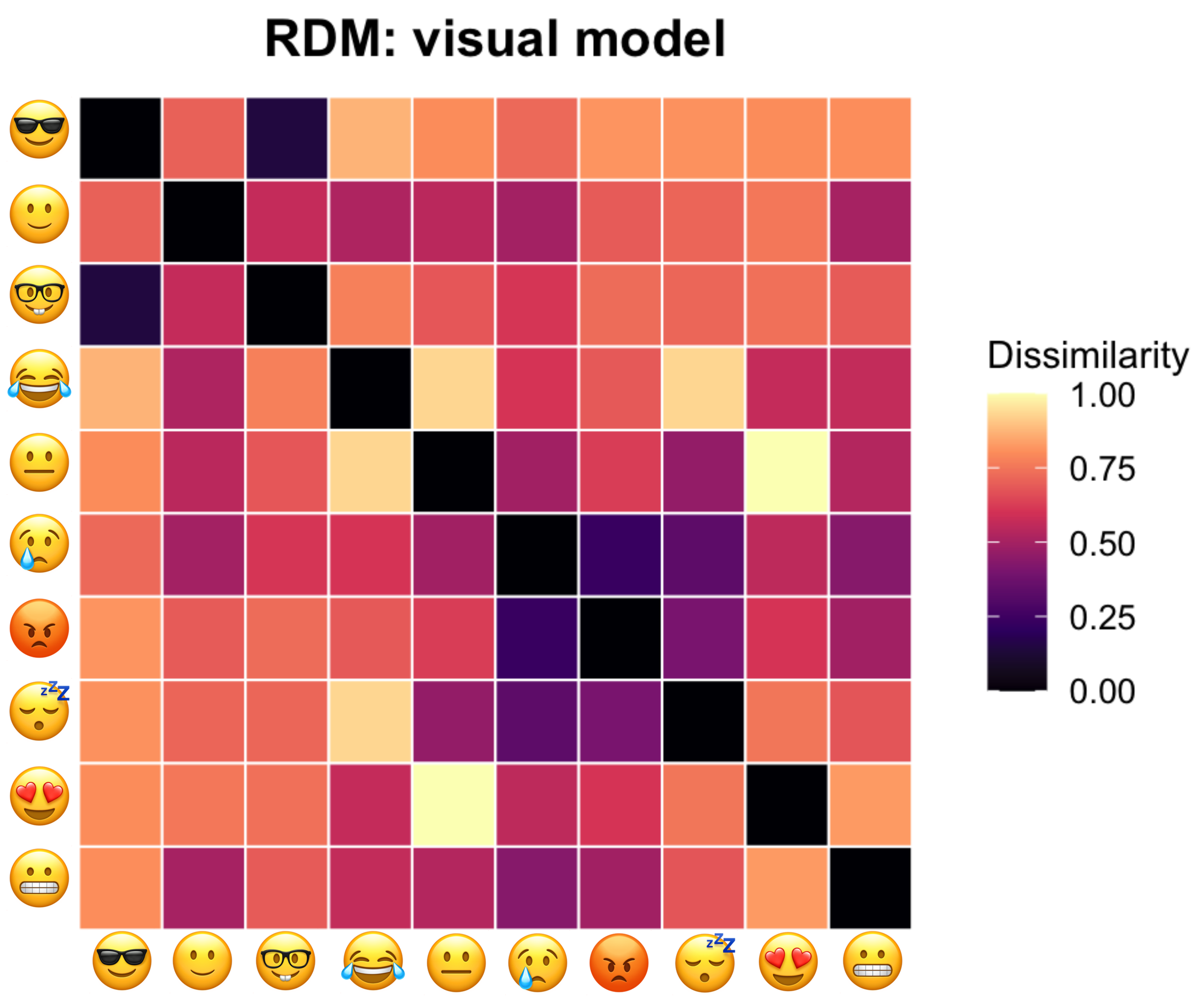
MDS
t-SNE
...



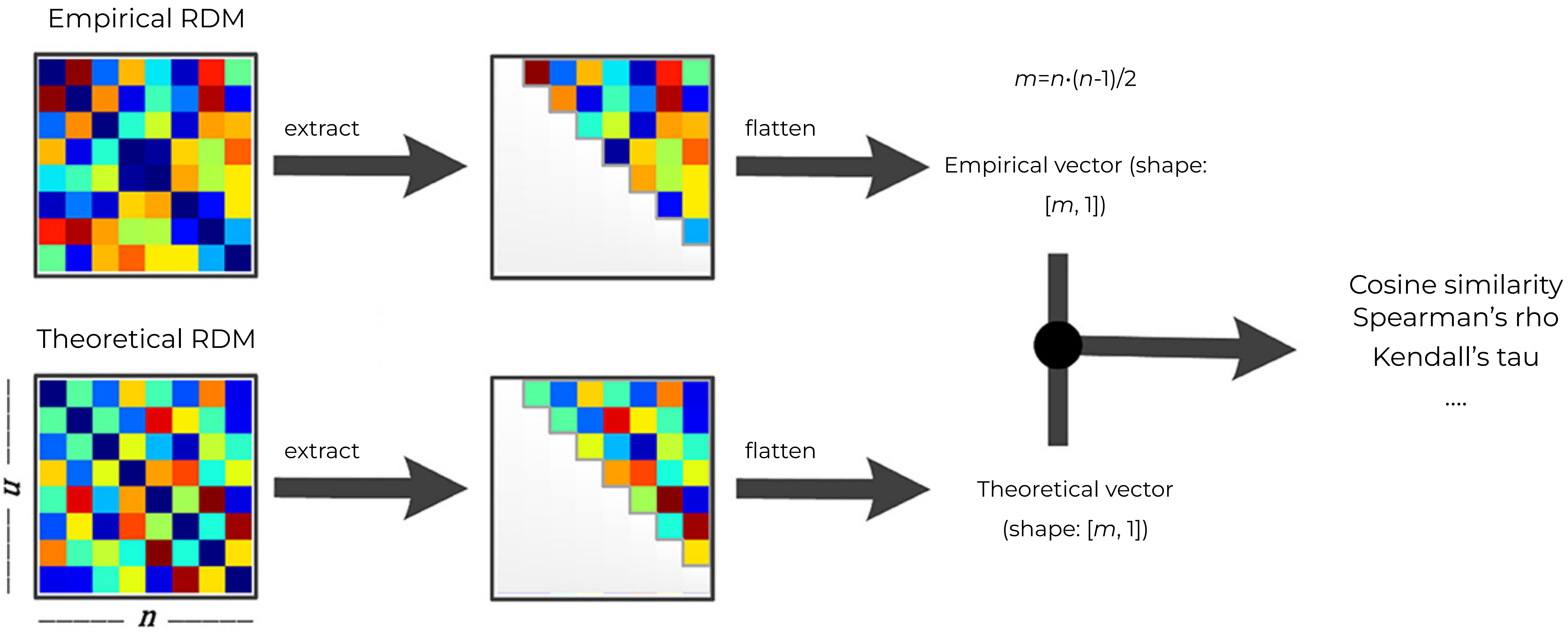
Building theoretical RDMs







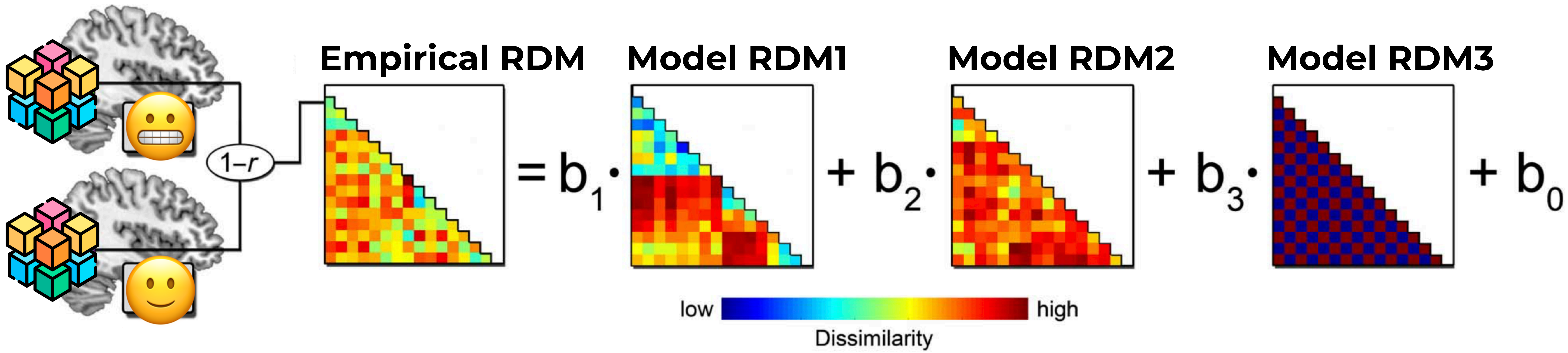
Model comparison (model-fit estimates)



(Make sure to use the same order for both vectors)

https://rsatoolbox.readthedocs.io/en/stable/comparing.html

Model comparison (model-fit estimates)

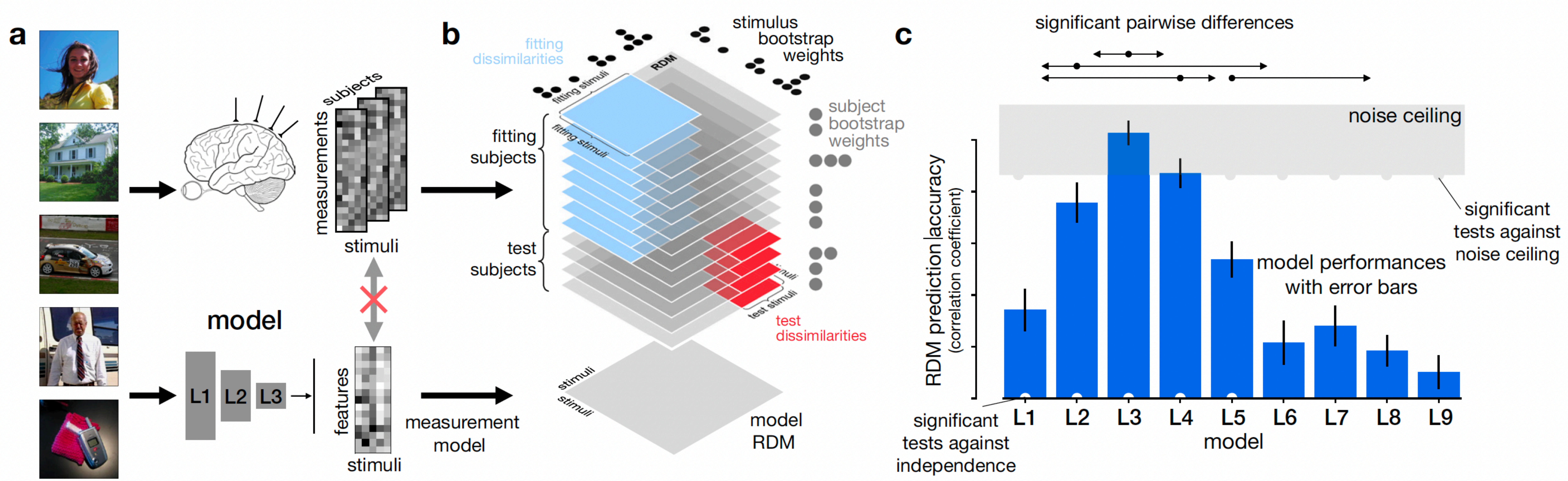


The empirical RDM is the outcome we want to predict and the theoretical RDMs are the predictors

Adapted from Proklova et al. 2016

Statistical inference

Test the reliability and generalizability of the models, when they emerge over time.

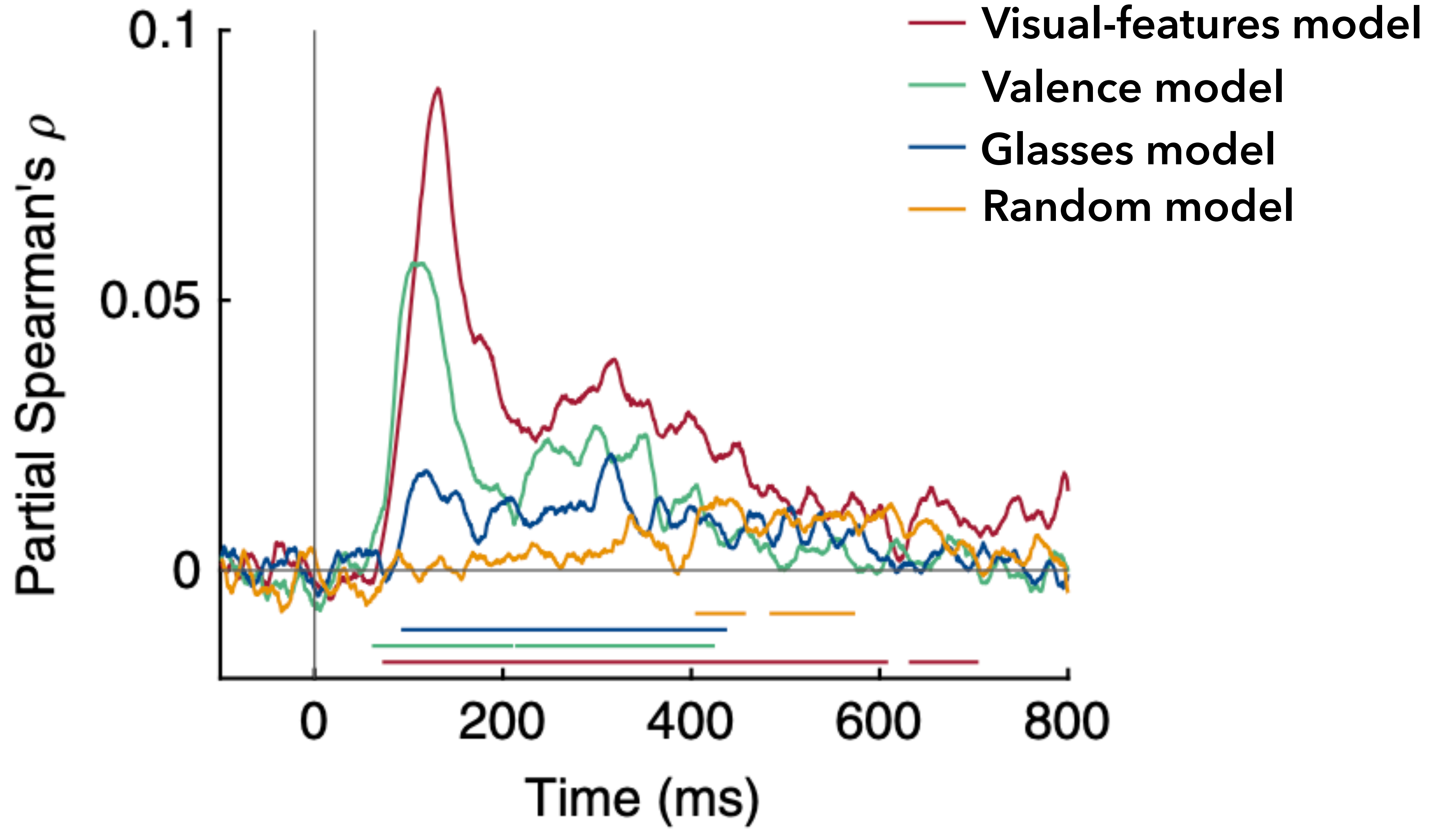


Diedrichsen & Kriegeskorte, 2017 ; Dobs et al., 2019; Schutt et al., 2022 ; See: https://mne.tools/mne-rsa/stable/auto_examples/tutorials/03_statistics_tutorial.html

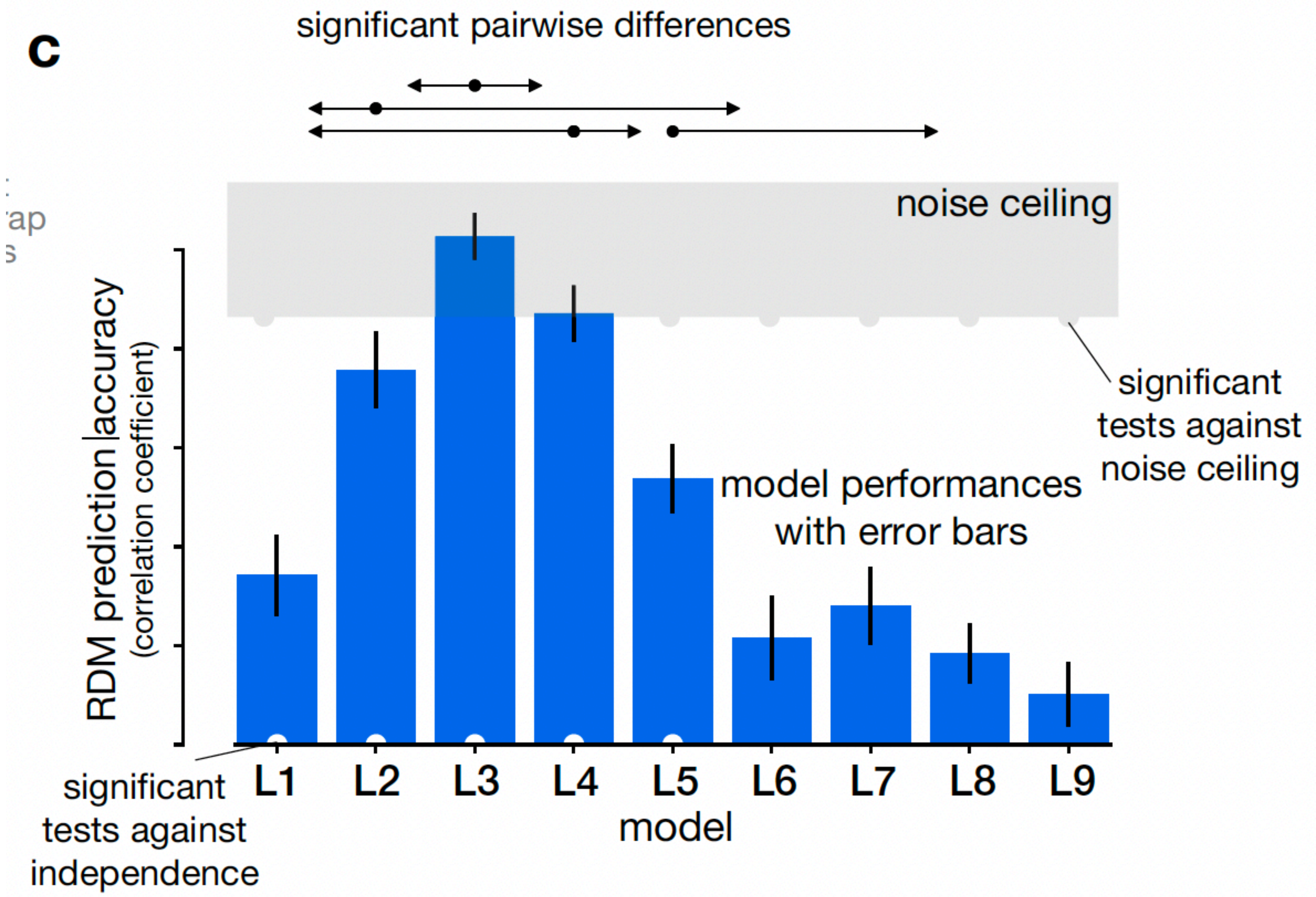
Statistical inference

Test the reliability and generalizability of the models, when they emerge over time.

Cluster-based permutations test



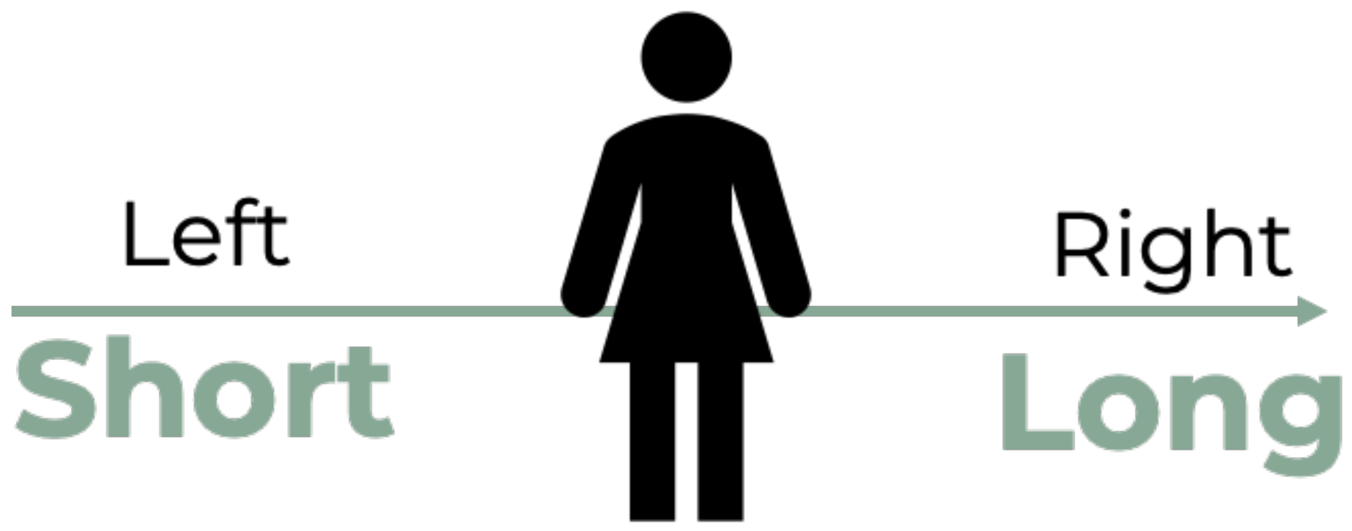
Noise ceiling



Et voilà

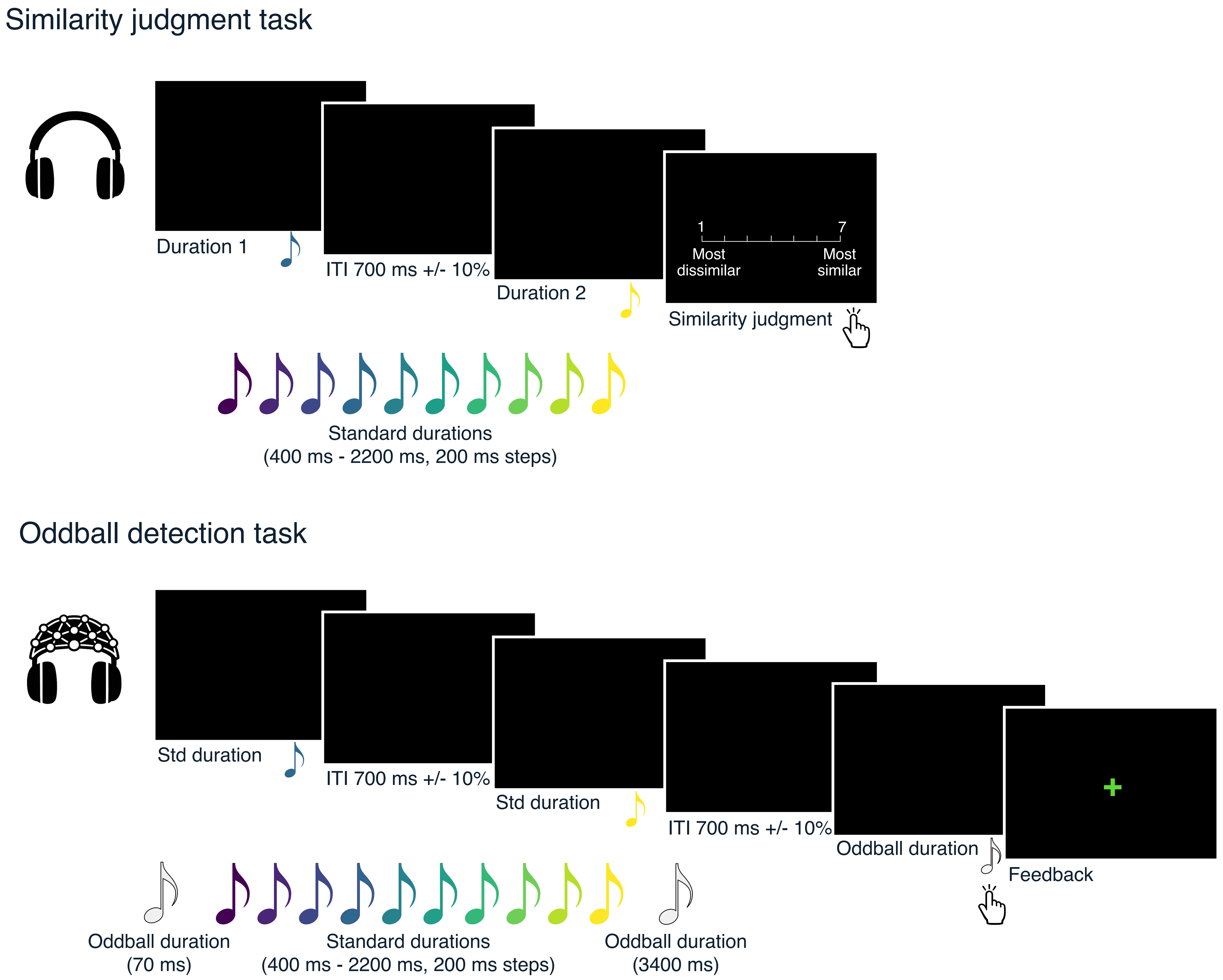
Diedrichsen & Kriegeskorte, 2017 ; Dobs et al., 2019; Schutt et al., 2022 ; See: https://mne.tools/mne-rsa/stable/auto_examples/tutorials/03_statistics_tutorial.html

Investigating the (geometrical) structure of duration representations



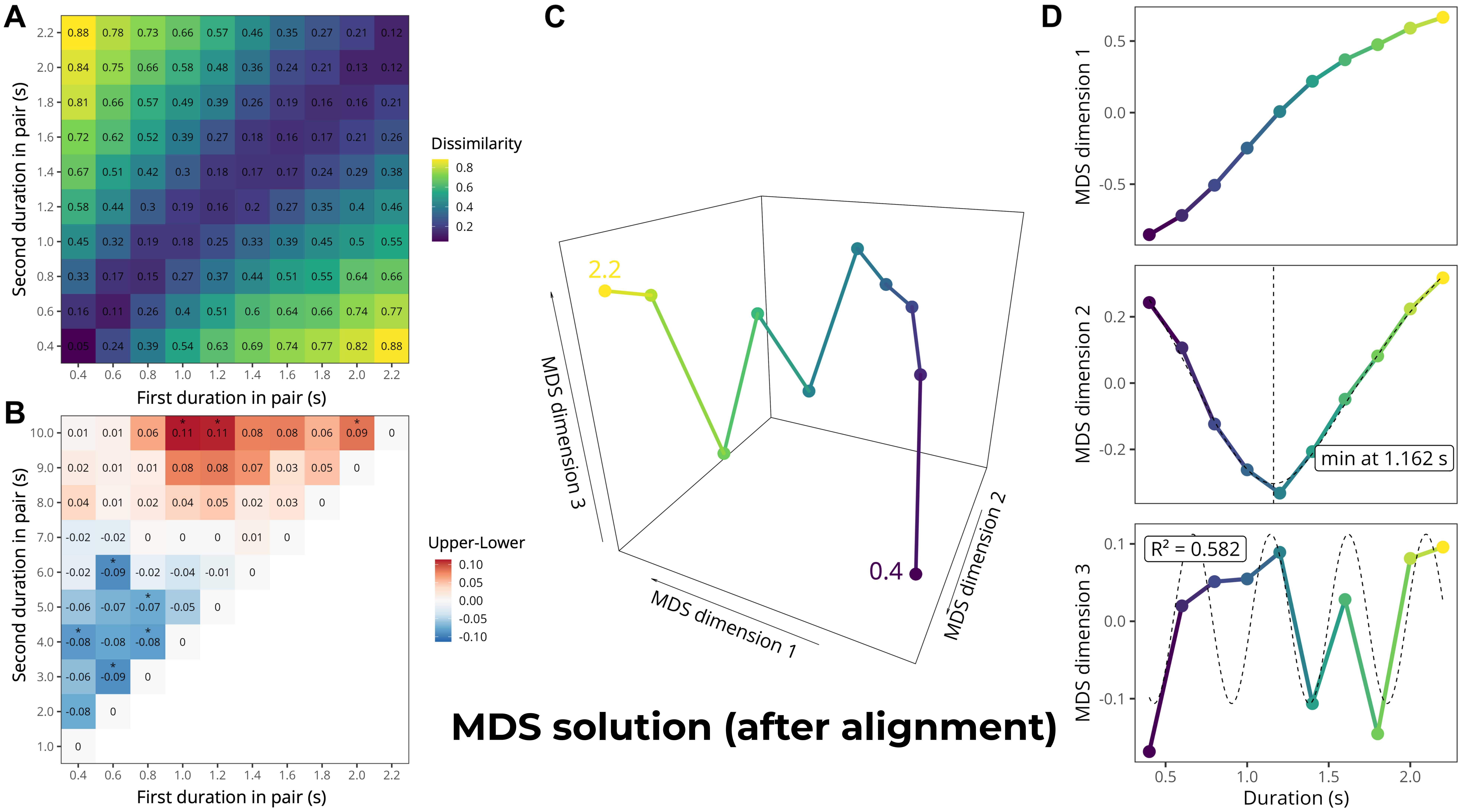
Bonato et al. (2012), Vallesi et al. (2008), Vicario et al. (2008)
Deheane et al. (1993), Deheane (2003); Loetscher et al., (2010)

Is duration represented as a simple mental timeline, or as a richer geometrical structure?



Building empirical RDMs

Behavioral RDMs



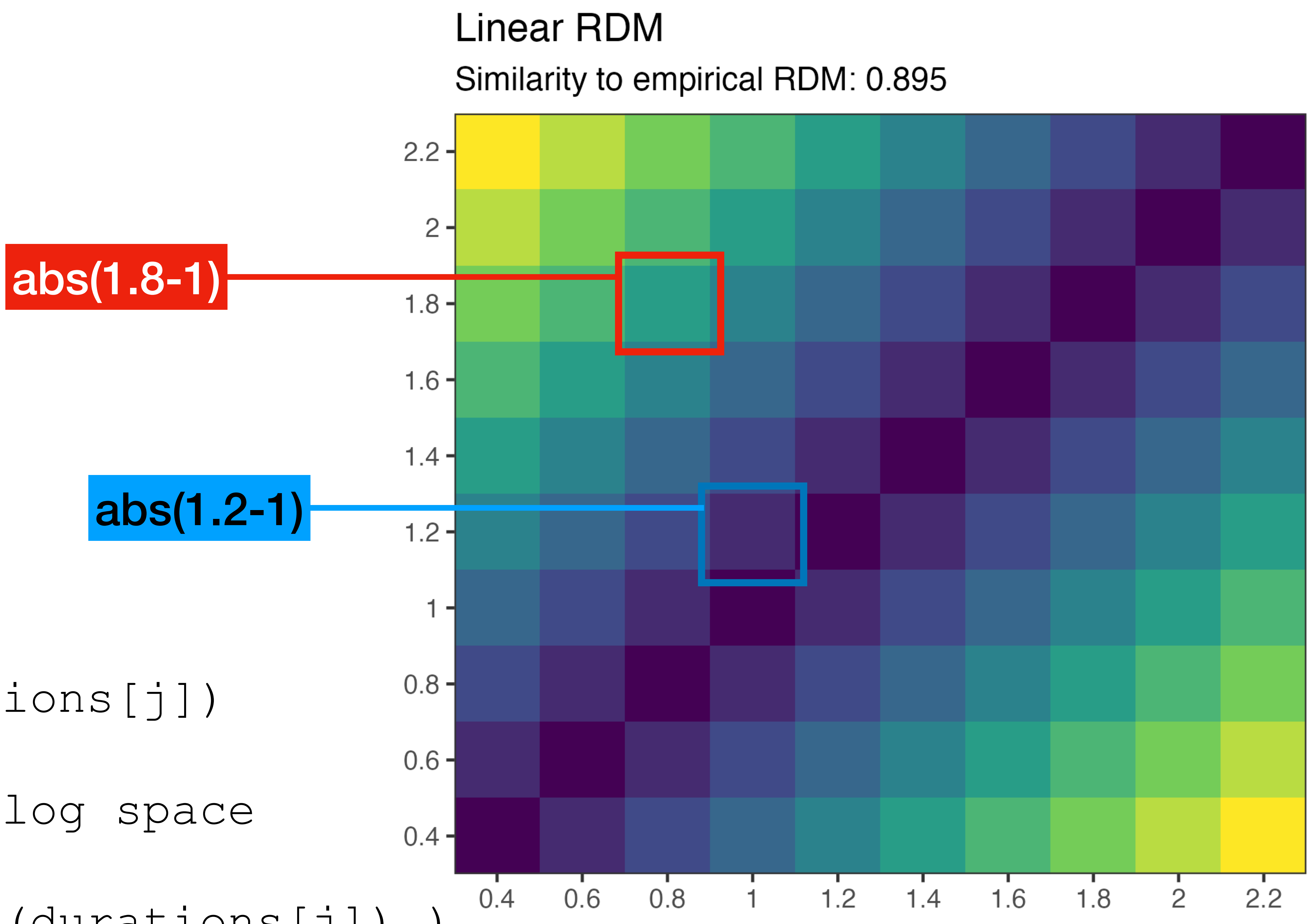
MDS solution (after alignment)

Building theoretical RDMs

Each element of the RDM contains the similarity predicted by each model.

For example, for the linear model, each element of the RDM simply contains the (absolute) difference of durations (in seconds).

```
# linear model: absolute difference

rdm_linear[i, j] <- abs(durations[i] - durations[j])

# logarithmic model: absolute difference in log space

rdm_log[i, j] <- abs(log(durations[i]) - log(durations[j]) )

# power-law model

rdm_power[i, j] <- abs(durations[i]^alpha - durations[j]^alpha)
```



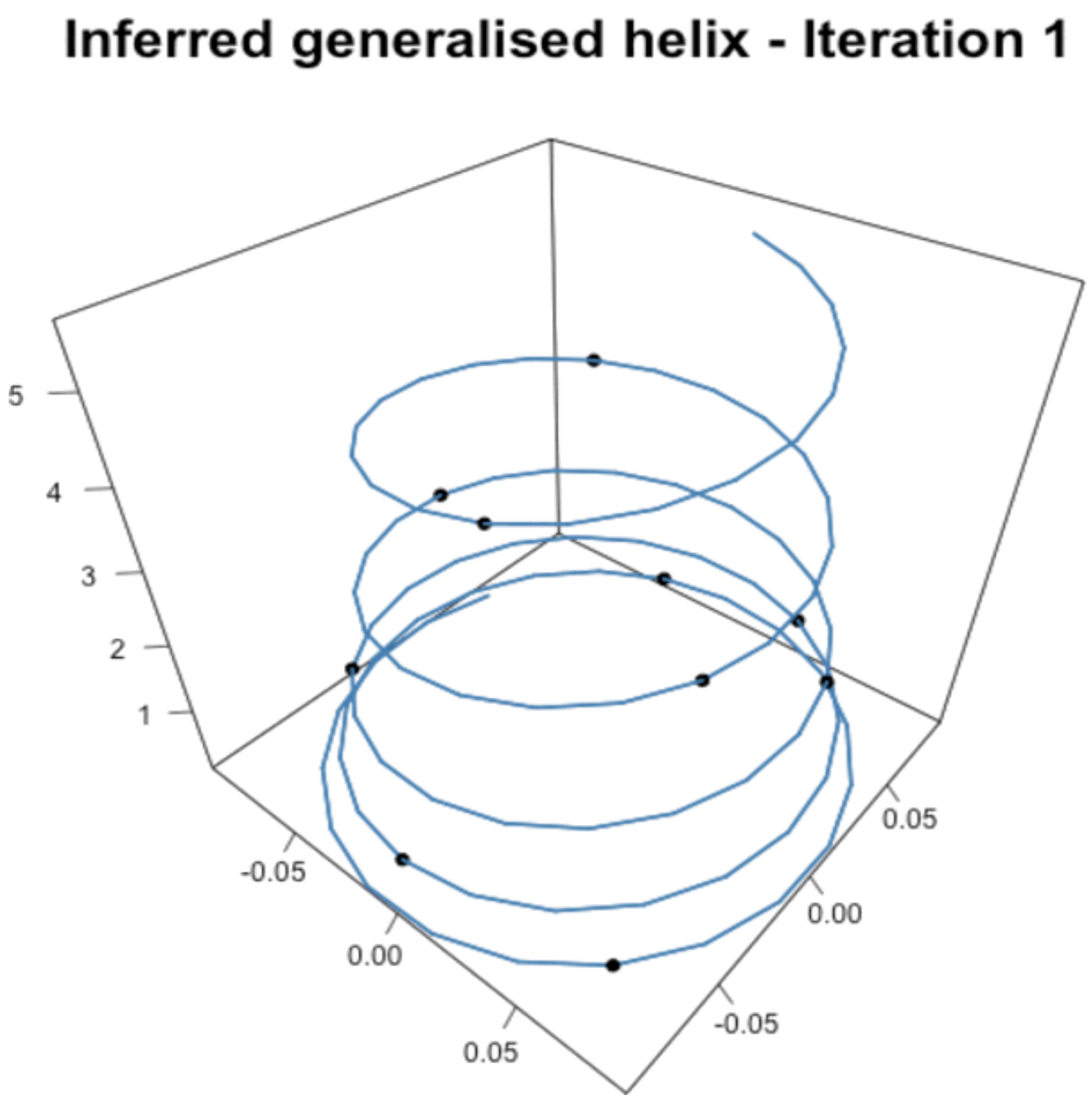
$$S = k \cdot I^{\alpha}$$

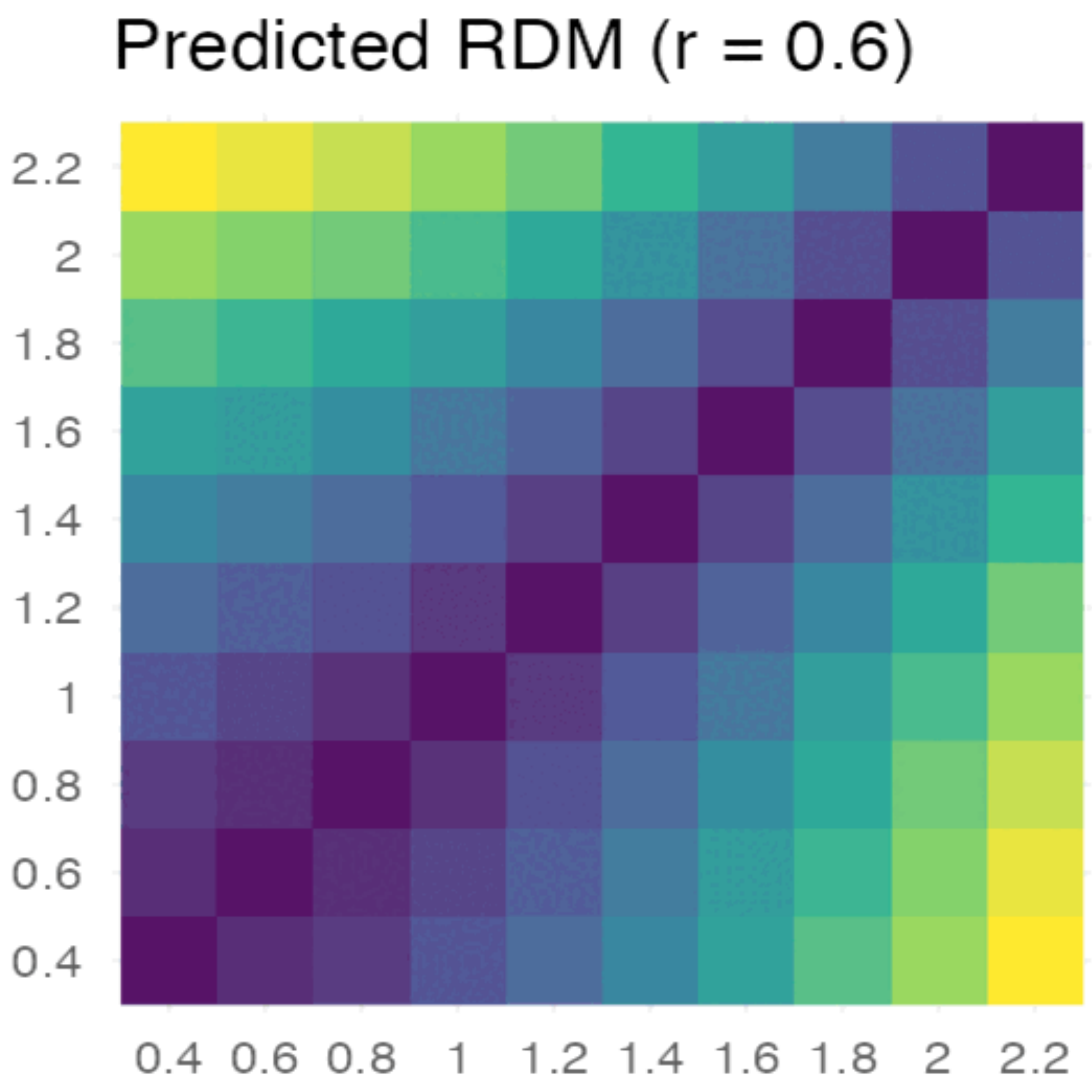
Note that alpha is optimised for the power-law model (cf. next slides).

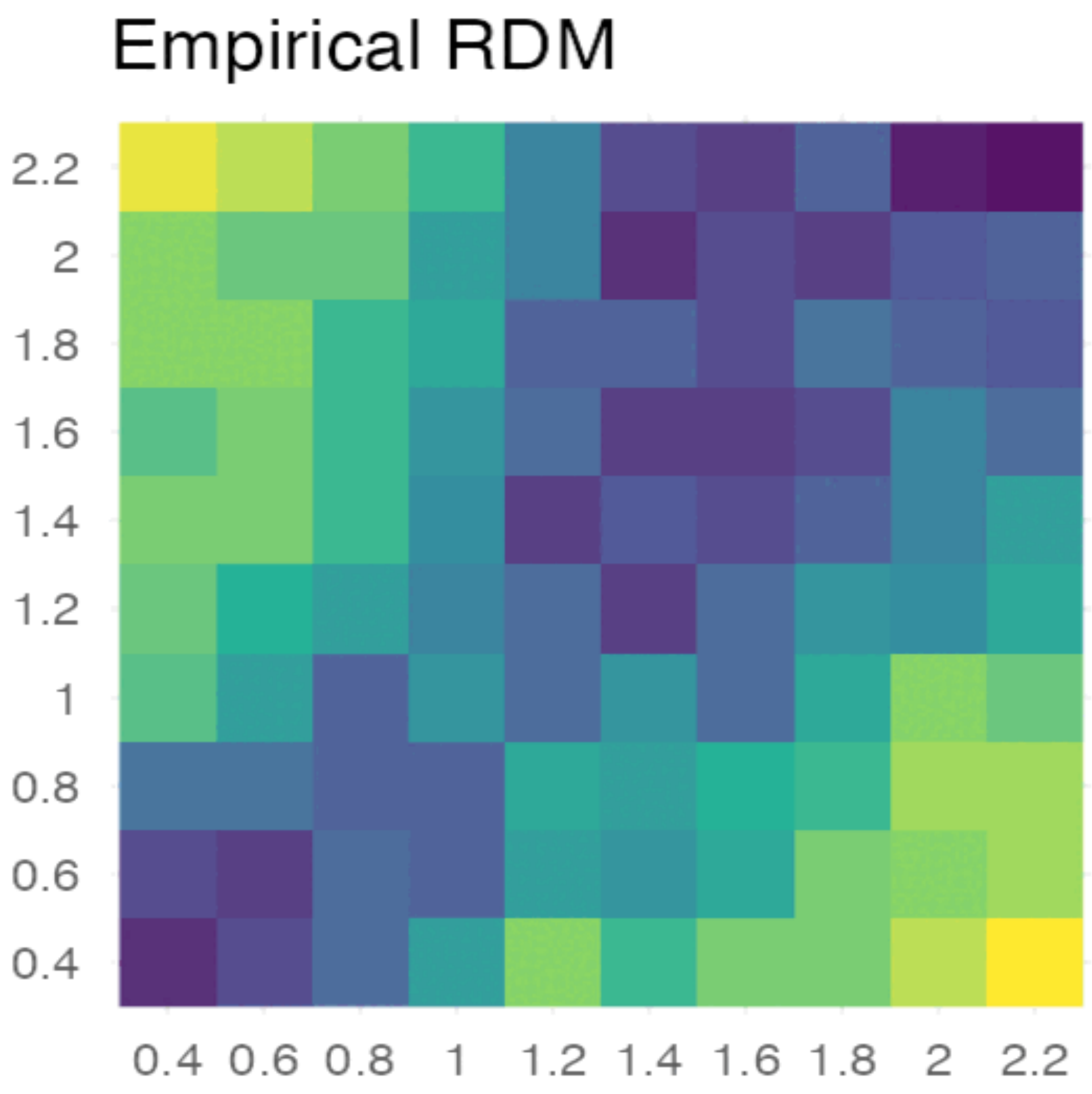
Optimisation steps

Initialise parameters values randomly

Look for the parameter values of the helix (left) that **maximise the correlation** between the theoretical RDM (middle) and the empirical RDM (subjective similarity ratings, right)



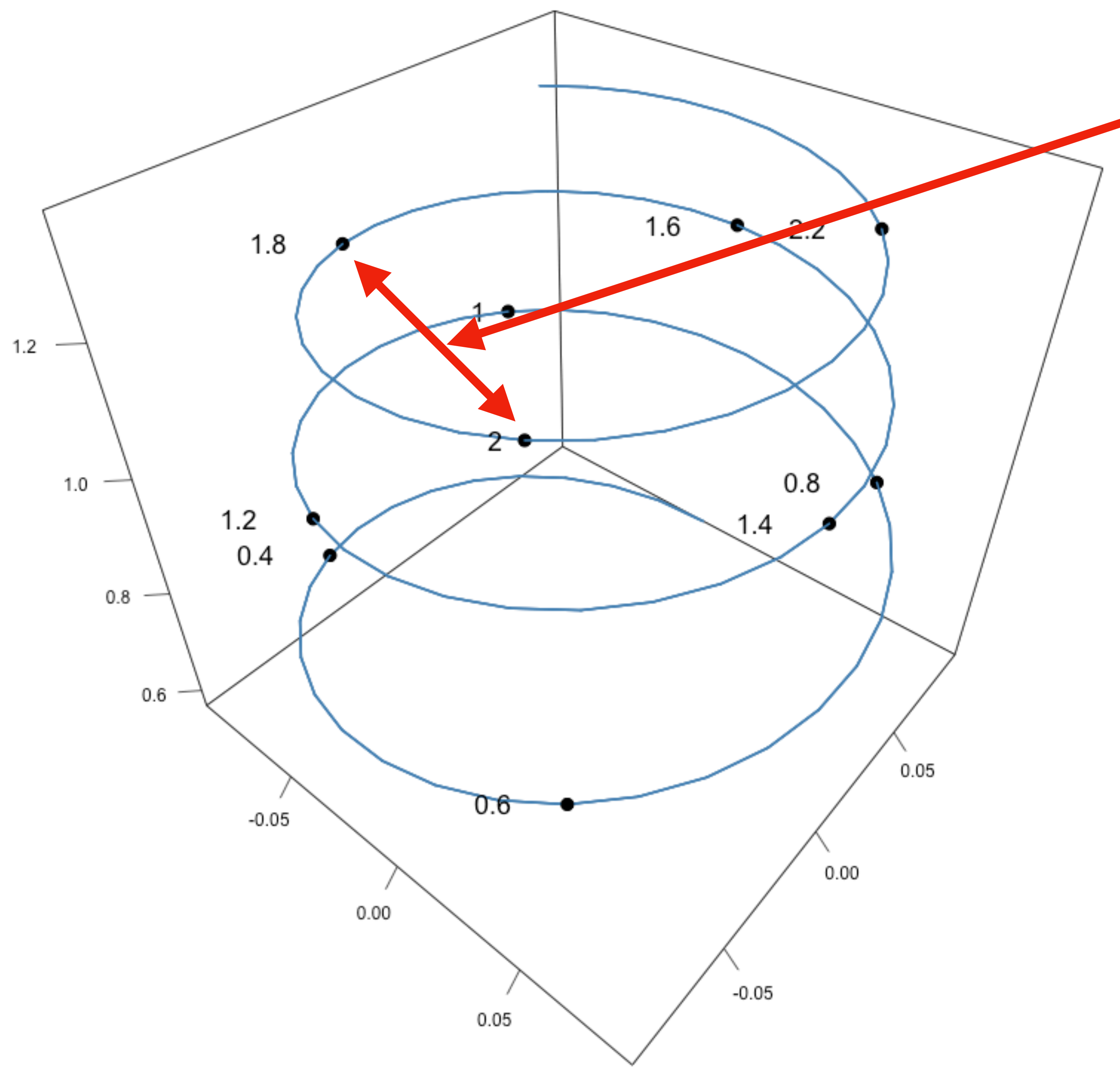


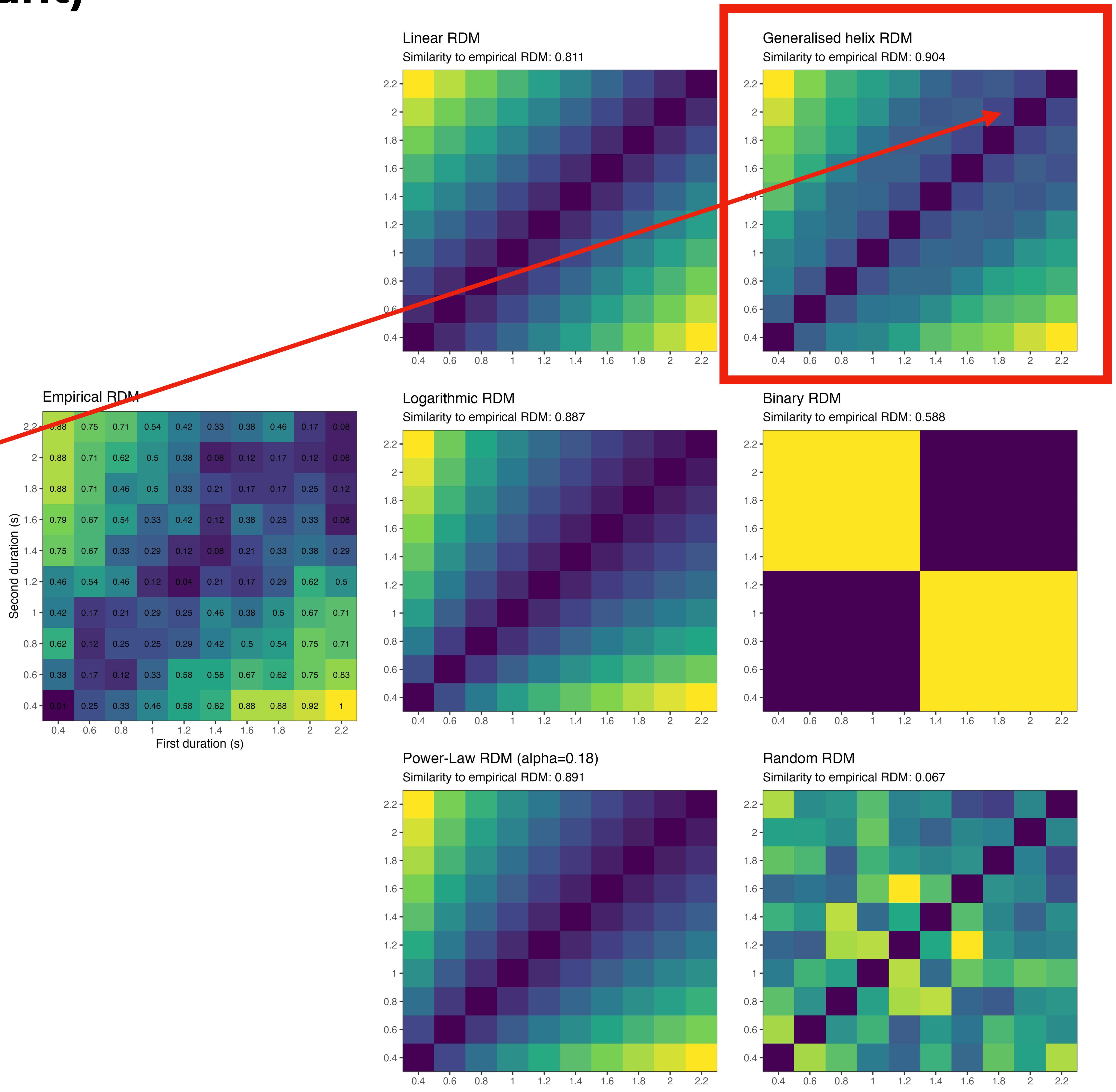


Exemplary behavioral results (1 participant)

- The helix RDM is best fitting the empirical RDM, but with 2 free parameters.

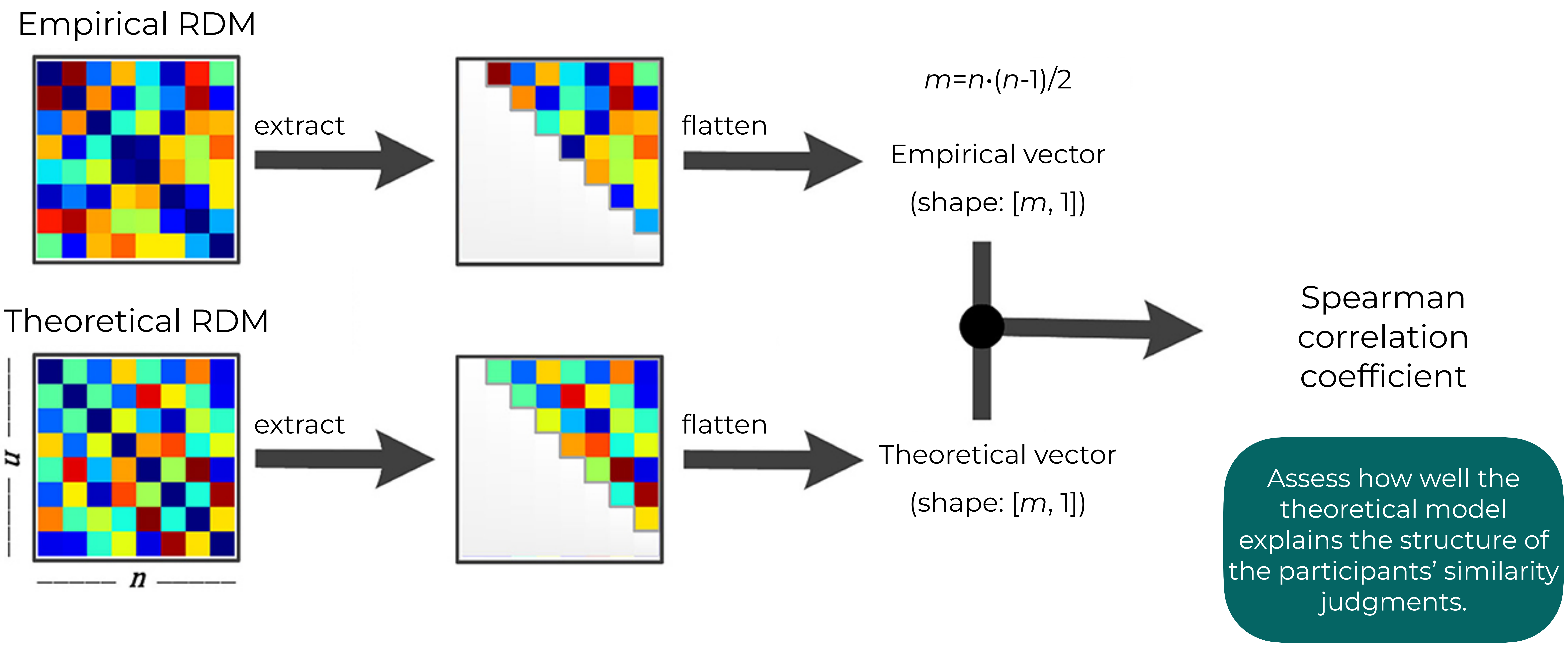
- The power-law model is very good as well, with only 1 free parameter.



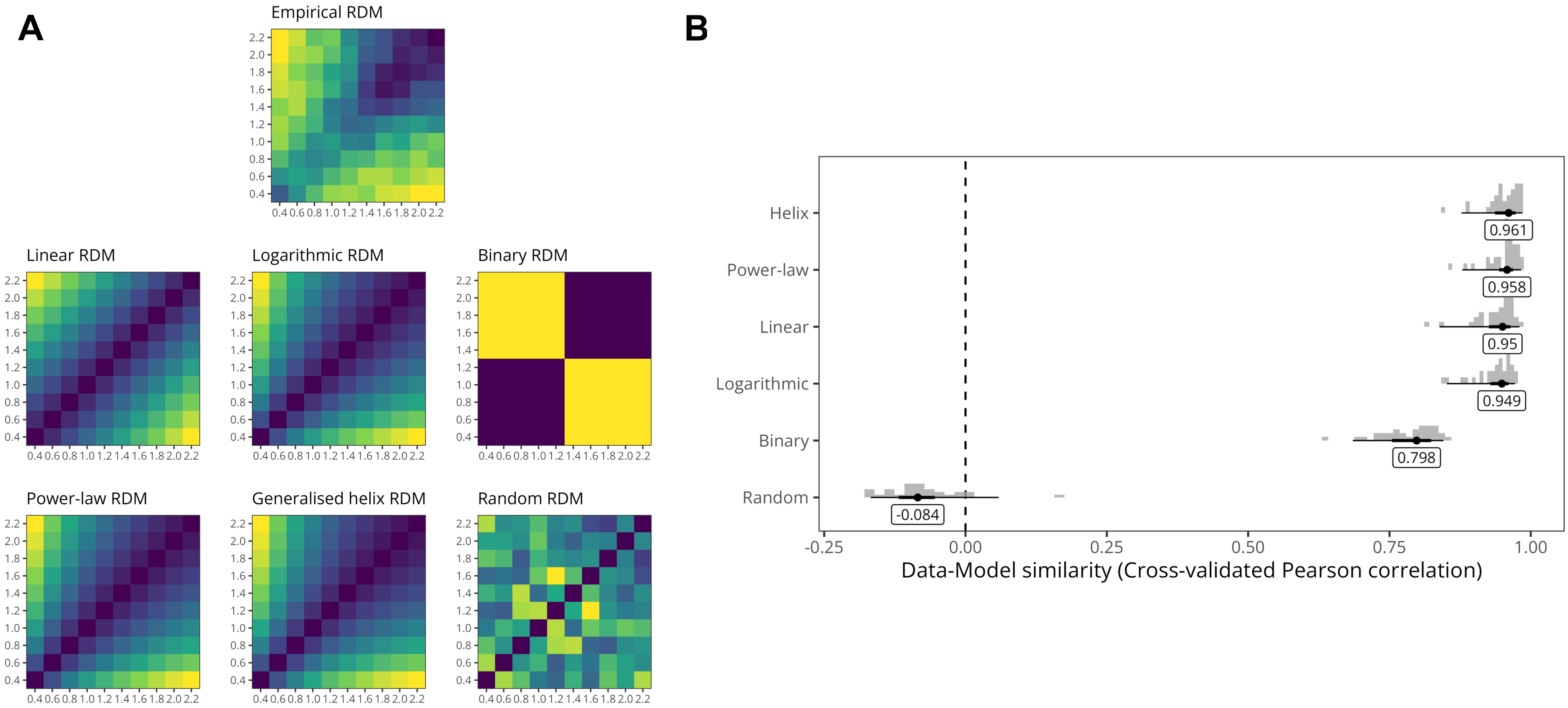


Quantifying similarity between subjective similarity judgments and theoretical models

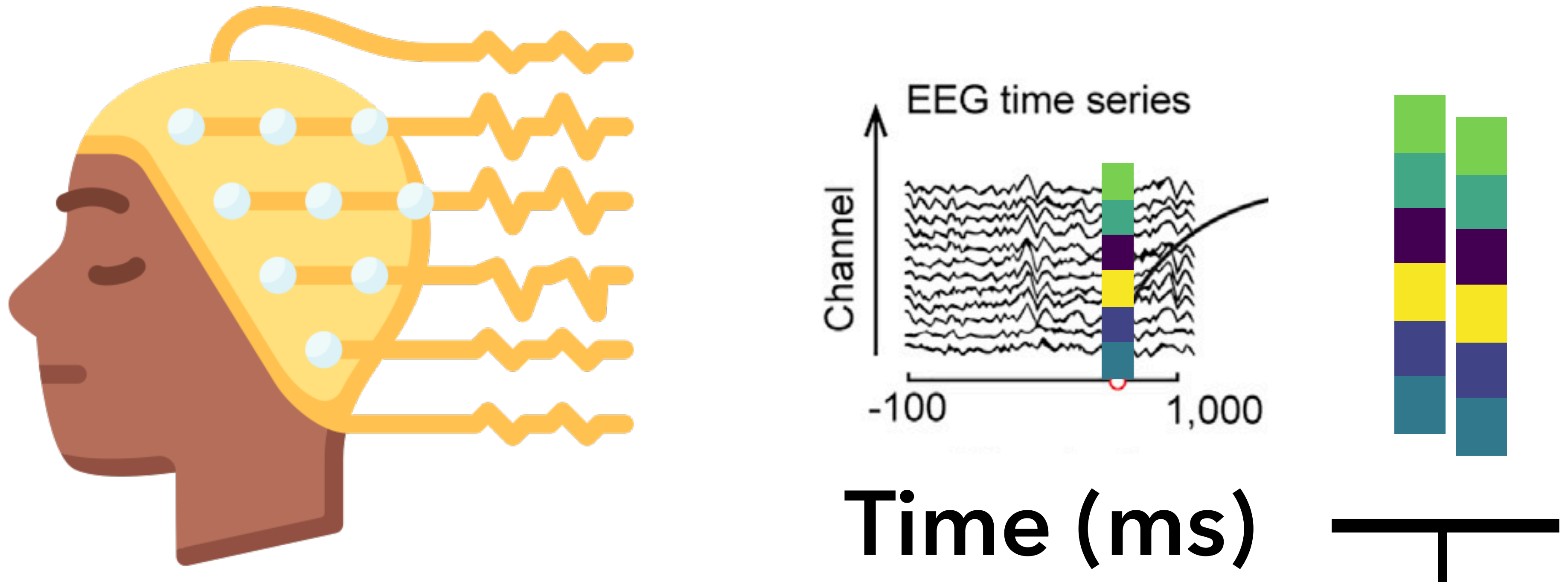
Representational dissimilarity matrices (RDMs) of shape $n \cdot n$ (for n durations)

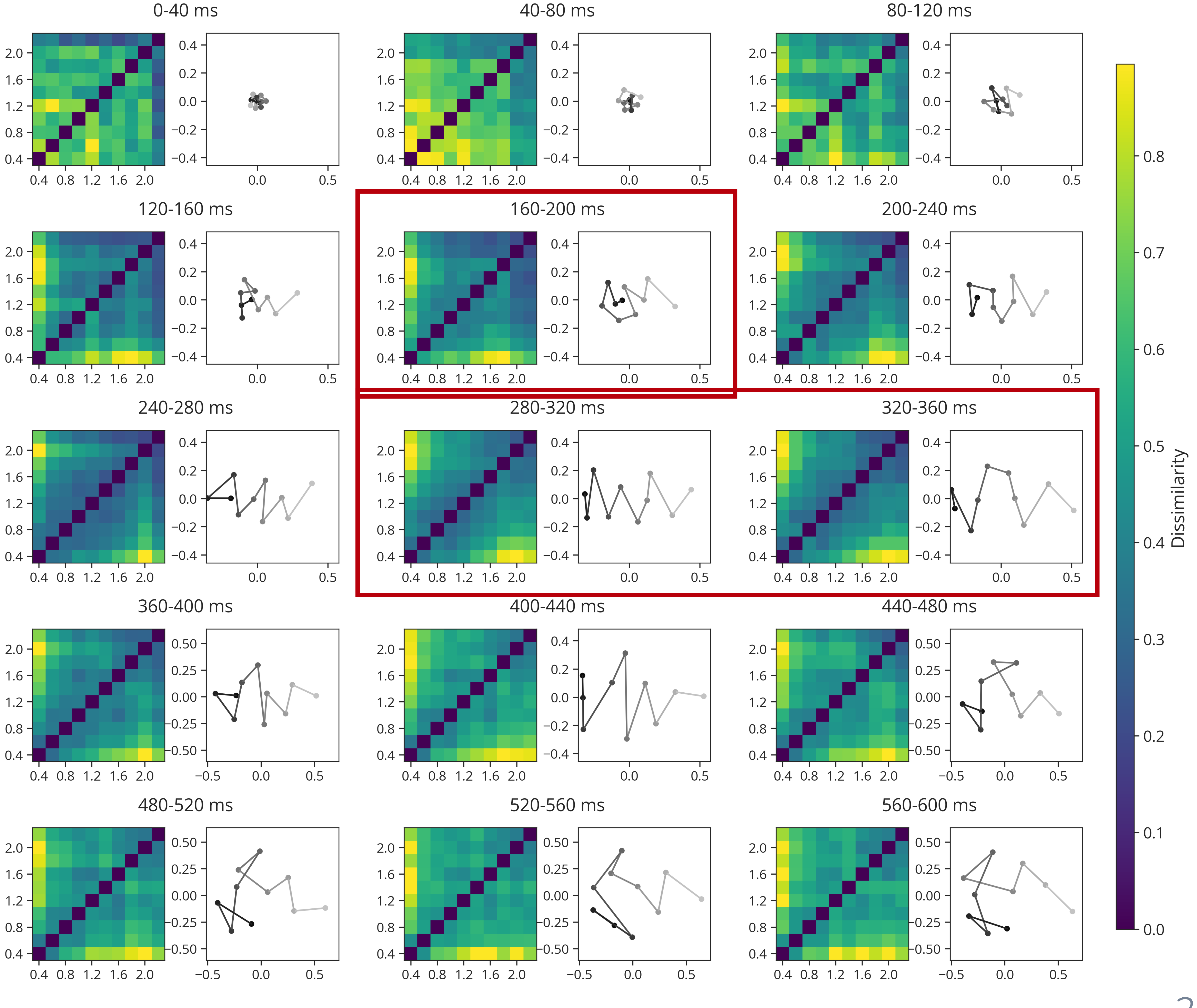


Model comparison

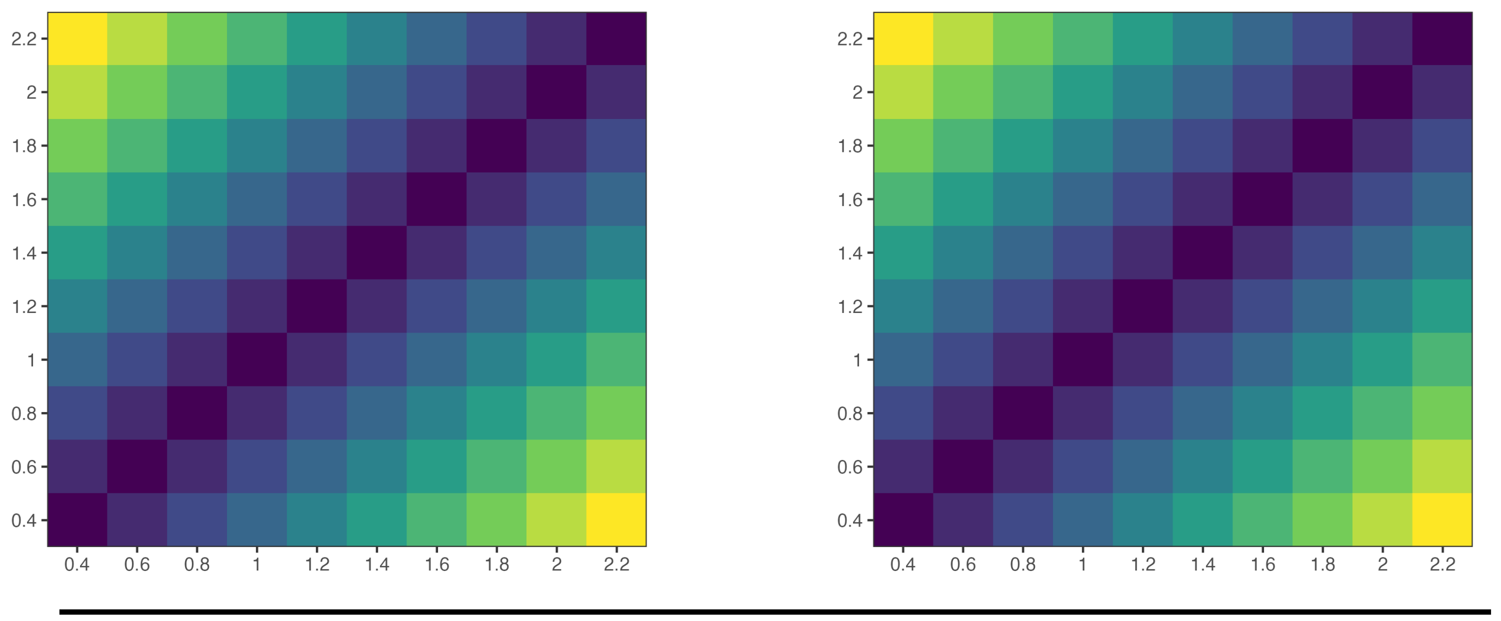


Building empirical (EEG) RDMs





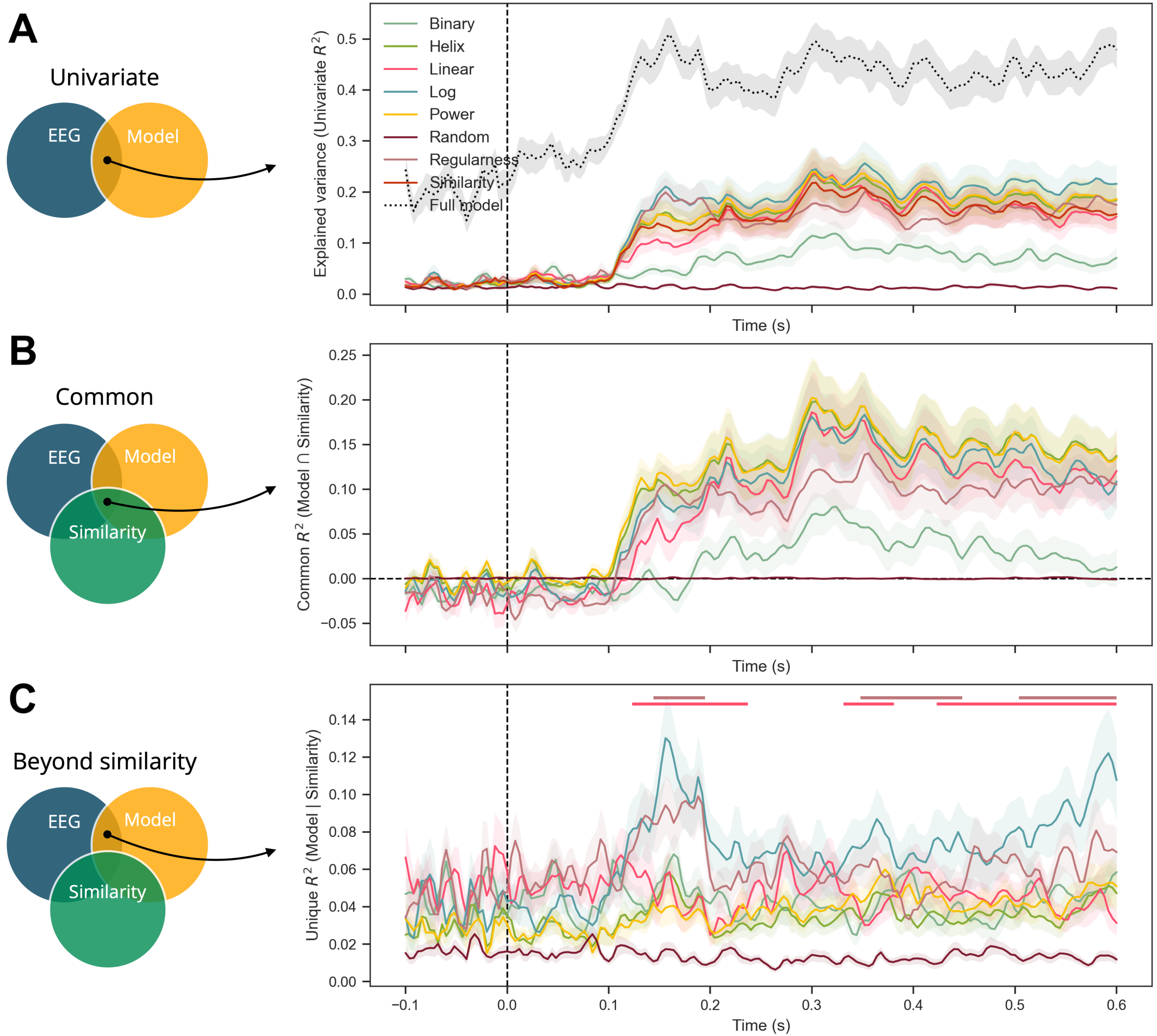
Cross-validated euclidean distance

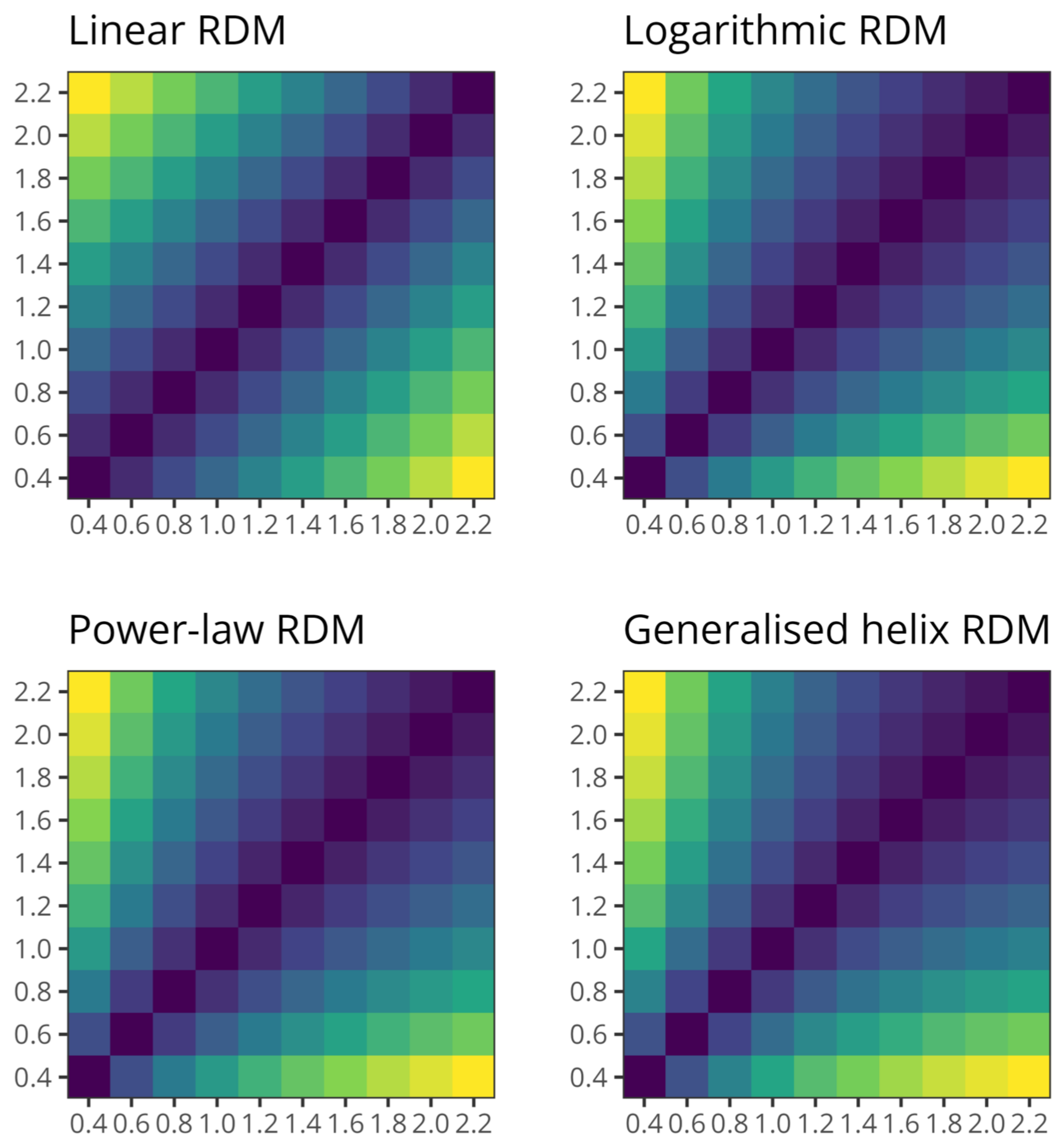


t1 **tn**

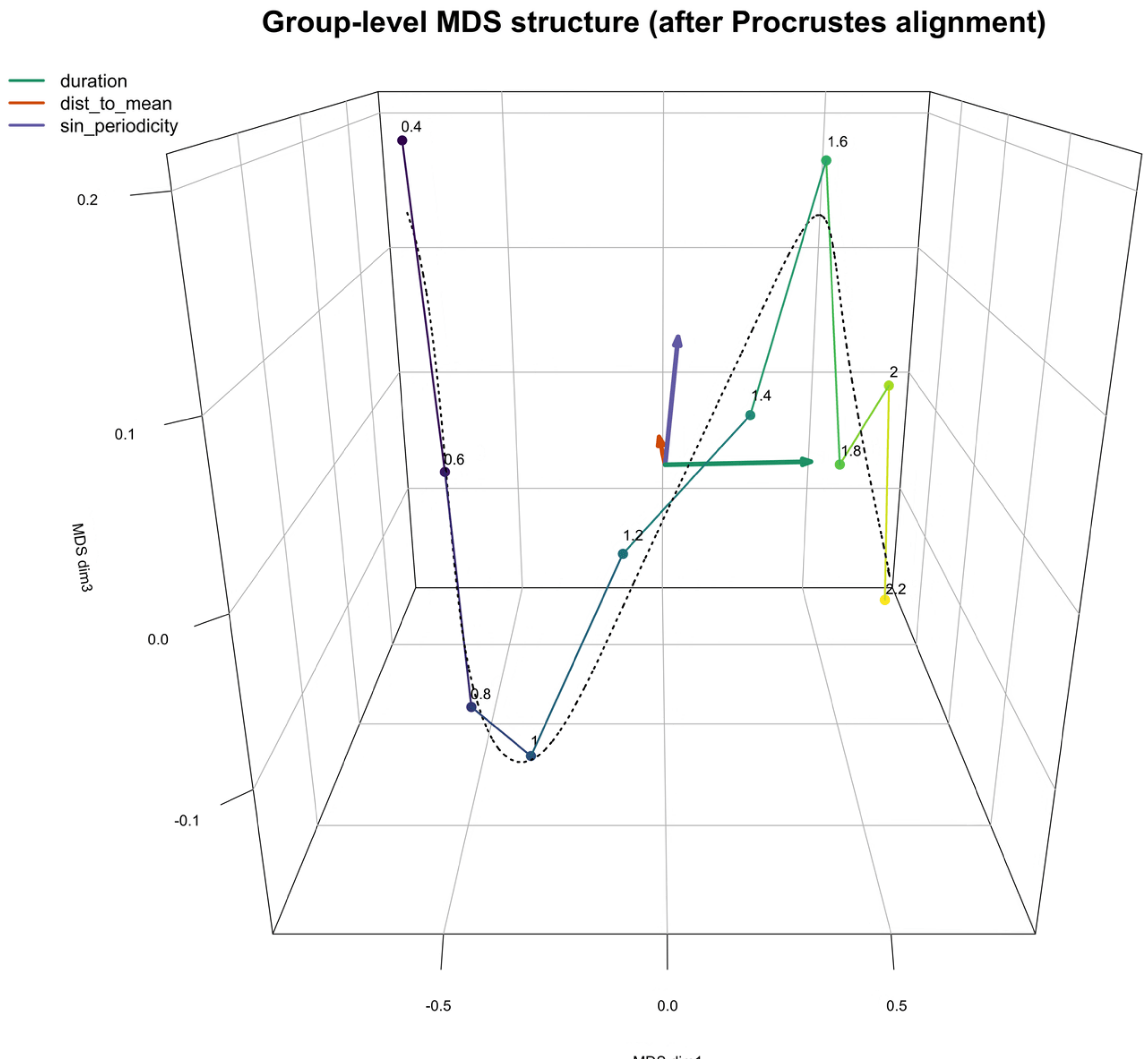
Time-resolved RSA and commonality analyses

How well this post-offset EEG geometry could be explained by our theoretical models?





TAKE HOME MESSAGE



RSA asks **what structure organizes representations**: it tests whether brain, behaviour, or model data share similar patterns of dissimilarity.

… and it does by turning different data types (modalities) and hypothesis into a c**ommon format**: RDMs, which capture how conditions (or items) relate to each other.

RSA is powerful, but… conclusions depend on meaningful conditions, appropriate dissimilarity measures, and proper statistical inference.

REFERENCES

Bobadilla-Suarez, S., Ahlheim, C., Mehrotra, A., Panos, A., & Love, B. C. (2018). Measures of neural similarity. Neuroscience. https://doi.org/10.1101/439893
Cichy, R. M., & Oliva, A. (2020). A M/EEG-fMRI Fusion Primer : Resolving Human Brain Responses in Space and Time. Neuron, 107(5), 772-781. https://doi.org/10.1016/j.neuron.2020.07.001
Diedrichsen, J., & Kriegeskorte, N. (2017). Representational models : A common framework for understanding encoding, pattern-component, and representational-similarity analysis. PLOS Computational Biology, 13(4), e1005508. https://doi.org/10.1371/journal.pcbi.1005508
Dobs, K., Isik, L., Pantazis, D., & Kanwisher, N. (2019). How face perception unfolds over time. Nature Communications, 10(1), 1258. https://doi.org/10.1038/s41467-019-09239-1
Edelman, S. (1998). Representation is representation of similarities. Behavioral and Brain Sciences, 21(4), 449-467. https://doi.org/10.1017/S0140525X98001253
Edelman, S., Grill-Spector, K., Kushnir, T., & Malach, R. (1998). Toward direct visualization of the internal shape representation space by fMRI. Psychobiology, 26(4), 309-321. https://doi.org/10.3758/BF03330618
Grasso, C. L., Nalborczyk, L., & Van Wassenhove, V. (2026, mars). Uncovering the representational geometry of durations. https://doi.org/10.64898/2026.03.29.715088
Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., & Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. eLife, 7, e32816. https://doi.org/10.7554/eLife.32816
Kawakita, G., Zeleznikow-Johnston, A., Takeda, K., Tsuchiya, N., & Oizumi, M. (2023). Is my « red » your « red »? : Unsupervised alignment of qualia structures via optimal transport [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/h3pqm
Kawakita, G., Zeleznikow-Johnston, A., Tsuchiya, N., & Oizumi, M. (2023). Comparing Color Similarity Structures between Humans and LLMs via Unsupervised Alignment (arXiv:2308.04381). arXiv. http://arxiv.org/abs/2308.04381
Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. Frontiers in Systems Neuroscience. https://doi.org/10.3389/neuro.06.004.2008
Marjieh, R., Griffiths, T. L., & Jacoby, N. (2023). Pitch is Not (Always) a Helix : Probing the Structure of Musical Pitch Across Tasks and Experience. https://doi.org/10.1101/2023.06.13.544763
Proklova, D., Kaiser, D., & Peelen, M. V. (2016). Disentangling Representations of Object Shape and Object Category in Human Visual Cortex : The Animate–Inanimate Distinction. Journal of Cognitive Neuroscience, 28(5), 680-692. https://doi.org/10.1162/jocn_a_00924
Shepard, R. N. (1980). Multidimensional Scaling, Tree-Fitting, and Clustering. Science, 210(4468), 390-398. https://doi.org/10.1126/science.210.4468.390
Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. Psychological Review, 89(4), 305-333. https://doi.org/10.1037/0033-295X.89.4.305
Shepard, R. N. (1987). Toward a Universal Law of Generalization for Psychological Science. Science, 237(4820), 1317-1323. https://doi.org/10.1126/science.3629243
Shepard, R. N., & Chipman, S. (1970). Second-order isomorphism of internal representations : Shapes of states. Cognitive Psychology, 1(1), 1-17. https://doi.org/10.1016/0010-0285(70)90002-2
Shepard, R. N., Kilpatric, D. W., & Cunningham, J. P. (1975). The internal representation of numbers. Cognitive Psychology, 7(1), 82-138. https://doi.org/10.1016/0010-0285(75)90006-7
Viganò, S., Bayramova, R., Doeller, C. F., & Bottini, R. (2024). Spontaneous eye movements reflect the representational geometries of conceptual spaces. Proceedings of the National Academy of Sciences, 121(17), e2403858121. https://doi.org/10.1073/pnas.2403858121
Xie, S., Hoehl, S., Moeskops, M., Kayhan, E., Kliesch, C., Turtleton, B., Köster, M., & Cichy, R. M. (2022). Visual category representations in the infant brain. Current Biology, 32(24), 5422-5432.e6. https://doi.org/10.1016/j.cub.2022.11.016

RESOURCES

https://mne.tools/mne-rsa/stable/auto_examples/tutorials/01_sensor_level_tutorial.html

https://rsatoolbox.readthedocs.io/en/latest/distances.html

https://nikokriegeskorte.org/2019/01/09/whats-the-best-measure-of-representational-dissimilarity/

https://mne.tools/mne-rsa/stable/auto_examples/tutorials/03_statistics_tutorial.html

https://mne.tools/mne-rsa/stable/auto_examples/core/plot_cross_validation.html

https://lnalborczyk.github.io/blog/2025-01-07-commonality/

https://www.flaticon.com/search?word=adults
(for logos representations)