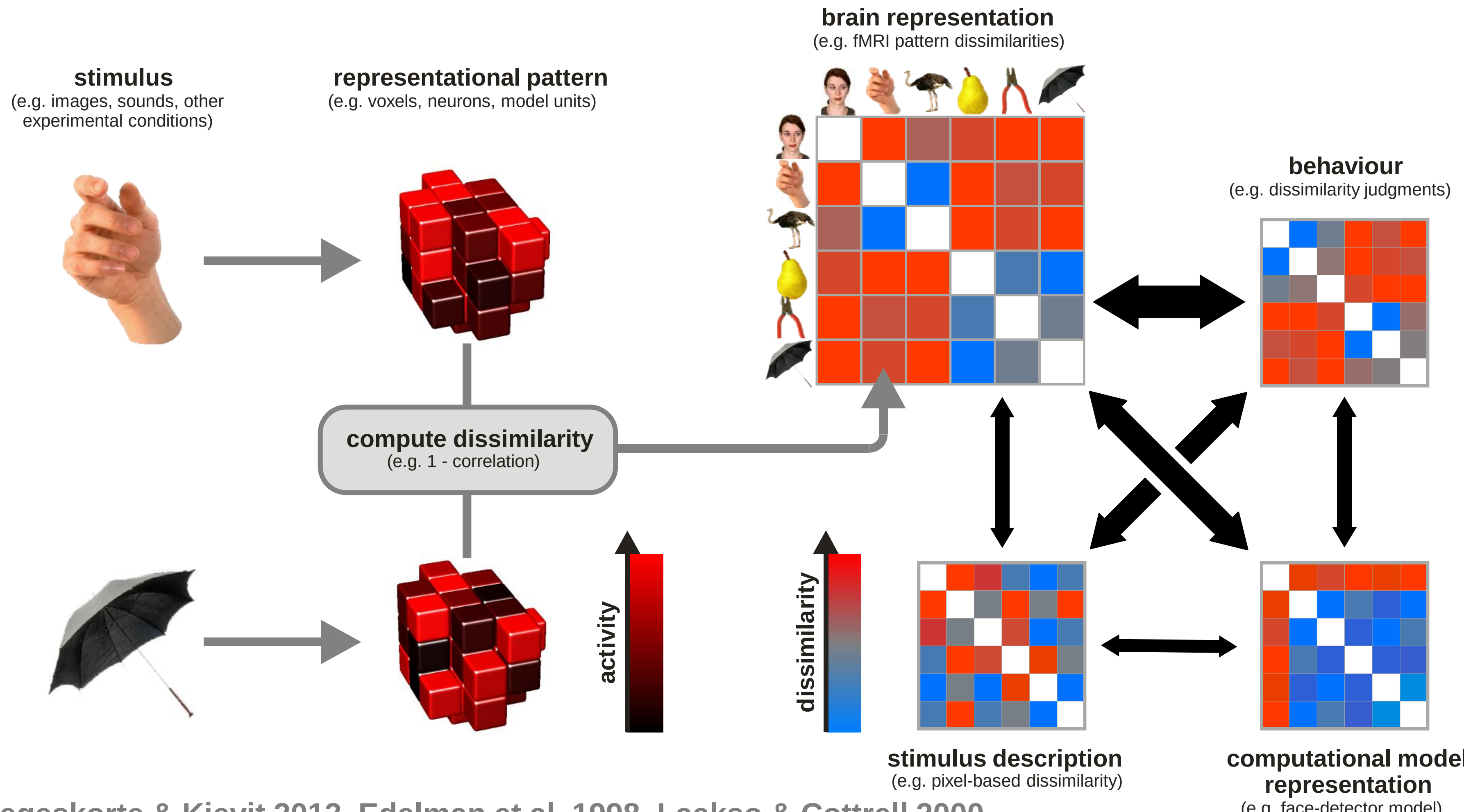
Representational similarity analysis

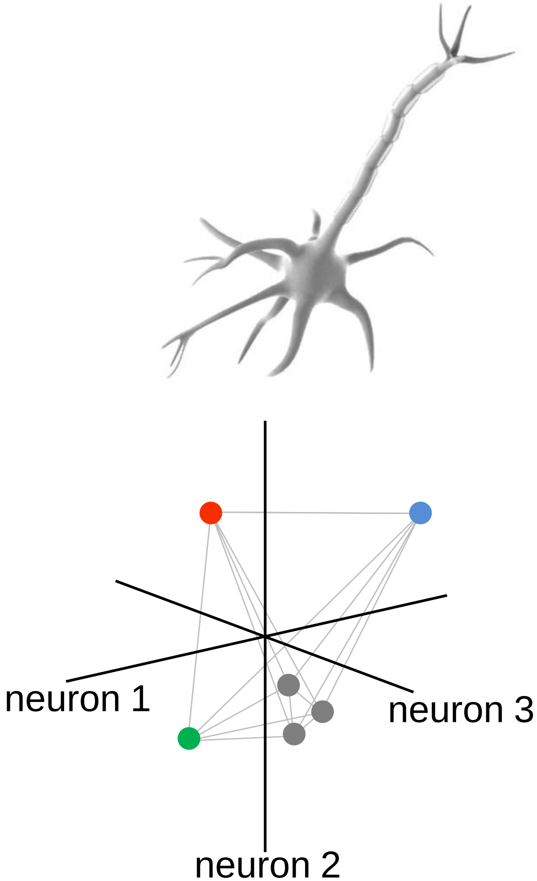
Nikolaus Kriegeskorte
MRC Cognition and Brain Sciences Unit
Cambridge, UK

Representational similarity analysis



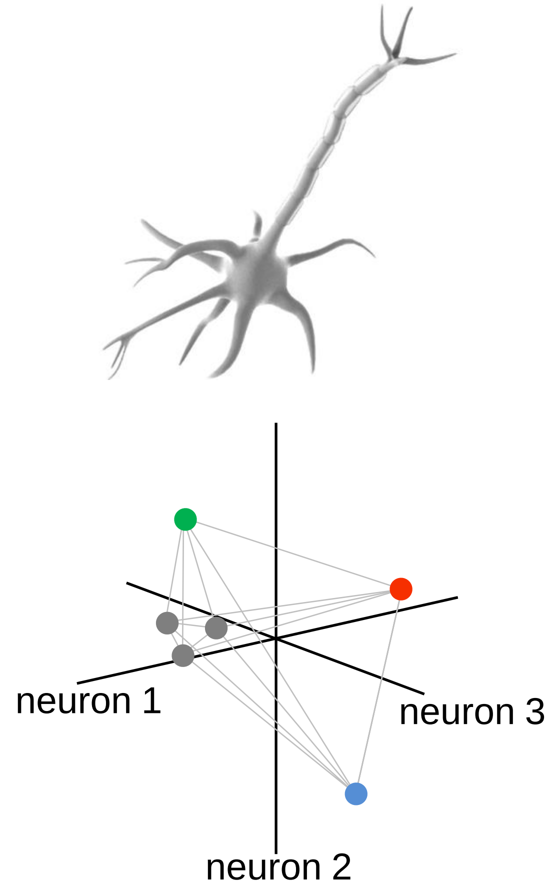
Kriegeskorte & Kievit 2013, Edelman et al. 1998, Laakso & Cottrell 2000,
Op de Beeck et al. 2001, Haxby et al. 2001, Aguirre 2007, Kriegeskorte et al. 2008, Diedrichsen et al. 2011

Why investigate representational geometries?



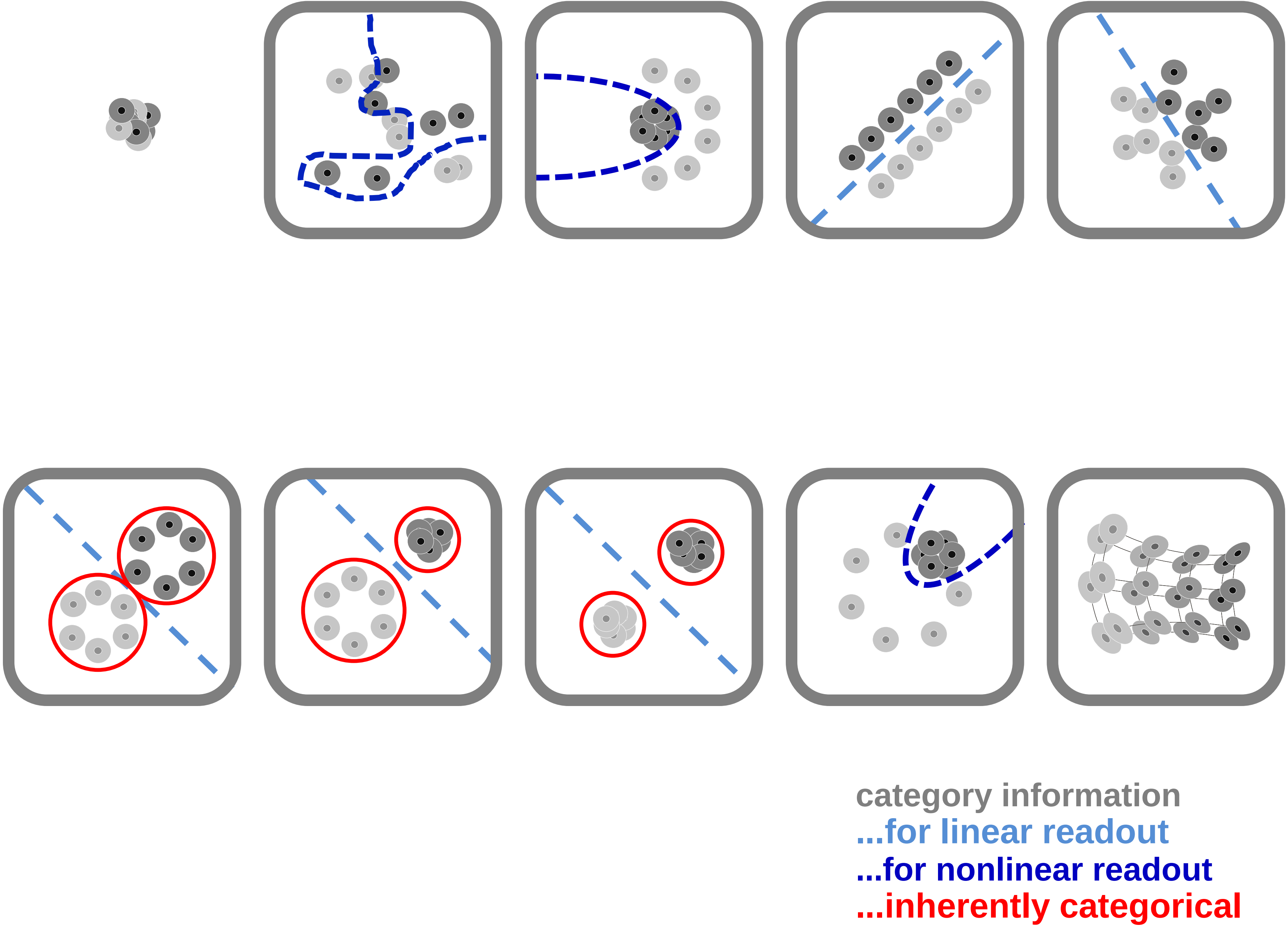
downstream neurons
can read out the same information from these codes

same geometry
→ same information
→ same format

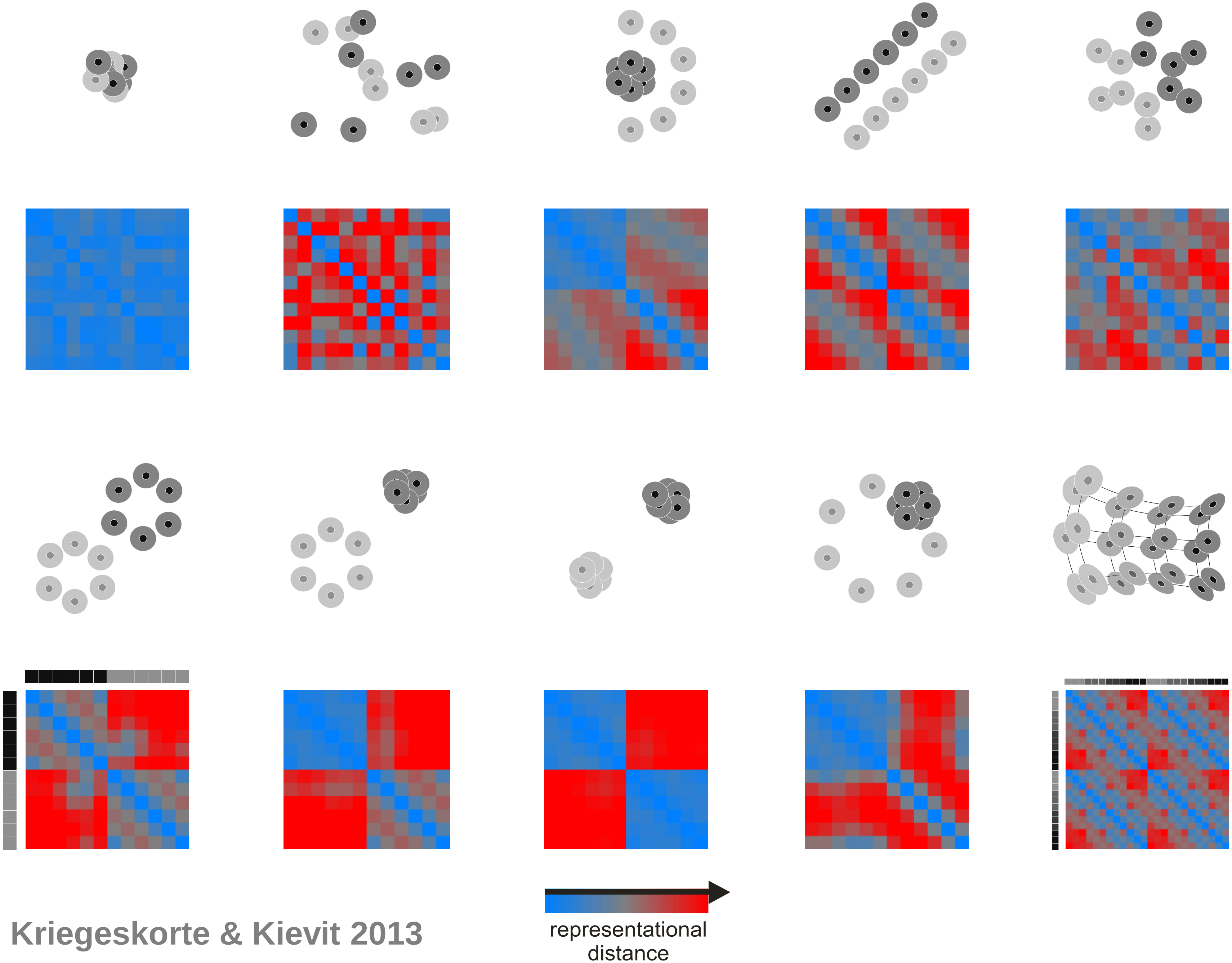


Representational geometry

The geometry of the points in a high-dimensional
response pattern space, which are thought to represent particular stimuli.

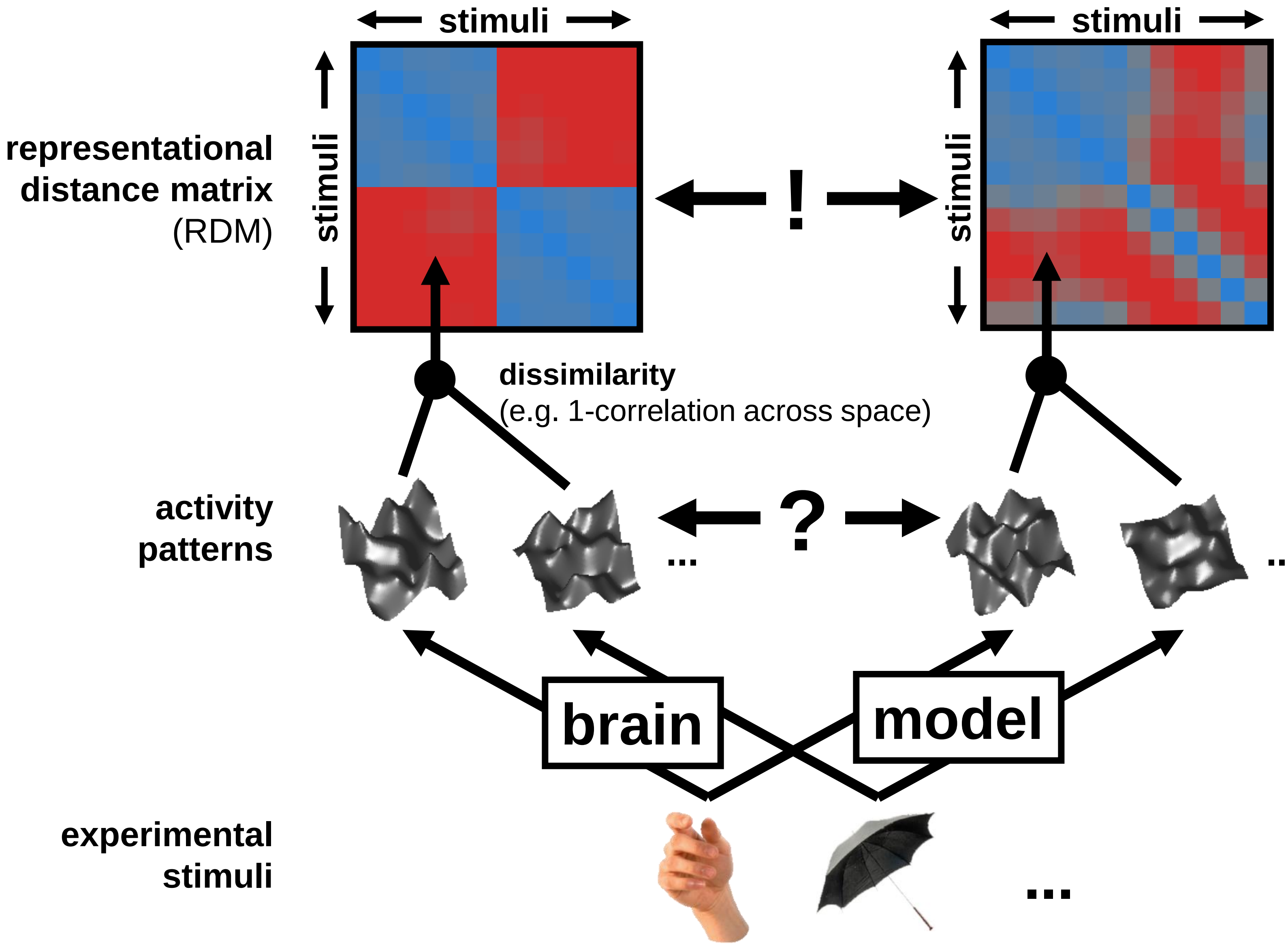


Kriegeskorte & Kievit 2013

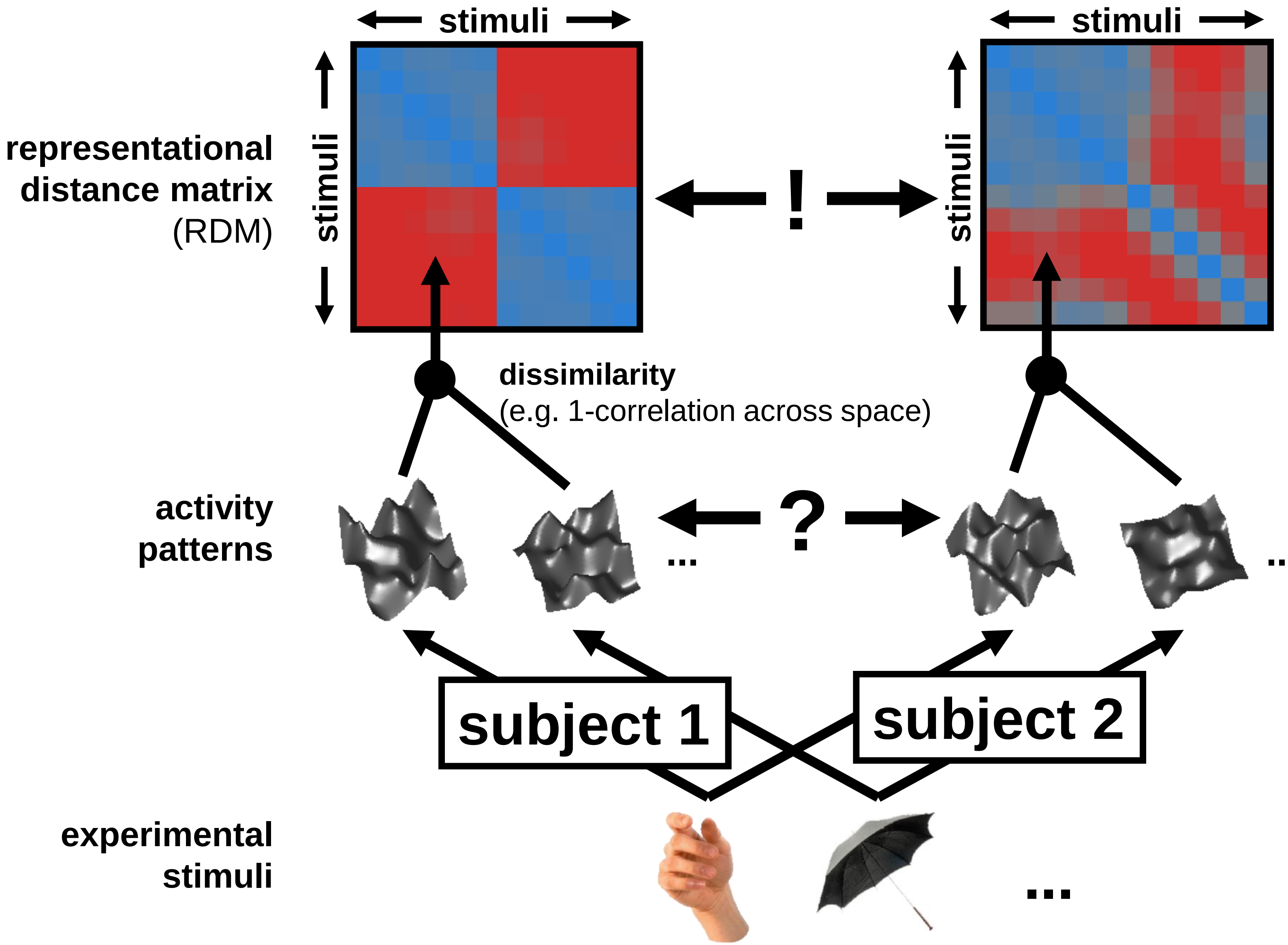


Kriegeskorte & Kievit 2013

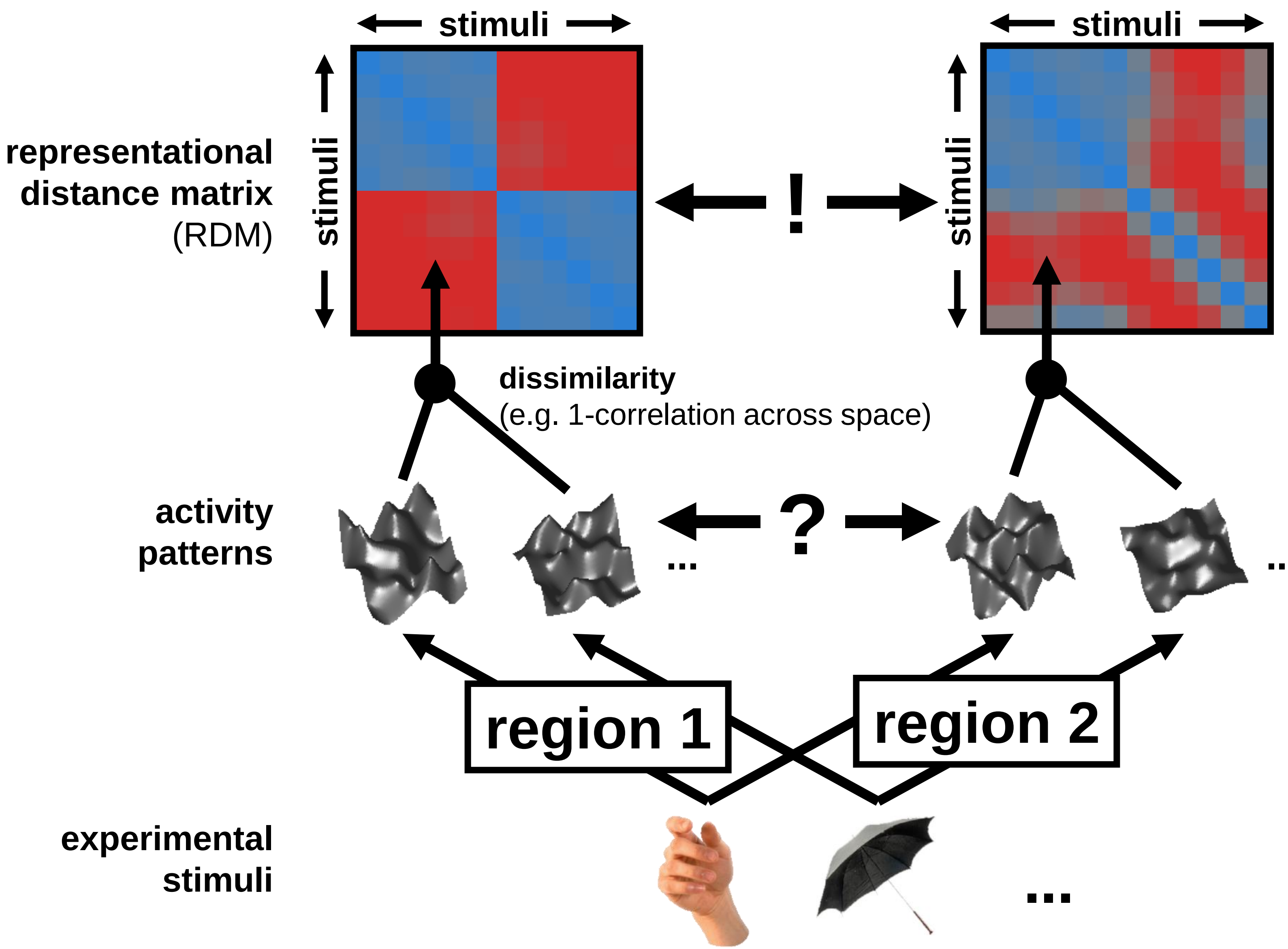
The representational similarity trick



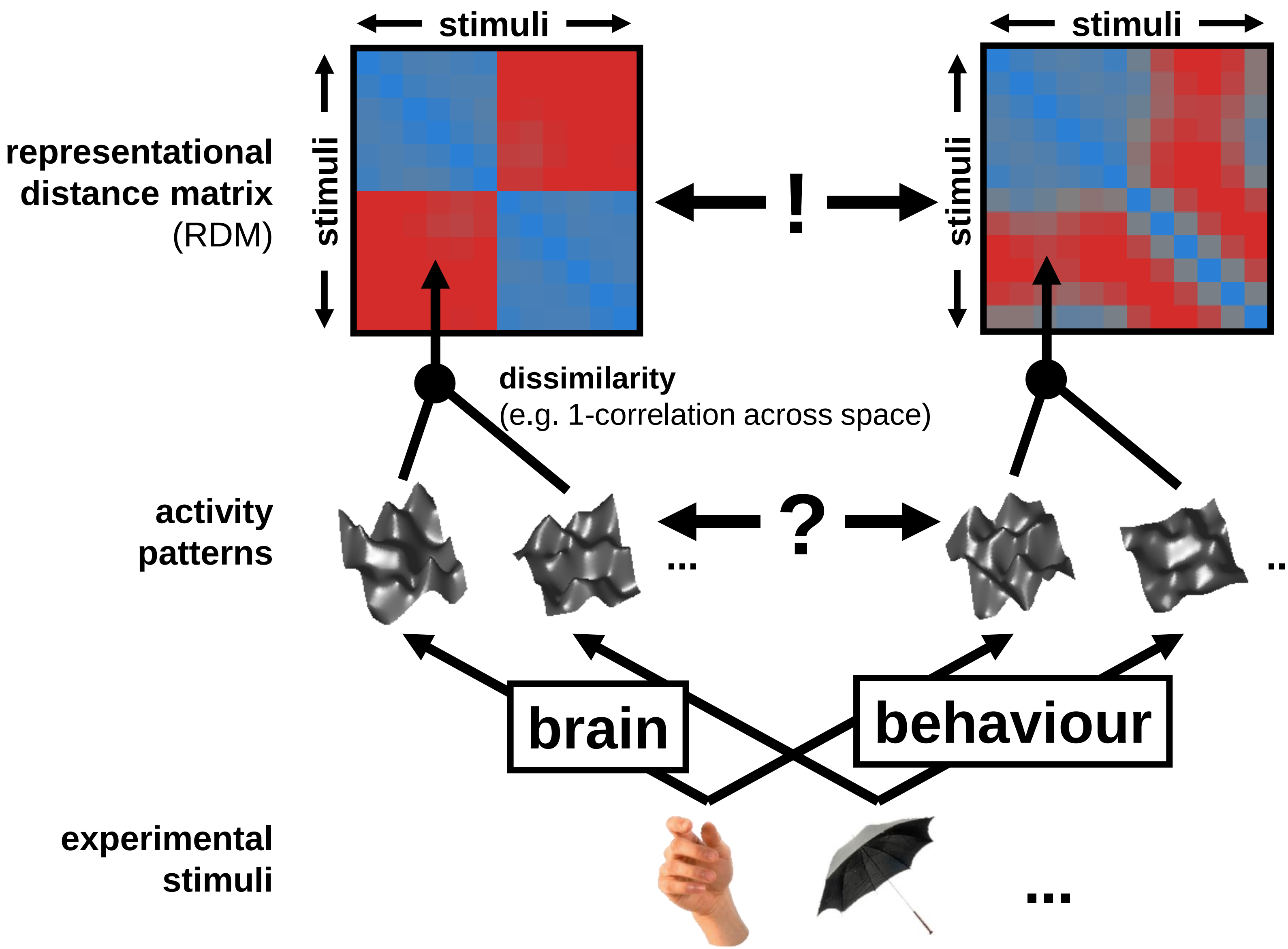
The representational similarity trick



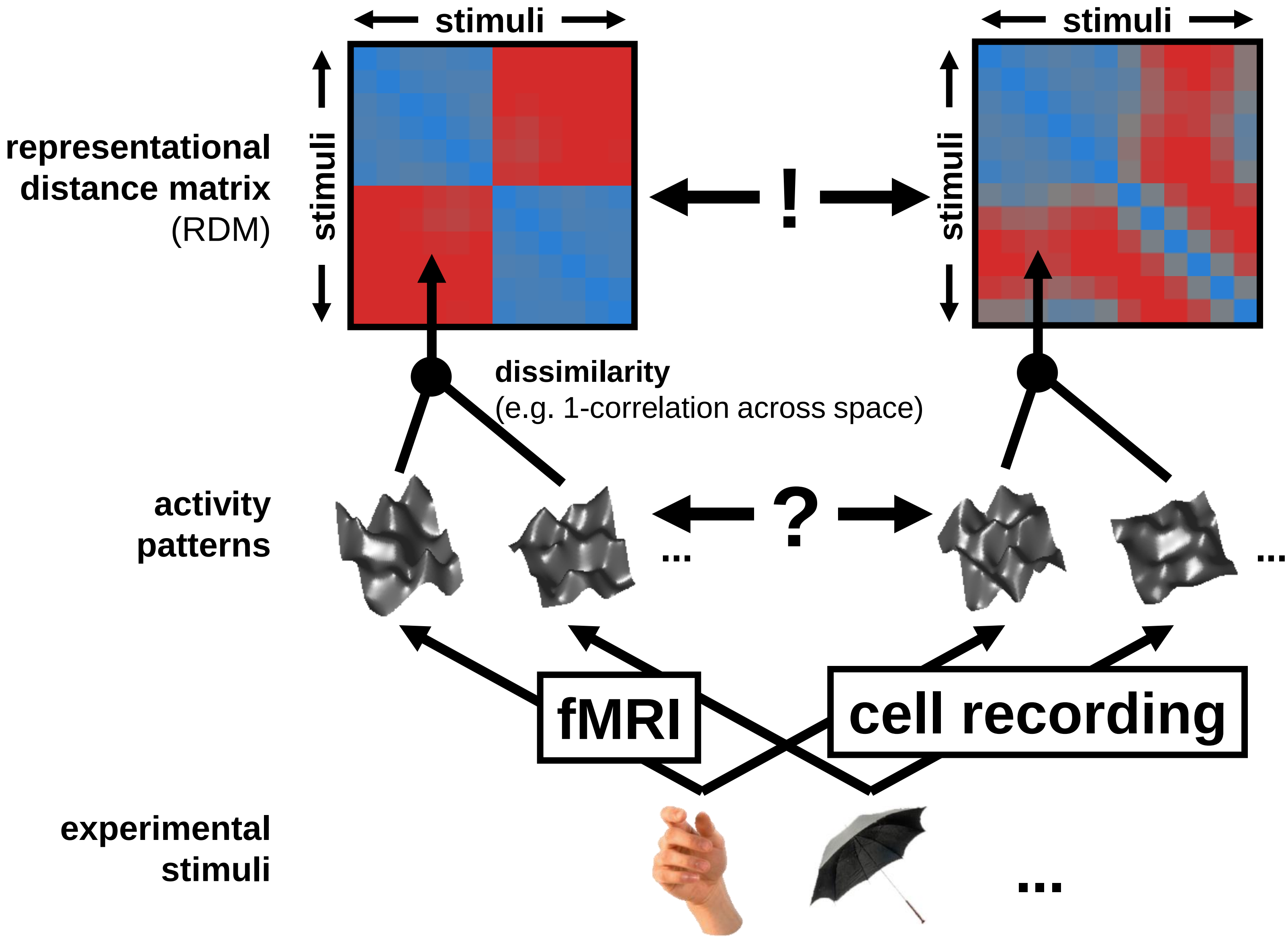
The representational similarity trick



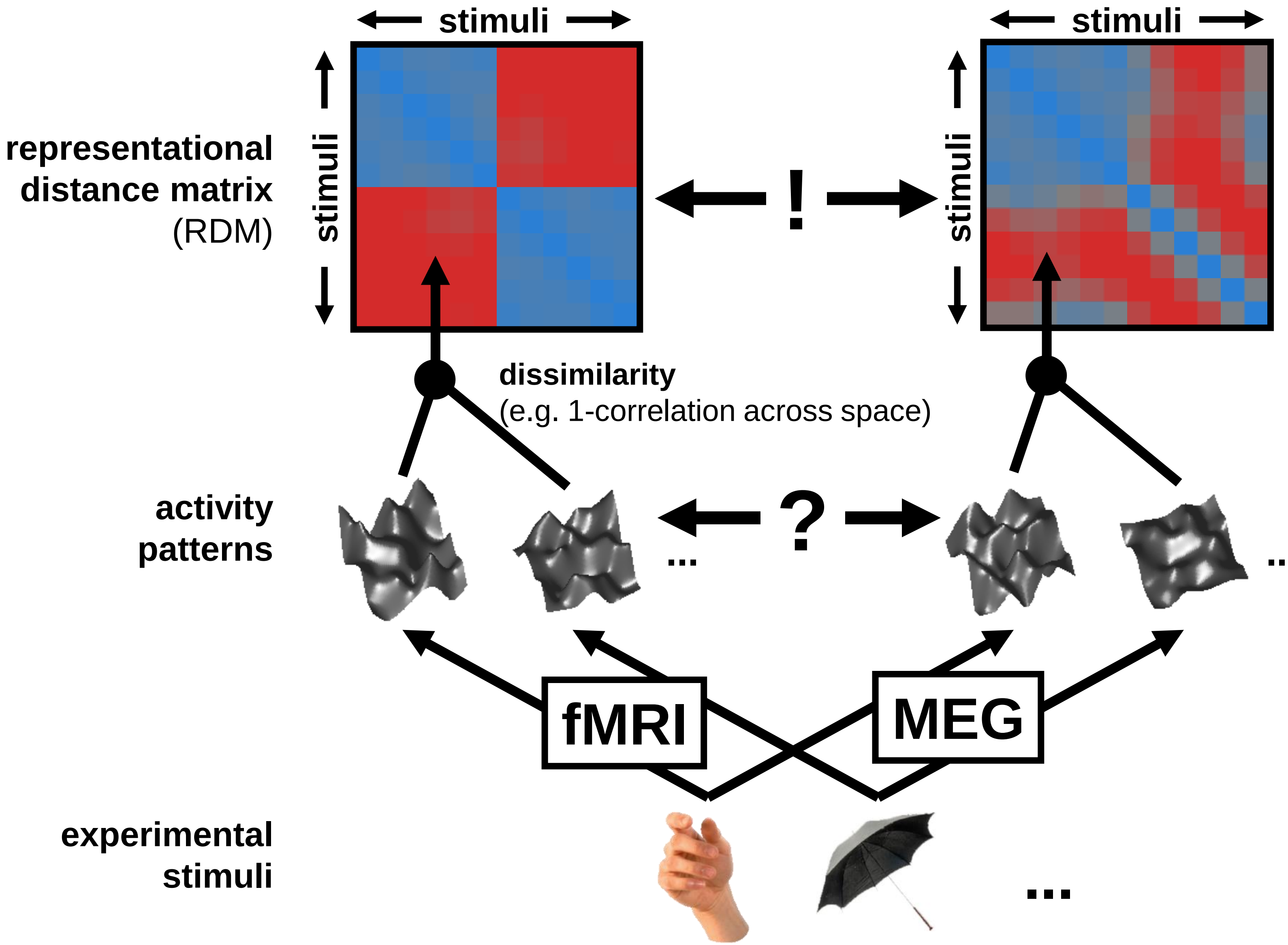
The representational similarity trick



The representational similarity trick

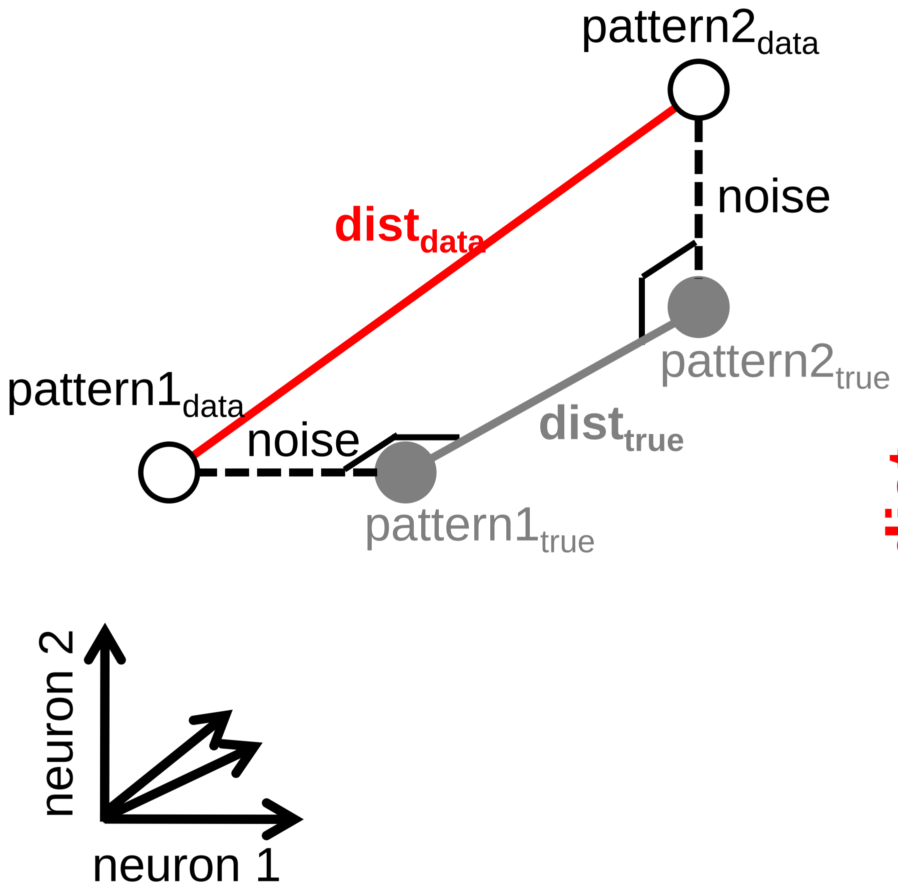


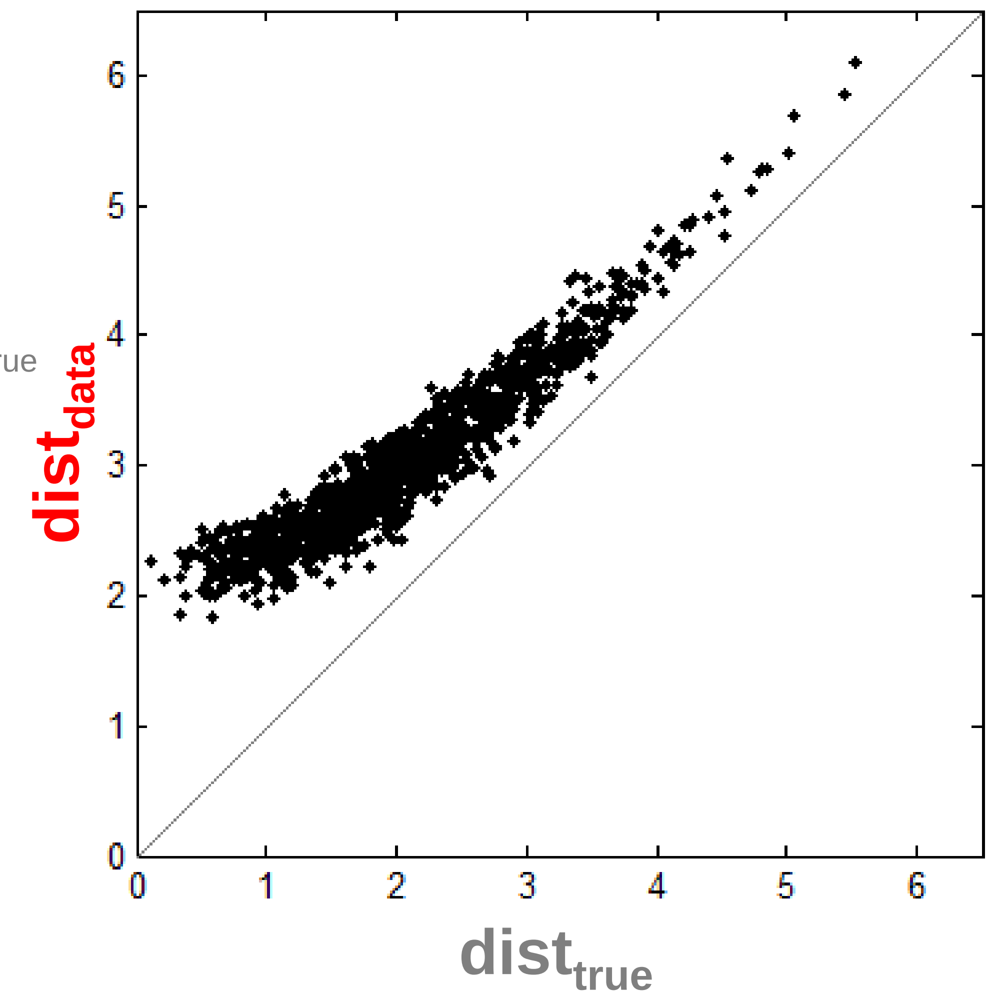
The representational similarity trick



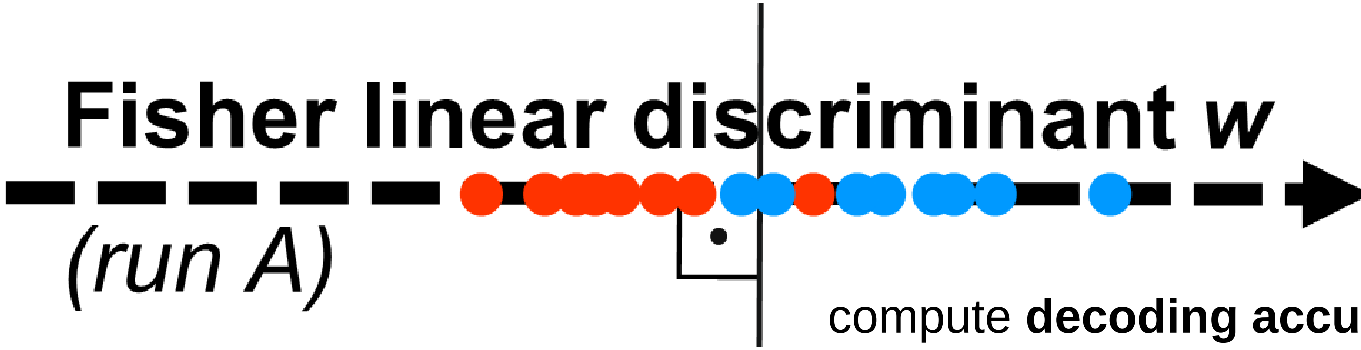
How can we best measure representational distances?

Distance estimates are positively biased



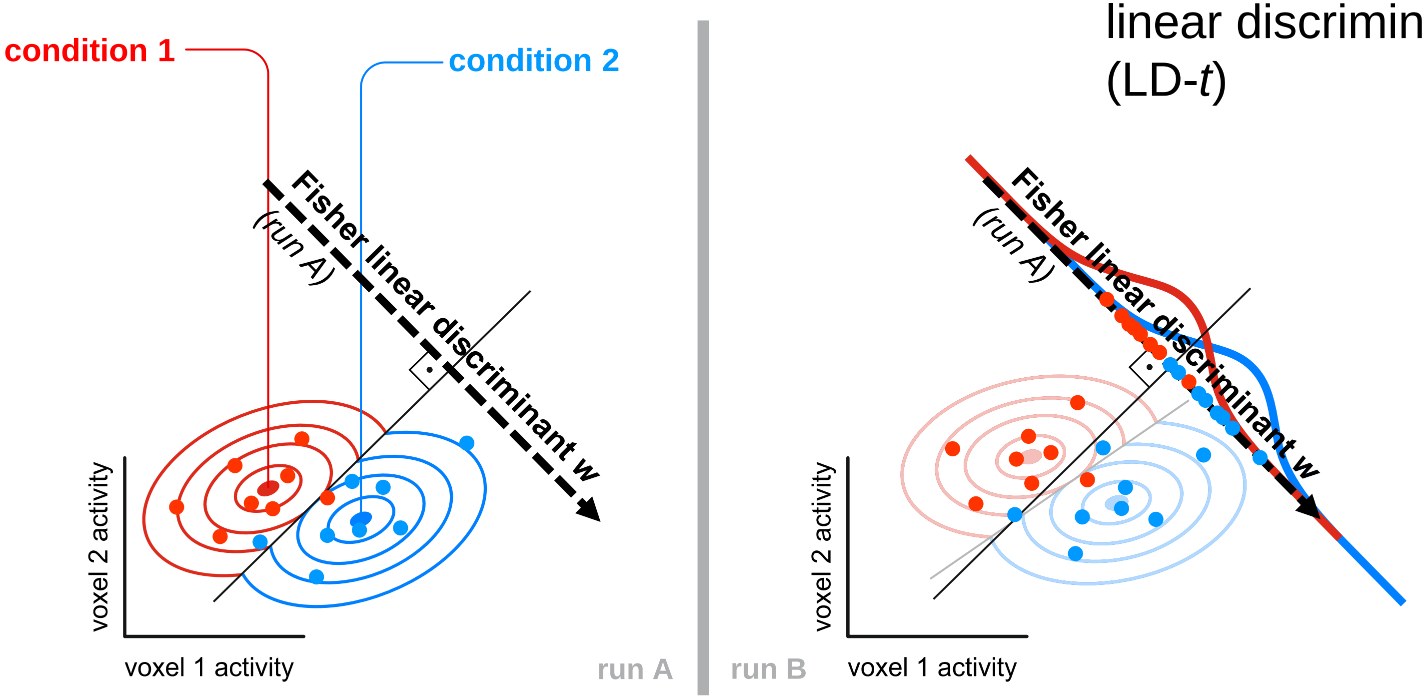


Distances are positively biased
– just like training-set decoding accuracies!



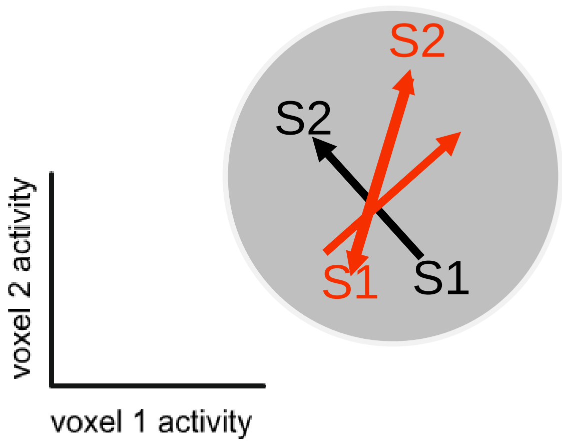
compute **decoding accuracy**?
or just do a ***t* test**?

linear discriminant *t* value
(LD-*t*)



Kriegeskorte et al. 2007

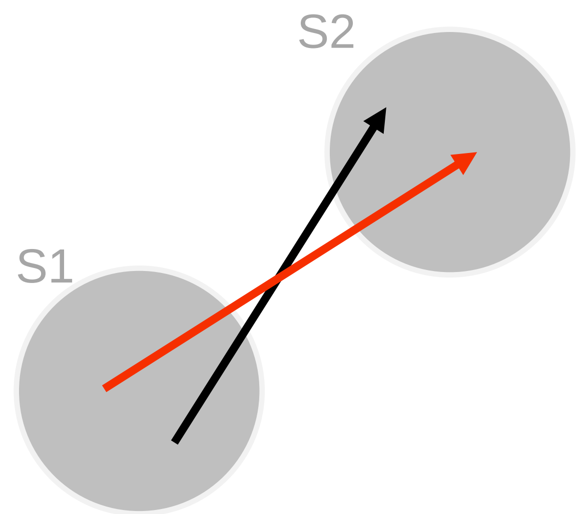
Unbiased distance estimates through crossvalidation



true distance = 0

average angle = 90°

E(inner product) = 0



true distance = 1

average angle < 90°

E(inner product) > 0

data set A

data set B

The linear discriminant contrast (LDC) is a crossvalidated variant of the Mahalanobis distance

Mahalanobis distance (single data set)

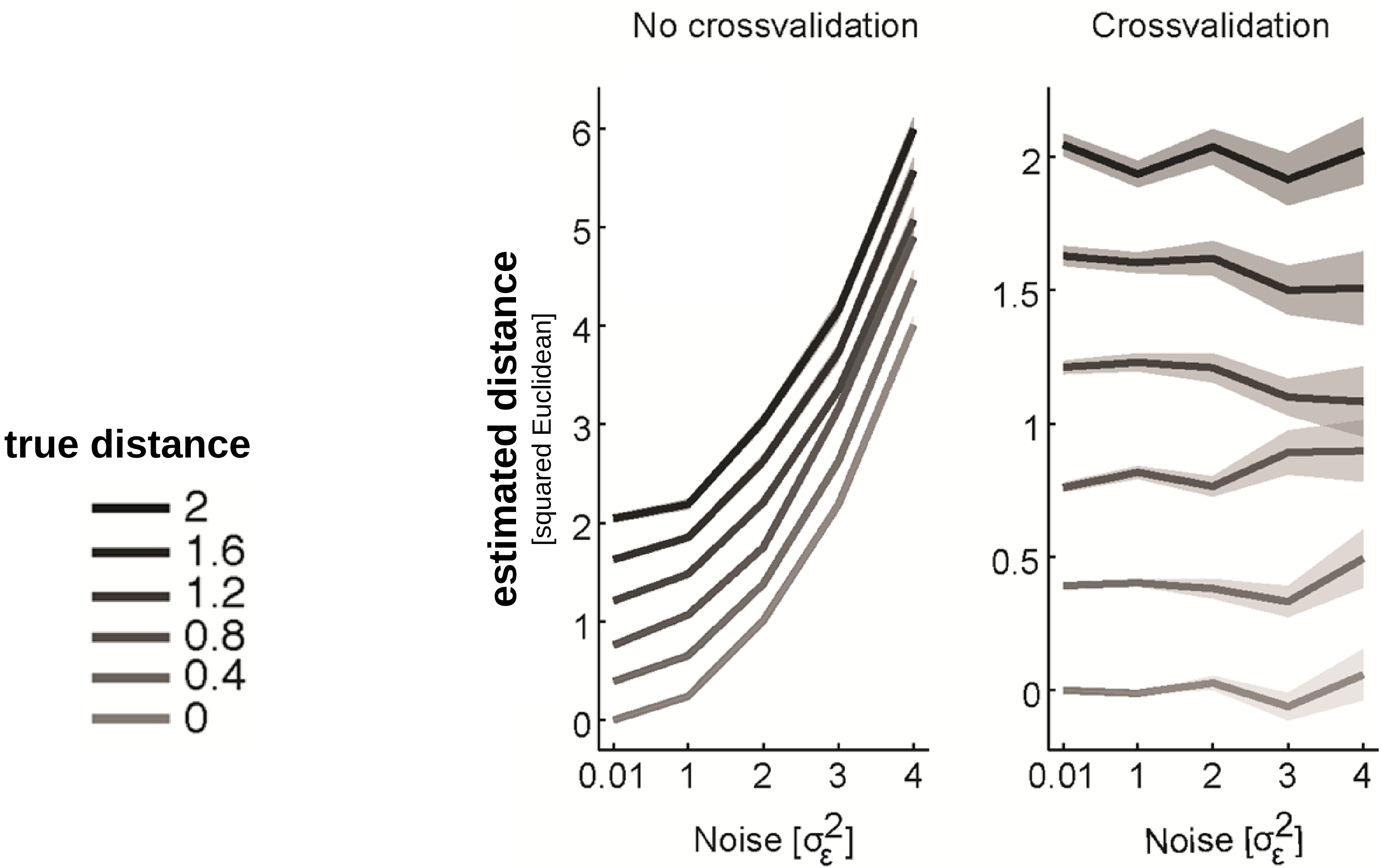
training set $(\mathbf{p2} - \mathbf{p1})^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{p2} - \mathbf{p1})$

Fisher linear discriminant contrast (crossvalidated)

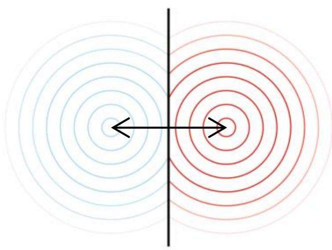
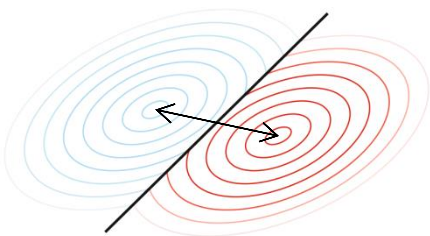
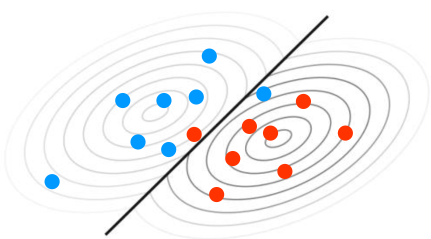
training set $(\mathbf{p2} - \mathbf{p1})^{\mathrm{T}} \mathbf{\Sigma}^{-1} (\mathbf{p2}' - \mathbf{p1}')$ test set

Nili et al. (2014)

Crossvalidation removes the bias of distance estimates



Walther et al. (in revision)

no crossvalidation	crossvalidation
Euclidean distance	Centroid connection discriminant contrast
Mahalanobis distance	**Fisher linear discriminant contrast**
Training-set decoding accuracy	Test-set decoding accuracy

covariance-blind

ceiling-limited and quantized

positively biased

The best of both worlds...

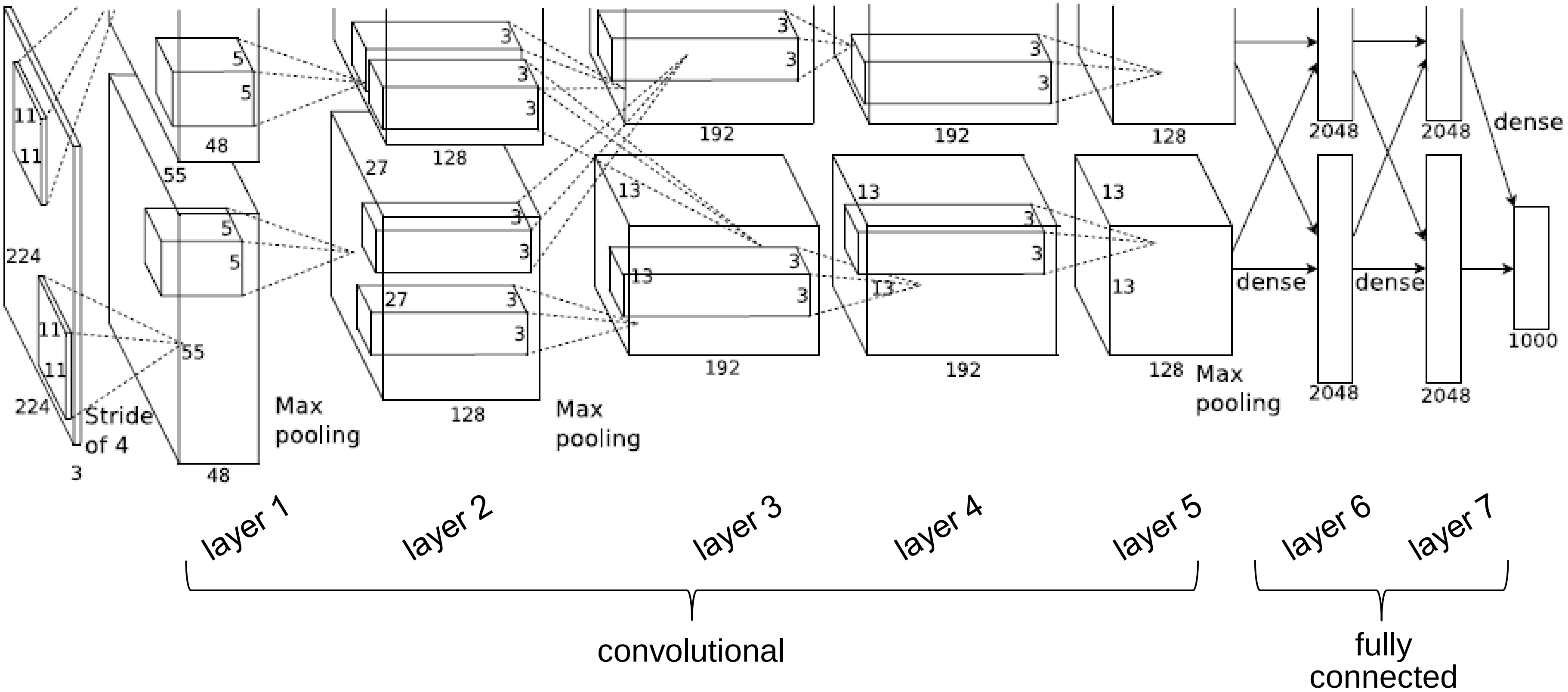
Multivariate statistics	**Machine learning**
multinormal distribution	pattern classifiers
continuous measures of multivariate separation	crossvalidation
inference relying on multinormality	nonparametric inference procedures

How can we test computational models?

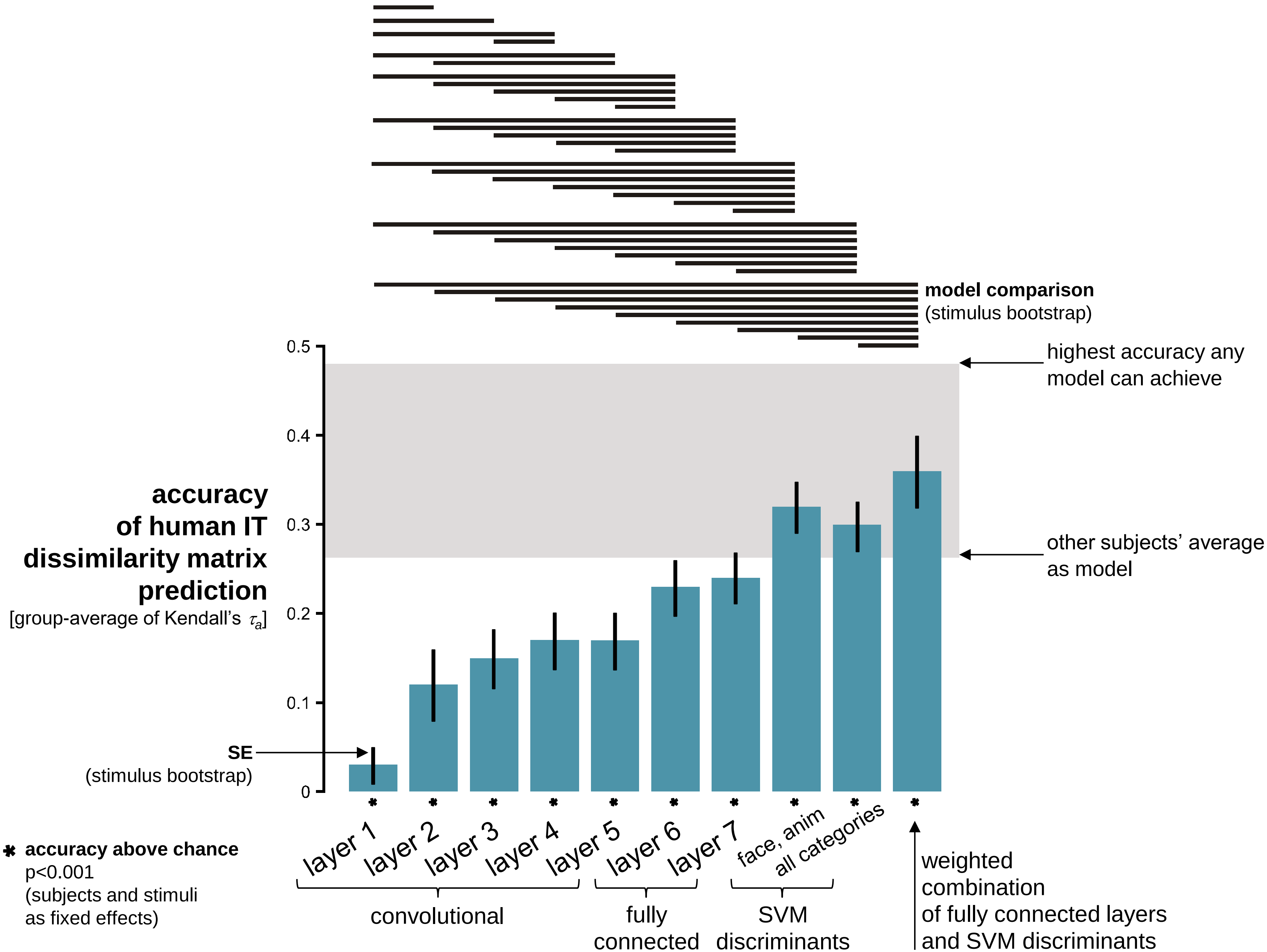
Deep convolutional neural network

- state of the art in computer vision
- trained with stochastic gradient descent
- supervised with 1.2 million category-labeled images
- 60 million parameters and 650,000 neurons

Is this network functionally similar to the brain?



Krizhevsky et al. 2012



Khaligh-Razavi & Kriegeskorte (2014), **Nili** et al. 2014 (RSA Toolbox)

Key insights

Representational geometries encapsulate the *content* and *format* of brain representations.

Representational geometries can be characterised by representational dissimilarity matrices (RDMs).

RDMs can easily be compared between brains and models, individuals and species, different brain regions, and brain and behaviour.

We can statistically compare multiple computational models and assess whether they fully explain the measured brain response patterns.