Geometric analysis of perceptual spaces

J. Victor, M. Conte
Weill Cornell Medical College

Topological Data Visualization
University of Iowa
June 2025

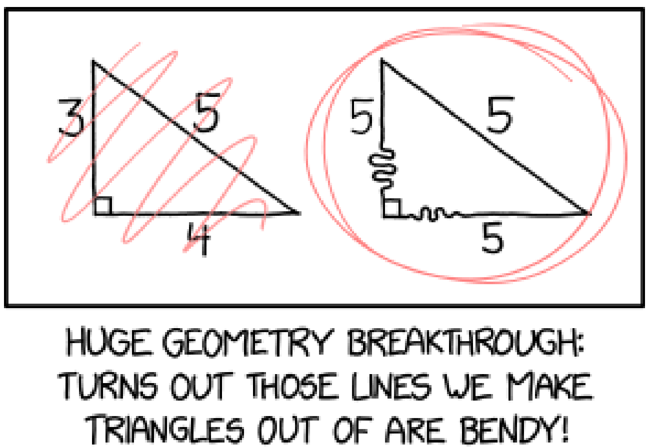
Support: NIH EY07977, NSF 2014217

What is a perceptual space?

- A model of a mental workspace
  - Points are stimuli within a domain (e.g., colors, faces, musical genres…)
  - Distances between points correspond to perceptual dis-similarity
- Why do we care?
  - Understand classification, generalization, learning
  - Understand the neural underpinnings of behavior and perception: *similar percepts should have similar neural representations*
- So it’s crucial to understand the geometry of similarity

Outline

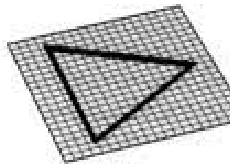
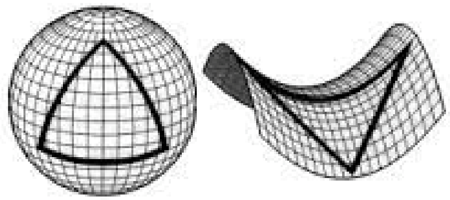
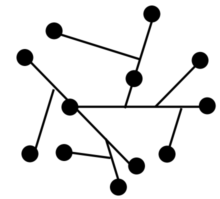
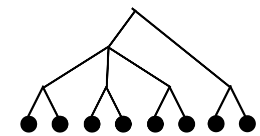
- What kinds of models do we need to consider for perceptual spaces?



https://xkcd.com/2706

- Testing these models experimentally
  - Low-level (features) and high-level (semantic) content
  - The influence of task
- A complementary analytic strategy
- Open questions

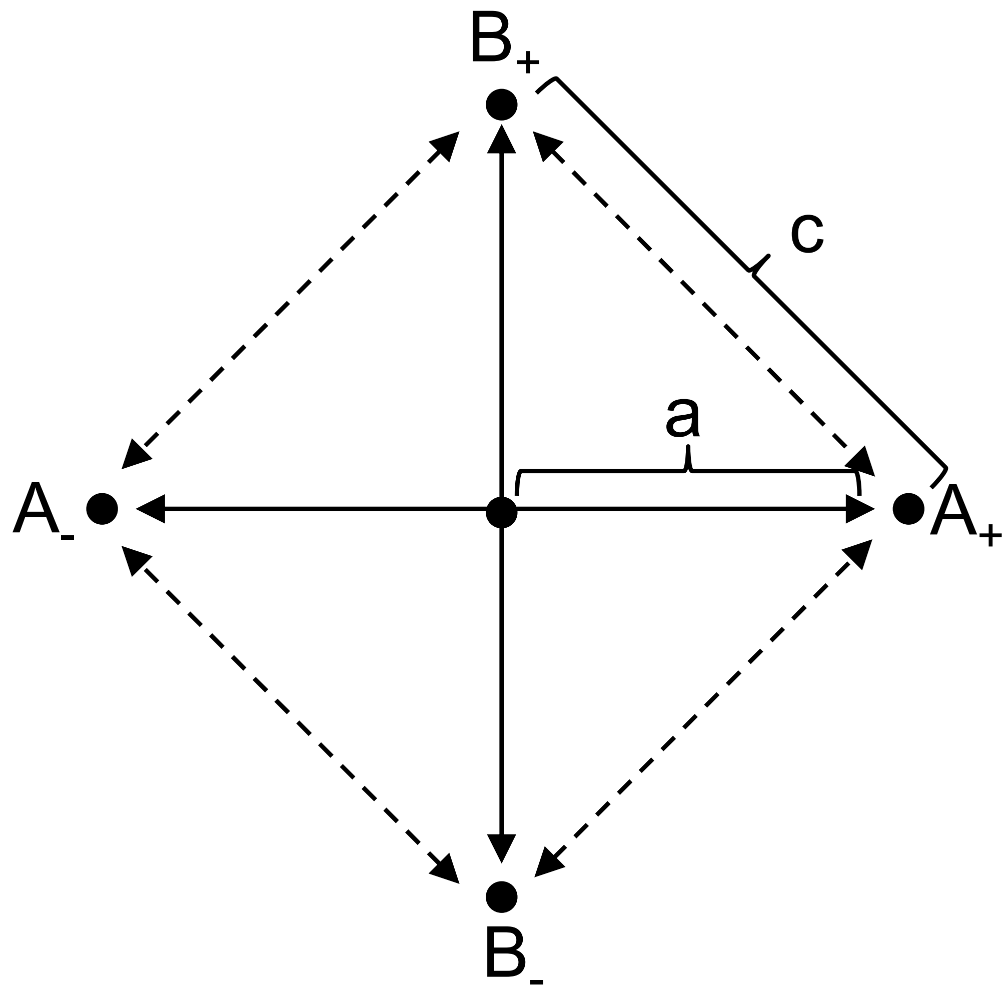
The Zoo

	Euclidean	*Locally Euclidean*	*Minkowski*	*Addtree*	*Ultrametric*
	Everyday life	Distances on curved surfaces	Manhattan distances	Mileage (no loops)	Hierarchy
					
coordinates?	✓	local	✓	✗	✗
continuous?	✓	✓	✓	✓ ✗	✗
Pythagorean?	✓	local	✗	✗	✗
Isotropic?	✓	✓	✗	✗	✗
Inequalities					
triangle	✓	✓	✓	✓	✓
“four-point”	✗	✗	✗	✓	✓
ultrametric	✗	✗	✗	✗	✓

And inhomogeneity.

And mixtures.

Toy Scenario



Euclidean: $c^2 = 2a^2$

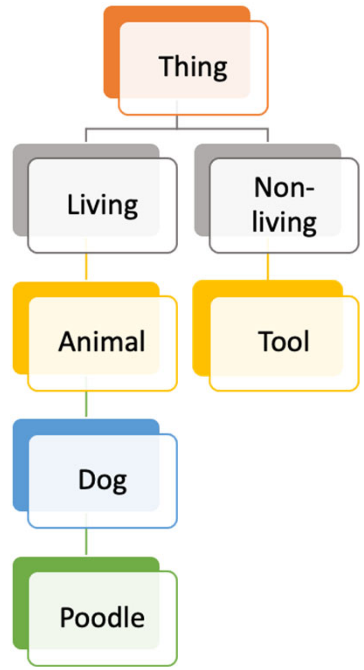
Spherical: $c^2 < 2a^2$

Hyperbolic: $c^2 > 2a^2$

Comparing c and a constrains the geometry.

Are there qualitative differences between perceptual spaces?

color lies in a continuous domain



objects are often categorical

A range of stimulus domains

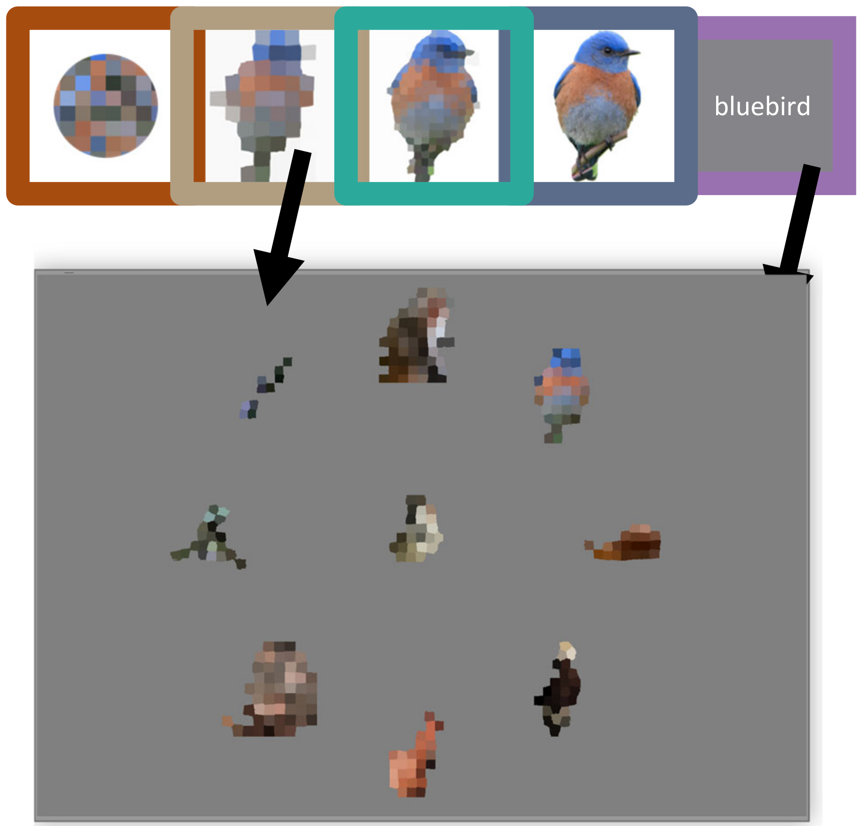
colored textures

common animal names

Stimuli correspond across domains.

Collecting similarity judgments

Subjects click each of 8 comparison stimuli in order of their similarity to the central reference



*One trial yields a ranking of 8 similarities to the central reference, i.e., (8*7)/2=28 comparison pairs.*

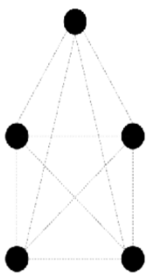
Waraich, S.A., Victor, J.D., (2022), J. Vis. Exp. (181)

Design Details

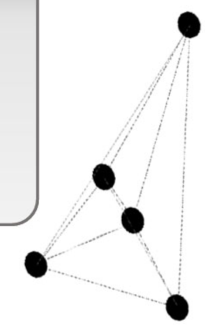
- In each domain
  - 37 stimuli
  - 222 unique trials
    - Designed to include all (reference , comparison) pairs
    - Designed to include some (reference, comparison) pairs in two contexts
    - Otherwise “frozen” randomization
  - One trial yields 28 distance comparison pairs
    - 222 trials x 28 distance comparison pairs = 6216
    - << all possible d(A,B) vs d(A,C) comparisons [N(N-1)(N-2)/2=23310]
    - << all possible d(A,B) vs d(C,D) comparisons [N(N-1)(N-2)(N-3)/8=198135]
    - But large enough to constrain models
  - Each unique trial repeated 5 times
    - Allows estimation of choice probability
- 5 domains, 11-12 subjects per domain
  - 10 hrs/subject/domain

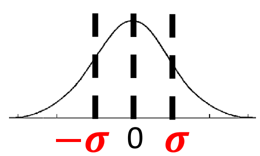
Inferring geometry from similarity judgments

Win/lose heuristic → standard multi-dimensional scaling



adjust point coordinates → Calculate pairwise distances → Predict trial-by-trial response using decision model → Calculate log-likelihood of observed responses → Get best coordinates





D = d(owl, cow) - d(owl, frog) + noise

$$P(d(owl, cow) < d(owl, frog) = P(D < 0)$$

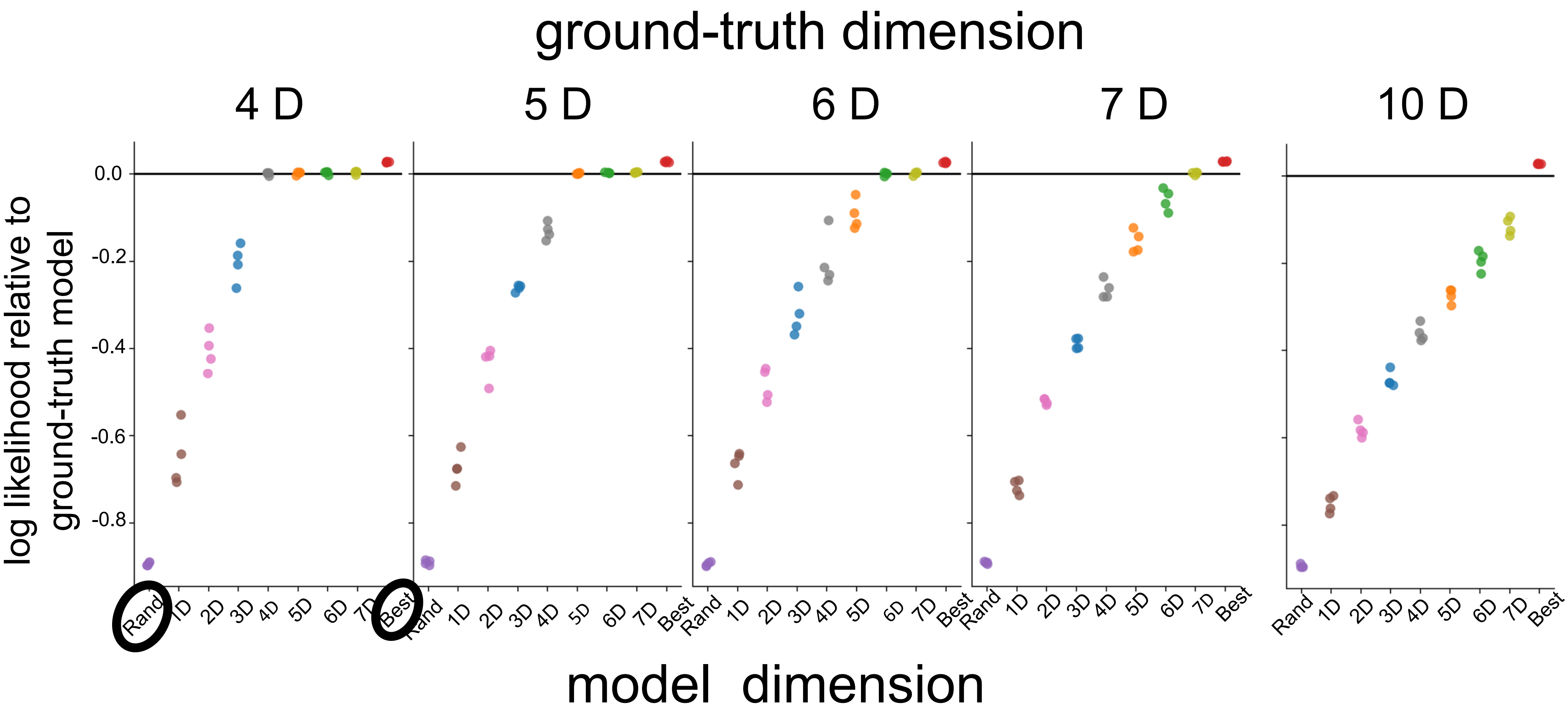
$$LL = \frac{1}{N_0} \sum_{r,i,j} \left(c_{r,i,j} \log_2 p_{r,i,j} + \left(N_{repeats} - c_{r,i,j}\right) \log_2\left(1 - p_{r,i,j}\right)\right)$$

r: reference
i,j: comparison stimuli

Comparison	Empirical Choice Probability	Model Probability
d(owl, mouse) < d(owl, elephant)	4/5	0.87
d(owl, cow) < d(owl, frog)	1/5	0.15
...	...	...

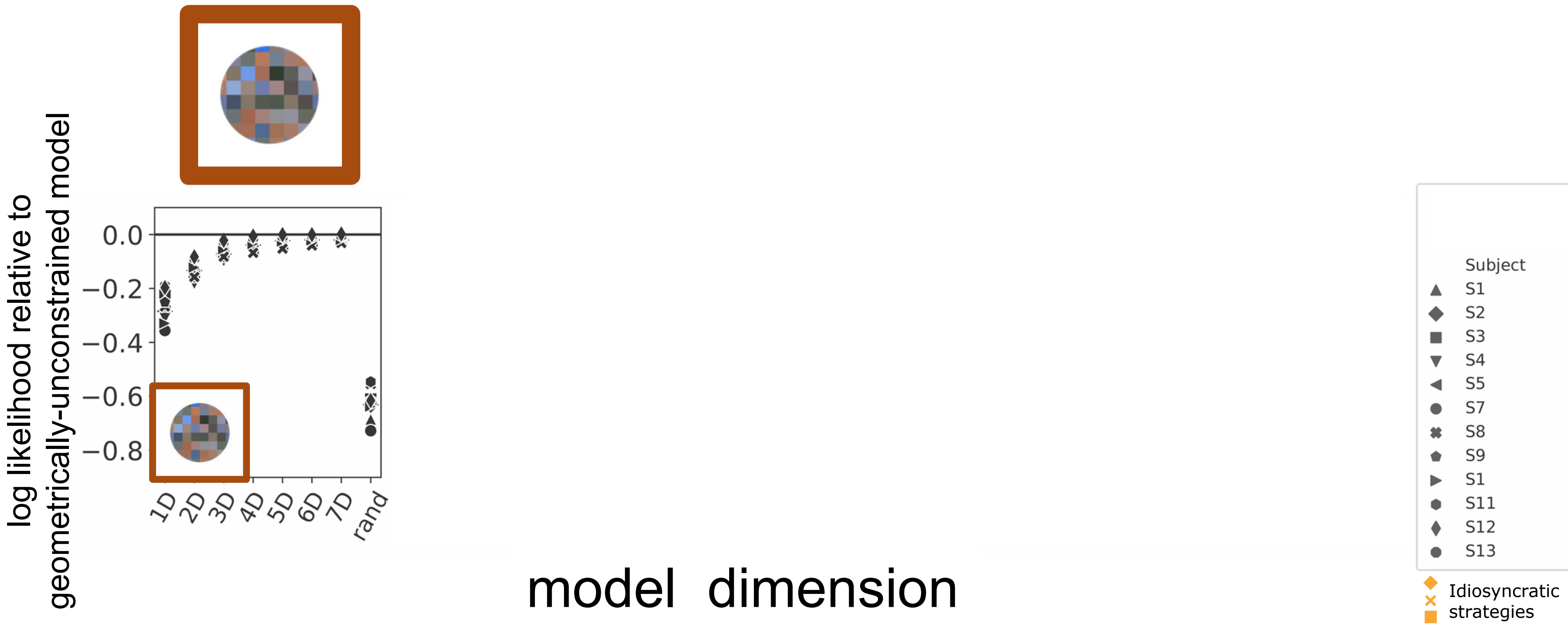
Distances are measured w.r.t. a noise parameter σ

Validation: numerical simulations



The analysis works for at least 7 dimensions.

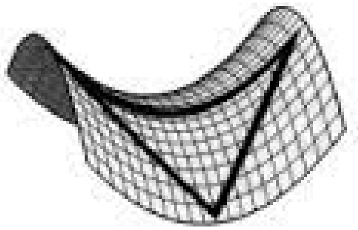
Results across the five domains



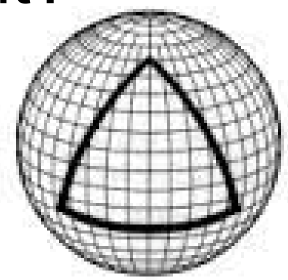
No clear difference in
dimensionality across domains.

Do the domains differ in curvature?

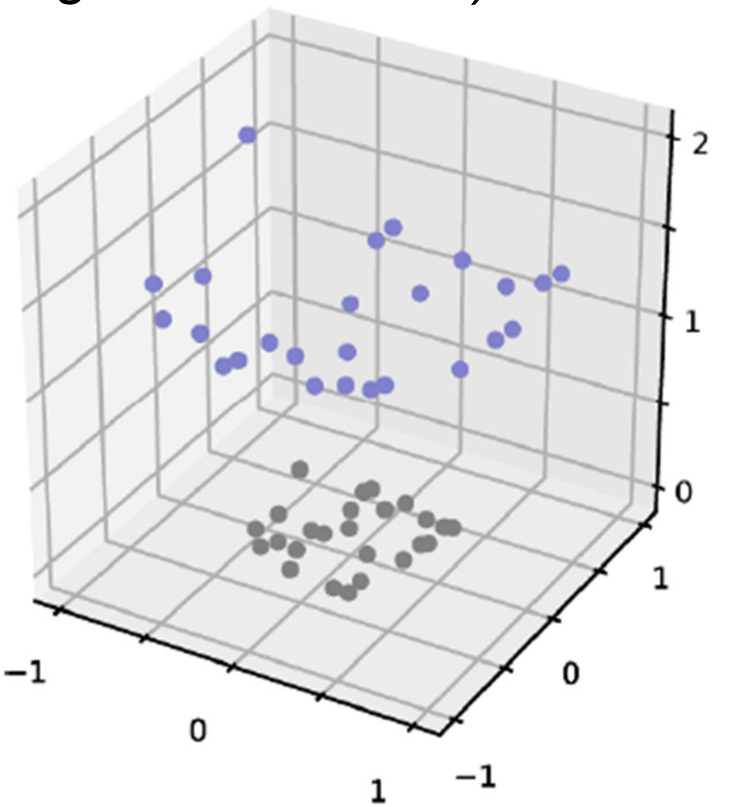
Does projecting the best Euclidean model onto a curved surface improve the fit?

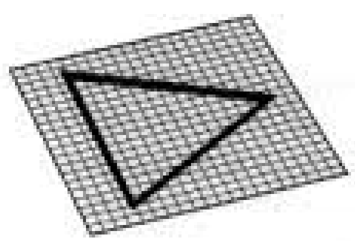


Hyperbolic
(negative curvature)

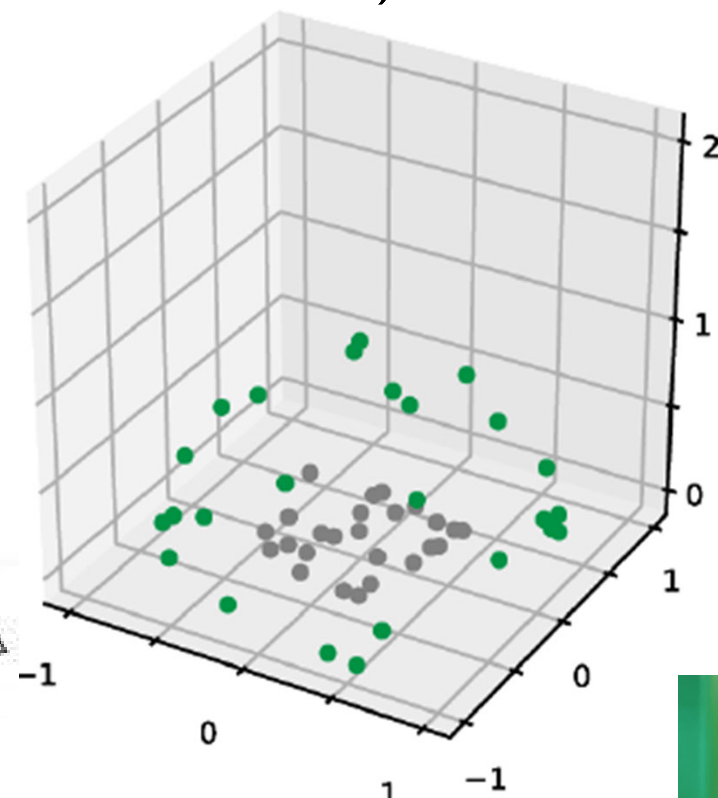


Spherical
(positive curvature)

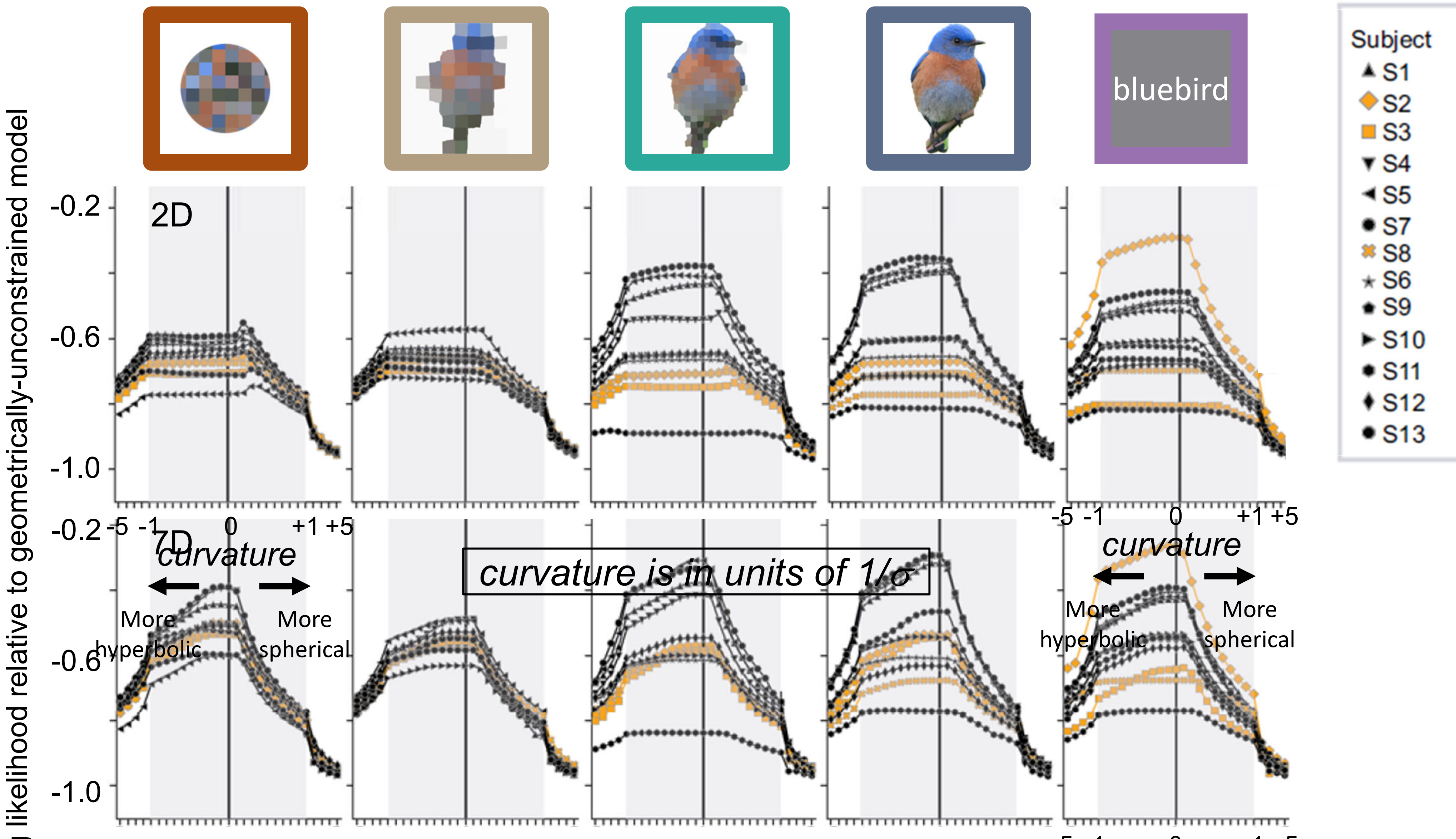




Euclidean

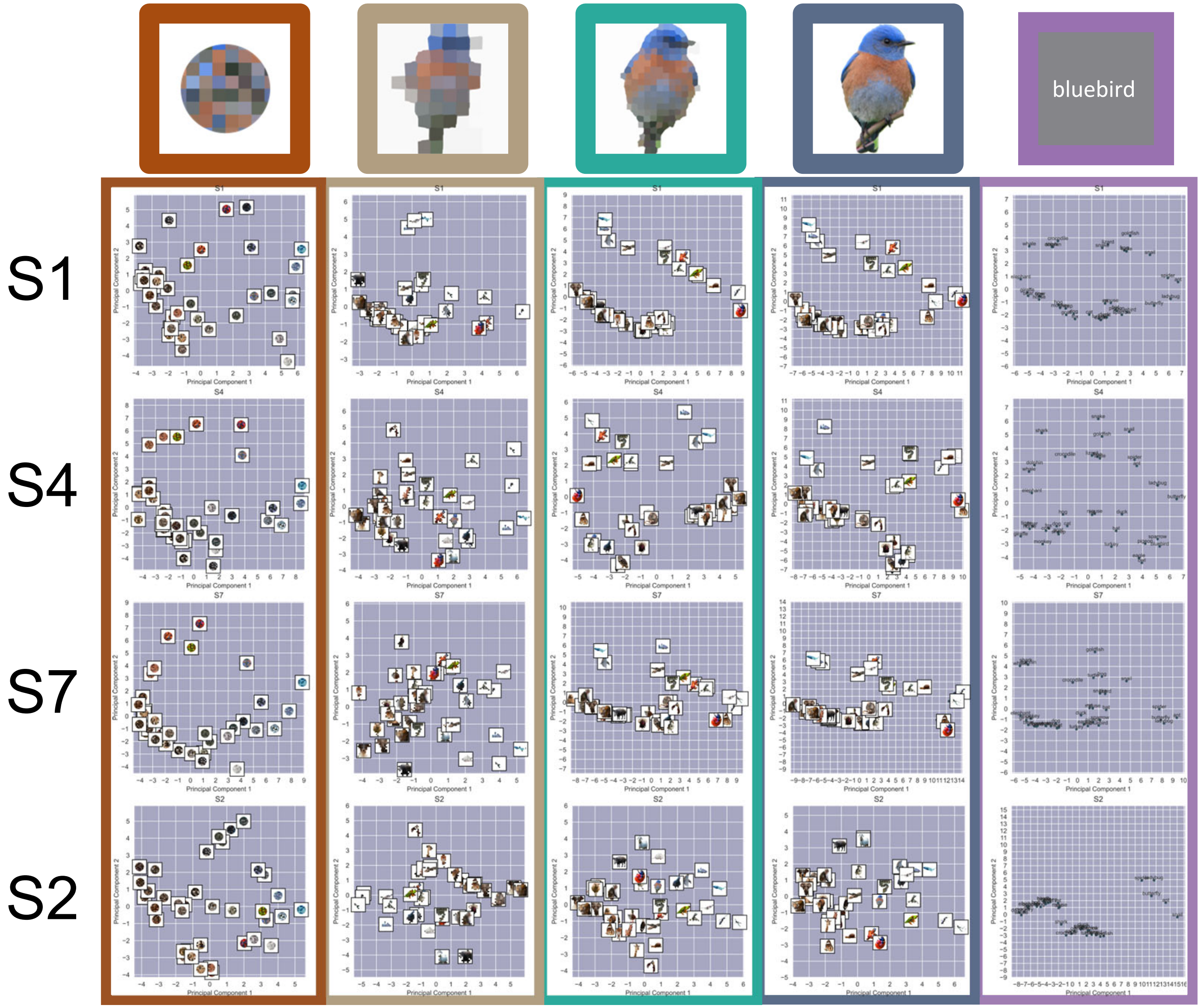


No

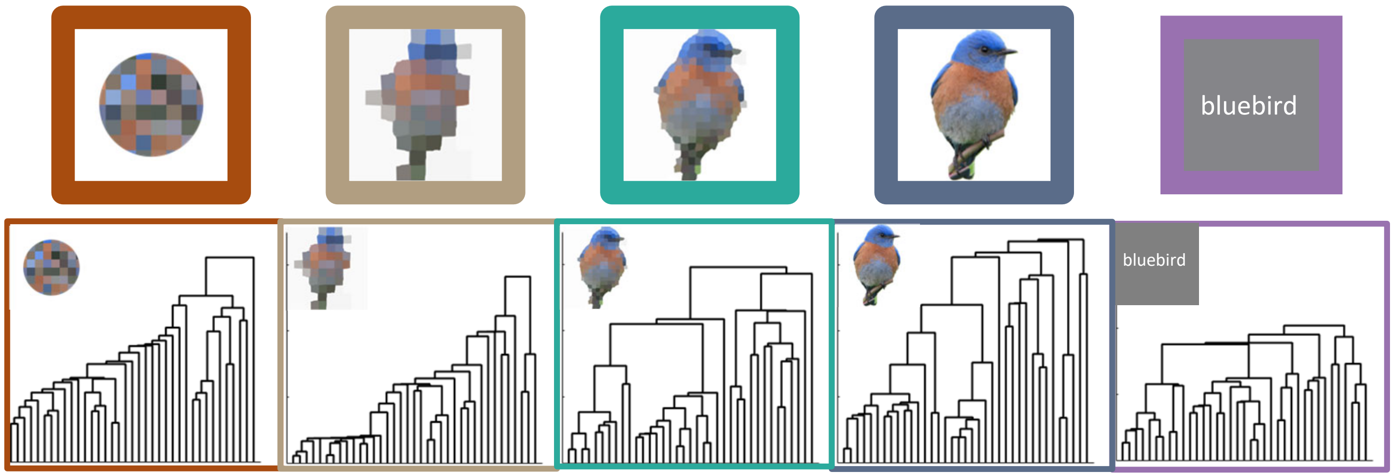


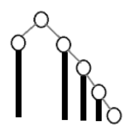
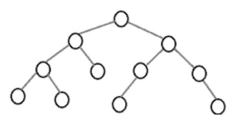
bluebird
Subject
S1
S2
S3
S4
S5
S7
S8
S6
S9
S10
S11
S12
S13
log likelihood relative to geometrically-unconstrained model
2D
7D
-0.2
-0.6
-1.0
-5 -1 0 +1 +5
curvature
More hyperbolic
More spherical
curvature is in units of 1/σ

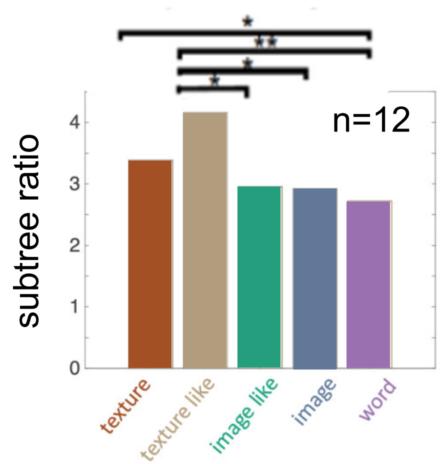
How are the points arranged?

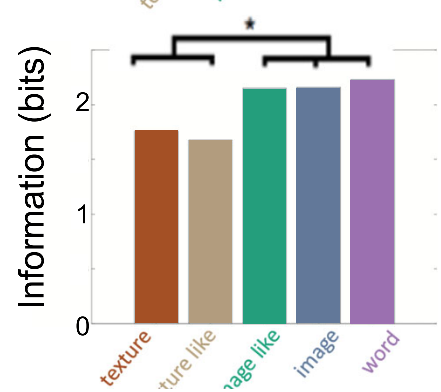


Analyze by hierarchical clustering



	 unbalanced tree	 balanced tree
ratio of subtree sizes	>>1	~1
information halfway down tree	low	high



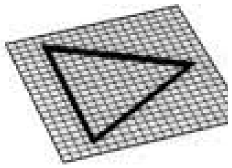
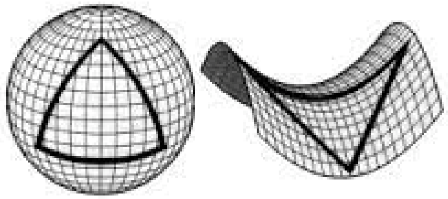
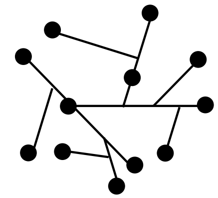
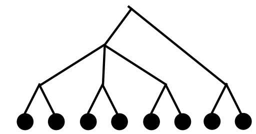


So far:

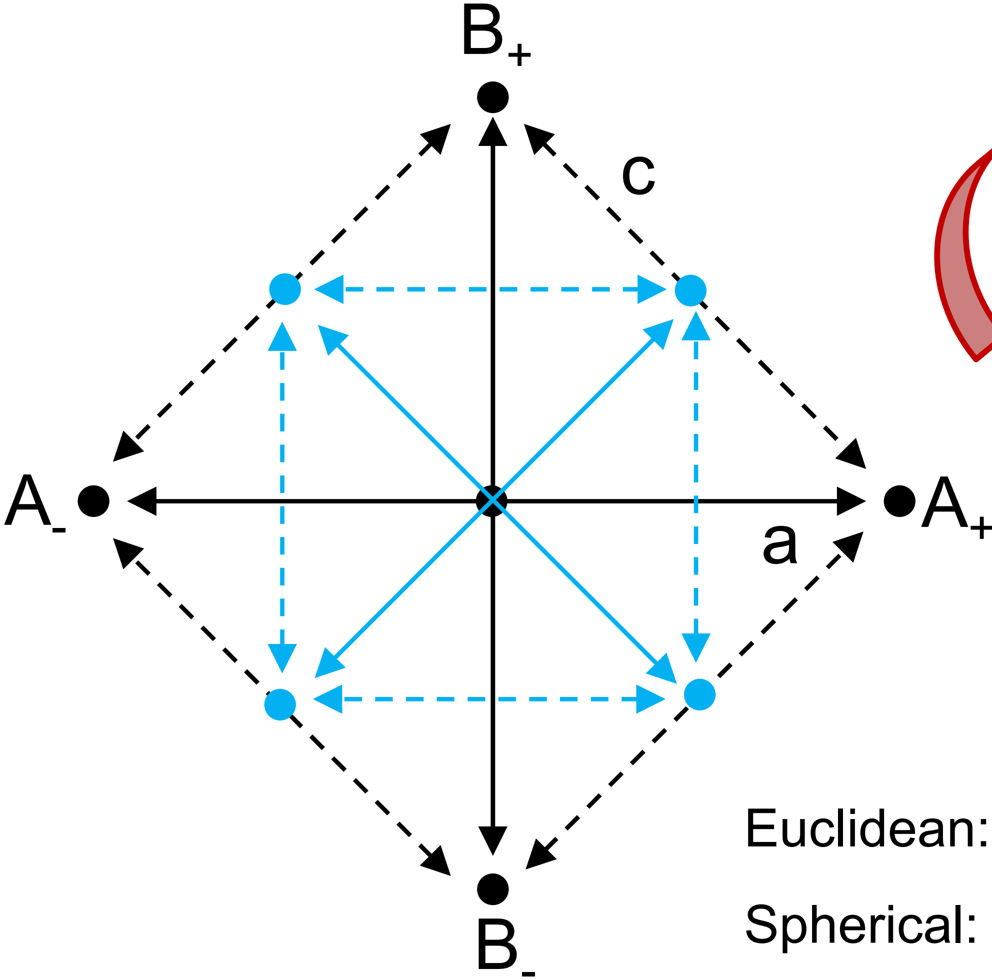
- A way to acquire and analyze similarity judgments
  - Euclidean models seem OK
  - Domains differ in geometry, but need to look at (relatively) subtle aspects

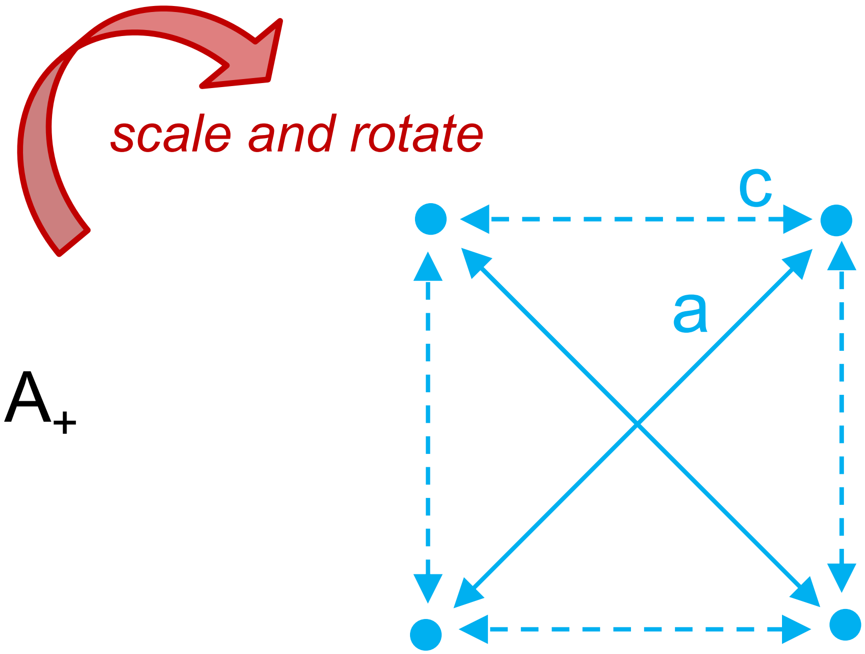
- Can we make better use of domain structure?

The Zoo

	Euclidean	*Locally Euclidean*	*Minkowski*	*Addtree*	*Ultrametric*
	Everyday life	Distances on curved surfaces	Manhattan distances	Mileage (no loops)	Hierarchy
					
coordinates?	✓	local	✓	✗	✗
continuous?	✓	✓	✓	✓ ✗	✗
Pythagorean?	✓	local	✗	✗	✗
Isotropic?	✓	✓	✗	✗	✗
Inequalities					
triangle	✓	✓	✓	✓	✓
"four-point"	✗	✗	✗	✓	✓
ultrametric	✗	✗	✗	✗	✓

Back to the Toy Scenario





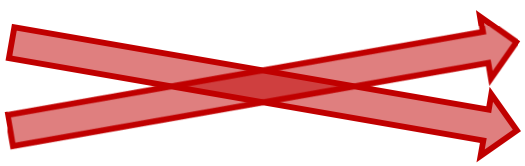
Euclidean: $c^2 = 2a^2$

Spherical: $c^2 < 2a^2$

Hyperbolic: $c^2 > 2a^2$

Minkowski (p>2): $c^2 < 2a^2$

Minkowski (p<2): $c^2 > 2a^2$



Minkowski (p<2): $c^2 < 2a^2$

Minkowski (p>2): $c^2 > 2a^2$

Rank order never helps to distinguish; in all cases, c>a!

We need quantitative distances, but also midpoints.

What we need is a perceptual space of high dimension, in which we can find midpoints.

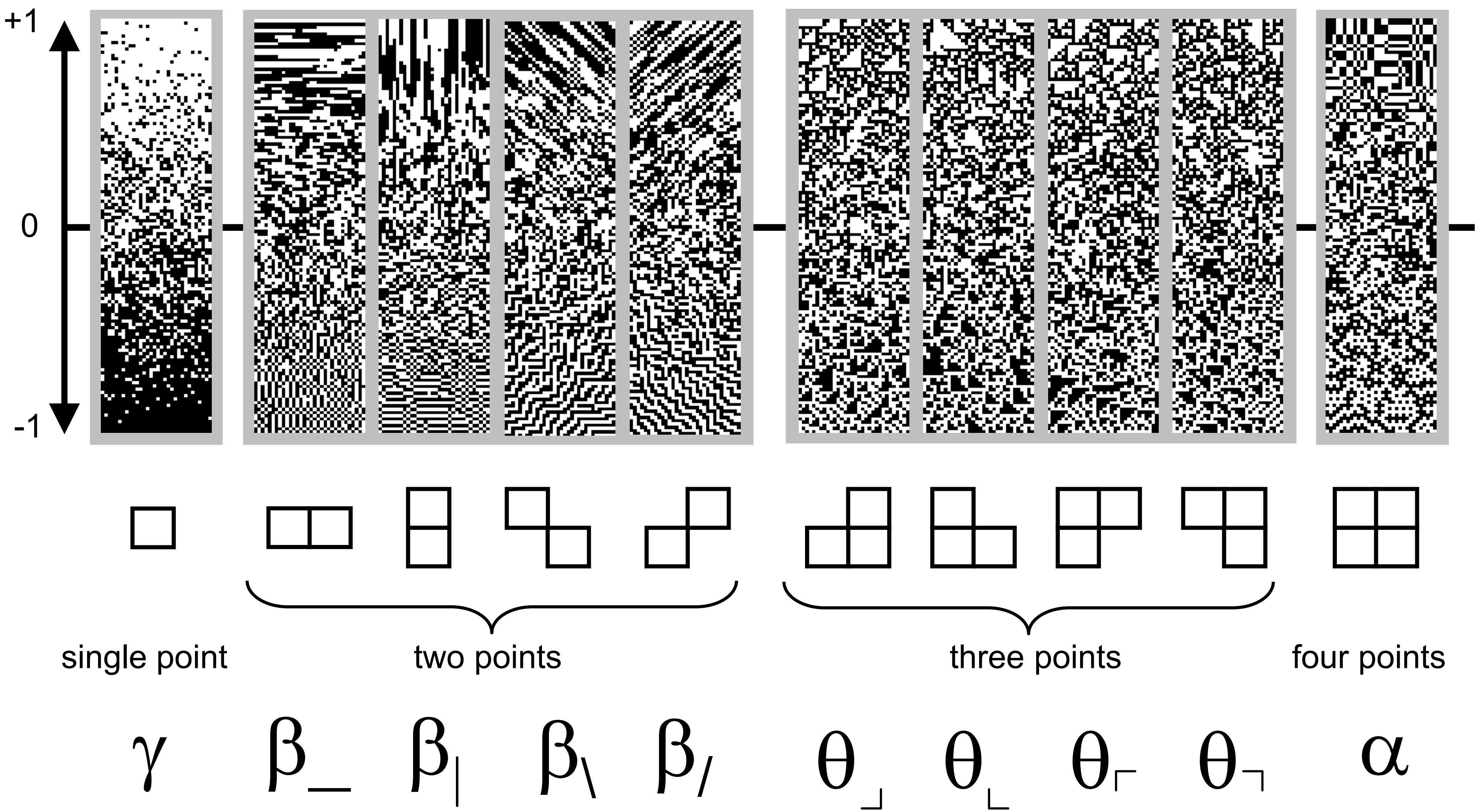
Visual textures: A good test case

- Functionally important
  - Segmentation
  - Material estimation

- Technical advantages
  - High-dimensional and continuous
  - Local image statistics can be independently controlled
  - Thresholds are well-characterized and consistent with a Euclidean geometry

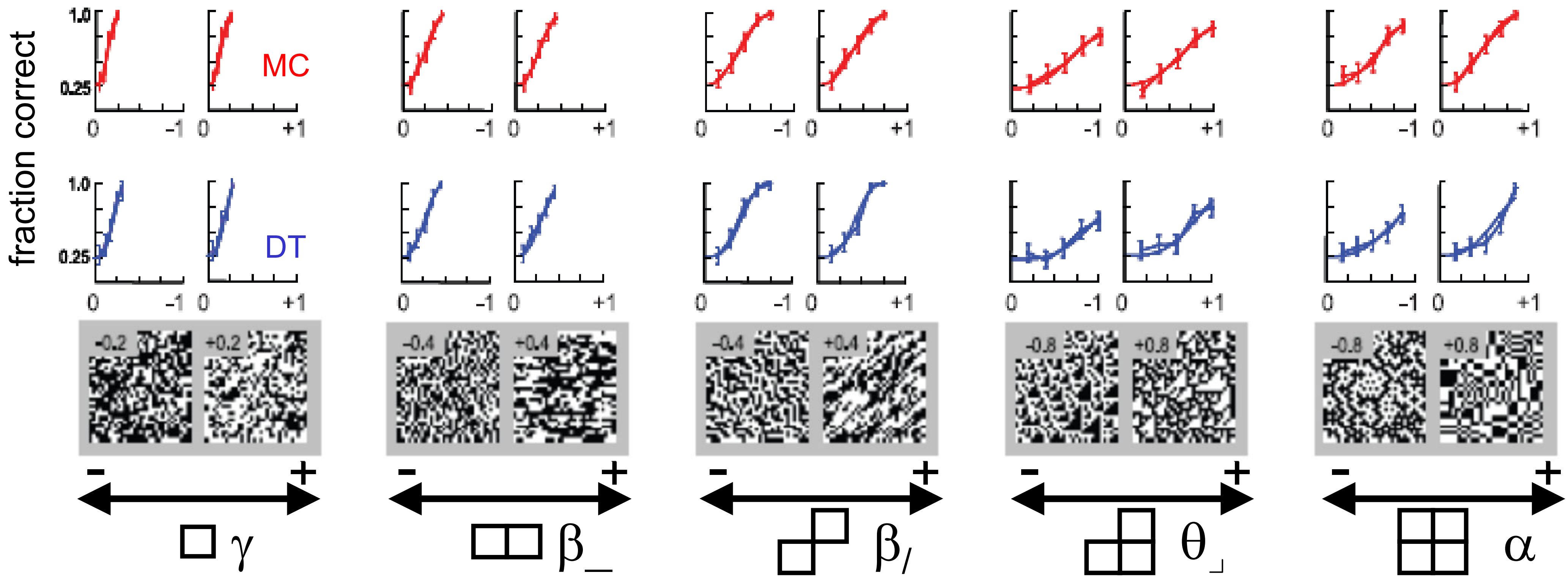
So things should be simple

A space of visual textures: 10 degrees of freedom



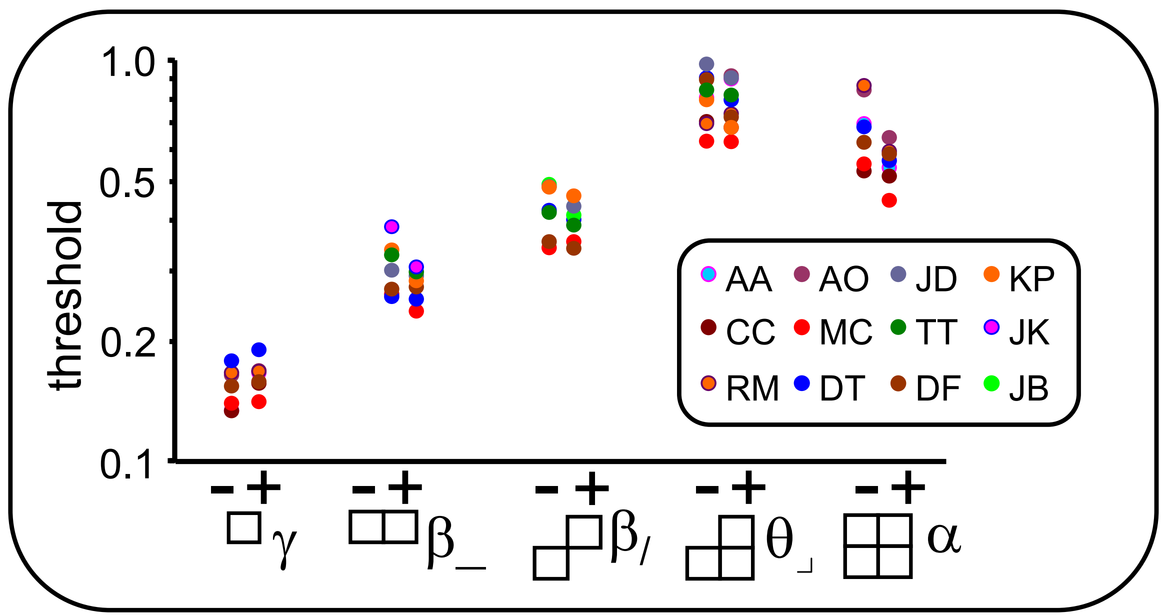
Victor & Conte, JOSA A 2012

Sensitivity is selective, and similar across observers

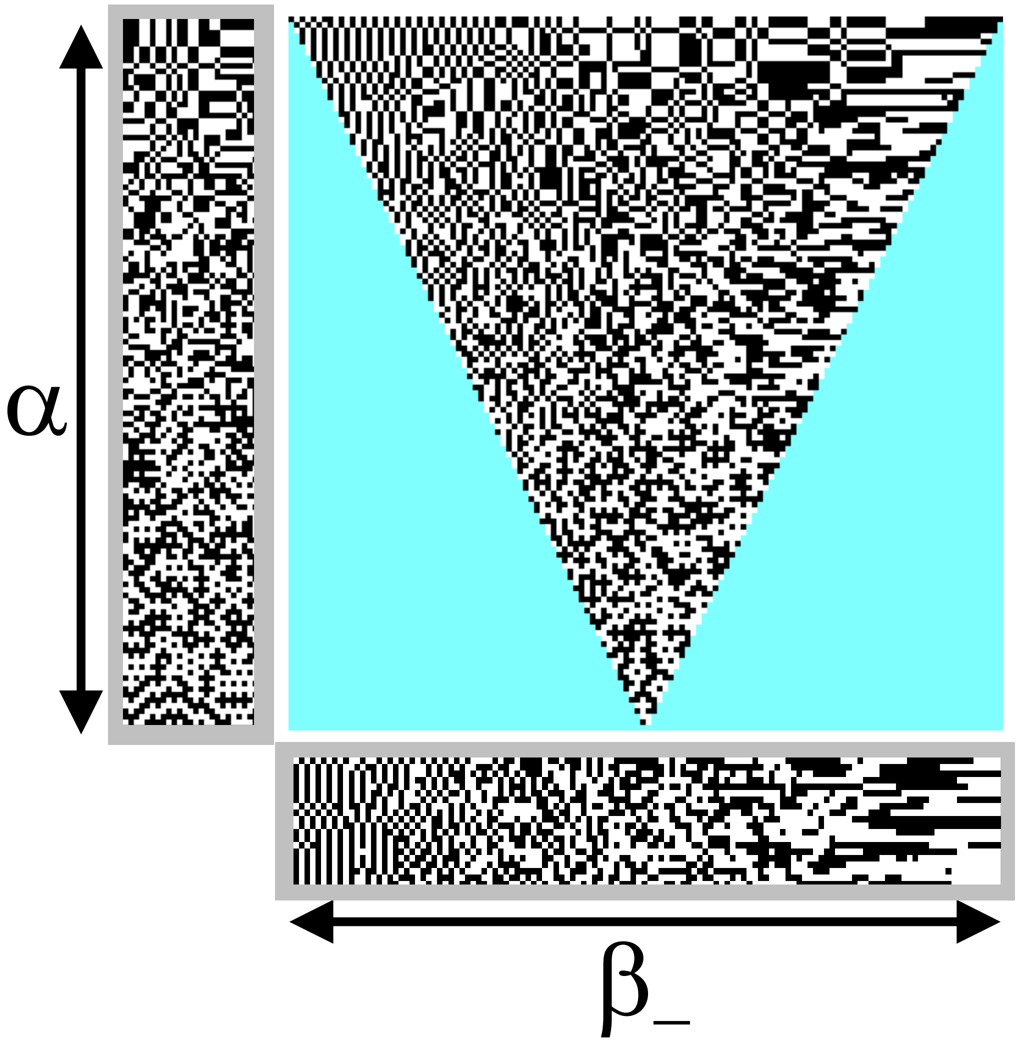


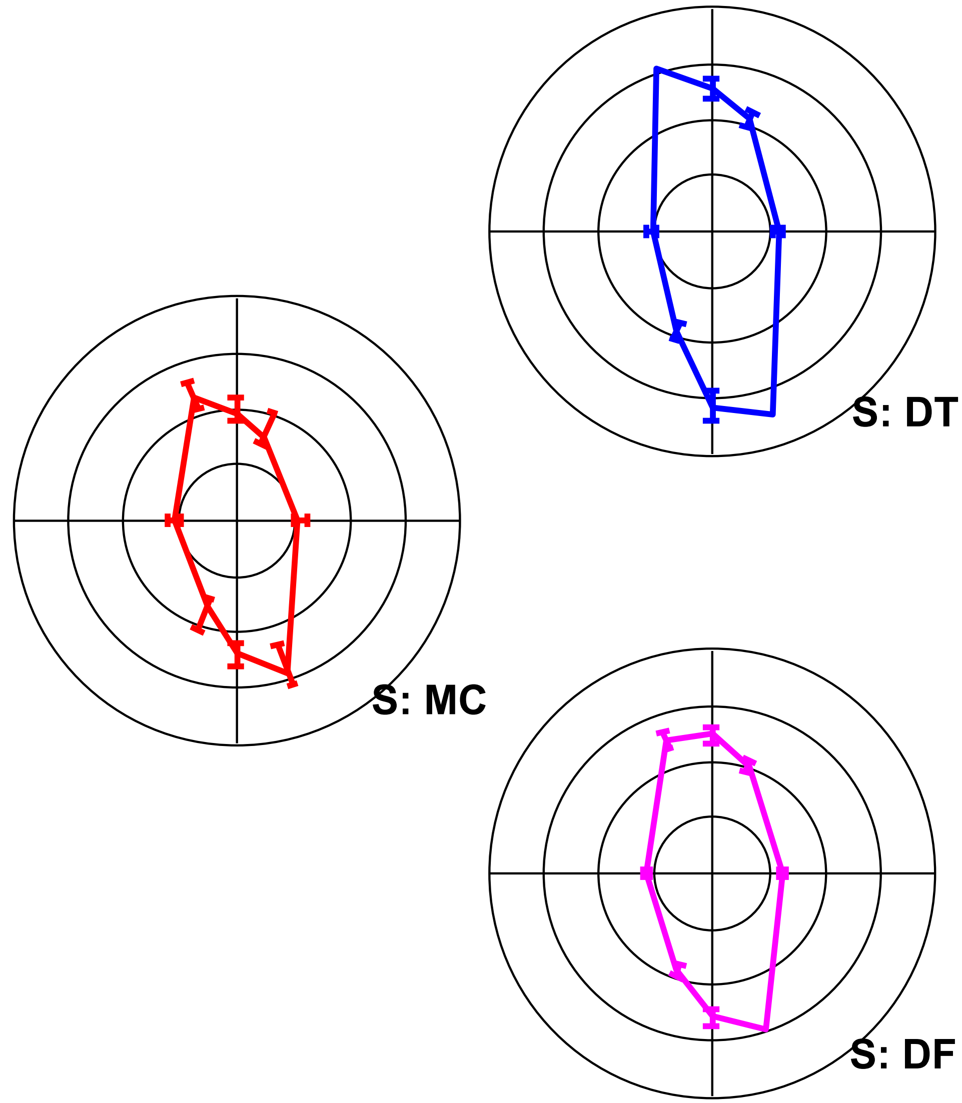
12 normal subjects show a consistent pattern of sensitivities

(we now have data on 26 subjects)



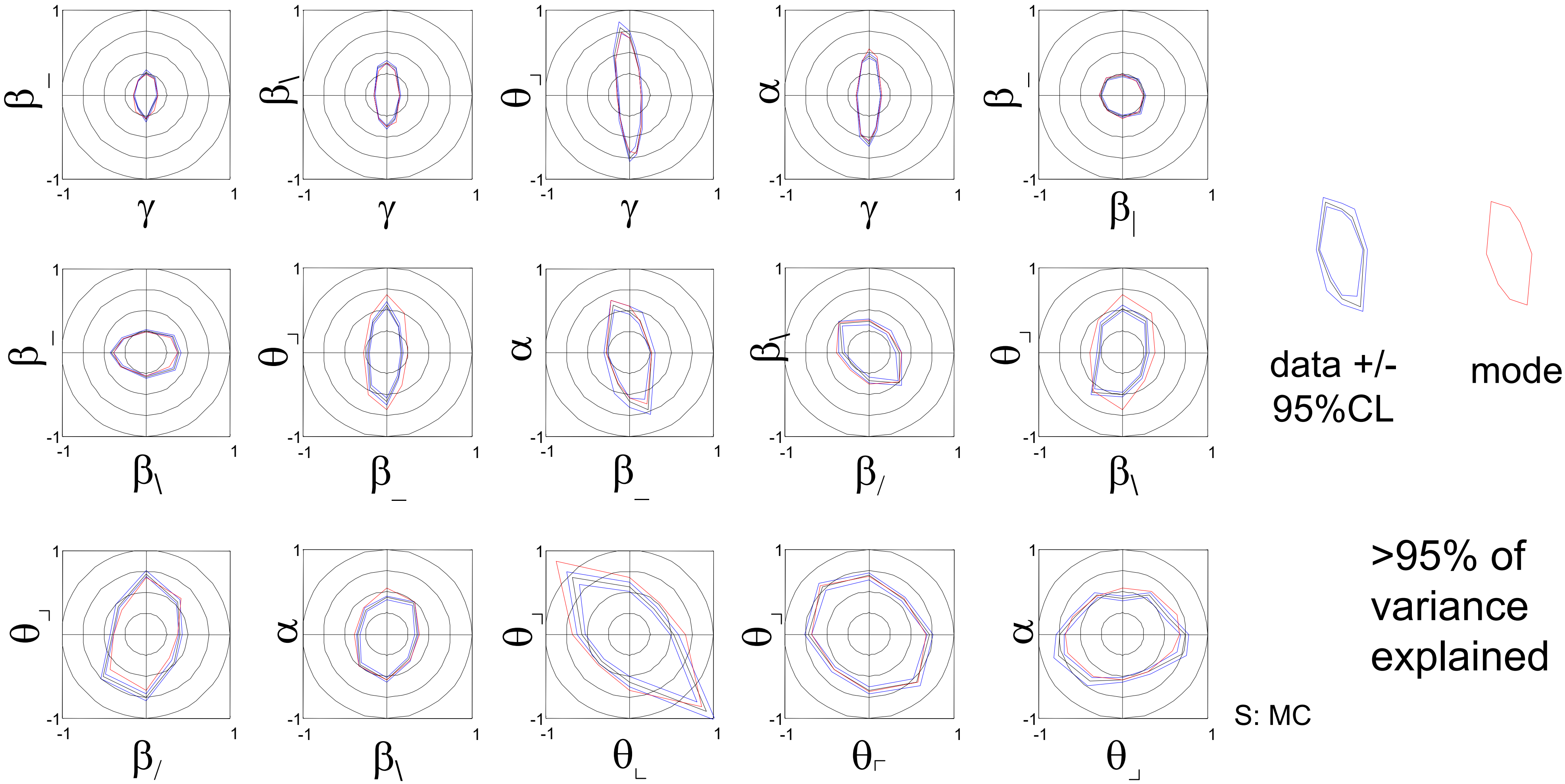
Pairwise interactions





Victor, Thengone, Rizvi, & Conte, Vision Res 2015

A quadratic model accounts for perceptual thresholds



In each plane, isodiscrimination contours are approximately elliptical.

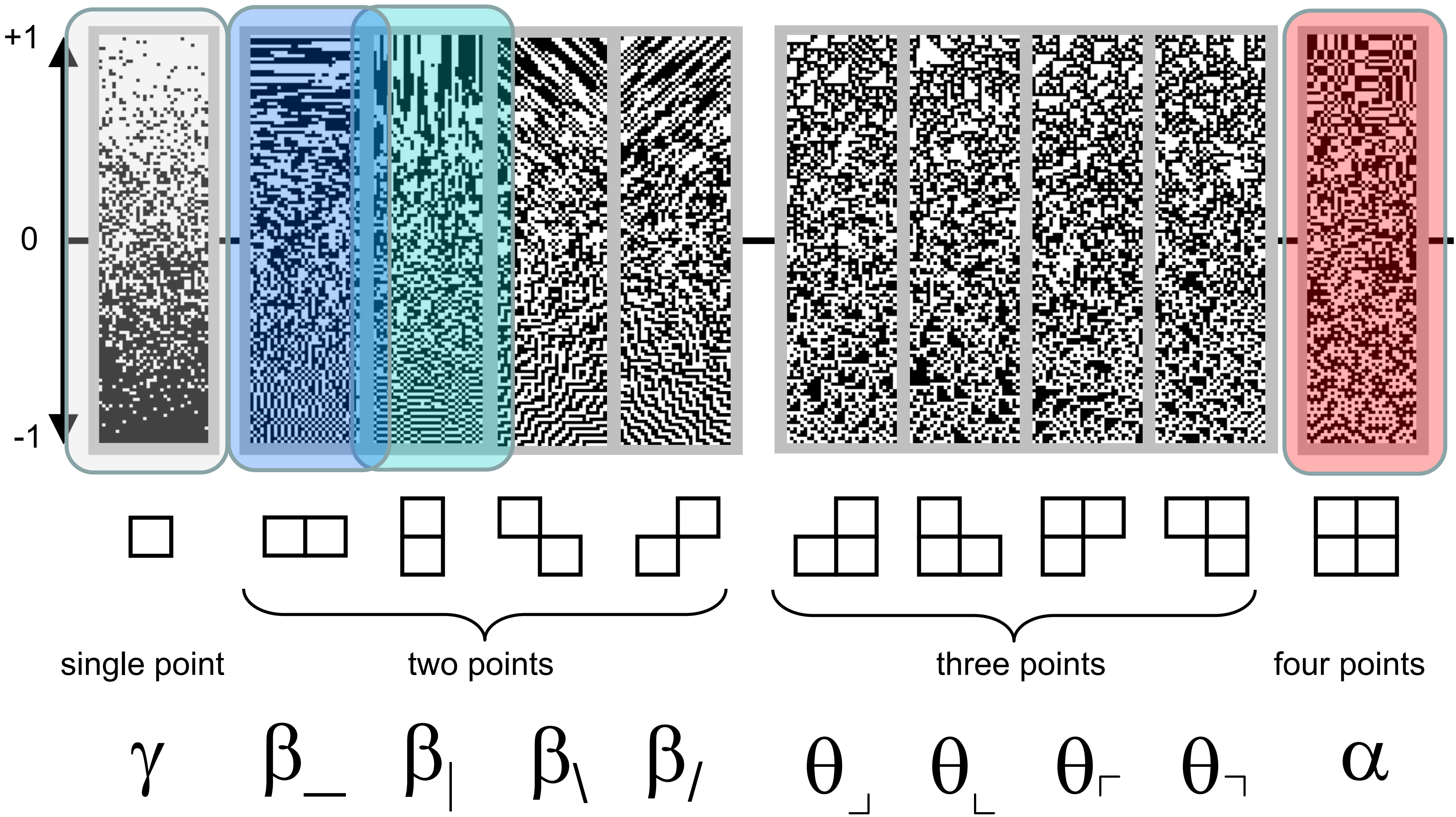
$$\text{Distance to threshold} = \sqrt{\sum_{i,j} Q_{i,j} c_i c_j}$$

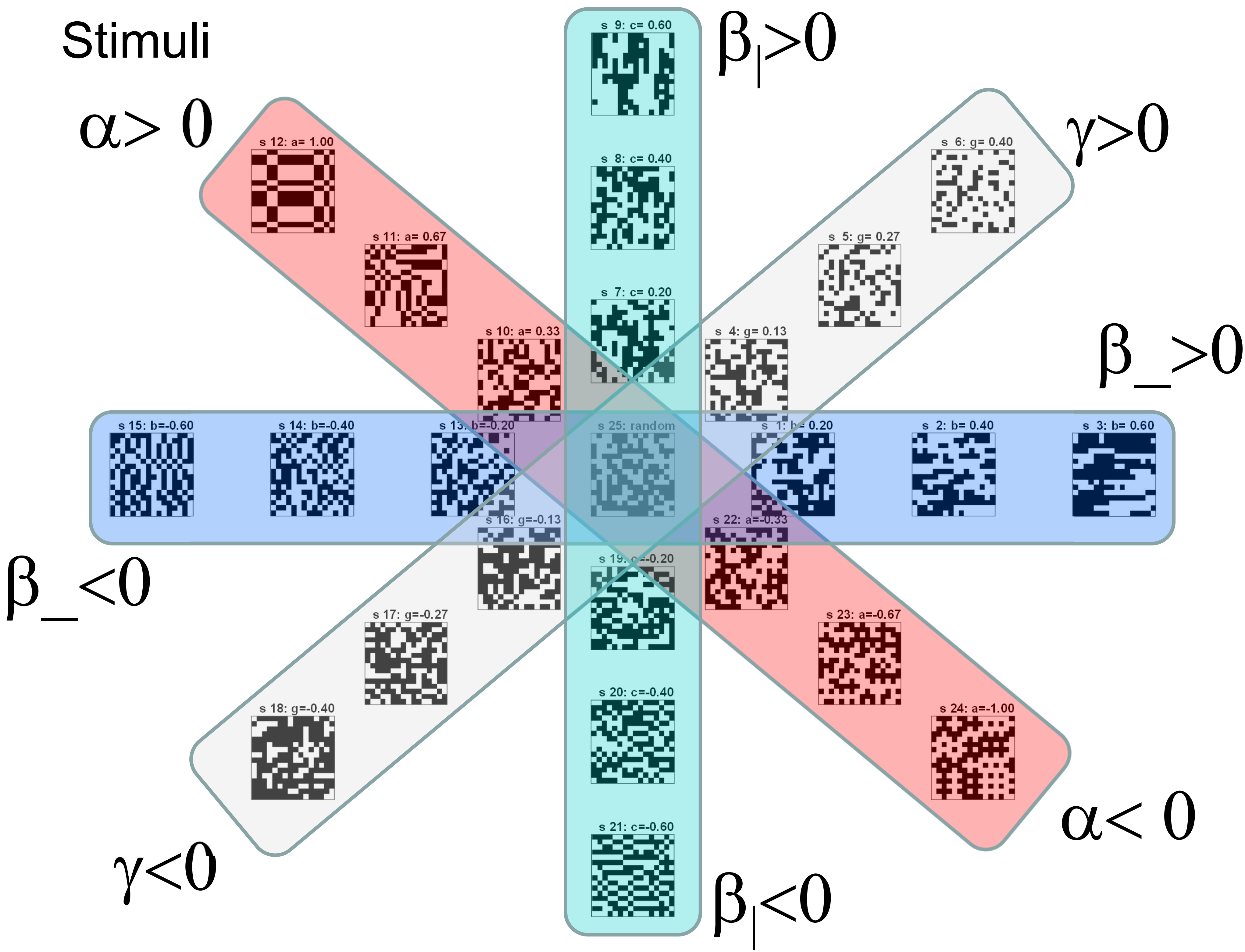
c_i : the coordinates
$Q_{i,j}$: the metric

Victor, Thengone, Rizvi, & Conte, Vision Res 2015

What about suprathreshold similarity?

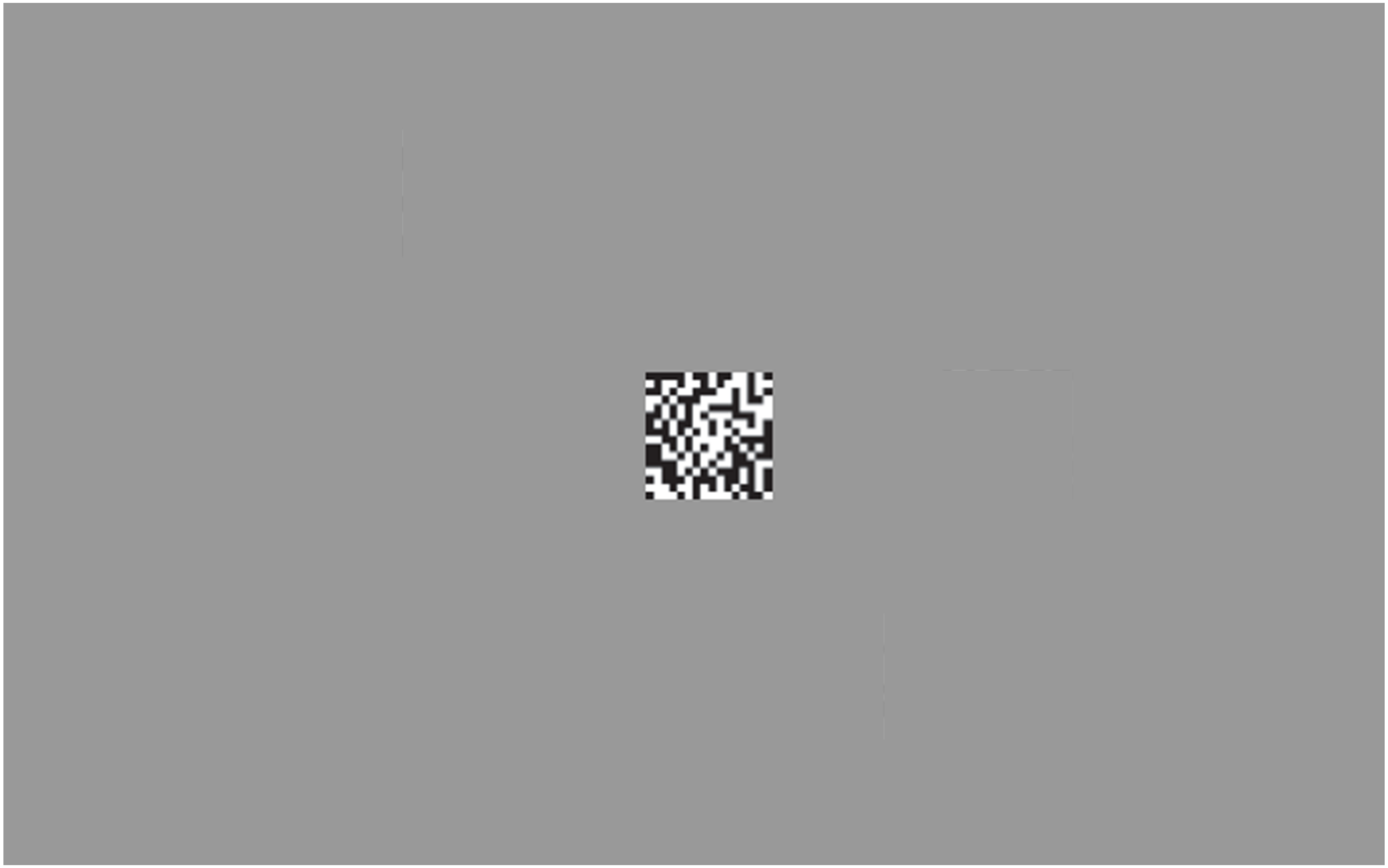
Select four approximately orthogonal coordinates…





Stimuli
α> 0
β|>0
γ>0
β_>0
β_<0
γ<0
β|<0
α< 0
s 9: c= 0.60
s 8: c= 0.40
s 7: c= 0.20
s 12: a= 1.00
s 11: a= 0.67
s 10: a= 0.33
s 6: g= 0.40
s 5: g= 0.27
s 4: g= 0.13
s 15: b=-0.60
s 14: b=-0.40
s 25: random
s 2: b= 0.40
s 3: b= 0.60
s 16: g=-0.13
s 22: a=-0.33
s 17: g=-0.27
s 23: a=-0.67
s 18: g=-0.40
s 24: a=-1.00
s 20: c=-0.40
s 21: c=-0.60

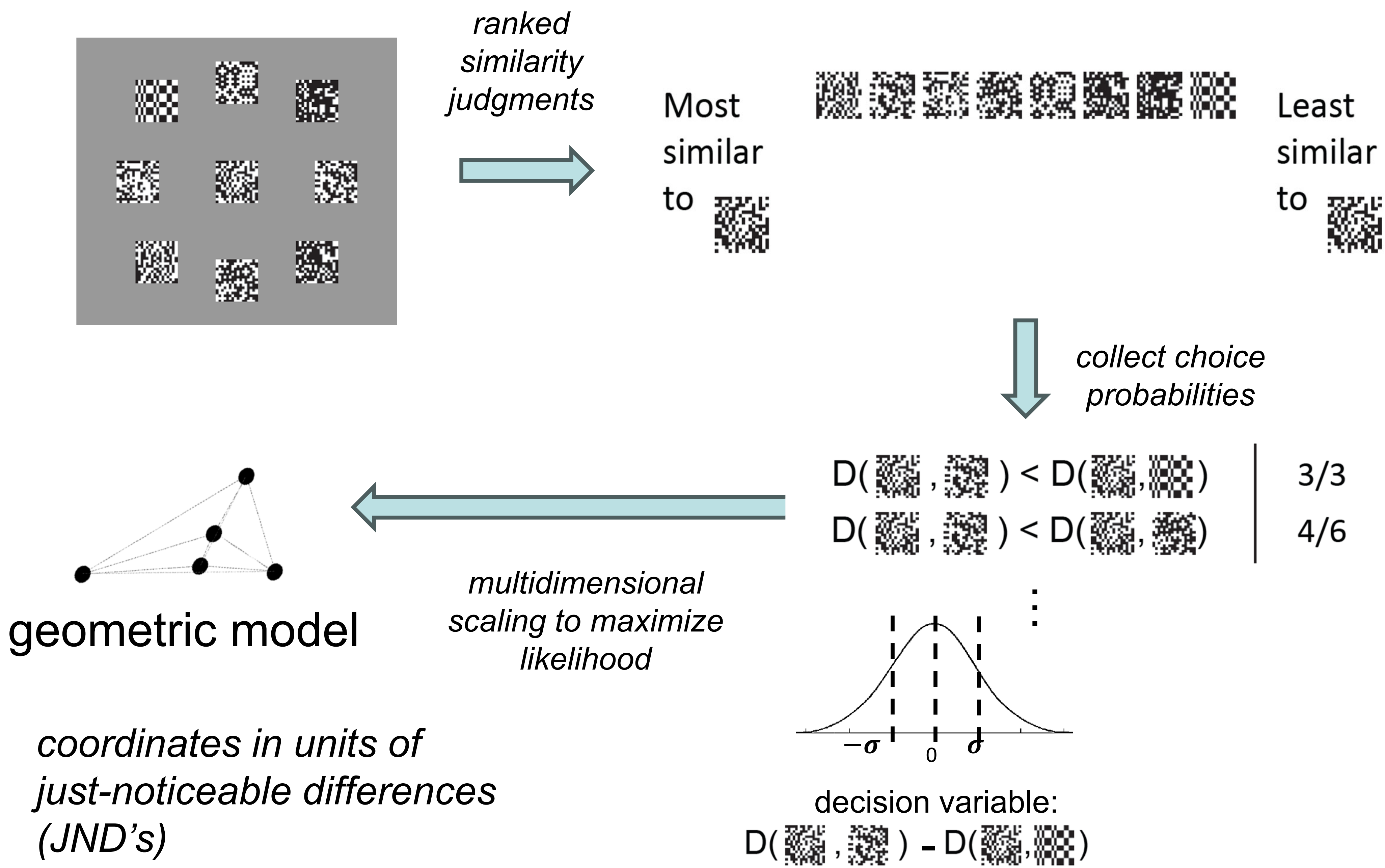
Collecting similarity judgments



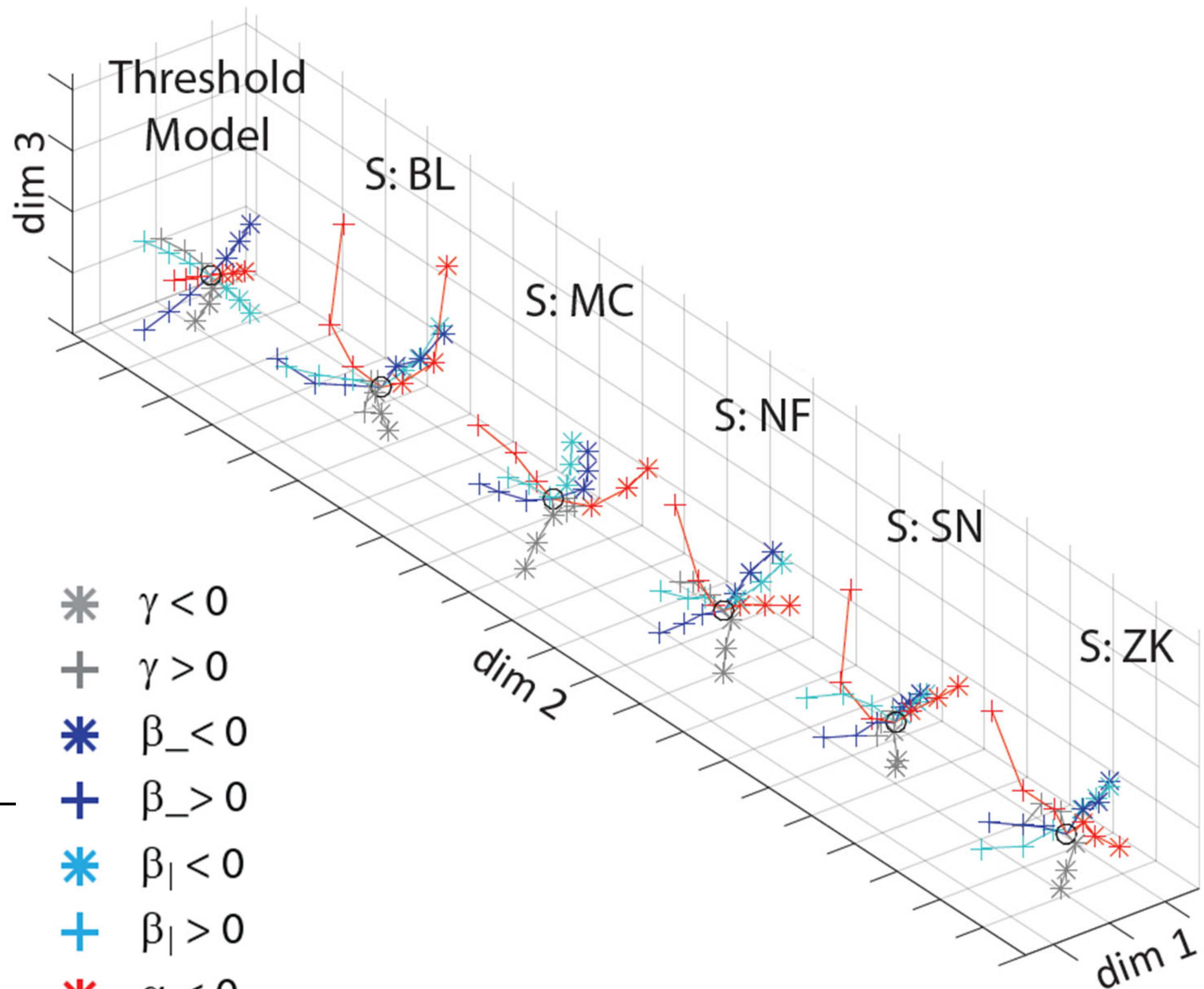
*One trial yields a ranking of 8 similarities to the central reference, i.e., (8*7)/2=28 comparison pairs.*

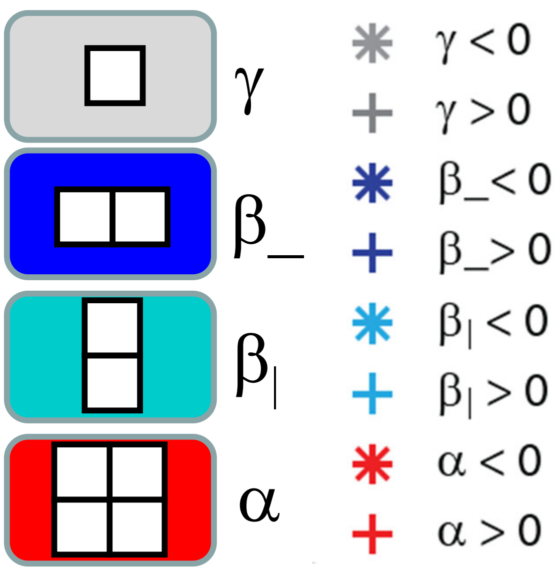
Waraich, S.A., Victor, J.D., (2022), J. Vis. Exp. (181)

Inferring geometry from similarity judgments



Similarity judgments, 5 subjects

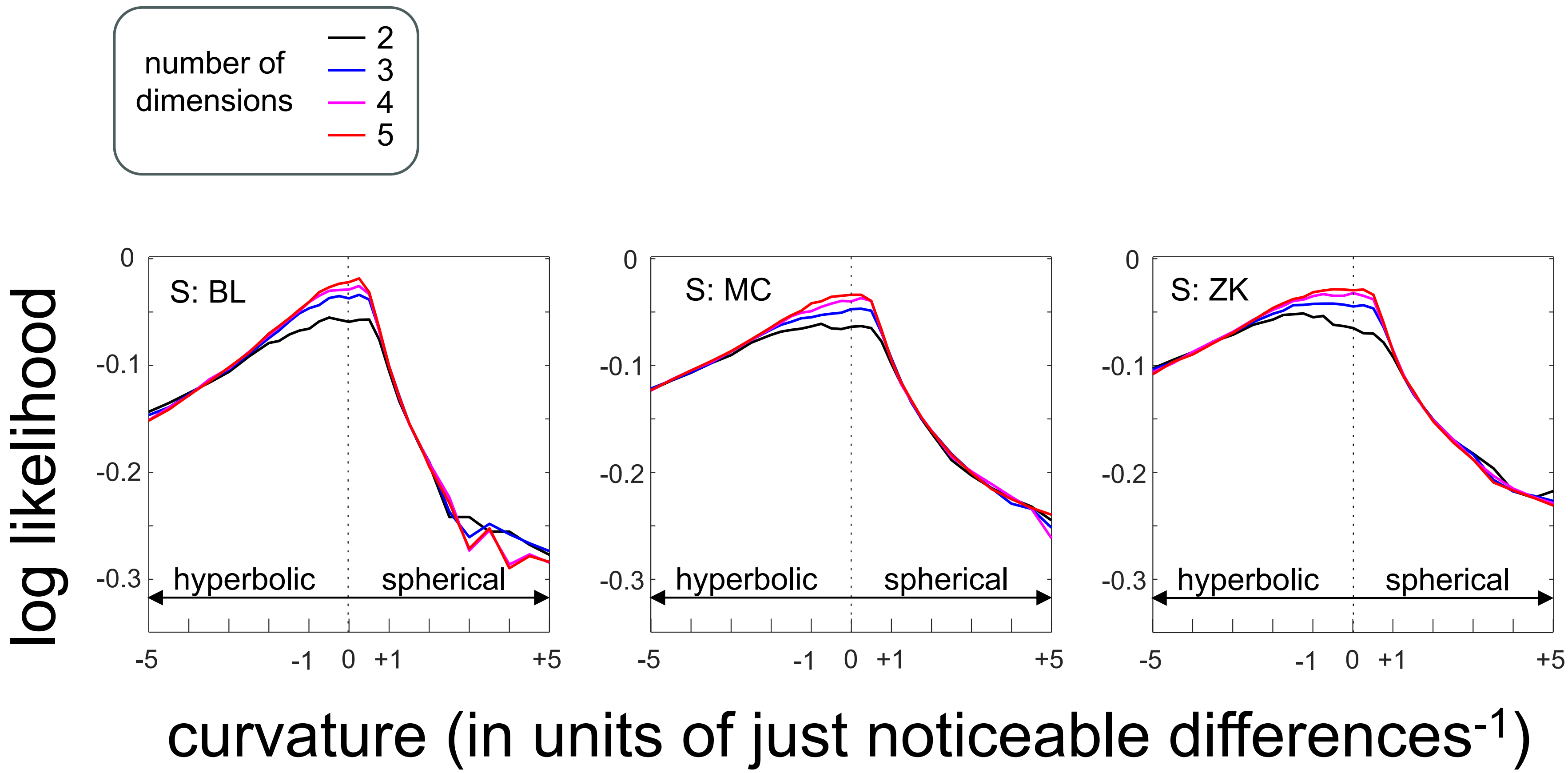




each unit is 2 JND

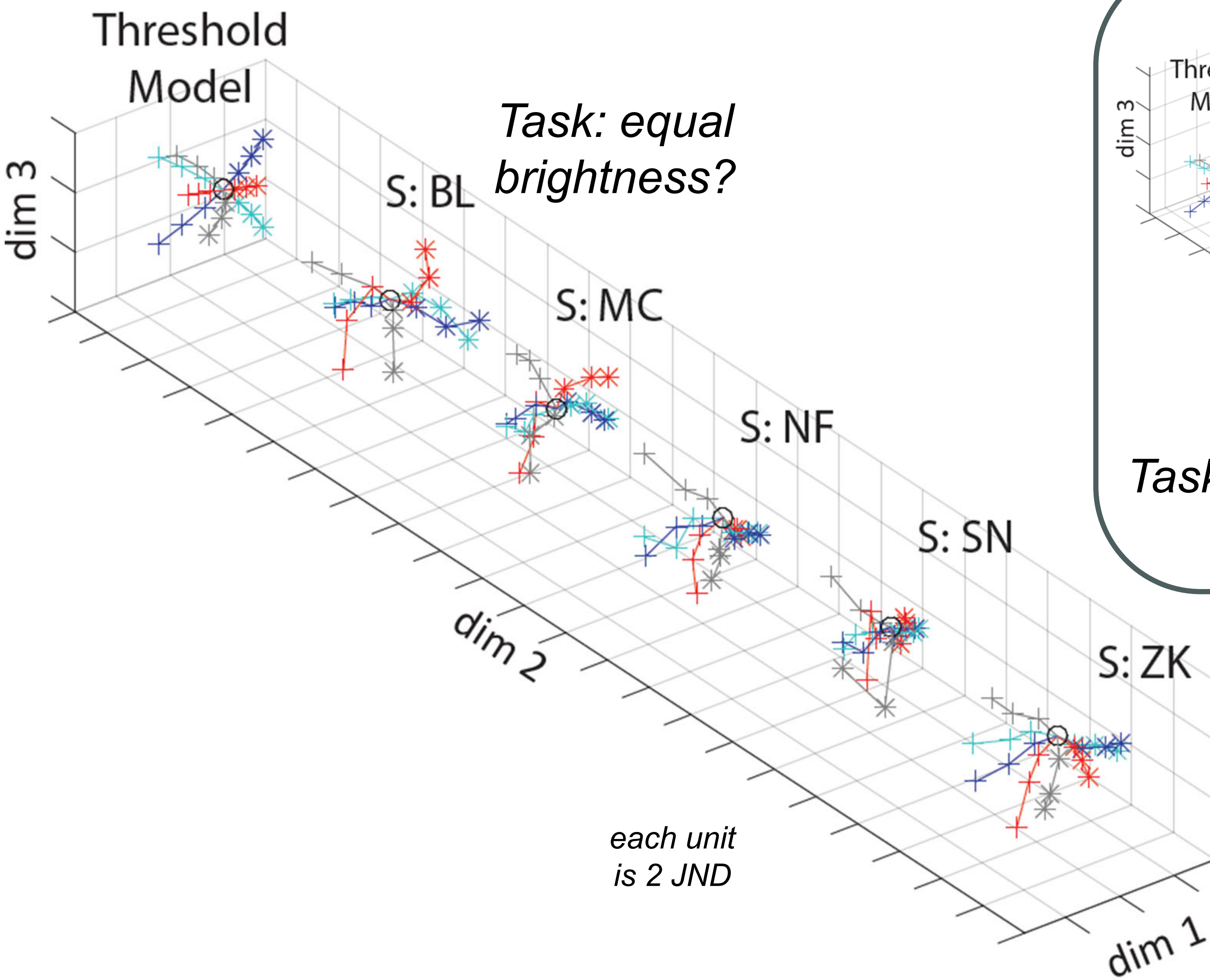
For suprathreshold similarity, lengths and angles are distorted, and axes are bent.

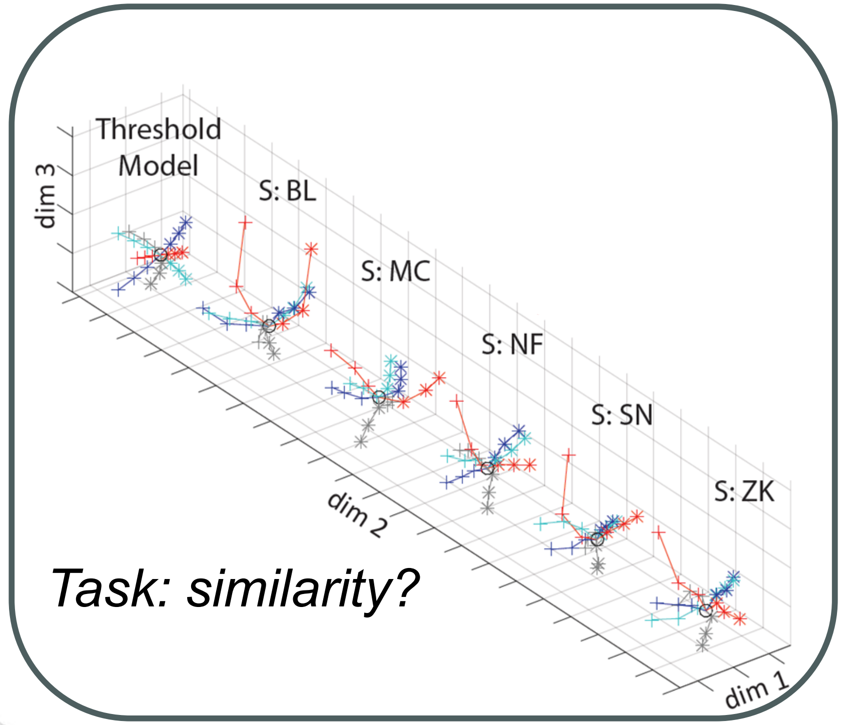
No evidence for global curvature

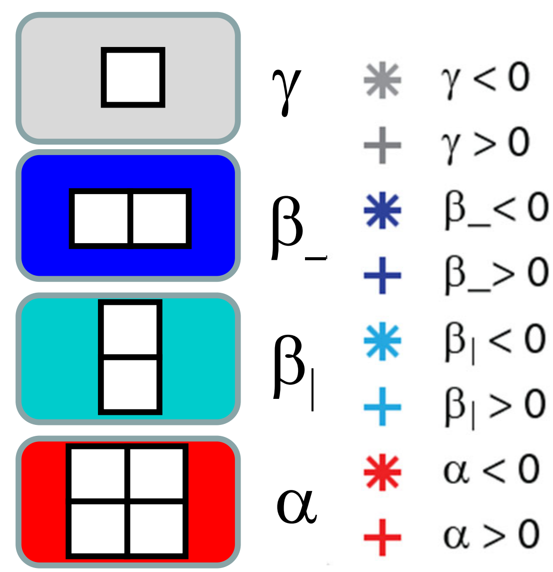


What about just brightness?

Brightness judgments: 5 subjects

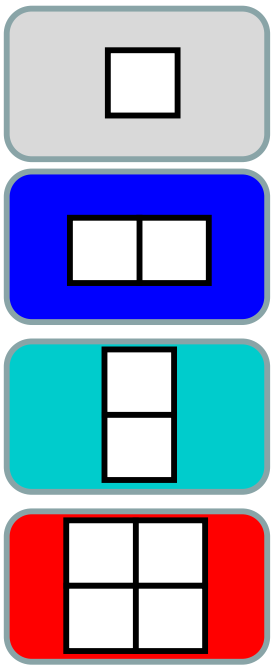






The luminance (γ) axis is longer, and the fourth-order (α) axis is shorter. But not one-dimensional.

Data summary: three tasks

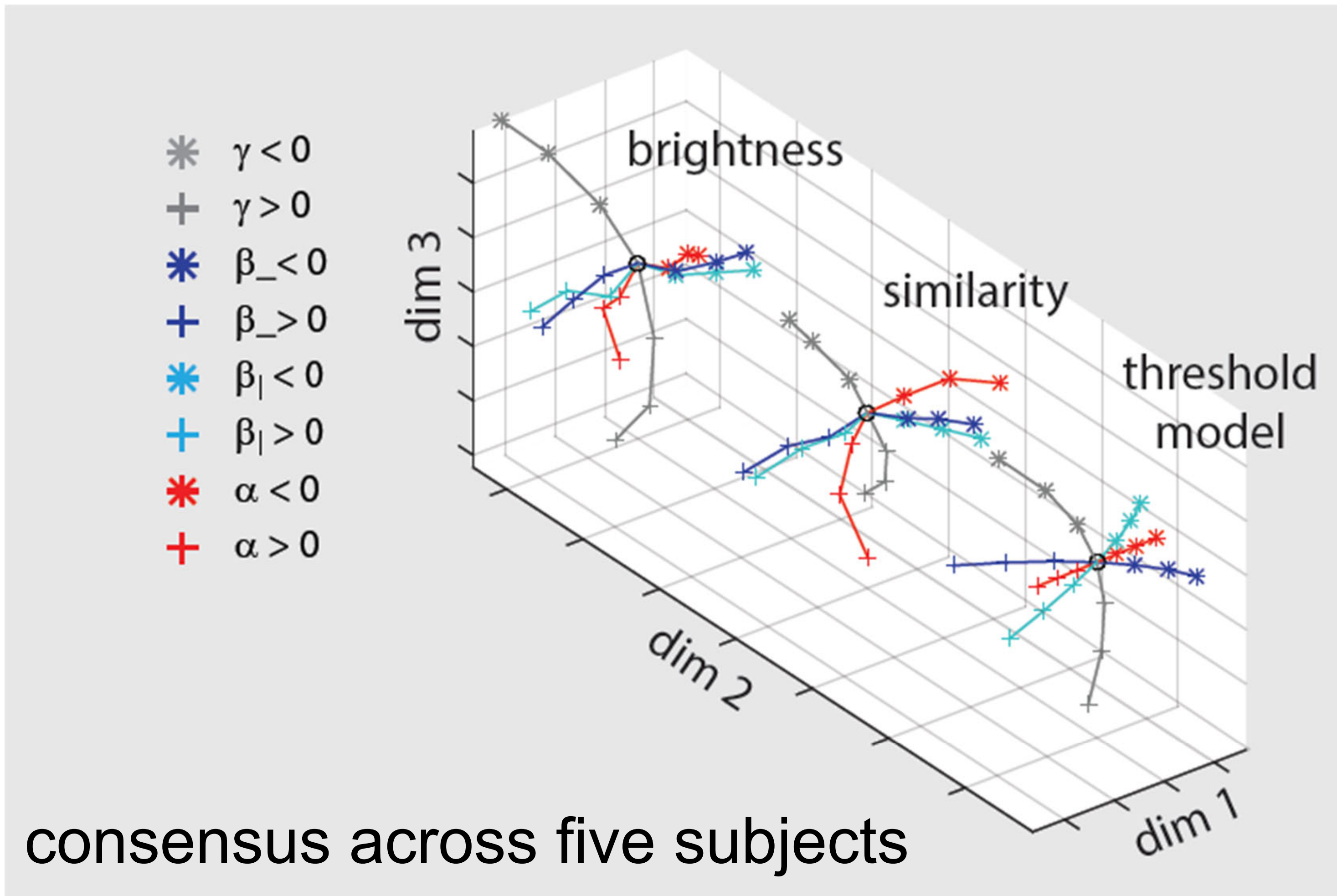


γ

β_

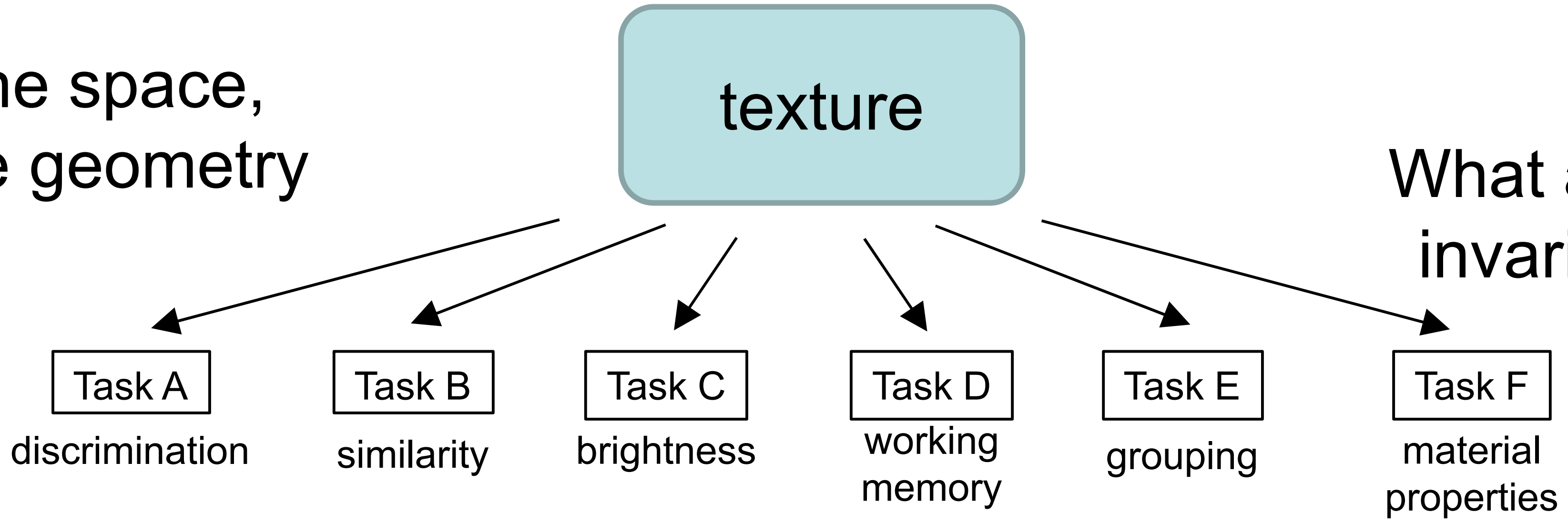
$\beta_{|}$

α



Two viewpoints

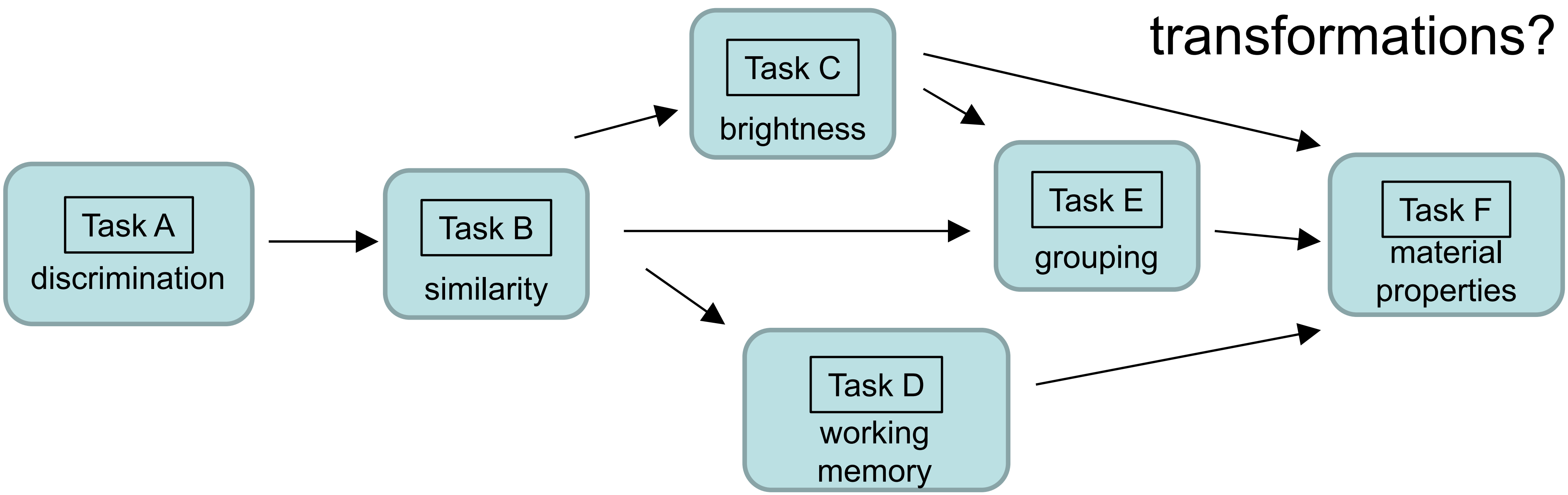
One space,
one geometry



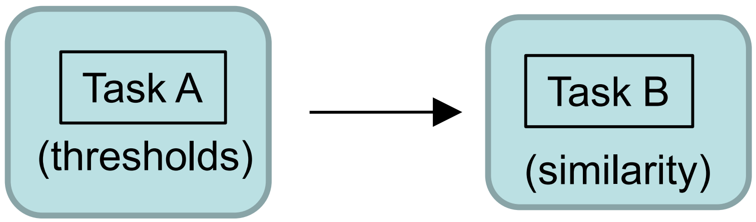
What are the invariants?

Multiple tasks, multiple geometries

What are the transformations?



Geometric transformations correspond to well-recognized neural operations



- distances are disproportionate
- rays are not orthogonal
- trajectories bend at the origin

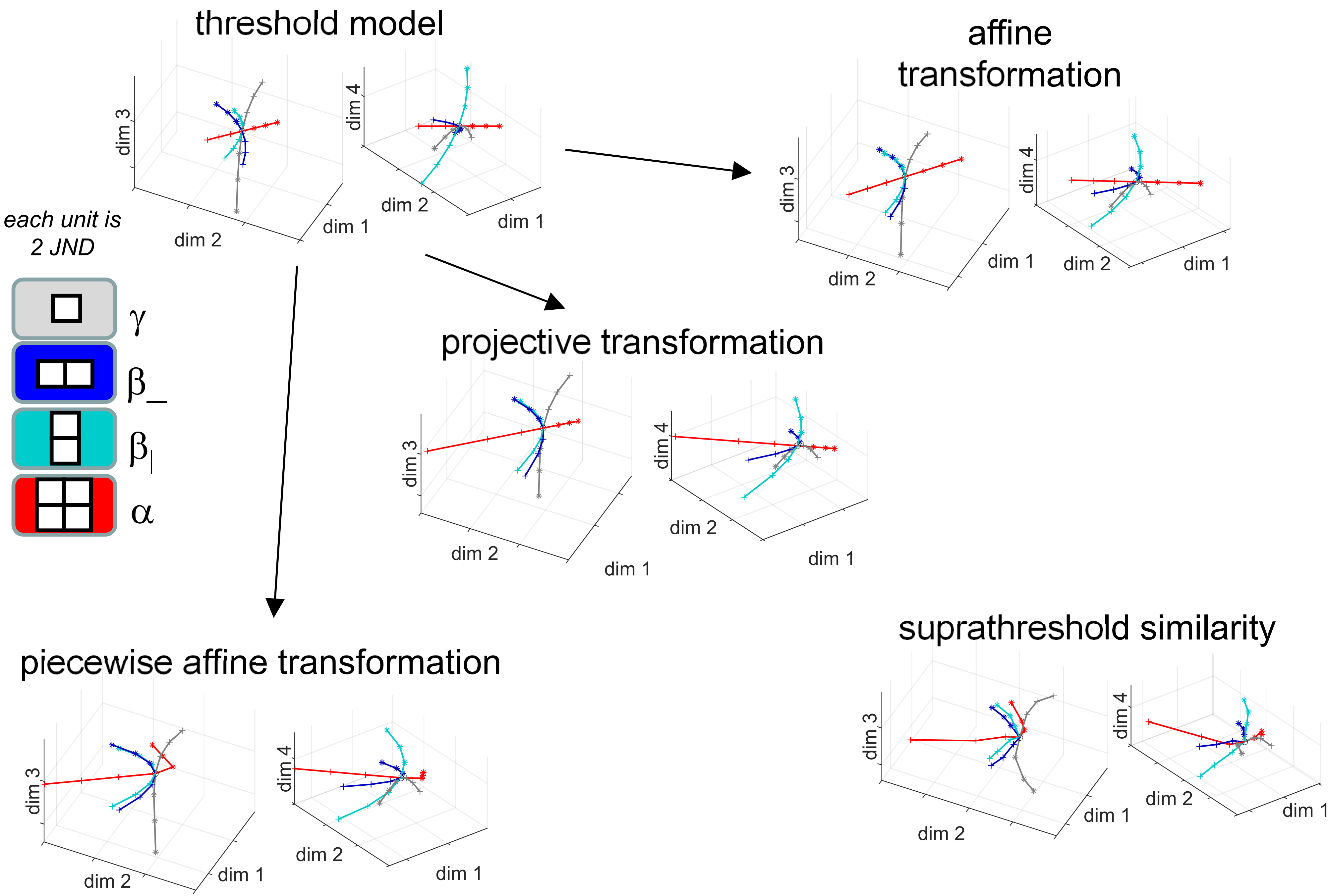
Affine: $y_k = \sum_j T_{kj} x_j$ Gain changes

Projective: $y_k = \dfrac{\sum_j T_{kj} x_j}{h + \sum_j U_j x_j}$ Divisive normalization

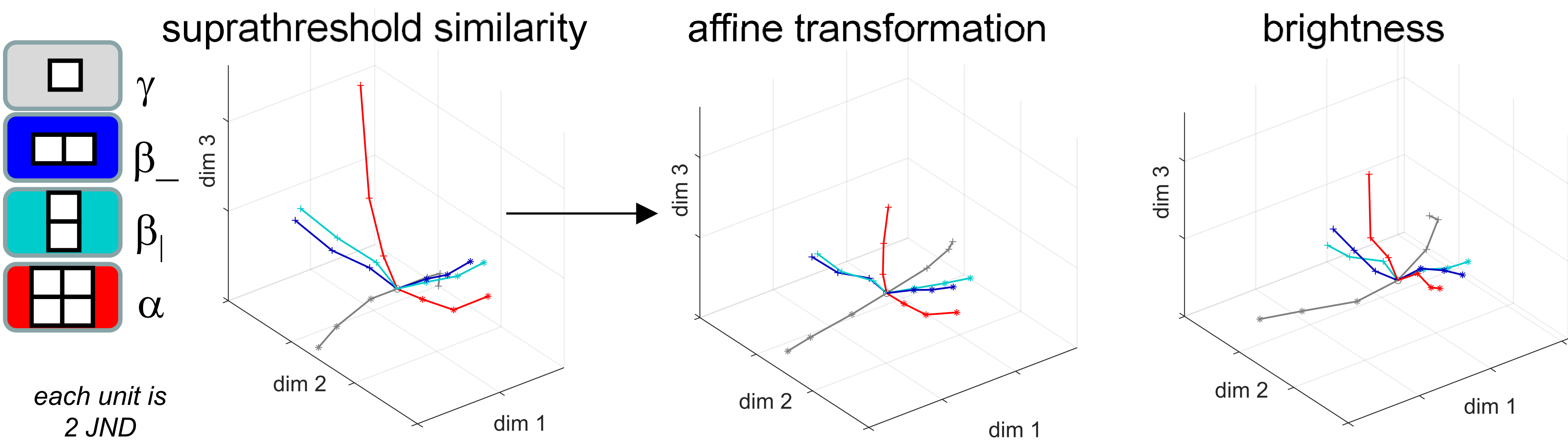
Piecewise linearity $y = a\max(x,0) + b\min(x,0)$ Thresholds

But which are needed to account for the data?

From threshold to similarity



From similarity to brightness



Interim Summary

For the domain of visual textures:

- Threshold and suprathreshold perceptual spaces are both Euclidean
- But their geometry differs greatly
  - Lines and angles are not preserved
  - *The transformation is approximately piecewise affine*
- Brightness comparisons result in gain changes, but not a collapse to one dimension

Pause

A complementary analytic strategy

Dispensing with numeric distances

We assumed that judgments reflect distances (d) – numbers – that can be added and multiplied.

Is d(spider, meatball) > d(flower, cat)+d(hat, tie)?

Is d(alligator, clothespin) > 2 x d(coffee, eggplant) ?

But what if judgments reflect dis-similarities (D) that can be ranked but not added or multiplied?

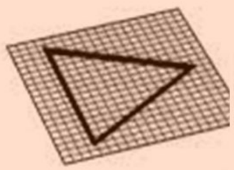
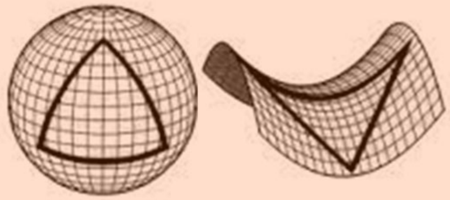
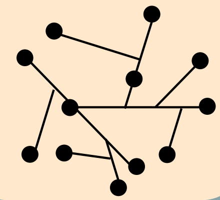
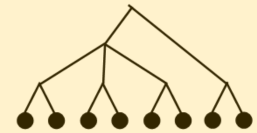
That is, we can ask if D(A,B)>D(A,C), but we can't ask by how much. Can we still characterize the perceptual space?

Formally: we assume that triadic judgments of dis-similarities (D) indicate the rank-order of underlying (but un-observable) distances (d):
$D(A,B)>D(A,C) \leftrightarrow d(A,B)>d(A,C)$.

Weaker than monotonicity … no claim that $D=f(d)$.

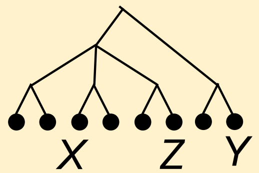
What kinds of inferences can we make about the space that generates d?

We can still test models!

	Euclidean	*Locally Euclidean*	*Minkowski*	*Addtree*	*Ultrametric*
	Everyday life	Distances on curved surfaces	Manhattan distances	Mileage (no loops)	Hierarchy
					
coordinates?	✓	local	✓	✗	✗
continuous?	✓	✓	✓	✓ ✗	✗
Pythagorean?	✓	local	✗	✗	
Isotropic?	✓	✓	✗		
Inequalities					
triangle	✓	✓	✓	✓	✓
“four-point”	✗	✗	✗	✓	✓
ultrametric	✗	✗	✗	✗	✓

Using ordinal relationships of dis-similarities

Ultrametric



Ultrametric inequality:

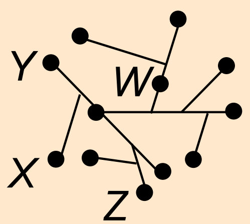
$d(X,Y) \leq \max\{d(X,Z), d(Y,Z)\}$

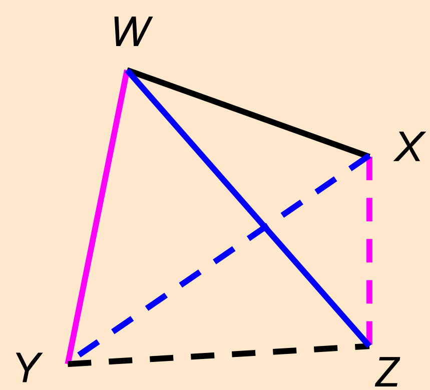
Of $d(X,Y)$, $d(X,Z)$, and $d(Y,Z)$, the largest two are equal.

Ordinal relationships preserved

Of $D(X,Y)$, $D(X,Z)$, and $D(Y,Z)$, the largest two are equal.

Addtree





Ordinal relationships preserved

Four-point condition:

Of $d(W,X)+d(Y,Z)$, $d(W,Y)+d(X,Z)$, and $d(W,Z)+d(X,Y)$, the largest two are equal.

Cannot have $d(W,X)>\max\{d(W,Y),d(W,Z)\}$ and $d(Y,Z)>\max\{d(X,Z),d(X,Y)\}$

Cannot have $D(W,X)>\max\{D(W,Y),D(W,Z)\}$ and $D(Y,Z)>\max\{D(X,Z),D(X,Y)\}$

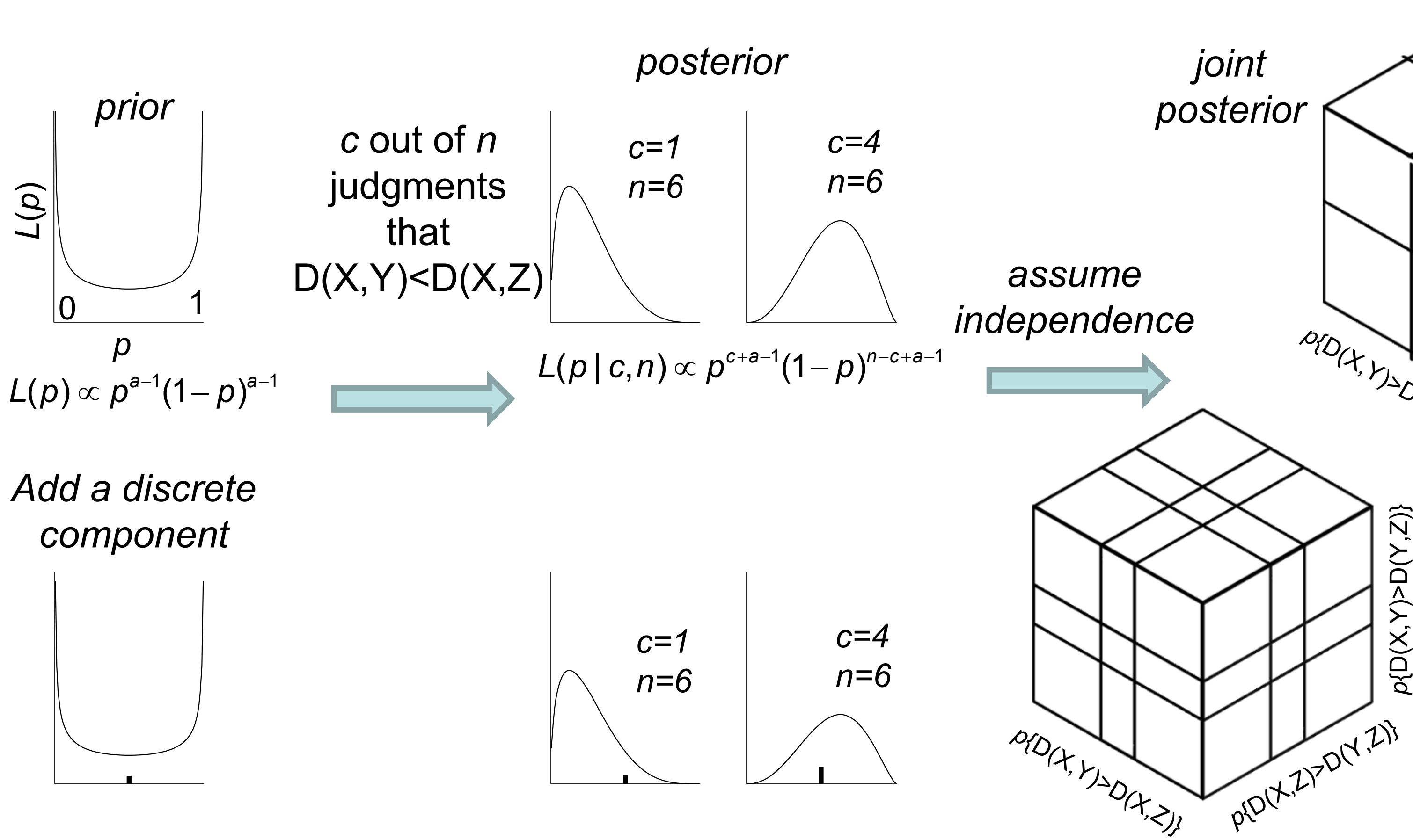
Implementation

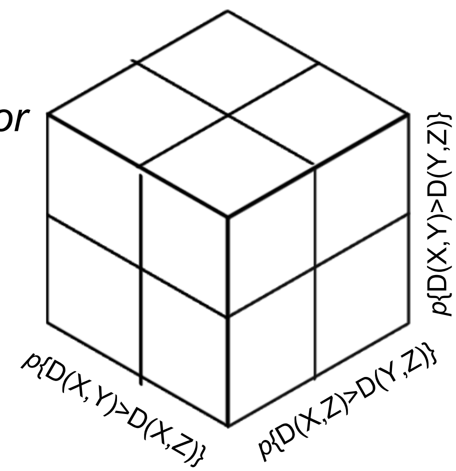
Typically, judgements are uncertain.

So the relationship between D(X,Y) and D(X,Z) is revealed by the probability that a subject will judge D(X,Y)>D(X,Z), i.e., the choice probability *p*{D(X,Y)>D(X,Z)}.

We assume that if *p*{D(X,Y)>D(X,Z)} is $\left\{\begin{matrix} >1/2, \\ =1/2 \\ <1/2, \end{matrix}\right\}$ then $\left\{\begin{matrix} D(X,Y)>D(X,Z) \\ D(X,Y)=D(X,Z) \\ D(X,Y)<D(X,Z) \end{matrix}\right\}$.

We use a Bayesian approach to estimate *p*{D(X,Y)>D(X,Z)} from the choice data:

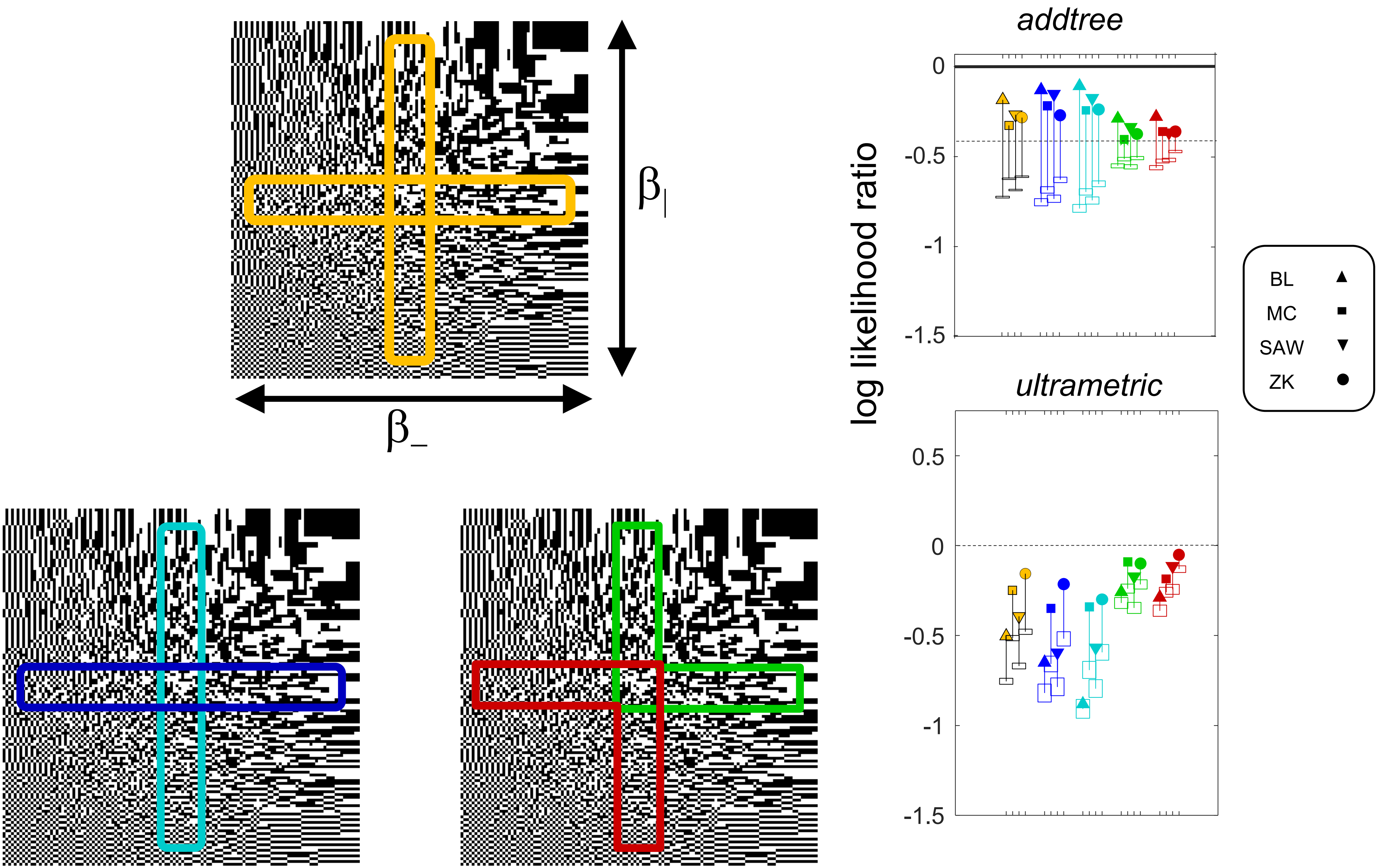




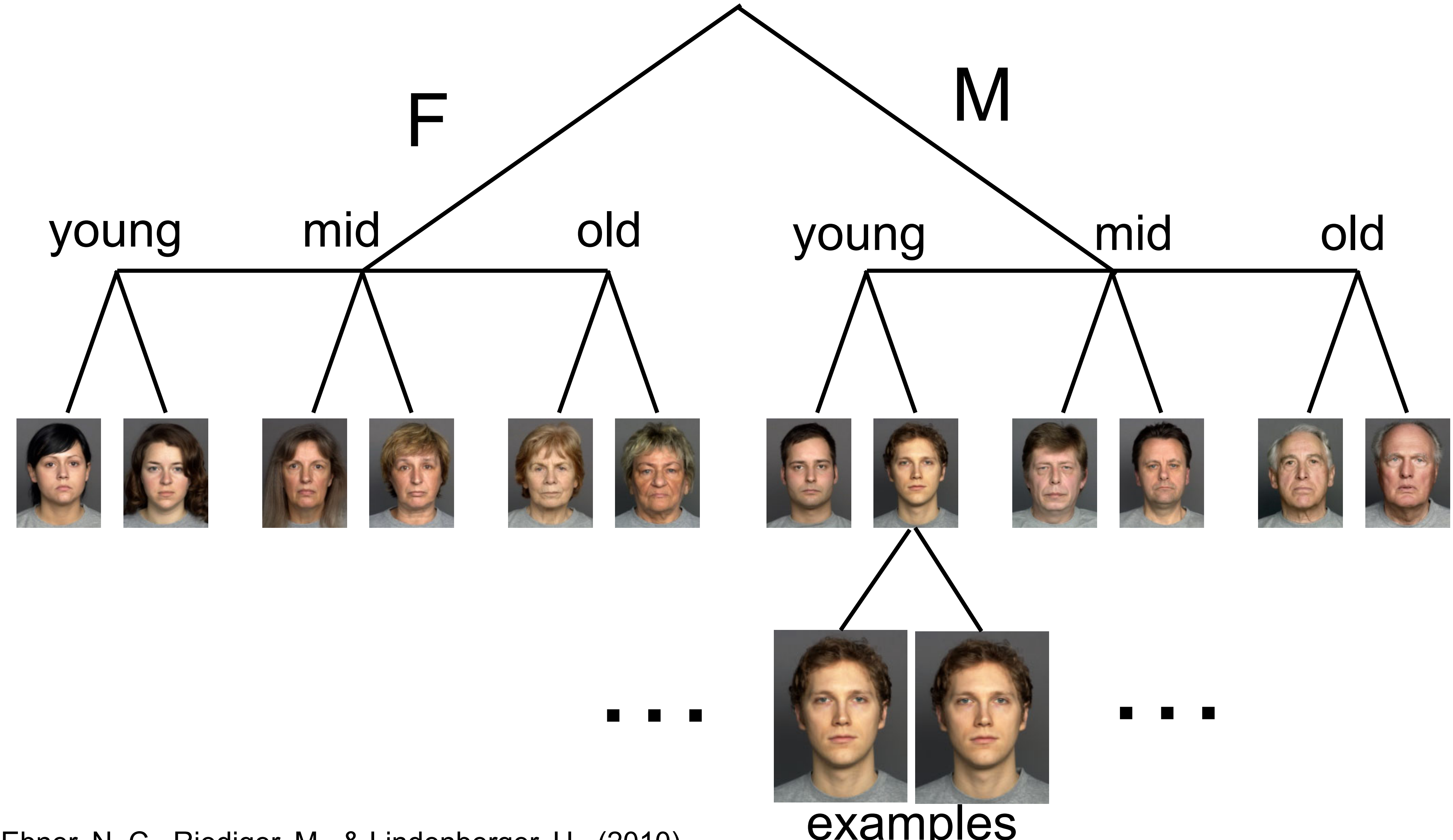
Only some of this volume is consistent with the ultrametric inequality.

What fraction of the joint posterior is in the allowed region?

Addtree test case: Texture

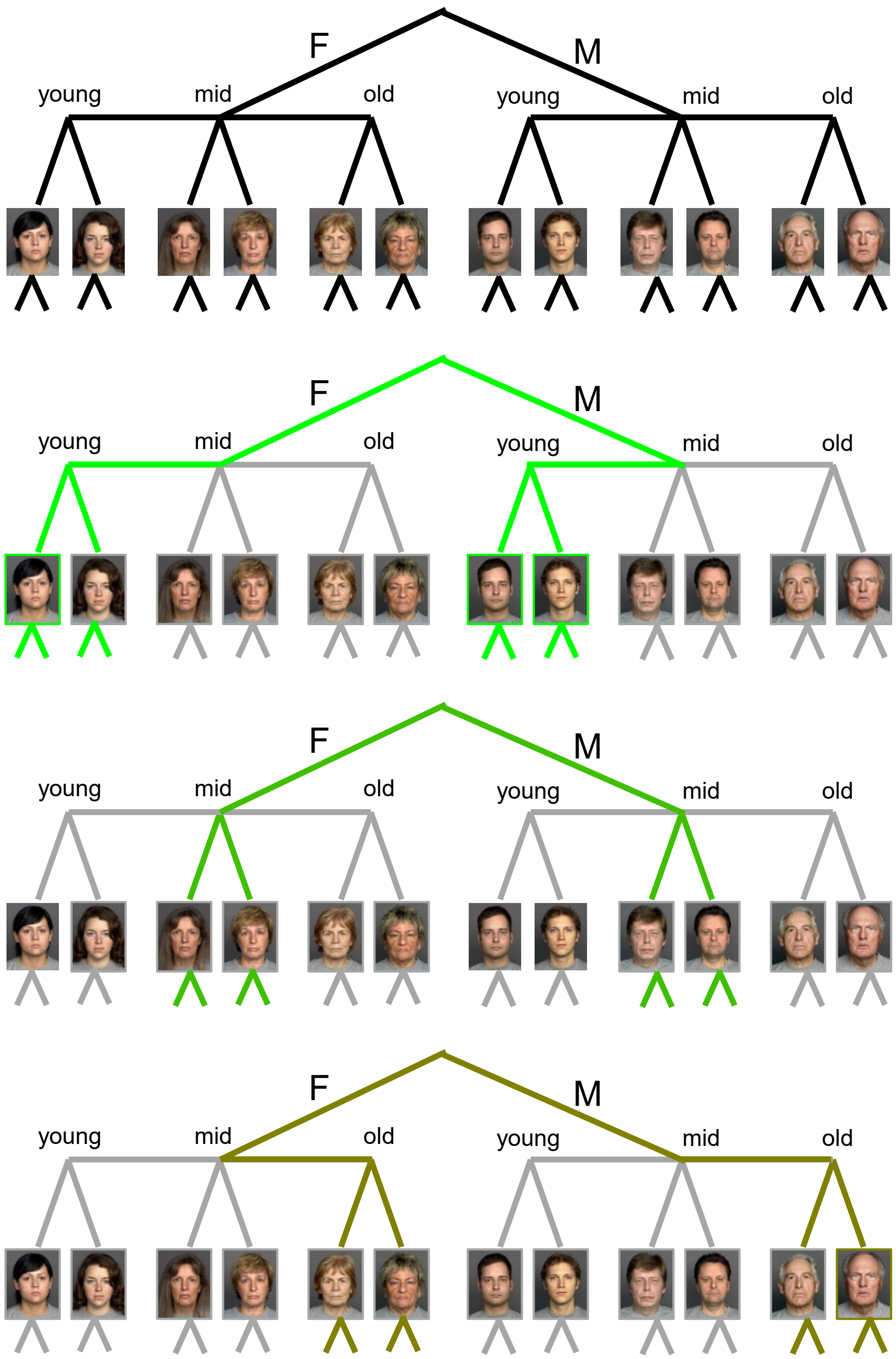


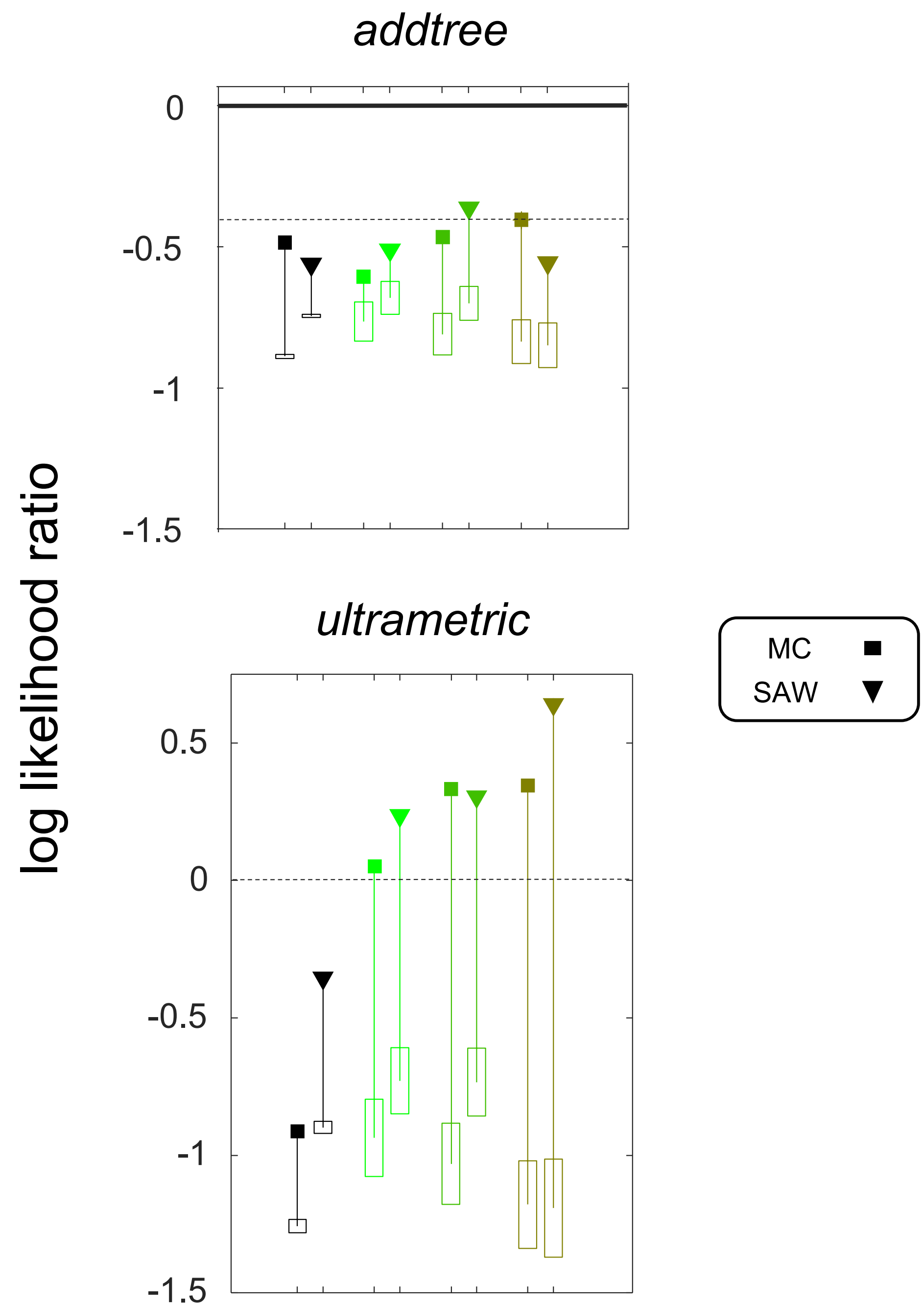
Ultrametric test case: Faces



Ebner, N. C., Riediger, M., & Lindenberger, U., (2010).
FACES. https://faces.mpdl.mpg.de/imeji/

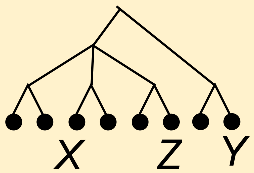
Ultrametric test case: Faces





From here…

Ultrametric



Ultrametric inequality:

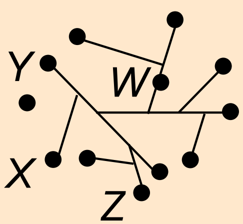
$d(X,Y) \leq \max\{d(X,Z), d(Y,Z)\}$

Of $d(X,Y)$, $d(X,Z)$, and $d(Y,Z)$, the largest two are equal.

Ordinal relationships preserved

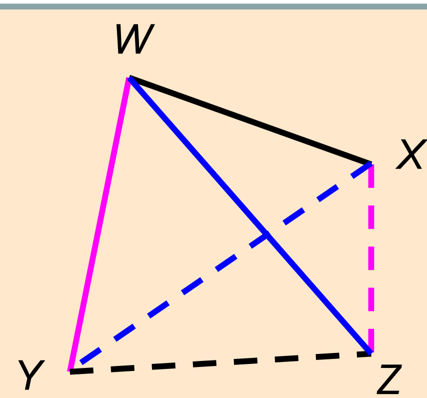
Of $D(X,Y)$, $D(X,Z)$, and $D(Y,Z)$, the largest two are equal.

Addtree



Four-point condition:

Of $d(W,X)+d(Y,Z)$, $d(W,Y)+d(X,Z)$, and $d(W,Z)+d(X,Y)$, the largest two are equal.



Cannot have
$d(W,X) > \max\{d(W,Y), d(W,Z)\}$
and
$d(Y,Z) > \max\{d(X,Z), d(X,Y)\}$

Ordinal relationships preserved

Cannot have
$D(W,X) > \max\{D(W,Y), D(W,Z)\}$
and
$D(Y,Z) > \max\{D(X,Z), D(X,Y)\}$

Can this approach be generalized to constrain cycle structure (or maybe planarity) based on combinations of inequalities of distances?

Summary

Methods

A practical approach to acquiring, and analyzing, similarity judgments that:

- Provides metrical information
- Allows inferences about geometry

Data

Perceptual spaces

- Are high-dimensional but sparsely populated
- Differ by degree of clustering rather than dimensionality or curvature
- Depend on task, in a way that meshes with well-recognized neural calculations

Questions

- How far can we go with rank-order judgments?
- Are we using the right kind of models?

Thank you