Computational Cognitive Science

Lecture 10: Concepts & Categorization

Benjamin Peters

School of Informatics

University of Edinburgh

October 16, 2025

Readings

- Battleday, Peterson, & Griffiths (2020) (link)

Categorization

We have learned about prototype and exemplar models of categorizations. Both try to account for how we categorize novel stimuli based on our past experiences.

- prototype models: the category membership of a new stimulus is calculated using a decision rule based on similarity to the protypes of candidate categories.
- exemplar models: decision rule based on summed similiarity to to all known members of the category.

Today

- Recap and and intuition of prototype and exemplar models.
- How can we test these models with real-world stimuli?

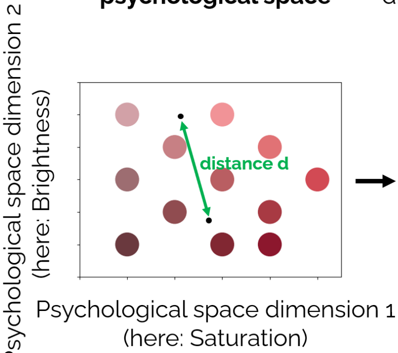
Categorization

Nosofsky (1987)

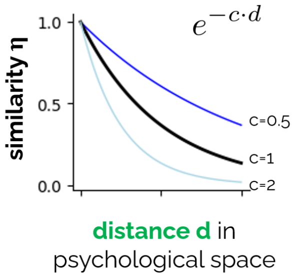
- In each trial, participants saw one of 12 colors and needed to classify the color into categories A or B. They then received feedback.
- After a learning period. The classification responses are then taken as data to evaluate categorization models.

Categorization

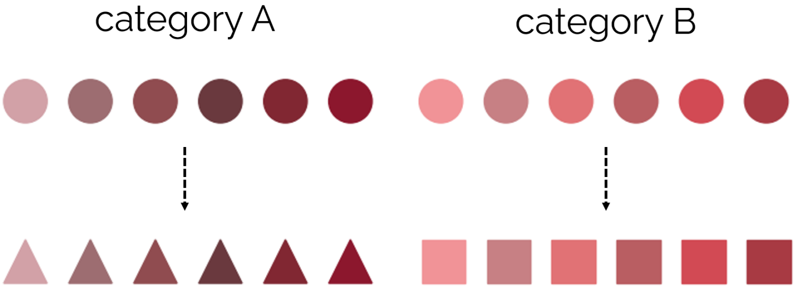
Stimuli are represented in **psychological space**



Similarity η is a function of distance in psychological space

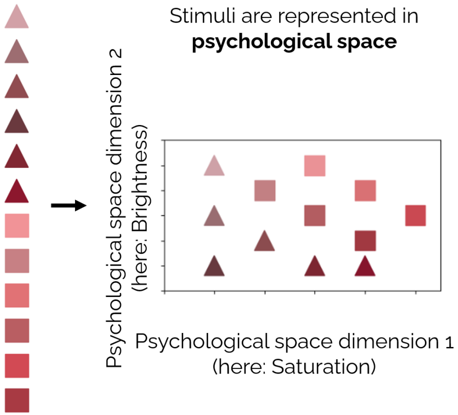


Categorization



We can visualize the category as triangle (A) or square (B). **The participants never saw a triangle or square.**

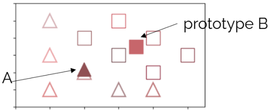
Categorization



Categorization

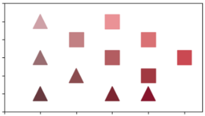
Learned representation

Prototype model



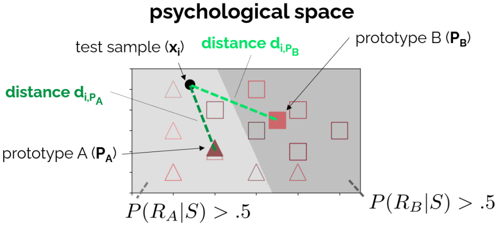
One prototype (centroid of training examples) per category in memory.

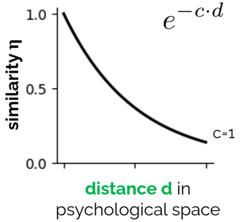
Exemplar model



All training examples are stored in memory.

Categorization: protoype model





similarities

$$\eta_{i,P_A} = \exp[-c\overbrace{\sqrt{(x_{i,1} - P_{A,1})^2 + (x_{i,2} - P_{A,2})^2)}}^{d_{i,P_A}}]$$

similarity of $\mathbf{x_i}$ to prototype $\mathbf{P_A}$

$$\eta_{i,P_B} = \exp[-c\overbrace{\sqrt{(x_{i,1} - P_{B,1})^2 + (x_{i,2} - P_{B,2})^2)}}^{d_{i,P_B}}]$$

similarity of $\mathbf{x_i}$ to prototype $\mathbf{P_B}$

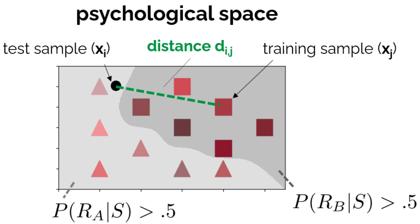
response model

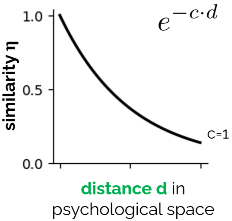
$$P(R_A|S) = \frac{B_A \eta_{i,P_A}}{B_A \eta_{i,P_A} + (1 - B_A)\eta_{i,P_B}}$$

probability of responding A to stimulus S

B_A response bias

Categorization: exemplar model





similarities

$$\eta_{i,j} = \exp[-c\underbrace{\sqrt{(x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2)}]}_{d_{i,j}}$$

similarity of $\mathbf{x}_i$ to training sample j

response model

$$P(R_A|S) = \frac{B_A \sum_{j \in A} \eta_{i,j}}{B_A \sum_{j \in A} \eta_{i,j} + (1 - B_A) \sum_{j \in B} \eta_{i,j}}$$

probability of responding A to stimulus S

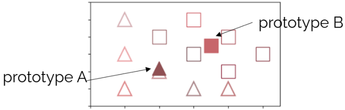
B_A response bias

Categorization: protype vs exemplar model

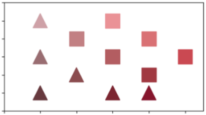
Prototype model

Exemplar model

Learned representation



One prototype (centroid of training examples) per category in memory.

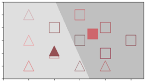


All training examples are stored in memory.

Categorization

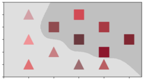
$$\eta_{i,P_A} = \exp[-c\sqrt{(x_{i,1} - P_{A,1})^2 + (x_{i,2} - P_{A,2})^2})]$$

$$P(R_A|S) = \frac{B_A \eta_{i,P_A}}{B_A \eta_{i,P_A} + (1 - B_A)\eta_{i,P_B}}$$



$$\eta_{i,j} = \exp[-c\sqrt{(x_{i,1} - x_{j,1})^2 + (x_{i,2} - x_{j,2})^2})]$$

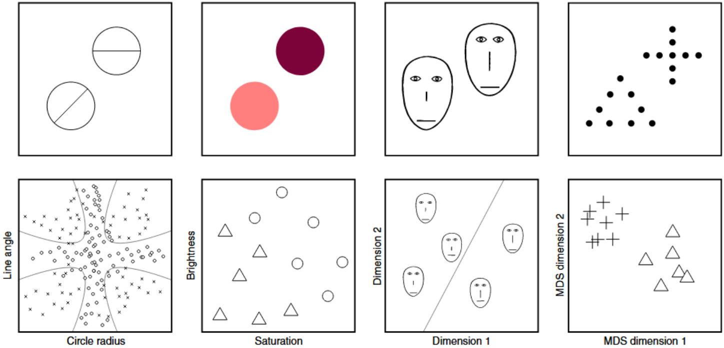
$$P(R_A|S) = \frac{B_A \sum_{j \in A} \eta_{i,j}}{B_A \sum_{j \in A} \eta_{i,j} + (1 - B_A) \sum_{j \in B} \eta_{i,j}}$$



Categorization: interim summary

Compared to the prototype model, the exemplar model can express much more complex decision boundaries.

Categorization



Studies have used highly controlled stimuli, which incurs very simplified stimuli. They showed much support for exemplar models.

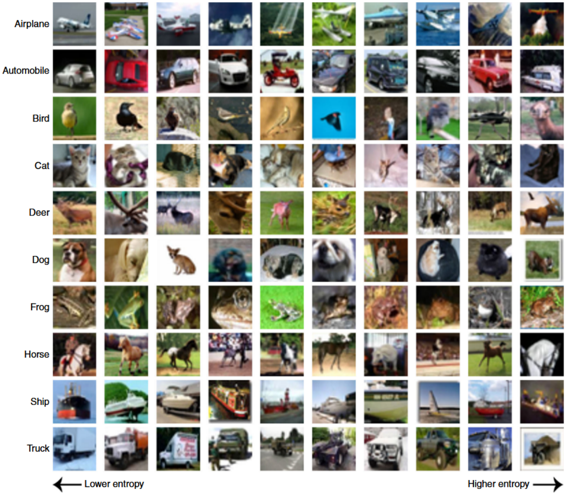
Categorization

What is not known is whether exemplar models account for how we categorizate real-world stimuli

- Possibly, categorization experiments on simplified stimuli do not capture how we categorize stimuli of our every-day environment (ecological validity).
- Real-world stimuli are high-dimensional as opposed to controlled stimuli. Dimensionality matters for computing similarity.

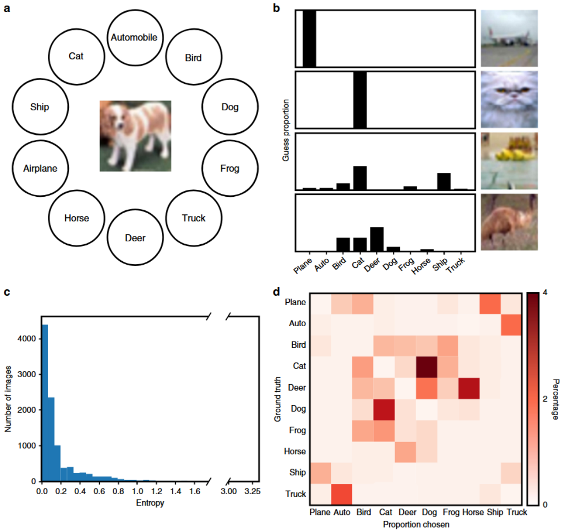
Are the exemplar models still better for real-world stimuli?

Categorization of real-world stimuli



The authors test categorization on images of a machine learning dataset (CIFAR10), which contains 10 categories of objects.

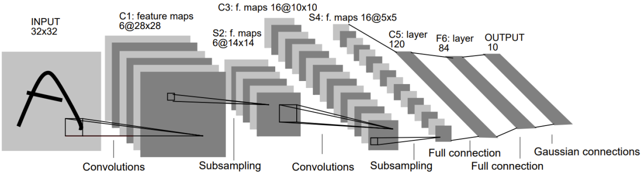
Human behaviour

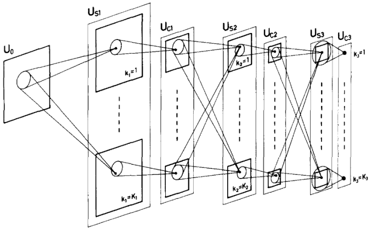


10, 000 images, 50 repetitions each (across 2570 participants).

Stimulus space

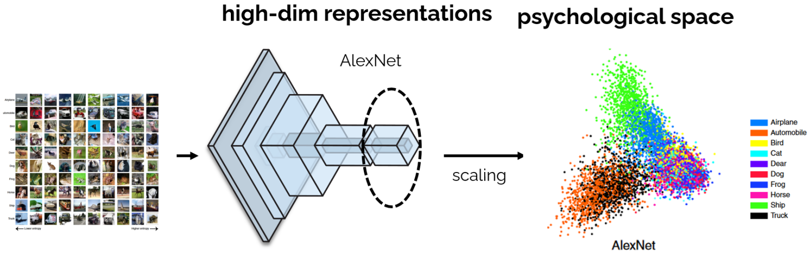
The authors chose the latent dimensions of convolutional neural networks (CCNs), pre-trained on image classification.





Top: LeNet (LeCun et al., 1998). Bottom: Neocognitron (Fukushima, 1980)

Psychological space



2D projection of the high-dimensional psychological space (for visualization)

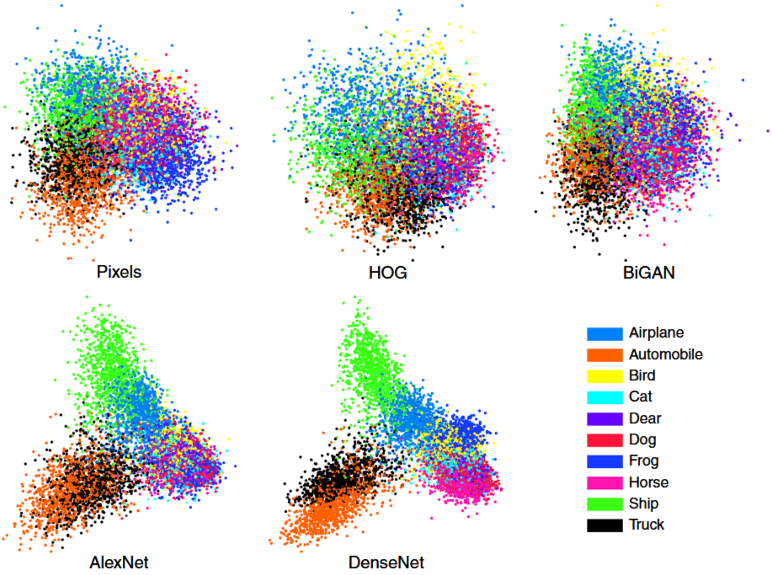
A K-dimensional vector (AlexNet: $K = 4096$) of the last layer of a CNN trained on ImageNet serves as the stimulus representation. After optional scaling (see next slide), this is our psychological space and we can compute psychological distance and transform them into similarities (as before).

Scaling

Each image is represented by a K-dimensional vector (AlexNet: K=4096). Some dimensions are more relevant than others. The authors consider transformations of the stimulus space before computing the Eucliden distance as psychological distance.

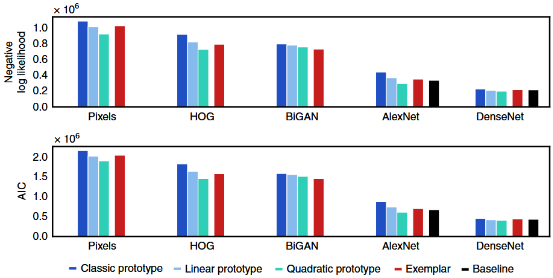
- classic: no transformation
- linear: stretch and rotate such that the variations across images are equal in all directions (whitening)
- quadratic: as linear, but transform for each category.

Other stimulus spaces



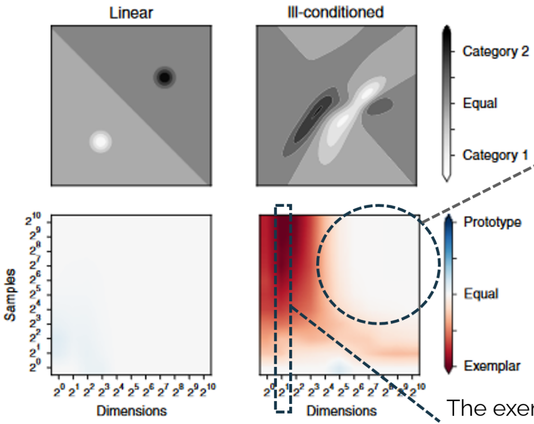
Evaluation

Now compare the predicted categorization from classic prototype and exemplar model for different stimulus representations with human behaviour via log-likelihood and AIC.



- No substantial benefit for exemplar over prototype models.
- Large effect of the stimulus representations.

Exemplar and prototype models in high dimensions



The advantage of the exemplar model for complex problems disappears in high dimensions.

color: which model is better?

The exemplar model outperforms the prototype model in 2D.

Summary

- We revisted exemplar and prototype models and introduced the intution of decision boundaries in psychological space.
- Models have so far been tested in abstracted, low-dimensional stimulus spaces. Here, exemplar models often outperform prototype models.
- In high-dimensional, real-world stimulus spaces, both models perform comparably.

Studying real-world complexity and behaviour:

- A fundamental challenge of the brain is to deal with the complexity of the world. This challenge shapes the mechanisms of the mind.
- Principles gleaned from abstracted tasks may not scale to real-world settings.

References

- Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2020). Capturing human categorization of natural images by combining deep networks and cognitive models. **Nature Communications, 11**(1), 5418.

- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. **Biological Cybernetics, 36**(4), 193–202.

- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. **Neural Computation, 1**(4), 541–551.

- Nosofsky, R. M. (1987). Attention and learning processes in the identification and categorization of integral stimuli. **Journal of Experimental Psychology: Learning, Memory, and Cognition, 13**(1), 87–108.