292—B16

Models of object categorization reflect multiple categorization strategies

Robert J. Peters (1), Fabrizio Gabbiani (2), and Christof Koch (1)

(1) Computation and Neural Systems, Caltech, Pasadena, CA 91125

(2) Baylor College of Medicine, Houston, TX 77030

1. Introduction

1.1 Visual categories

This work is all about ***categorization***. Categorization can be seen as one of the fundamental aspects of higher-order brain function. It is about attaching meaning(s) or label(s) to a piece of sensory input. Members of a category are referred to as ***exemplars*** of that category.

Often we associate many categories with a given item. Consider all the categories that can apply to the image to the left:

Oak tree
deciduous tree
tree
plant
green thing
source of firewood
food (for a termite)

Of all the possible types of categories that could be attached to an object, the type we consider here is ***visual categories***–the kind in which the exemplars share some set of visual features. Further, we focus specifically on ***subordinate-level*** categories, in which the exemplars not only share similar visual features, but are constrained to some particular arrangements of those features.

For exemplar, male and female human faces might belong to the same basic-level visual category (since all faces share similar features), but belong to different subordinate-level categories, since male and female faces have different arrangements of those features.

On the other hand, while we might associate certain faces with the categories "Democrat" or "Republican", these are not visual categories, since it is (arguably) not possible to distinguish these categories on the basis of visual features of the face.

1.2 Important Questions

In order to understand the mechanisms underlying subordinate-level visual object categorization in humans, we (and others) would like to know the answers to some Important Questions:

***What** is the representation?*
(edges, colors, distances?)

***How** is the representation used?*
(how do single neurons compute categories?)

***Where** is the representation?*
(what brain regions are used?)

The present study addresses only the second question, of how the representation is used, and leaves aside the other two.

References

[Maddox1993] Maddox WT & Ashby FG. *Percept & Psychophys* 53(1):49-70 (1993).

[Nosofsky1986] Nosofsky RM. *J Exp Psych: General* 115(1):39-57 (1986).

[Peters2000] Peters RJ et al. *J Cog Neurosci* Suppl:72-72 (2000).

[Reed1972] Reed SK. *Cognitive Psychology* 3:382-407 (1972).

Acknowledgements

This work was supported by grants from NSF-ERC and HHMI.
I thank my wife Kasi for her constant support.

2. Psychophysics

2.1 Stimuli in Feature Space

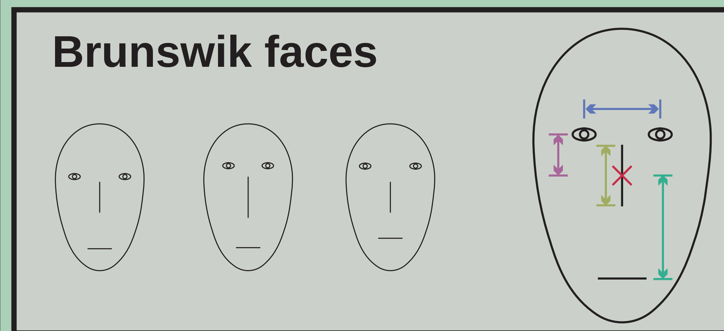
Objects are defined (by us, the experimenters) to exist in a mathematical ***feature space***, in which each variable parameter of the objects is represented by one dimension.

The objects' parameters (i.e., the dimensions of the feature space) were never explicitly revealed to our human subjects.

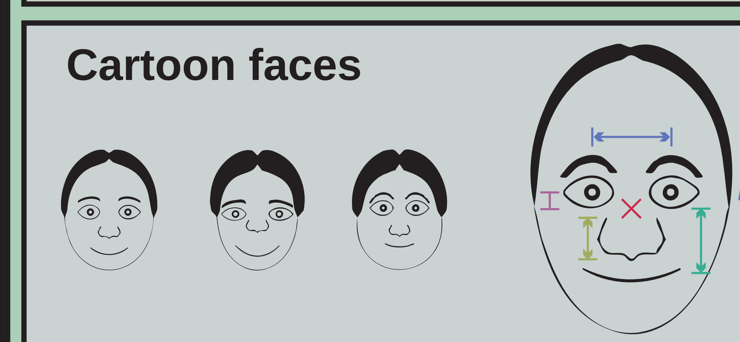
We define mathematical models of categorization that operate in this feature space.

Although the Important Question "what is the representation?" asks what exactly is the feature space and how is it constructed by the brain, we do not address this directly here. Nevertheless, previous work [Peters2000] using multidimensional scaling suggests that, at least for the type of stimuli used here, subjects' natural psychological feature space is very similar to the physical feature space in which we define the objects. Thus, for our present purposes, we simply use the physical feature space.

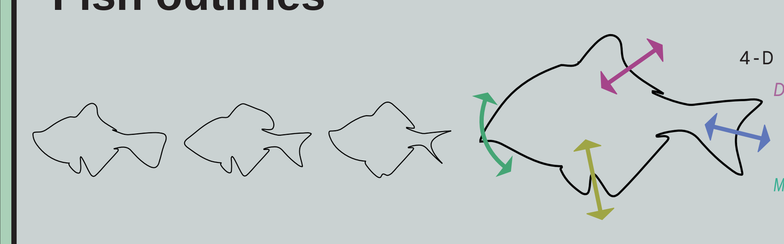
Brunswik faces



Cartoon faces

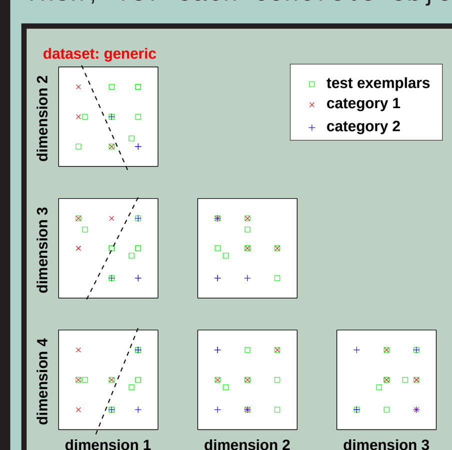


Fish outlines



2.2 Categories in Feature Space

The categories were first defined in a generic 4-D feature space, and along each dimension the features were quantized to three values, so that the entire set of objects occupied a 3x3x3x3 lattice. Each set of objects contained an equal number of ***training exemplars*** for each of two categories, as well as an additional number of ***test exemplars***. The arrangement of the training and test exemplars in 4-D feature space is depicted in a set of plots (below left), which show projections of the exemplars onto different 2-D cross-sections of feature space. Then, for each concrete object type, we took the four features associated with that type and plugged them into the generic category arrangement (below right). Note that with the Brunswik faces, we used several different ways of plugging the features into the generic category arrangement.



Dotted lines indicate possible linear boundaries that separate the categories in different 2-D subspaces

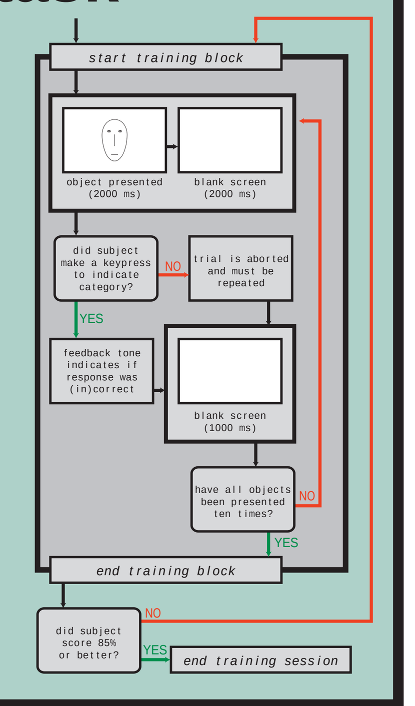
Object type	Dimension 1	2	3	4
Brunswik faces	EH	ES	NL	MH
Brunswik faces	NL	MH	EH	ES
Brunswik faces	MH	EH	NL	ES
Cartoon faces	EH	ES	NL	MH
Fish outlines	TF	LF	DF	MA

2.3 Categorization task

Training phase:
In a two-alternative forced-choice (2-AFC) task with auditory feedback (see diagram at right), subjects learned to categorize the training exemplars through trial-and-error. Each training exemplar was shown ten times in a training block. Subjects performed training blocks until they reached 85% accuracy in a single block.

Testing phase:
In a similar 2-AFC task, but without audio feedback, subjects categorized both the training exemplars and the test exemplars. In the testing block, each exemplar was shown five times. The result of the testing phase was a set of frequencies with which the subjects categorized the exemplars into the two categories; this set of frequencies was subsequently used to fit different computational models of categorization.

Each subject repeated the training/testing cycle in separate sessions on different days.

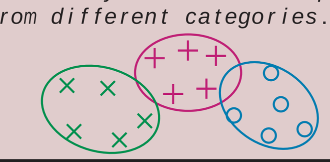


3. Models

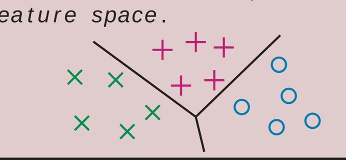
3.1 Categorization strategies

Intuitively, there are two strategies that people are likely to use for perceptual categorization.

exemplar-based strategy:
New exemplars are categorized according to their relative similarity to stored exemplars from different categories.

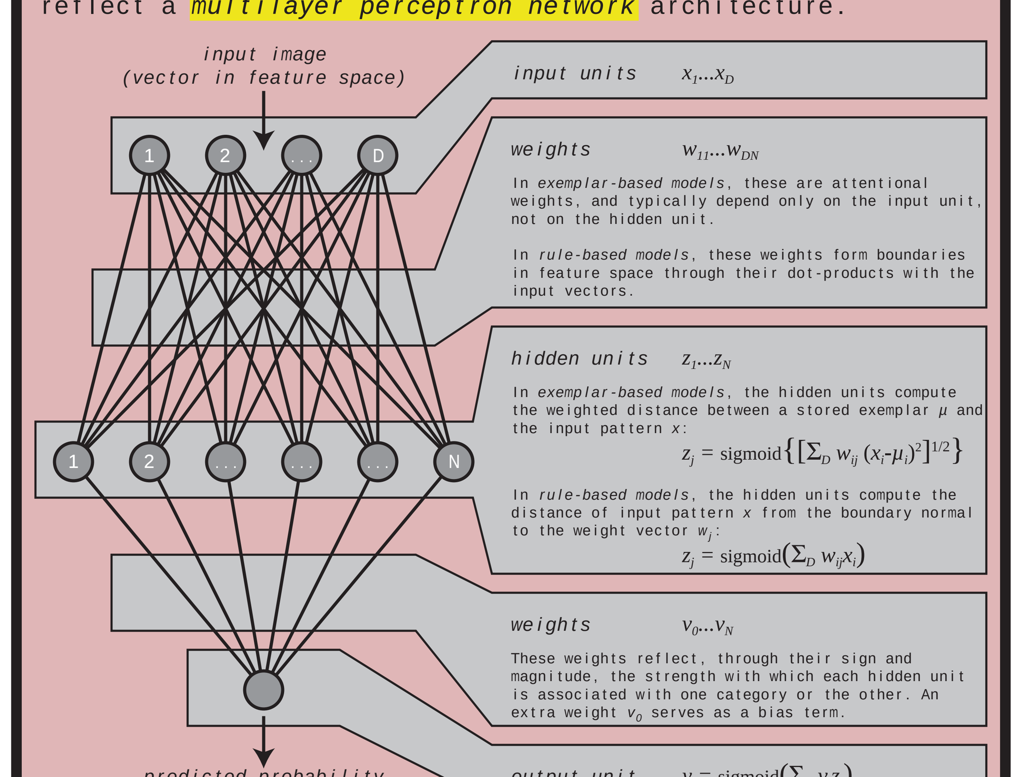


rule-based strategy:
New exemplars are categorized according to mathematical rules or boundaries that operate in feature space.



3.2 Categorization models

A ***categorization model*** takes as its input a vector in feature space, and along with an output of a probability for categorizing the input vector into one of two categories. Such a model is then judged by how well its categorization probabilities match those of human observers, over a set of possible inputs. Two main types of models correspond to the two types of categorization strategies: ***exemplar-based models*** include the generalized context model [GCM, Nosofsky1986] and weighted prototype similarity model [WPSM, Reed1972], and imply a ***radial-basis function network*** architecture; ***rule-based models*** include general boundary models such as the probit linear boundary model [PBI, Maddox1993], and reflect a ***multilayer perceptron network*** architecture.

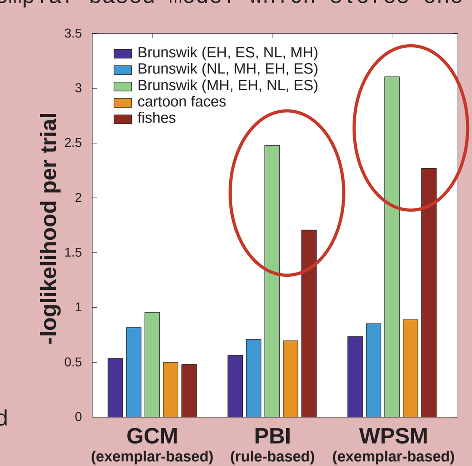


Keywords

categorization
exemplar
visual category
subordinate-level category
feature space
training exemplar/test exemplar
exemplar-based strategy/model
rule-based strategy/model
radial basis function network
multilayer perceptron network
GCM (generalized context model)
WPSM (weighted prototype model)
PBI (probit linear model)
loglikelihood
attentional weights

3.3 Model fitting

We tested three models: the ***GCM*** (exemplar-based model which stores all training exemplars), the ***WPSM*** (exemplar-based model which stores one prototype per category), and the ***PBI*** (rule-based model which uses one linear boundary). The free parameters of the models were fitted to maximize the model ***loglikelihood*** using the Nelder-Mead downhill simplex algorithm. This was done both for individual subject data and for pooled group data. As the plot at right reveals, we observed different patterns of model fits among the object sets. For two sets of Brunswik faces and the cartoon faces, the GCM fit best, with the PBI performing nearly as well, and the WPSM slightly worse. However, for the third set of Brunswik faces and the fish outlines, the PBI and WPSM performed far worse than the GCM. Why?

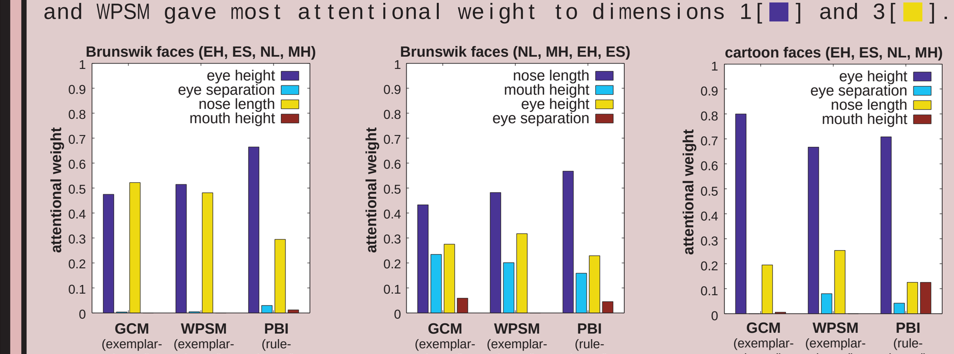


Minus the loglikelihood per trial, for the model fits obtained for three models on five datasets. Smaller bars indicate better fits.

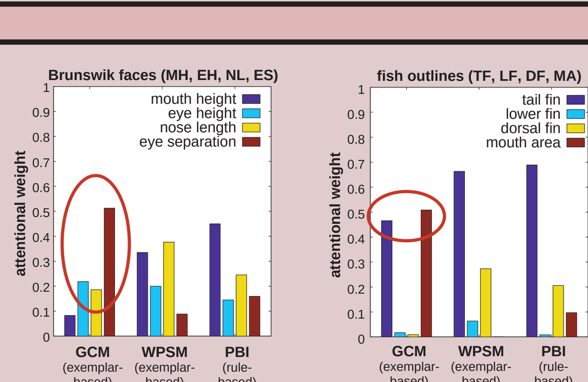
3.4 Attentional weights

The input units and the hidden units in the neural network architecture are connected by a set of ***attentional weights*** that are intended to reflect subjects' relative allocation of attentional resources to the different features. With the configuration of categories used in this experiment, the optimal attentional strategy is to devote attentional resources only to the features that correspond to generic dimensions 1 and 3, since the categories are best separated in the subspace spanned by these dimensions. However, we found that the models fitted to subjects' data did not always reflect this optimal strategy. Instead, we observed a qualitative distinction among the datasets that parallels the distinction in the goodness-of-fit results.

For two sets of Brunswik faces and the cartoon faces, the GCM, PBI, and WPSM gave most attentional weight to dimensions 1[■] and 3[■].

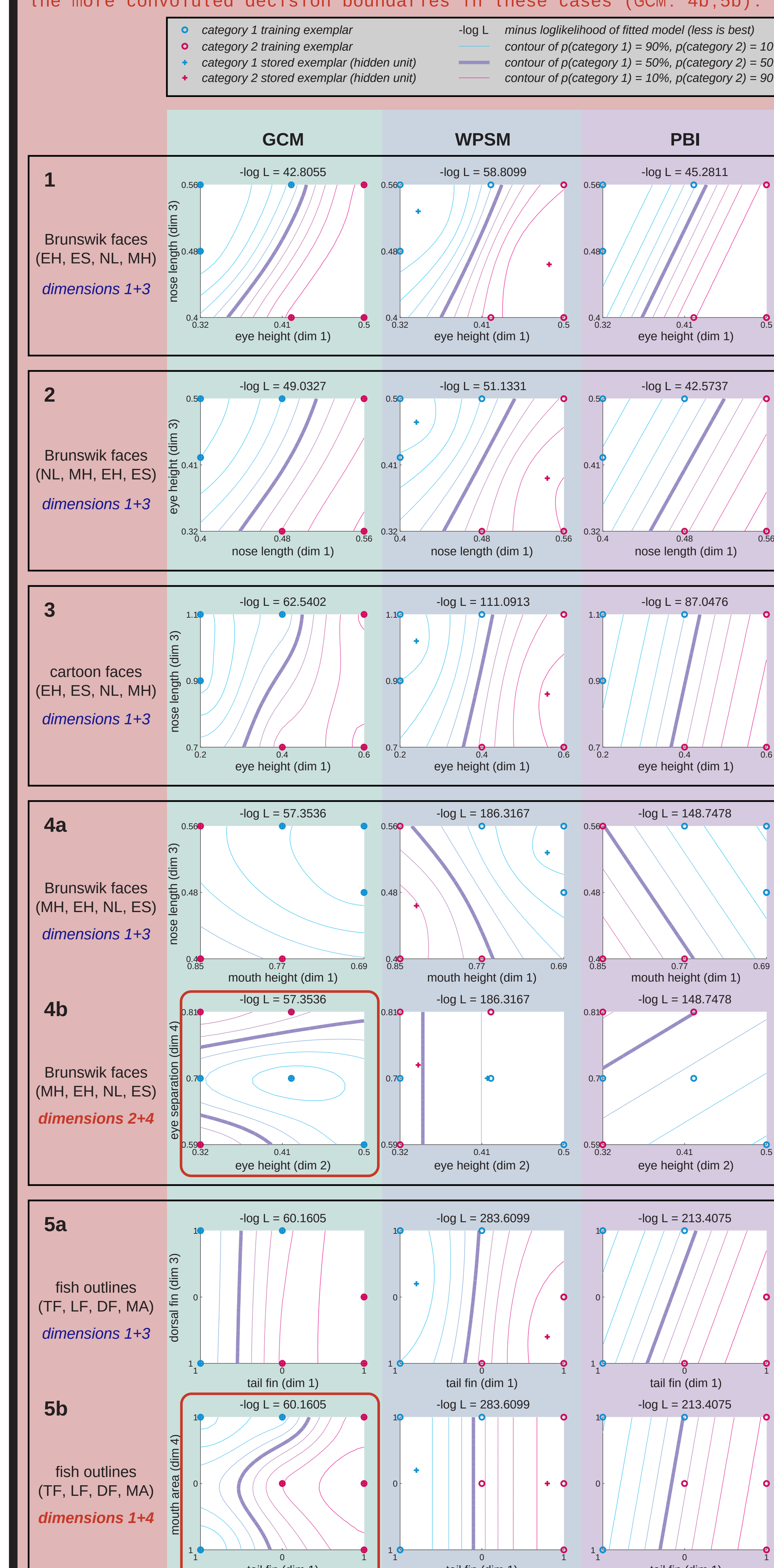


However, for the third Brunswik face set and the fish outlines, only the PBI and WPSM (which obtained poor fits in these cases) devoted most weight to dimensions 1 and 3, whereas the GCM allocated weight to different dimensions.



3.5 Decision surfaces

Finally, we considered the decision surfaces of the different models when fitted to subjects' data. In the plots below, these surfaces are depicted below by 2-D slices of their isoprobability contours. Again, in the first three datasets (1,2,3), subjects' categorization behavior is well described by a boundary in the dimension 1/3 plane, and all of the models are able to capture this trend. However, for the last two sets, subjects' categorization decisions do not follow this pattern (4a,5a); instead they categorize using a different pair of dimensions, and only the GCM is able to match this behavior. Note the more convoluted decision boundaries in these cases (GCM: 4b,5b).



Summary

When faced with an open ended categorization task, human observers do not form decision rules that are oriented in arbitrary directions in feature space, even if valid such rules are available; otherwise, our observers would have generated identical results for all of the object sets in these experiments.

If the features that are highly informative about category membership are not sufficiently salient, observers will prefer to use a more ad hoc exemplar-based approach which allows them to attend to other more salient, but less informative, features.