Learning and Judging the Attributes of Natural Objects

Wanling Zou, Sudeep Bhatia

Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

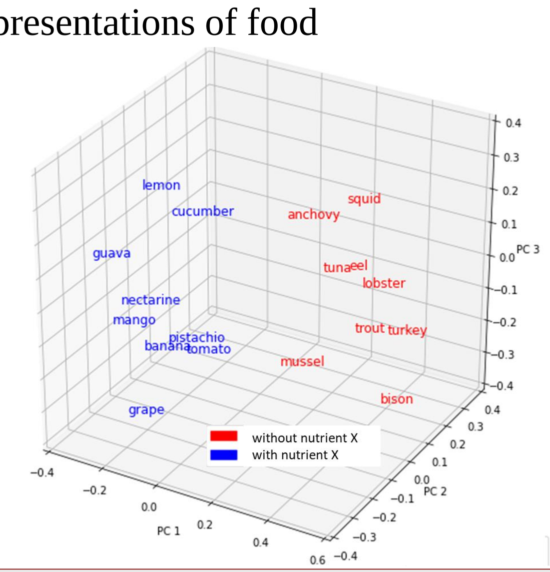
Join live chat at:
https://upenn.zoom.us/j/5481122840

Abstract

Category learning is a major research topic in psychology. Yet theories of learning are difficult to apply to problems in business and policy, as they typically involve artificial experimental paradigms with highly stylized stimuli. We combine category learning theories with rich representations for natural objects derived from large-scale digital data. Our approach achieves high accuracy rates in predicting how people learn and judge new attributes of foods, occupations, and countries, and can also be used to identify the learning environments that impair judgment. In doing so, it shows how existing psychological theories can be used to predict and improve everyday cognition and behavior.

Introduction

- Recent applications of categorization models
  - Rock image classification (Nosofsky et al., 2018)
    - Input: Multidimensional scaling (MDS) solutions of similarity ratings of rock image
    - Model: the Generalized Context Model (GCM, Nosofsky, 1986)
  - Plane vs. bird image classification (Battleday et al., 2017)
    - Input: Hidden layers from convolutional neural network (CNN)
    - Model: GCM
- Open question
  - How do people learn new categories of existing natural objects with language labels?
  - Examples of real-world categories:
    - Food items that have or do not have a novel nutrient
    - Occupations that will be in high or low demand after a technological breakthrough
    - Countries where people will or will not be predisposed to a rare cancer
- One promising solution
  - Use vector space semantics as input into GCM
  - Pretrained Word2Vec model
    - Trained on a large corpus of Google news articles (Mikolov et al., 2013)
    - Vocabulary of 3M words defined in a 300-dimensional space
  - Example of Word2Vec representations of food



Modeling Approach

- Distance function

$$d_{ij} = \left[\sum \left|x_{im} - x_{jm}\right|^2\right]^{\frac{1}{2}}$$

- Similarity function

$$s_{ij} = e^{-c d_{ij}}$$

- Category probability

$$P(A) = \frac{\left(\sum_{j \in A} s_{ij}\right)^{\gamma}}{\left(\sum_{j \in A} s_{ik}\right)^{\gamma} + \left(\sum_{k \in B} s_{ik}\right)^{\gamma}}$$

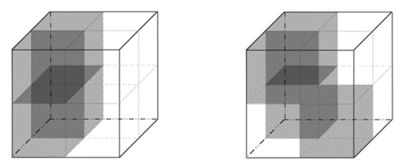
- Loss function

$$\sum I(\hat{y} \neq y)$$

- Note: Model was trained on the 50 training stimuli without human response.

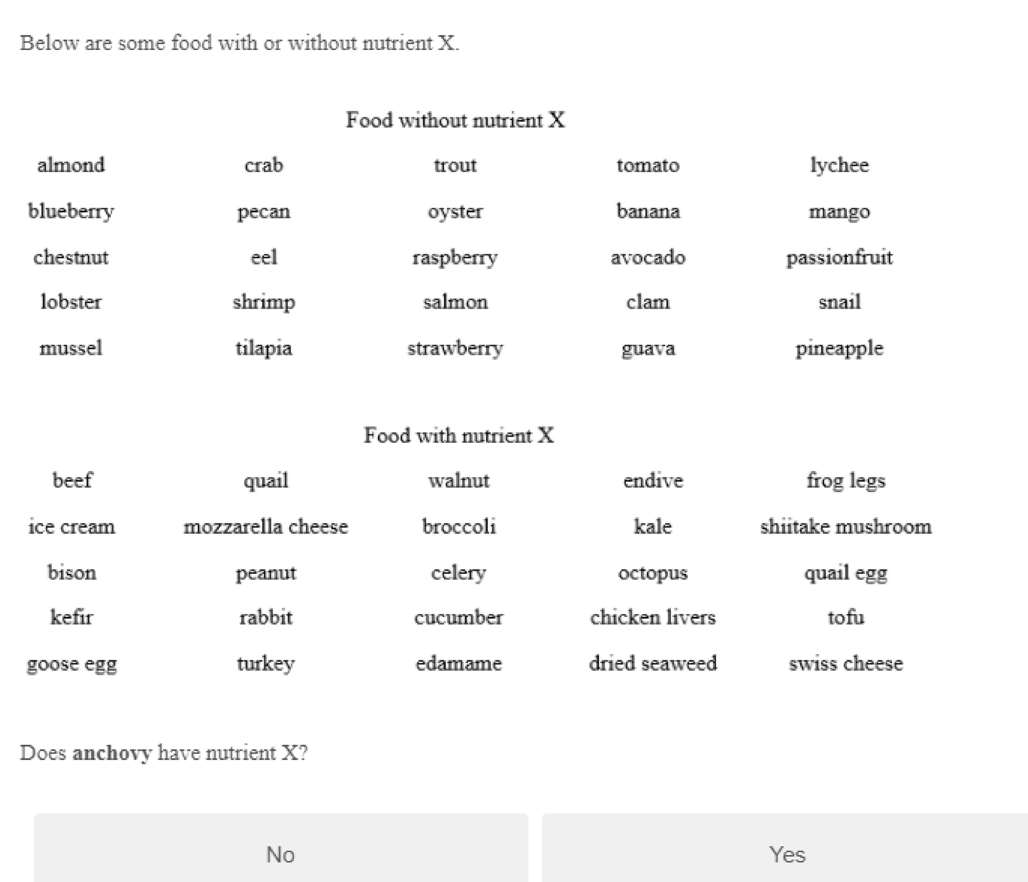
Experimental Design

- 100 participants in each study recruited from Prolific Academic
- 100 stimuli were split into 2 equal-size categories based on three generating methods.
- Each method generated one simple and one complex category structure (between-subject).
- Category generating methods:
  - Studies 1A-1C, 4 & 5: k-means clustering solution with k = 2 (simple) and k = 10 (complex)
  - Study 2: type I (simple) and type V (complex) from Shepard et al. (1961)



I V

  - Study 3: real nutrient content – cholesterol (simple) and lutein (complex)
- Study outline:
  - Study 1A: 100 foods from two categories generated by k-means clustering
  - Study 1B: 100 countries from two categories generated by k-means clustering
  - Study 1C: 100 occupations from two categories generated by k-means clustering
  - Study 2: 100 foods from two categories generated by method from Shepard et al. (1961)
  - Study 3: 100 foods from two categories generated by real nutrient content
  - Study 4: test the effect of different training size (10 vs. 50) with food categories generated by k-means clustering
  - Study 5: test the effect of different compositions of training set (size of 10) with food categories generated by k-means clustering
- Procedure
  - Participants predicted category labels for 50 test items when viewing either 10 or 50 training items in each study.
  - Participants were incentivized by $1 if their accuracy reached the top 10%.
  - Sample instruction: "We recently discovered a nutrient X that may be found in some food items. In this study, we will show you food items that have or do not have this nutrient. Your task is to predict whether some other food items have this nutrient or not."



Below are some food with or without nutrient X.

Food without nutrient X

almond	crab	trout	tomato	lychee
blueberry	pecan	oyster	banana	mango
chestnut	eel	raspberry	avocado	passionfruit
lobster	shrimp	salmon	clam	snail
mussel	tilapia	strawberry	guava	pineapple

Food with nutrient X

beef	quail	walnut	endive	frog legs
ice cream	mozzarella cheese	broccoli	kale	shiitake mushroom
bison	peanut	celery	octopus	quail egg
kefir	rabbit	cucumber	chicken livers	tofu
goose egg	turkey	edamame	dried seaweed	swiss cheese

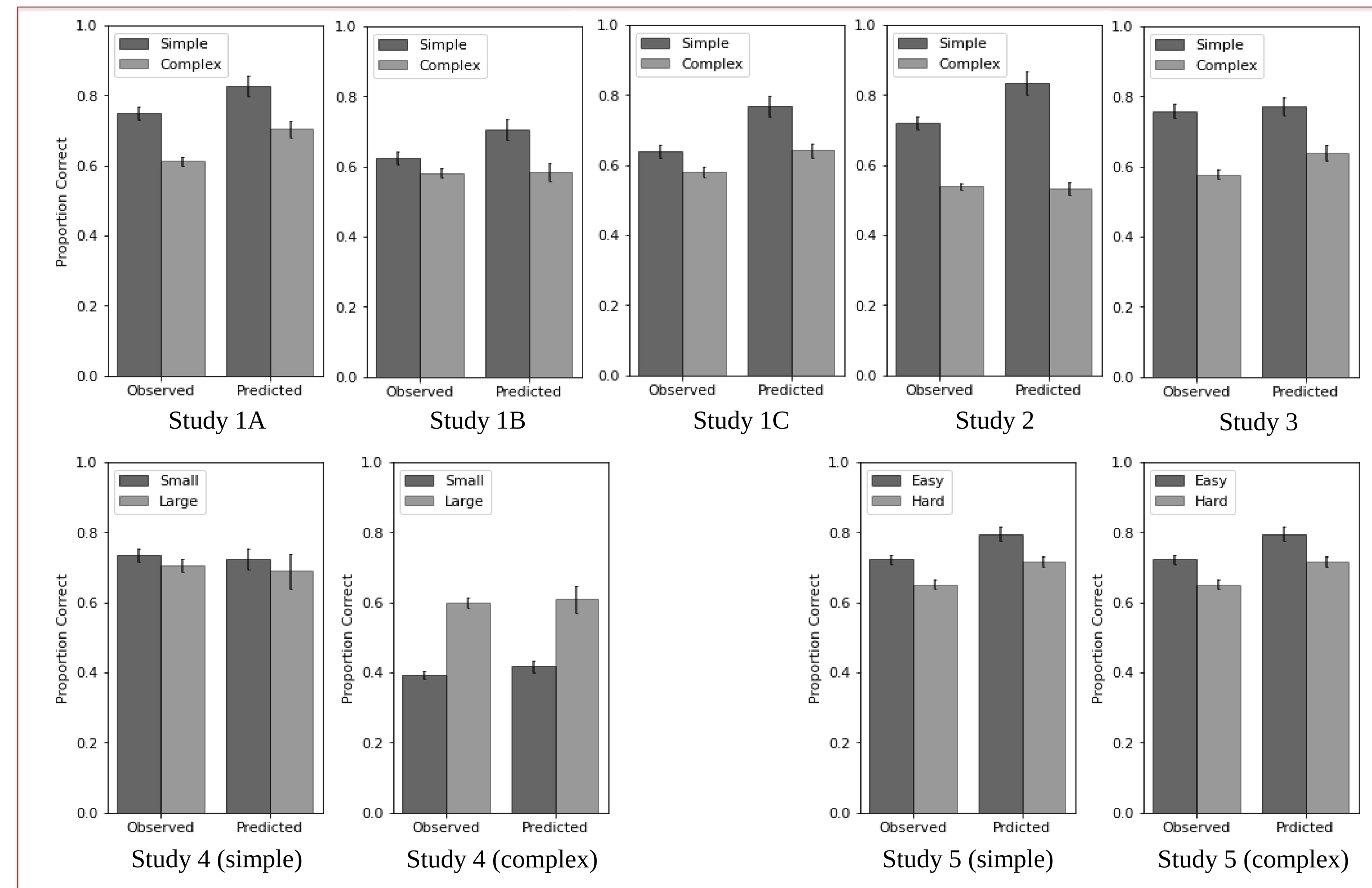
Does anchovy have nutrient X?

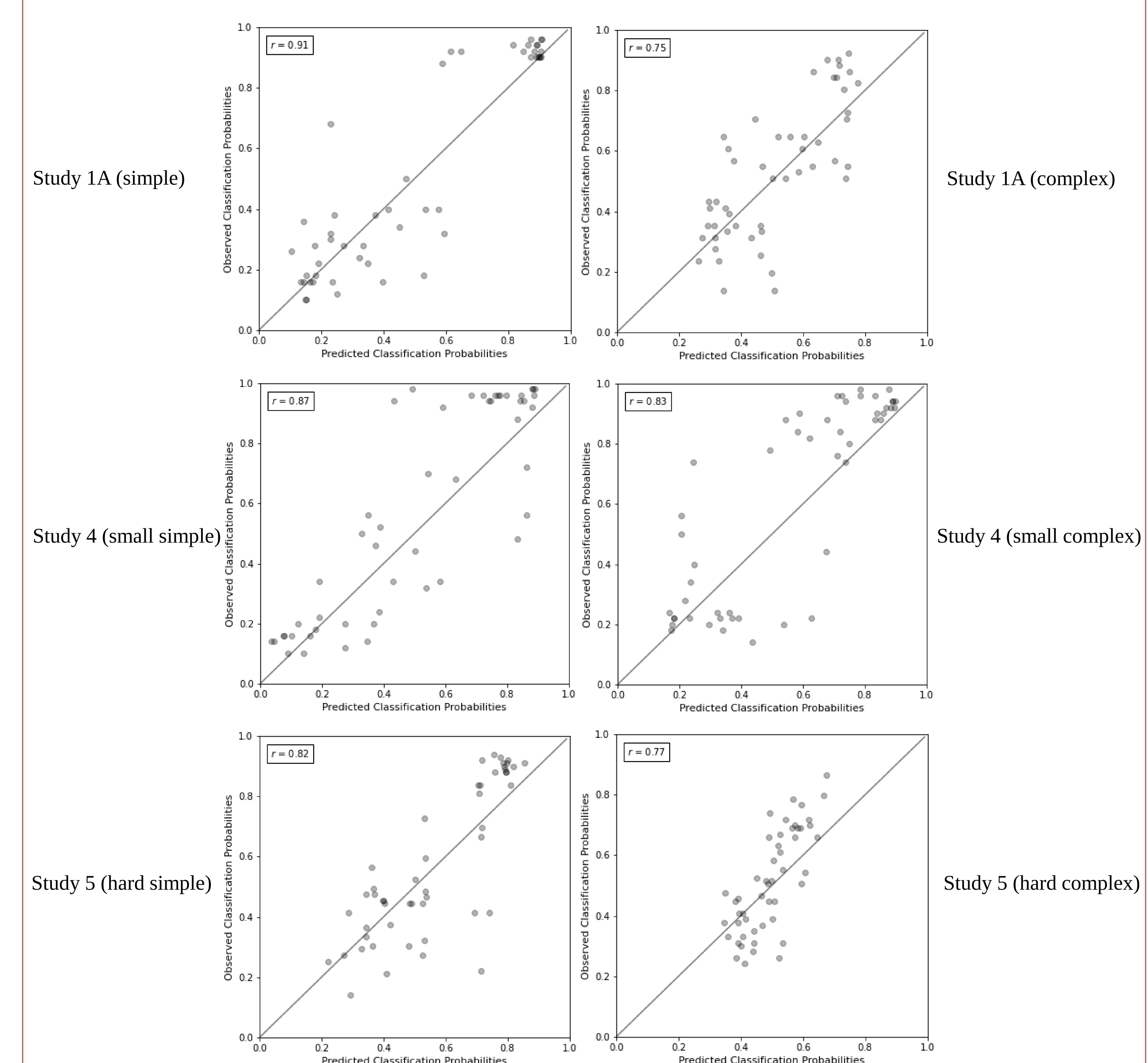
No Yes

Results

- The average human accuracy rates decreased as category structure got more complex and this result is consistent across different domains and different category generating methods (Studies 1A-1C, 2 & 3).
- The average human accuracy rates decreased as training size decreased (Study 4).
- The average human accuracy rates decreased as training set were harder to learn by the model (Study 5).
- In all studies, GCM-predicted accuracy rates mimicked the same pattern of the average human accuracy rates.
- In addition, GCM predicted item-level category probabilities accurately.

Results (Cont.)





- Word vector representations approximate human knowledge well.
- GCM provides a reasonable account for category learning.

Battleday, R. M., Peterson, J. C., & Griffiths, T. L. (2017). Modeling human categorization of natural images using deep feature representations. arXiv preprint arXiv:1711.04855.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111–3119).

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*(1), 39–57.

Nosofsky, R. M., Sanders, C. A., & McDaniel, M. A. (2018). A formal psychological model of classification applied to natural-science category learning. Current Directions in Psychological Science, 27, 129–135.