Learning Categories by Generating Examples

Kenneth J Kurtz, Nolan Conaway, Kimery Levering, Emily Eisenberg

Binghamton University, SUNY (USA)

We investigate a mode of category learning inspired by the distinction in machine learning between generative (learning about the basis of categories) and discriminative (learning how to tell categories apart) methods. Whereas the most commonly studied mode of human category learning (classification) is strongly discriminative, a strongly generative task is one in which the learner creates examples of categories. On each learning trial, a minimal featural cue is the starting point for building a complete member of a target category. The learner receives feedback on whether the generated example is a category member. In a 2x3 design, we manipulated the type of learning (generation vs. classification) for three different elemental category structures based on three binary features. Using a set of test measures, we found differences in the quality of category knowledge for the two learning modes that were consistent with the generative/discriminative framework.

Task effects in Category Learning

Classification learning is the most widely studied task in categorization research.

- Recent interest in alternative training modes
  - Feature inference (Yamauchi & Markman, 1998)
  - Observation (Levering & Kurtz, 2011)

Successful classification only requires knowledge of the *difference* between categories (*discriminative* learning; Ng & Jordan, 2001).

- We developed a novel training mode, where learners are asked to 'generate' examples of a target category.
  - *Will generate learning result in generative representations?*
  - *Can models of classification account for such learning?*

Generate learning is theoretically similar to feature inference

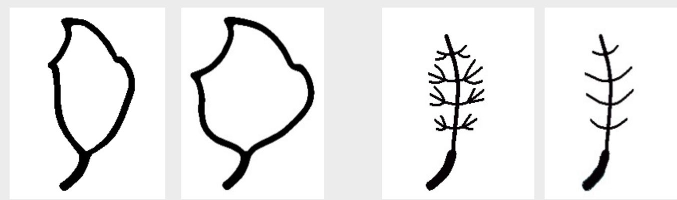
- Feature inference, but for more than one feature
- Generate is qualitatively different since learners 'make' examples

Previous work on generation of categories (Jern & Kemp, 2013).
No known prior research on generate task for category *learning*.

Generate Task

- Trials begin with a single feature.
- Subjects asked to complete it as a member of a target category.
- Image is updated to reflect the selection after each response.
- Feedback provided when the example is complete.

Click on the choices below to turn what you see into a Tannet leaf.



Classification Task

- Trials begin with completed examples.
- Subjects asked to guess the category that example belongs to.
- Participants receive feedback on their responses.

Click a button to select the correct category.

Lape   Tannet

Stimuli & Design

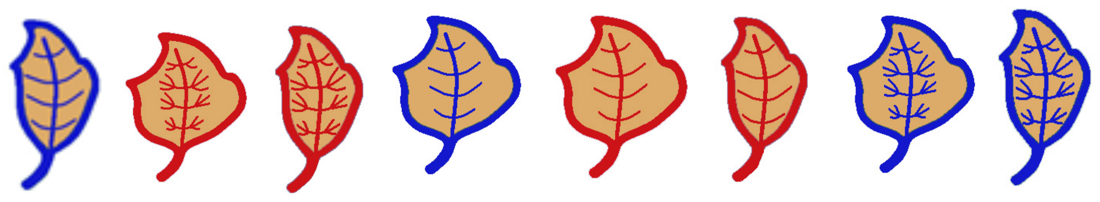
- Two novel leaf categories: Lape and Tannet
- Leaf images vary in three binary dimensions →
- 120 training trials

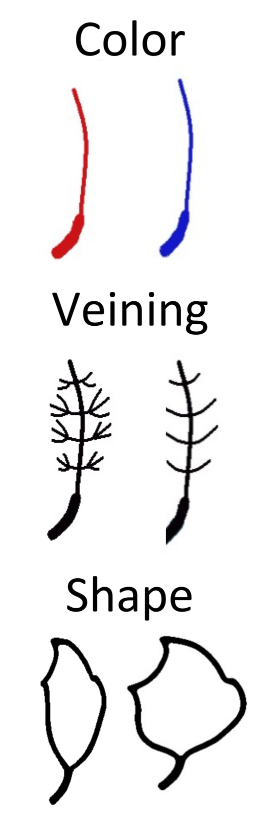
Type II focal dimensions: XOR on *color* and *veining*

Type III focal dimensions: *veining* and *shape* each support unidimensional rule plus exception

Type IV prototypes:

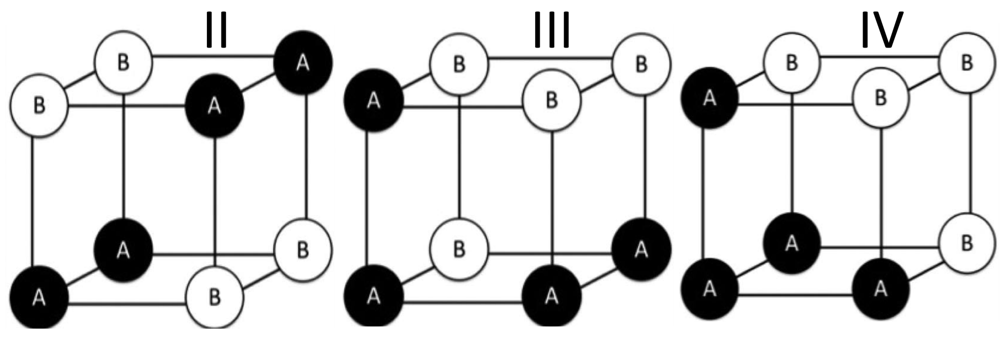
[red, hi veining, narrow], [blue, lo veining, wide]





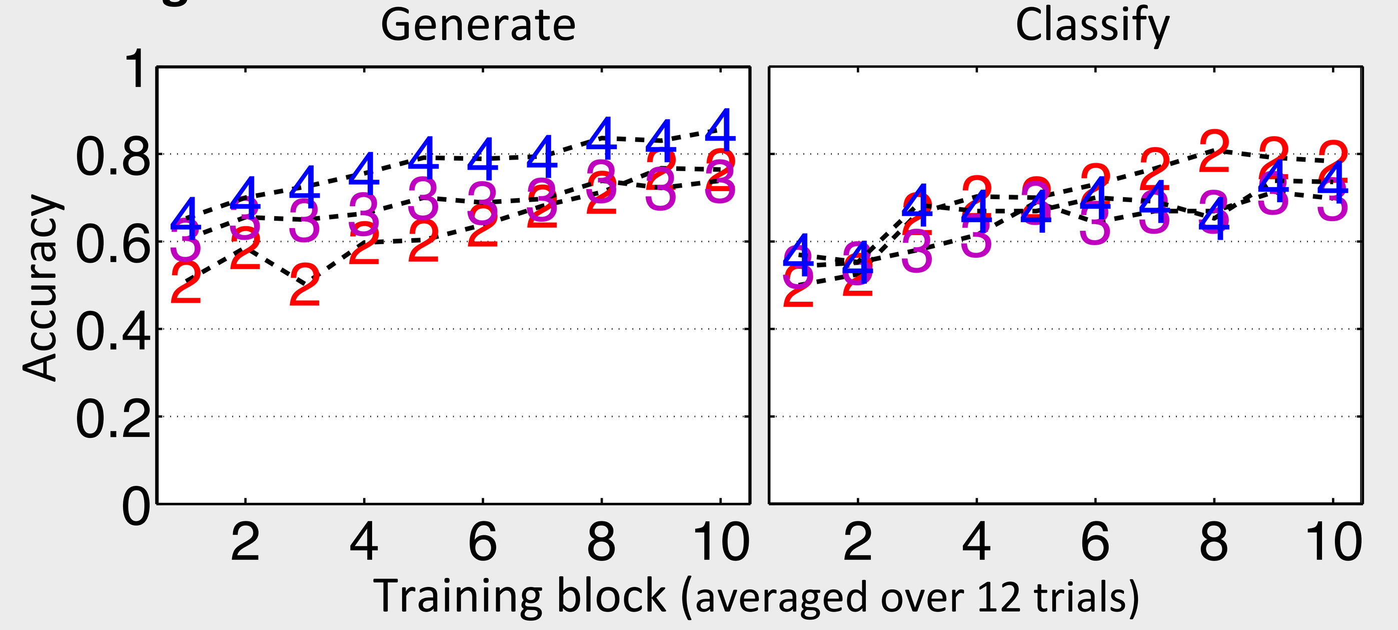
Shepard, Hovland, and Jenkins (1961) tested ease of learning of six elemental category structures (I < II < III,IV,V < VI)

Here we explore the generate learning mode using types II - IV.



	Training					Test Measures			
	F1	F2	F3	Class		F1	F2	F3	Class
Generate	?	?	1	Tannet	**Endorsement**	1	0	1	*Tannet?*
Classify	1	0	1	?	**Single Feat.**	1			Tannet *1-9?*
					Typicality	1	0	1	Tannet *1-9?*

Training Results



Classify results replicate:

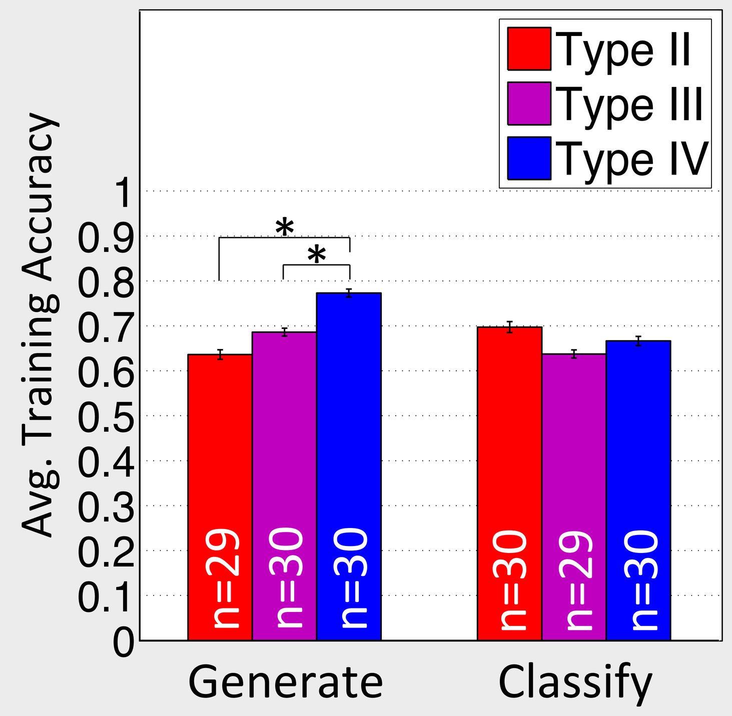
Kurtz et al. (2013): *II == IV*

Shepard et al. (1961): *III == IV*

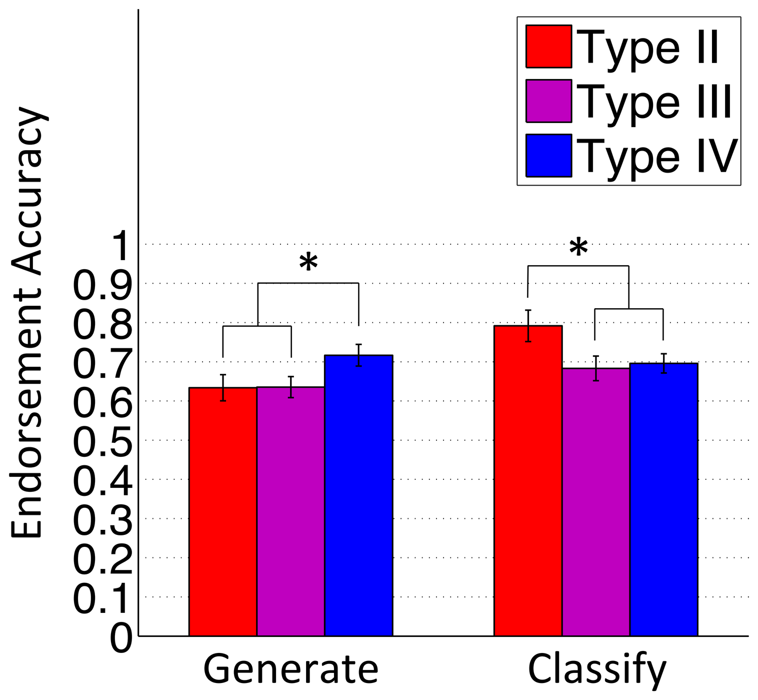
Generate results show key finding of Type IV advantage:

Type IV better than Type II (reverses trad SHJ order)

Type IV better than Type III (equal in trad SHJ order)



Endorsement Results



Classification Type II adv. (due to late learners)

Type II > Type III == Type IV

Generate Type IV adv.

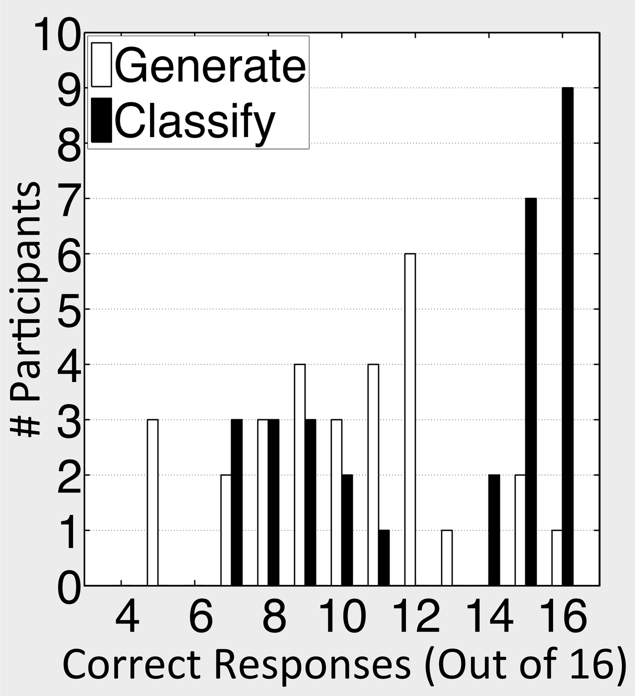
Type IV > Type II == Type III

Due to classify task similarity:

- Classify > Generate
- generate endorse < training

Evidence for differences in Type II learning

000
001
110
111

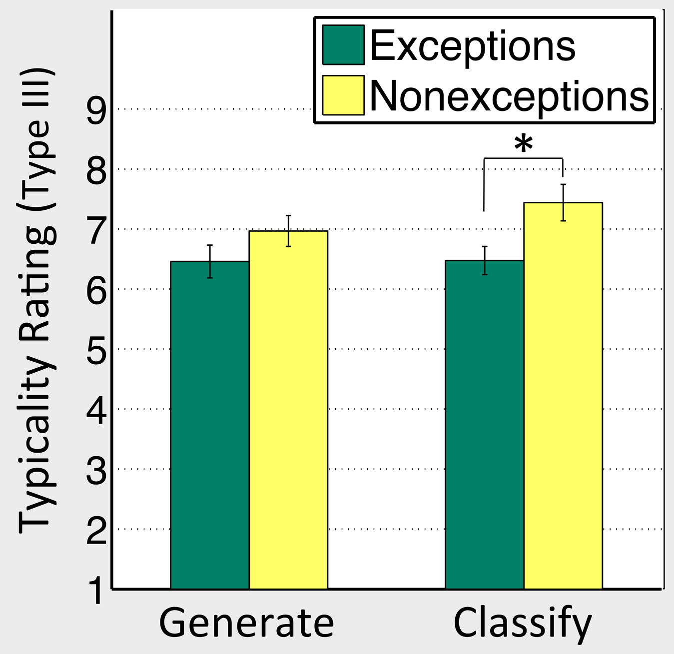


Endorsement Results

- Classify replicates bimodal distribution; Kurtz et al. (2013) [mean = 12.67]
- Bimodal distribution not seen in generate group [mean = 10.14]

Evidence for differences in Type III learning

000
001
010
101



Typicality Results

Items containing an exception feature were rated as less typical in classify

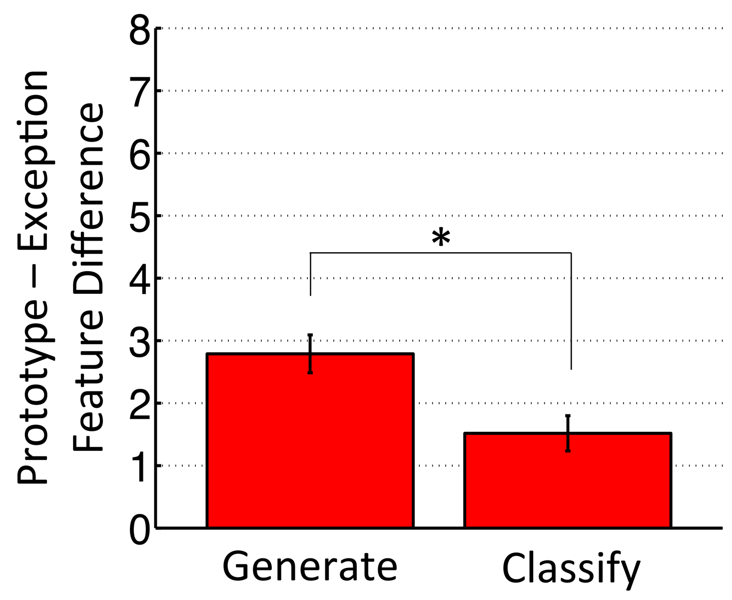
No difference in generate

Evidence of rule + exception strategy (Nosofsky et al., 1994) only in classify group

Evidence for differences in Type IV learning

000
001
010
100

Single Feature Test Results



Larger difference ratings between prototype and exception features in generate

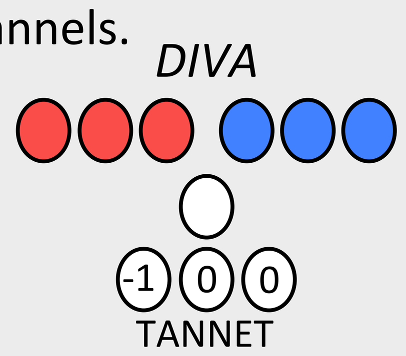
No accuracy difference between generate and classify

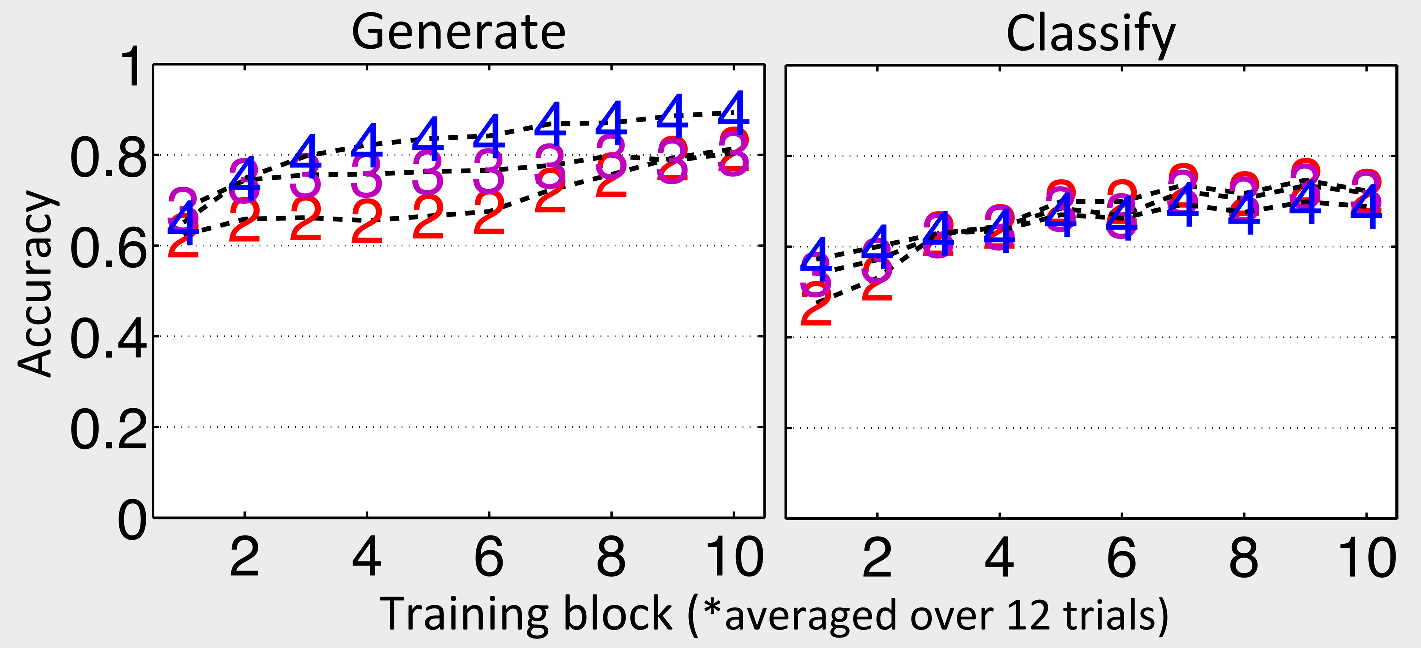
Evidence of rule strategies in classify and family resemblance in generate

Modeling Generate Learning using DIVA

DIVA (Kurtz, 2007) is a *DIV*ergent Autoencoder that learns to reconstruct inputs on dedicated category channels.

- DIVA takes as input a single feature
  (missing features coded as 0 in a [-1 1] space)
- Generates example based on reconstruction along targeted channel (pattern completion)
- Model is trained on the generated example





DIVA captures qualitative & quantitative patterns of learning
all fits < .017 MSQ

Discussion

- We developed and tested a new *generative* learning mode, inspired by a distinction proposed in machine learning research.
- Using a set of test phases, we conclude that generate learners differed from classification learners in speed of acquisition of the categories, as well as the type of knowledge learned.

Jern A & Kemp C (2013). A probabilistic account of exemplar and category generation. *Cognitive Psychology*, 66, 85-125.

Kurtz KJ. (2007). The divergent autoencoder (DIVA) model of category learning. Psychological Bulletin & Review, 14(4), 560–576.

Kurtz KJ, Levering KR, Stanton RD, Romero J, Morris SN (2012). Human Learning of Elemental Category Structures: Revising the Classic Result of Shepard, Hovland, and Jenkins (1961). *Journal of Experimental Psychology: Learning Memory & Cognition*, 39(2), 552-572.

Levering, KR & Kurtz KJ (2011). Observational category learning as a path to more robust generative knowledge. *Proceedings of the 33rd Annual Cognitive Science Society* (pp. 631-636). Boston, MA.

Ng AY & Jordan M. On discriminative vs. generative classifiers: A comparison of logistic regression and naive Bayes. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems,* volume 14, pages 841–848, Cambridge, MA, 2001. MIT Press.

Nosofsky RM, Palmeri TJ, & McKinley SC (1994). Rule-plus-exception model of classification learning. *Psychological Review*, 101(1), 53-79.

Shepard RN, Hovland HL, & Jenkins HM (1961). Learning and memorization of classifications. Psychological Monographs, 75 (13, Whole Number 517).

Yamauchi T., & Markman, AB (1998). Category learning by inference and classification. Journal of Memory & Language, 39, 124–148.