Now You Know It, Now You Don't

Asking the right question about category learning

Nolan Conaway and Kenneth J. Kurtz
Department of Psychology, Binghamton University

A key goal of category learning research is to describe how categories are represented. Essential to this research are measures that provide investigators insight into exactly what learners have gained from their training experience. In this paper, we review and explore three commonly used measures: A) ease of acquisition, B) generalization, and C) single feature classification. We report results of a category learning experiment in which these measures are compared side-by-side. We find that generalization and single feature classification data are the more informative measures; we also find a novel inconsistency between them. Specifically, many learners who generalize based on only a single dimension demonstrate robust knowledge of both dimensions during the single feature classification test. We discuss implications for methodology in the field, as well as for selective attention and theories of human category learning.

Accessing a category representation

Ease of acquisition (e.g., SHJ six types)

- block-by-block accuracy to show time course of learning of training items
- ubiquitous in psychology of category learning

Generalization (e.g., 5-4 problem)

- classification of unpresented items after training
- establishes knowledge beyond memorization

Single Feature Classification

- classification of partial training items ('single features') after training
- recently employed in inference learning (Anderson, Ross, & Chin-Parker, 2002)

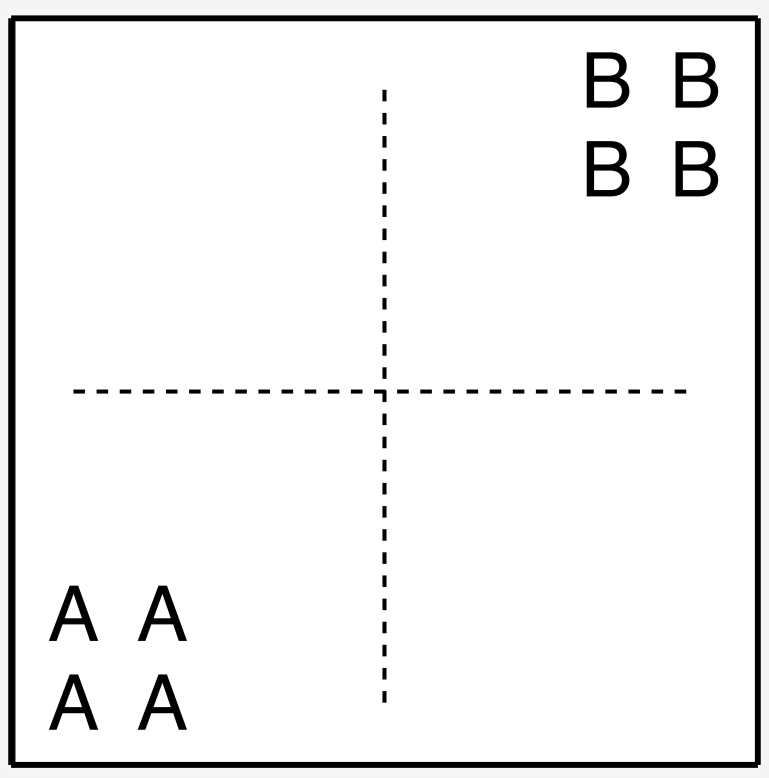
Despite wide use, this set of measures have not been compared side-by-side.

What aspects of category knowledge are effectively probed by each measure?

'Minimal Case' categories

A highly minimal category structure

Categories are separated on both dimensions (along a diagonal)

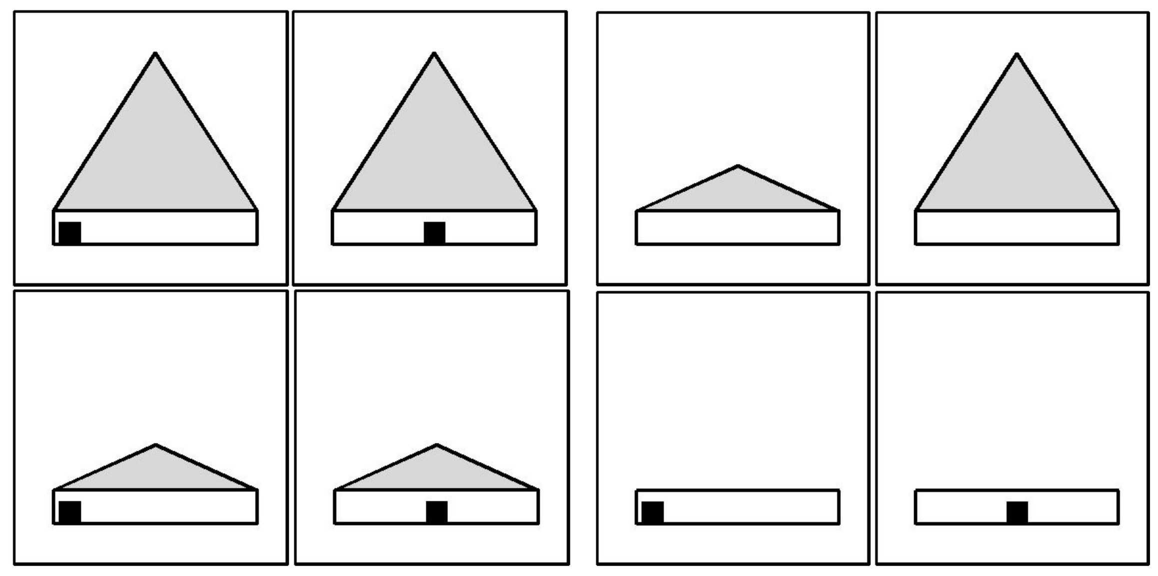


The potential basis for categorization can be:

* unidimensional rule on either dimension
* multidimensional/integrative (diagonal)

What do acquisition, generalization, and single feature classification tell us about how learners represent the minimal case categories?

Stimuli and Procedure



Full Feature Single Feature

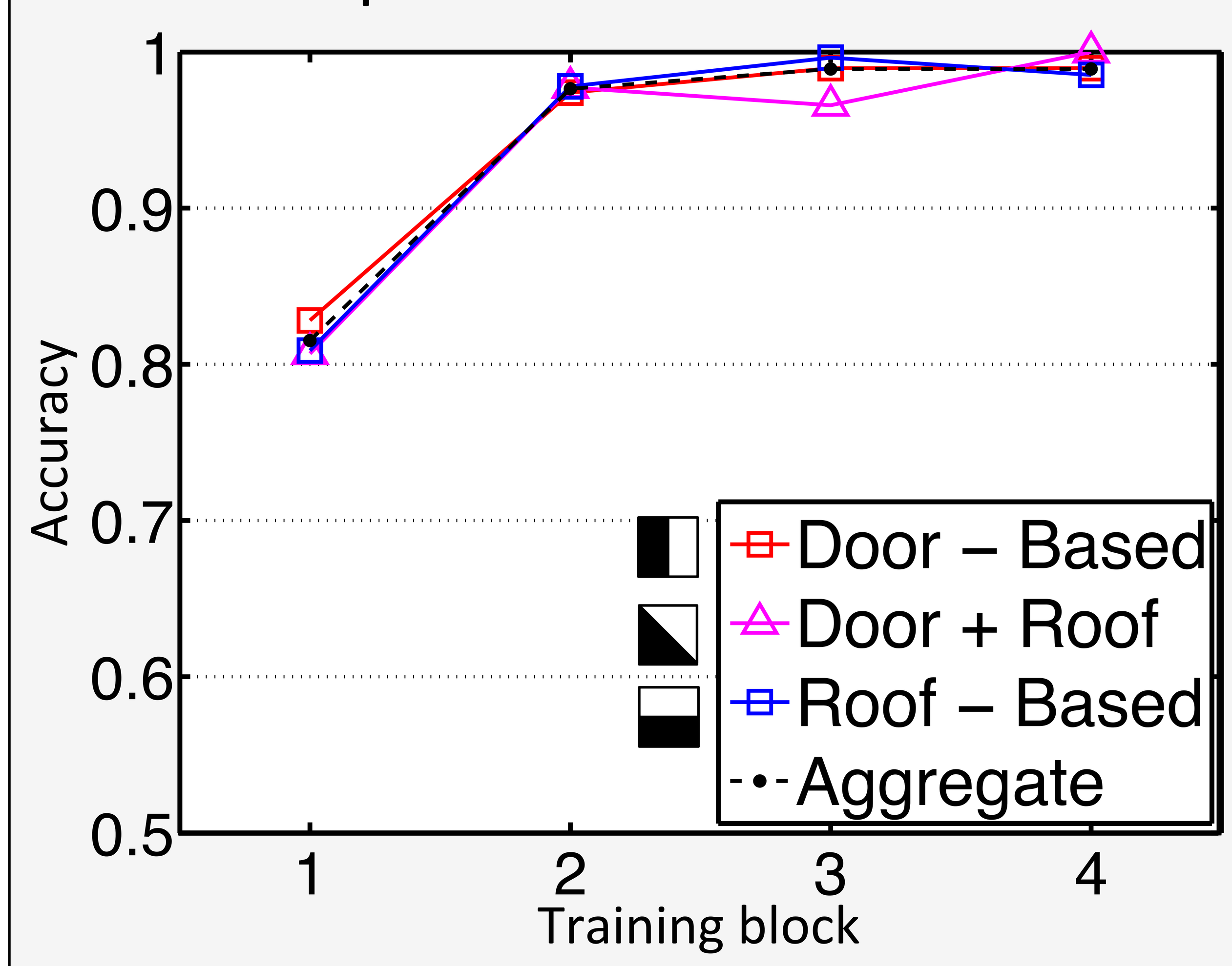
Stimuli

Houses varying in **Door** position and **Roof** height.
Generated evenly at 8 positions on each dimension

Procedure

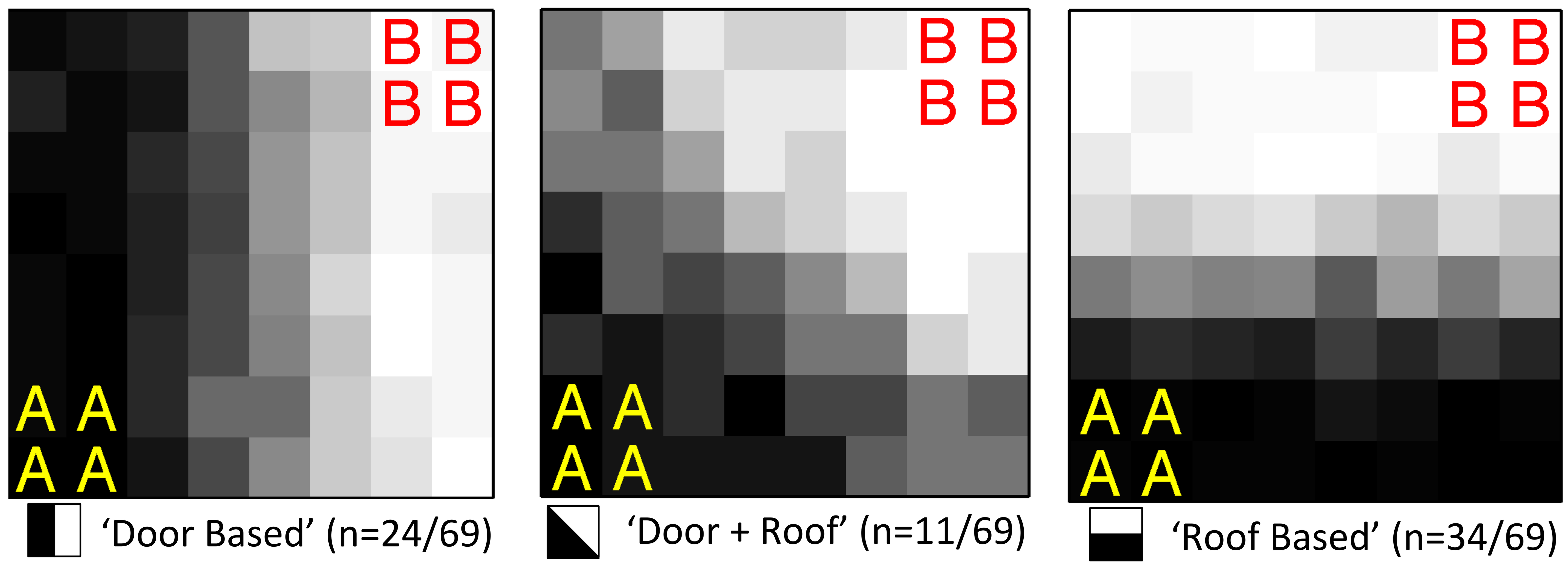
Training: 32 trials, (classification w/ feedback)
Generalization: 64 trials (classify, no feedback)
Single feature: 16 trials (classify, no feedback)

Ease of acquisition



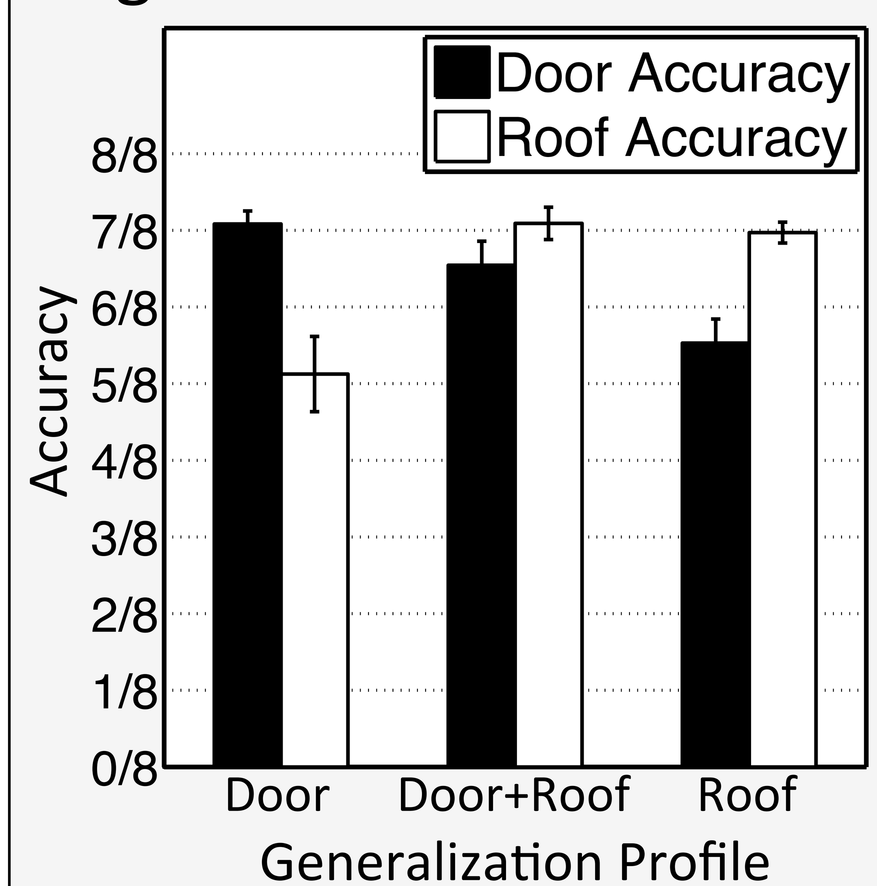
* Most learners acquired the categories within one block
* Acquisition data tell us little about the learned representation (perhaps due to a ceiling effect)

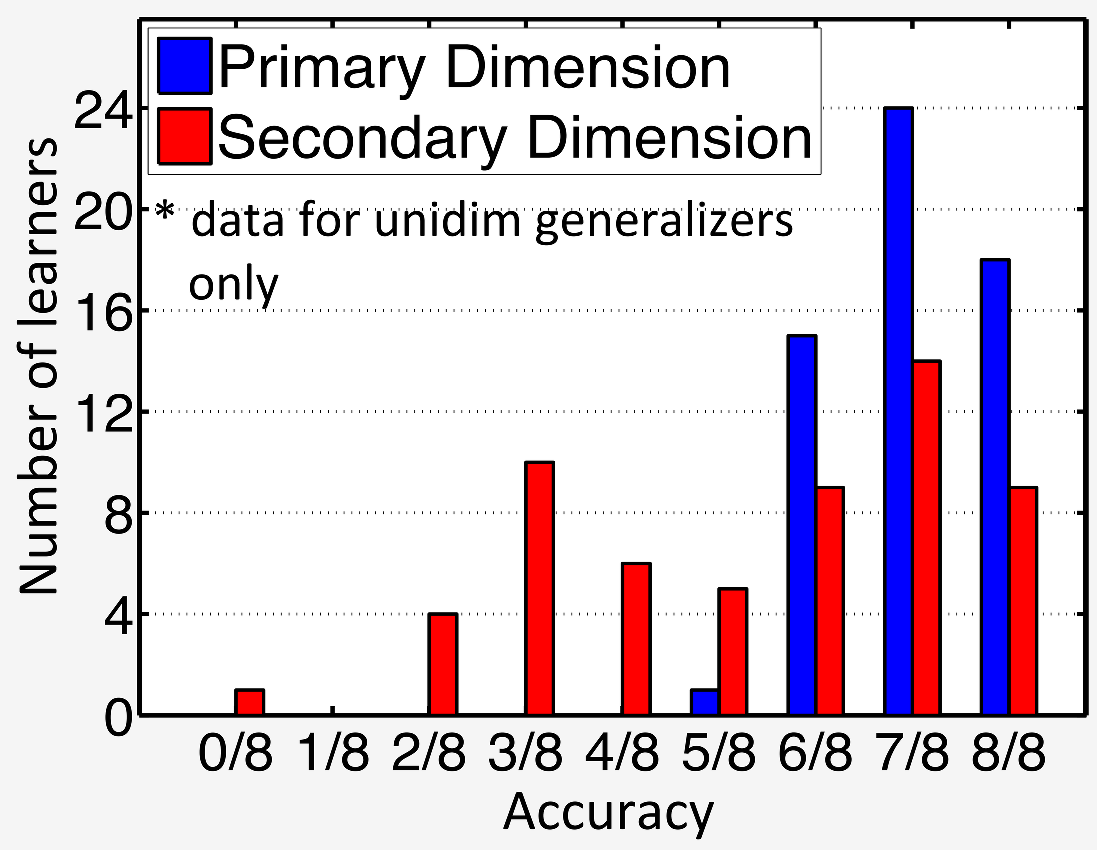
Generalization responses were profiled based on their match to three templates.



'Door Based' (n=24/69) 'Door + Roof' (n=11/69) 'Roof Based' (n=34/69)

Single Feature Classification





Primary Dim
Used as the basis for generalization

Secondary Dim
Not used as the basis for generalization

Note the bimodality in the secondary distribution

Unidim generalizers more accurate on selected dimension, but above chance on discarded dim ($ps < 0.001$)

32/58 unidim generalizers retained full knowledge of the dimension not used for generalization.

Many learners acquired knowledge of how the categories vary on a dimension they did not use as the basis for generalization

Implications for theory

**Inconsistent with traditional role of attention (optimization) in leading reference point models—attention not used to expedite acquisition (Kruschke, 1992)*

**Instead used to augment decision process, as in focusing mechanism developed for DIVA (Kurtz, 2007)*

Implications for methodology

**Generalization thought to provide a definitive depiction of learning (Shepard, 1987)*

**Instead it appears that generalization is subject to decision strategies*

Anderson, A. L., Ross, B. H., & Chin-Parker, S. (2002). A further investigation of category learning by inference. *Mem & Cog*, 30, 119–128.
Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psych. Review*, 99, 22-44.
Kurtz, K. J. (2007). The divergent autoencoder (DIVA) model of category learning. *Psych. Bulletin & Review*, 14, 560 –576.
Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, 237, 1317-1323.