Categories and Concepts - Spring 2019

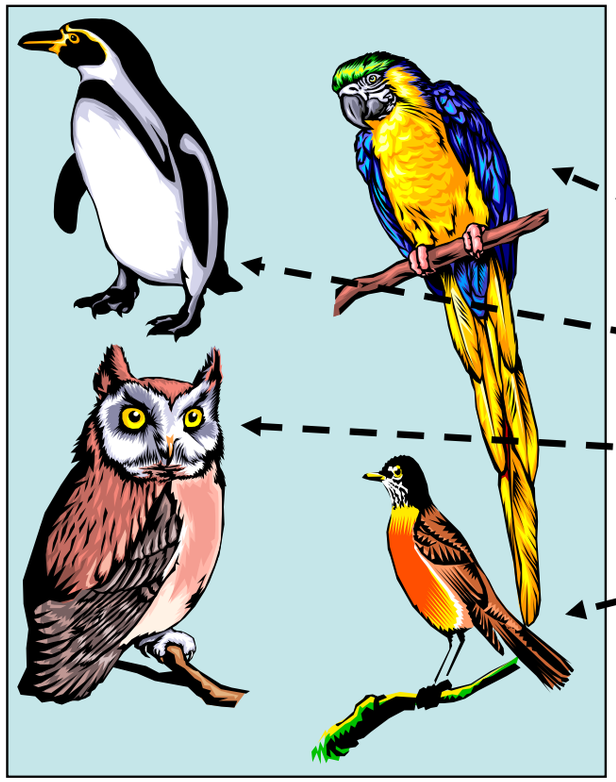
Concepts as theories, and the knowledge view

Brenden Lake

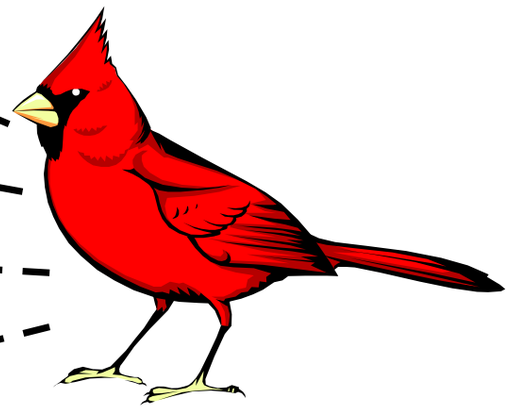
PSYCH-GA 2207

Review: major theories

exemplar theory



Birds
You've Seen



Bird?

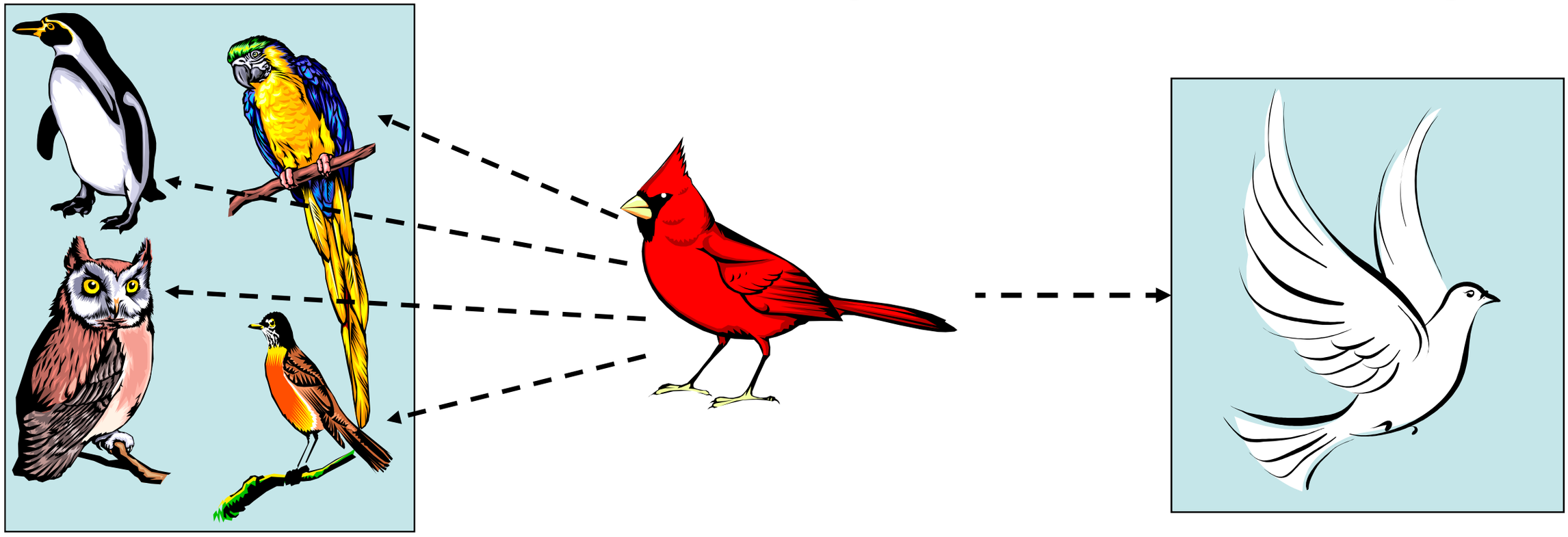
prototype theory

Prototypical
Bird

No summary representation. Representation is based on remembered category members.

Summary representation of typical properties or central tendency.

Both theories rely on “similarity”



Birds
You’ve Seen

Bird?

Prototypical
Bird

Relies on “similarity” of an item to exemplars

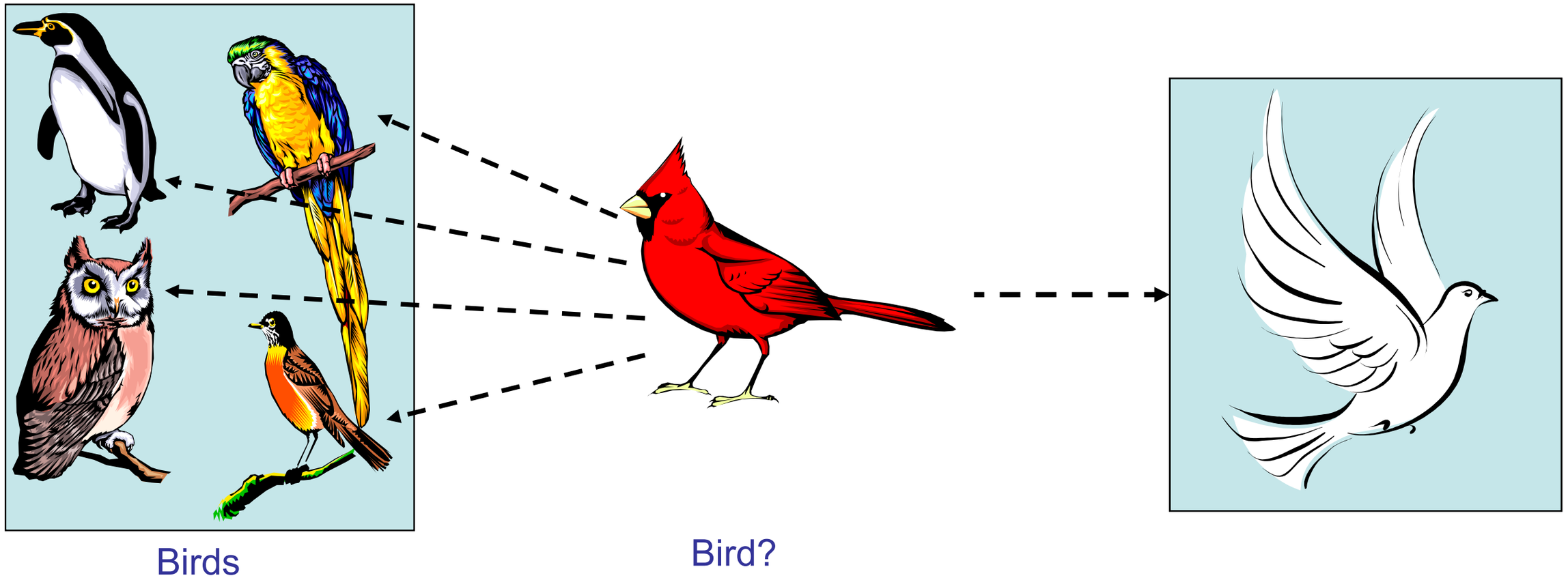
$$\textbf{sim}(y, C) = \sum_{x \in C} \textbf{sim}(y, x)$$

$$P(y \in C) = \frac{\textbf{sim}(y, C)}{\sum_{C'} \textbf{sim}(y, C')}$$

Relies on “similarity” of an item to prototype

$$P(y \in C) = \frac{\textbf{sim}(y, C)}{\sum_{C'} \textbf{sim}(y, C')}$$

Similar in what *respect*? What counts as a feature?



Birds You've Seen

Bird?

Prototypical Bird

- What counts as a feature? (Murphy & Medin, 1985)
  - To change the importance of age, we could include features for "around 10 years ago," "around 100 years ago," "1000 years ago", etc.
  - To change importance of size, we could include "smaller than the earth," "smaller than a country", "smaller than a city," etc.
- It is difficult to establish the "respects for similarity" (Medin, Goldstone, & Gentner, 1993, *Psych Rev)*

Whatever is selecting the features is doing the explanatory work

Flammable? ✓

Flammable? ✓

“Flammable” applies to many things, but we only associate it strongly to some things due to its *theoretical role* (Murphy & Medin, 1985)

Concepts are more than characteristic features or sets of examples

Would you classify this man as “drunk”? (Murphy & Medin)

What characteristic feature tells you this? Would you need a similar previous example in order to tell you this?

Ad hoc categories (Barsalou, 1983)




- e.g., “Things to carry out of a burning house”
  - [children, dog, photo albums, computer, etc.]

- “ways to escape being killed by the Mafia”
  - [changing your name, move to Montana, etc.]

- “things that could fall on your head"
- “things to take on a camping trip”
- “possible costumes to wear to a Halloween party”
- “places to look for an antique desk”

Ad hoc categories (Barsalou, 1983)

Ad hoc categories are like common categories in many ways..

- have a graded structure
- lead to reliable typicality ratings
- certain items being generated more in production tasks

- How would a prototype or exemplar model represent these categories?
- These properties of categories don't seem affected by how well the category has been learned, how many examples have been seen, or how well-established the class is in memory

Conceptual combination

Prototypes or sets of exemplars don't help to explain how categories combine

"pet"

"fish"

"apartment"

"dog"

?

?

"pet fish"

?

?

"apartment dog"

Concepts as theories, and the knowledge view (Murphy & Medin, 1985)

- **A concept is naive theory or a "mental explanation"**
- Explanations are the glue that holds concepts together
- Concepts are coherent to the extent that they fit people's background knowledge or naive theories about the world

- Don't take "theory" too seriously.
  - not a full-fledged scientific-like theory
  - Greg Murphy says it is better to talk about how knowledge influences concepts, and the term "theory theory" is an abomination
  - Some people do believe in "core theories," which are actually like real theories (e.g., Carey, Spelke, Gopnik)

Examples of Theories or General Knowledge That Might be Relevant to Concept Learning

- biological knowledge
- naive physics
- naive psychology
  - beliefs, desires, goals
  - psychological effects of different events
  - different personality types
- causal mechanics of various machines and artifacts

Note that each of these has many specific parts and pieces of knowledge.

Category Coherence

Concepts are more than characteristic features or sets of examples

Here are features of something that is not a concept:

- requires oil when heating
- can play music
- is written in Sanskrit
- chews food thoroughly before swallowing
- is transparent

Category Coherence

Such a concept does not exist, and if you tried to teach it to people, it would presumably be difficult to learn.

How do categorization theories address this? They don't, really.

- Classical view
  - Identifies a common feature(s), but doesn't constrain what it is ("likes pizza or is on the Sistine Shapel")
- Exemplar theory
  - any set of objects can be a category
- Prototype theory
  - incoherent list of features could be prototype

These accounts of concepts don't provide constraints on what concepts are natural or learnable

Barsalou (1985). JEP:LMC

Possible Determinants of Typicality

- Exemplar's similarity to central tendency (family resemblance)

- Exemplar's frequency of instantiation (how often do you think of the object as a category member?)

- Rating of proposed ideal qualities (most relevant for ad hoc categories, but reasonable goals given to common categories)
  - e.g. efficiency for "vehicles"
  - necessary of wearing for "clothing"

Barsalou (1985)

Predicting “Exemplar Goodness” (typicality) Ratings

<u>Common Categories</u> (e.g., vehicles, clothing, birds)

	Raw *r*	Partial *r*
Central Tend.	.63	.71
Frequency	.47	.36
Ideal	.46	.45

<u>Goal-Derived Categories</u> (e.g., birthday presents, foods to eat on a diet)

	Raw *r*	Partial *r*
Central Tend.	.38	.05 (ns)
Frequency	.72	.51
Ideal	.70	.44

Pazzani (1991). JEP:LMC

Between-subject condition

- learning which examples are in “category alpha” (control condition)
- learning “whether the balloon will be inflated when a person blows into it” (knowledge condition)

Stimuli

- photograph of person performing an action on balloon (stretch or dip)

Stimulus dimensions

- color—“yellow” or “purple”
- size of balloon—“large” or “small”
- age of person—“child” or “adult”
- action—“stretch” or “dip in water”

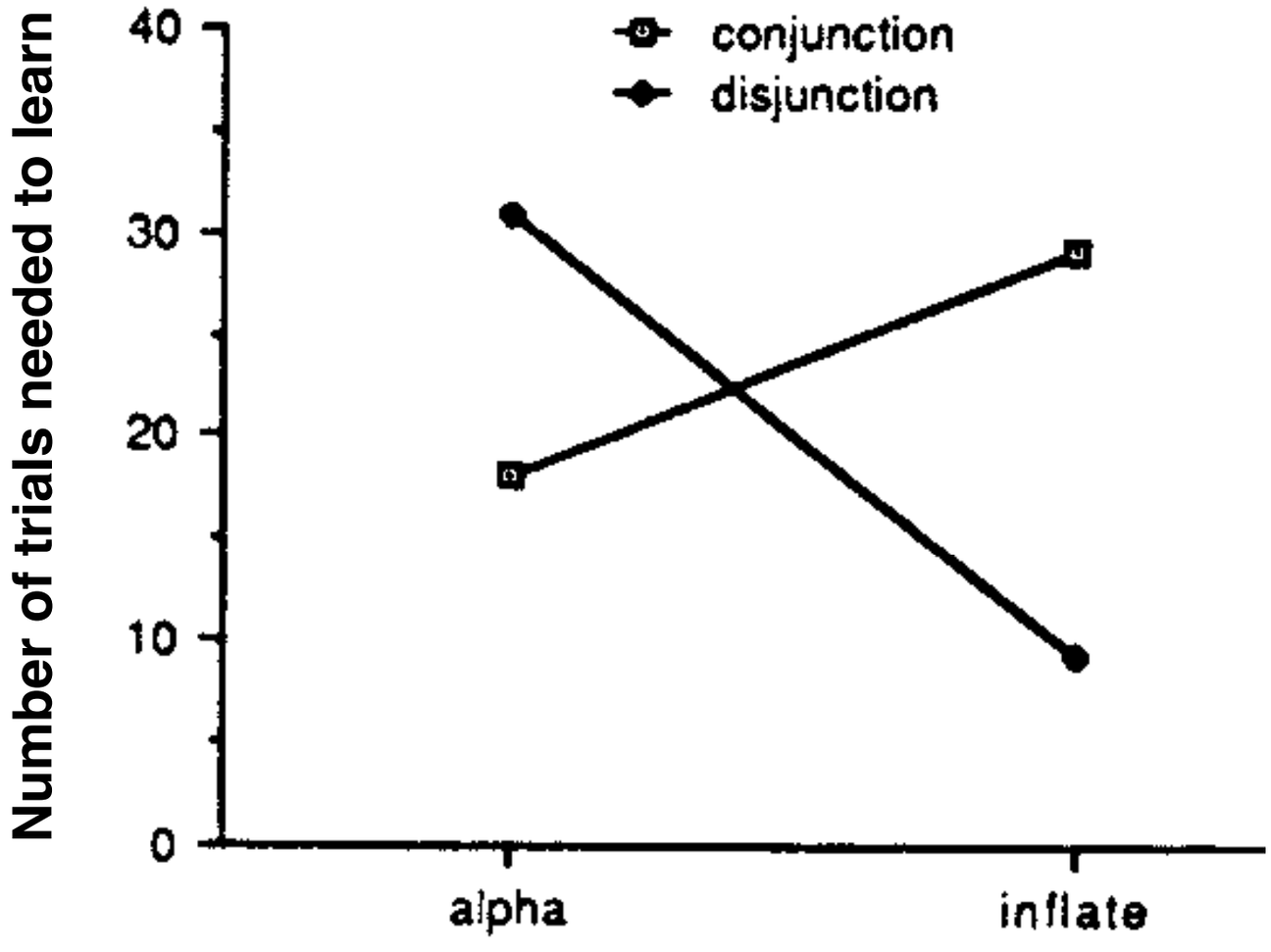
Rules

- Conjunctive: “small AND yellow”
- Disjunctive: “adult OR stretch”

(random photo from web)

Pazzani (1991)

Knowledge influences what is easy to learn. Conjunctive categories are usually easier to learn than disjunctive categories, but knowledge can make this the opposite



Murphy & Allopenna (1994). JEP:LMC

Prototypes of Two Categories: Knowledge Condition

Category 1
- Made in Africa
- Lightly insulated
- Green
- Drives in jungles
- Has wheels

Category 2
- Made in Norway
- Heavily insulated
- White
- Drives on glaciers
- Has treads

(Plus some nonpredictive features)

Murphy & Allopenna (1994)

Prototypes of Two Categories: Neural Condition

Category 1
- Green
- Manual
- Radial tires
- Air bags
- Vinyl seat covers

Category 2
- White
- Automatic
- Non-radial tires
- Automatic seat belts
- Cloth seat covers

(Plus some nonpredictive features)

Murphy & Allopenna (1994)

Results:

- Knowledge condition: learned in 2.2 blocks
- Neutral condition: learned in 4.1 blocks

- It's easy to learn categories that build a coherent model of the object
  - **not just meaningful features in isolation**

(here are more random photos from the web)

Linear separability

- Prototype models can learn linearly separable (LS) categories but not non-linearly separably (NLS) categories
- Exemplar models can learn either, and don't care
- People can learn either equally well (or poorly)

LINEARLY SEPARABLE CATEGORIES

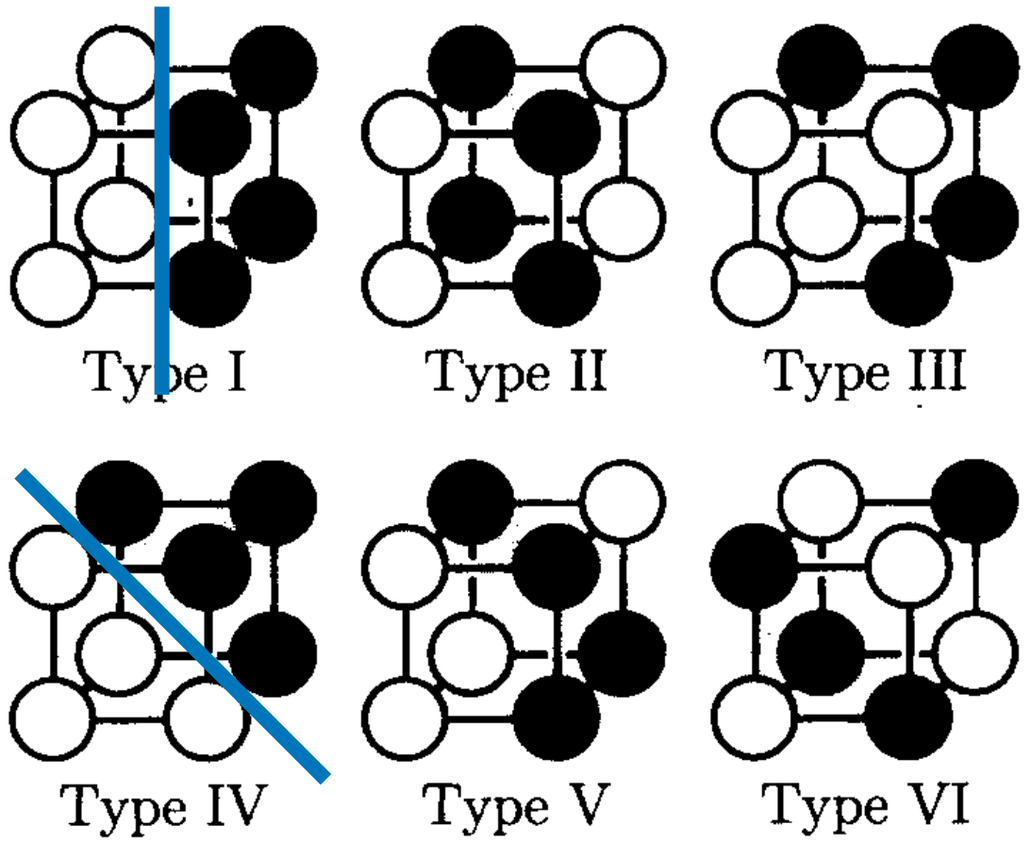
	CATEGORY A				CATEGORY B				
	DIMENSION					DIMENSION			
EXEMPLAR	D_1	D_2	D_3	D_4	EXEMPLAR	D_1	D_2	D_3	D_4
A_1	1	1	1	0	B_1	1	1	0	0
A_2	1	0	1	1	B_2	0	0	0	1
A_3	1	1	0	1	B_3	0	1	1	0
A_4	0	1	1	1	B_4	1	0	1	0

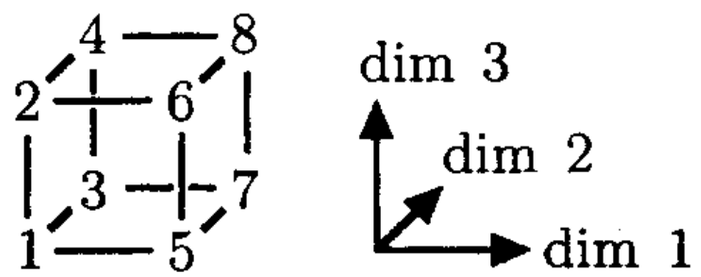
CATEGORIES NOT LINEARLY SEPARABLE

	CATEGORY A				CATEGORY B				
	DIMENSION					DIMENSION			
EXEMPLAR	D_1	D_2	D_3	D_4	EXEMPLAR	D_1	D_2	D_3	D_4
A_1	1	0	0	0	B_1	0	0	0	1
A_2	1	0	1	0	B_2	0	1	0	0
A_3	1	1	1	1	B_3	1	0	1	1
A_4	0	1	1	1	B_4	0	0	0	0

Linear separability: Another view

For Shepard, Hovland, and Jenkins stimuli, only Type I and Type IV are linearly separable.





Wattenmaker, Dewey, Murphy, & Medin (1986)

- Wattenmaker et al. (1986) investigates how category structure interacts with prior knowledge

- Some knowledge just emphasizes specific features
  - suggest summing of evidence

- Some knowledge emphasizes relations
  - suggests configural properties are important

Wattenmaker, Dewey, Murphy, & Medin (1986)

LS vs. NLS category structure

- 1's and 0's referred to honest or dishonest actions (trait condition)
- or else they had various traits mixed together (control)

Prediction

- People should want to sum up trait-consistent features, leading to LS categories being better
- "category A is people who are mostly dishonest"
  - only when there are consistent traits

Wattenmaker, Dewey, Murphy, & Medin (1986)

- It worked
- # of errors:

	Control	Trait
LS	49	**27 (easier)**
NLS	**44 (easier)**	37

- Knowledge can do more than highlight individual features (e.g. Pazzani or Murphy and Allopenna)
- Here (Ex 1), knowledge can make people behave more like a prototype model (or more like an exemplar model, Wattenmakers et al. Ex 3)
- Knowledge does not always strictly lead to easier learning, it depends on the category structure