Categories and Concepts- Fall 2023

Brenden Lake

PSYCH-GA 2207

Course website: https://brendenlake.github.io/CC-site/

Categories and Concepts- Fall 2023

Brenden Lake

PSYCH-GA 2207

Course website: https://brendenlake.github.io/CC-site/

Brenden Lake

Assistant Professor of Psychology and Data Science

brenden@nyu.edu

office hours: Thursdays, 4:30-5:30pm, 6 Washington Place, Room 589. (Also zoom by request in this slot)

https://cims.nyu.edu/~brenden
https://lake-lab.github.io/

Agenda for today

- What is this course?
- Mini-lecture / history lesson, and topic overview
- Introductions and getting to know each other

“Without concepts, mental life would be chaotic” (Smith & Medin, 1981)

chairs

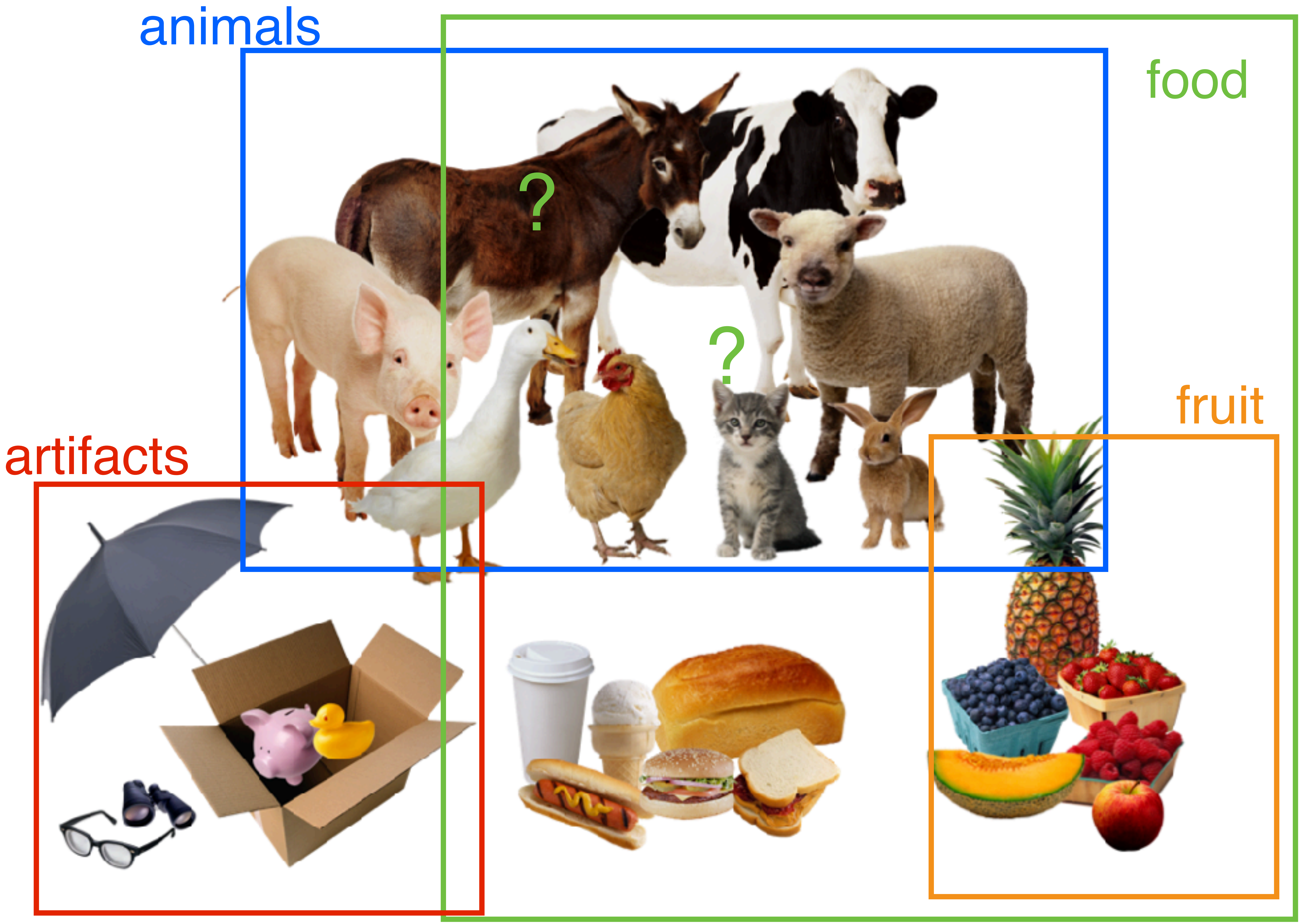
windows

desks

computer


WINNER
BEST
MUSICAL
BILLY
ELLIOT
WHERE
MAXIMUM
PERFORMANCE
LIVES
maxell
MAMMA
MIA!
Helped injured
opponent win.
SPORTSMANSHIP
Bank of America

Concepts can be hierarchical and cross-cutting



Concepts can be relational

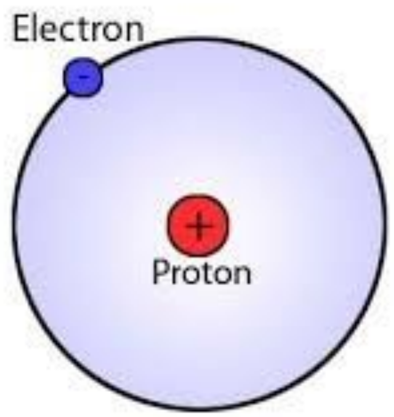
dog and owner ***playing catch*** with a ball

two people **sitting** on a bench and **talking**

people ***watching*** dogs running

Concepts can be abstract

electron



infinity

galaxy

wisdom,
cancer,
love,
peace,
money,
…

Why do we have concepts?

To make inferences: if you know an object is a "dog", you also know that:
-it is a mammal,
-it can bark,
-probably likes squirrels,
-likes to eat doggy chow,
-etc.

To process information efficiently:
You can summarize an event as "I saw a dog." You don't need to store all the pixels in the image for memory, reasoning, etc.

Why do we have concepts?

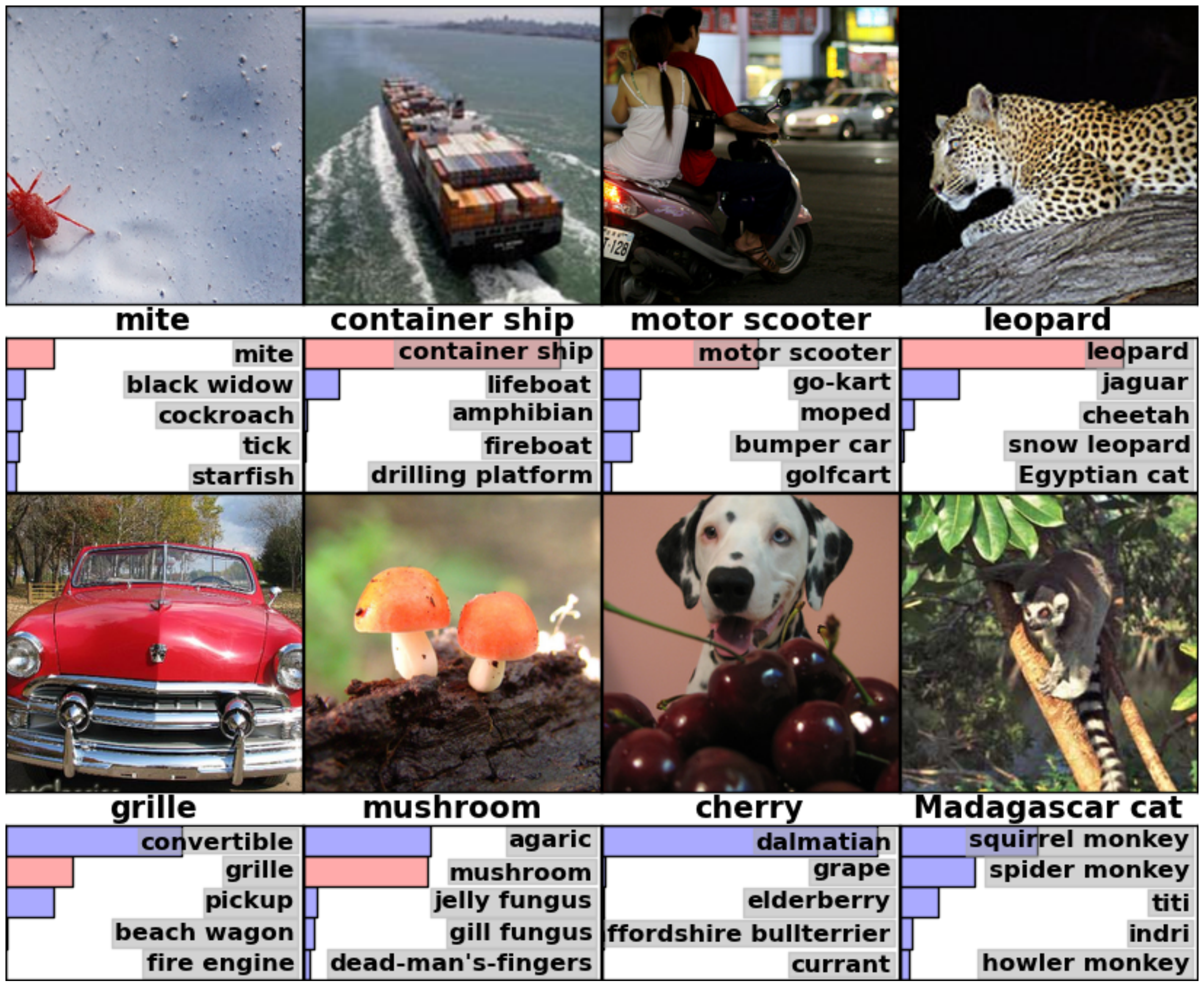
Example: Going to a movie in a new city

1) Identify the theater
2) Identify the box office
3) Buy a ticket
4) Sit quietly through the movie
5) When you see the credits role, get up and leave

None of this is possible without concepts and categories

Intelligent machines need concepts, too!

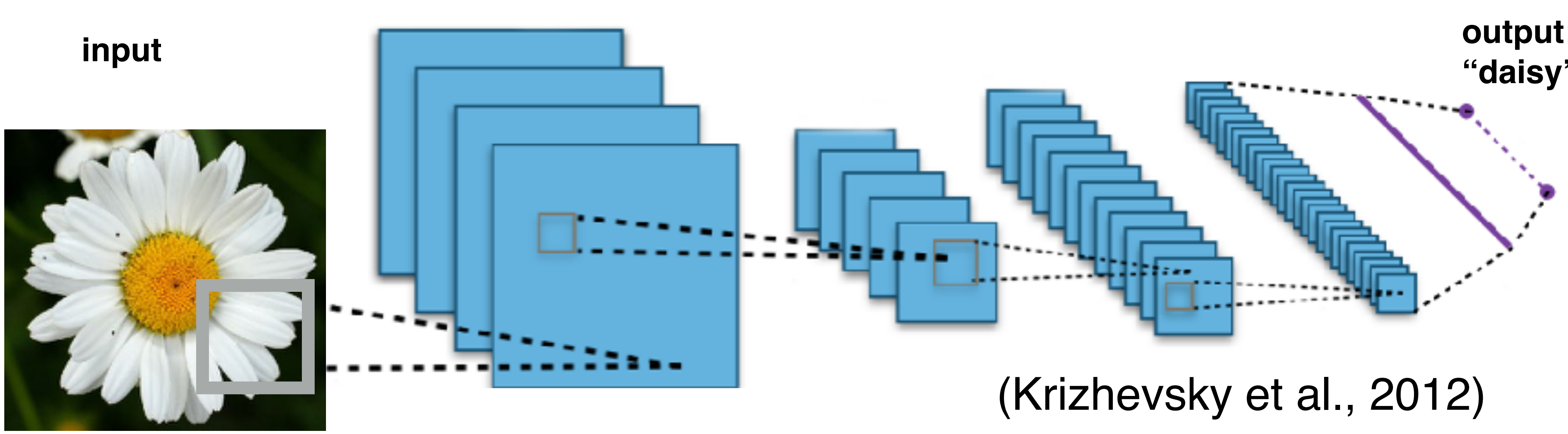
object recognition



finding similar images

probe **similar images…**

layers of feature maps



(Krizhevsky et al., 2012)

Categories vs. Concepts

- A **category** is a set of objects in the world that have some commonality
- A **concept** is a mental representation of a category
- Often one is really referring to both, and so you can use either term
- Psychology often asks what category people are using, as well as what is in their heads

Two strands in the psychology of concepts

1) Experimental psychology, learning theory, computational modeling

2) Knowledge representation and domain knowledge

Strand 1: Experimental psychology, learning theory, computational modeling

This strand focuses on:

- Underlying learning mechanisms
- Effects of category structure on learning and representation
- Laboratory experiments
- Interaction with attention, memory, conditioning, etc.

Strand 2: Knowledge representation and domain knowledge

This strand focuses on:

- Studying existing concepts: fish, weapons, personality types
- These studies cannot manipulate category structures directly— often don't have a "condition 1" and "2"
- Less emphasis on computational modeling and learning, more emphasis on knowledge representation
- Wider populations tested: children, different cultures, different levels of expertise
- Greater emphasis on how concepts relate to one another and to general knowledge

The Classical View

- Actually held classically (Aristotle's *On Categories*)
- View of the person on the street
- Can be identified in many early papers even when it is not explicitly described
- Assumed in some form or another in all work prior to 1970s

The Classical View

- **Concepts can be <u>defined:</u> there are necessary and sufficient conditions for category membership**

- *bachelor*: an unmarried man

- *prime number*: a number not divisible by any number besides “1” and itself

- *triangle*: a polygon with three sides

- **Important implication:** membership is all-or-nothing. All members of a category are equally good members.

Hull’s (1920) thesis

- Supposed Chinese characters
- Subjects saw many different characters and had to learn to provide different names for them
- 12 names and 12 examples per name (i.e., categories)

Category “oo” or “ta”?

oo

Category “oo” or “ta”?

ta

Category "oo" or "ta"?

ta

Category “oo” or “ta”?

oo

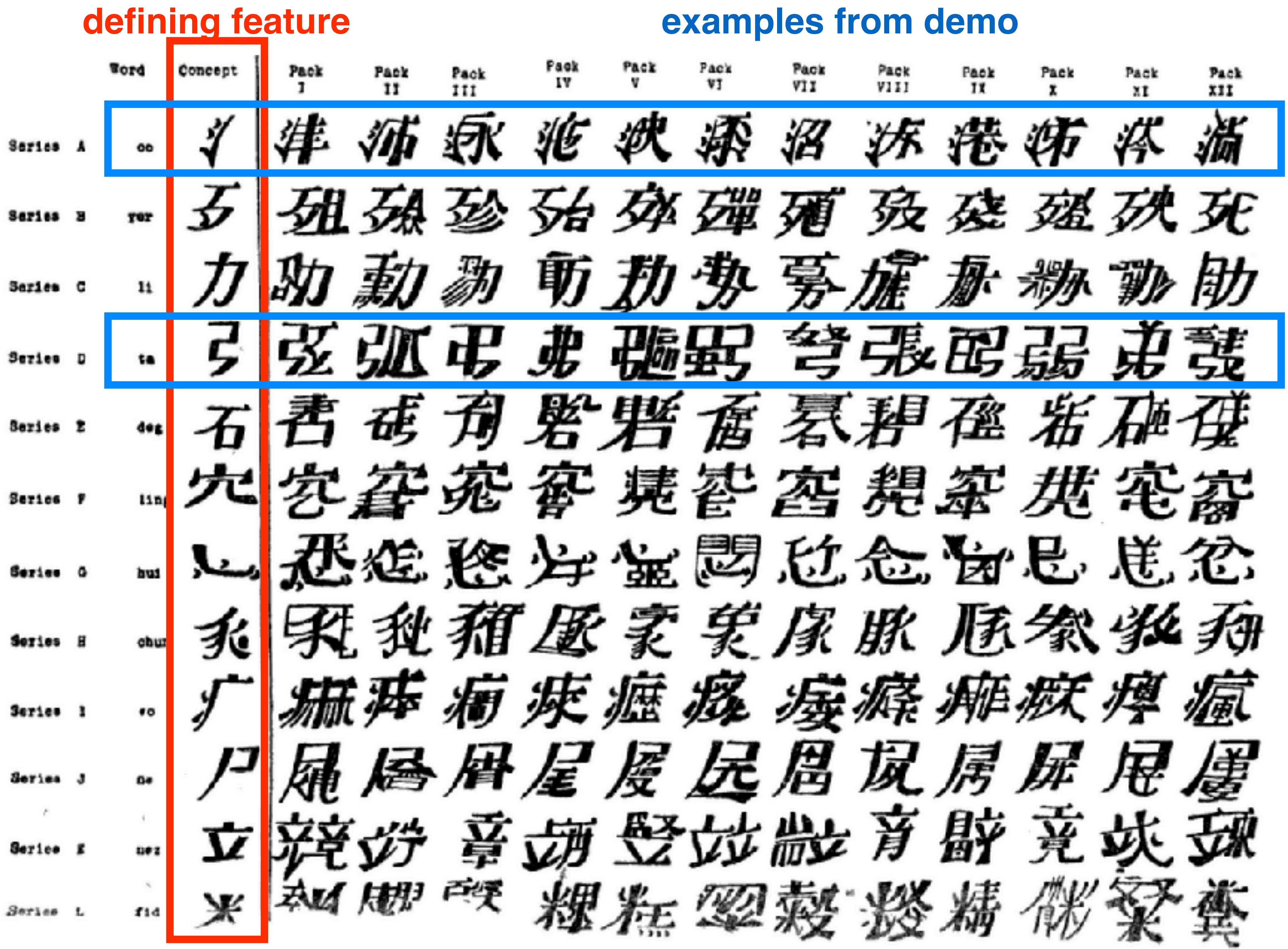
Category “oo” or “ta”?

oo

Category “oo” or “ta”?

ta

The classical view is clear from Hull's writing: "All of the individual experiences which require a given reaction must **contain certain characteristics which are at the same time common to all members of the group** requiring this reaction and which are NOT found in any members of the groups requiring different reactions."



Experiments were harder to run in 1920…

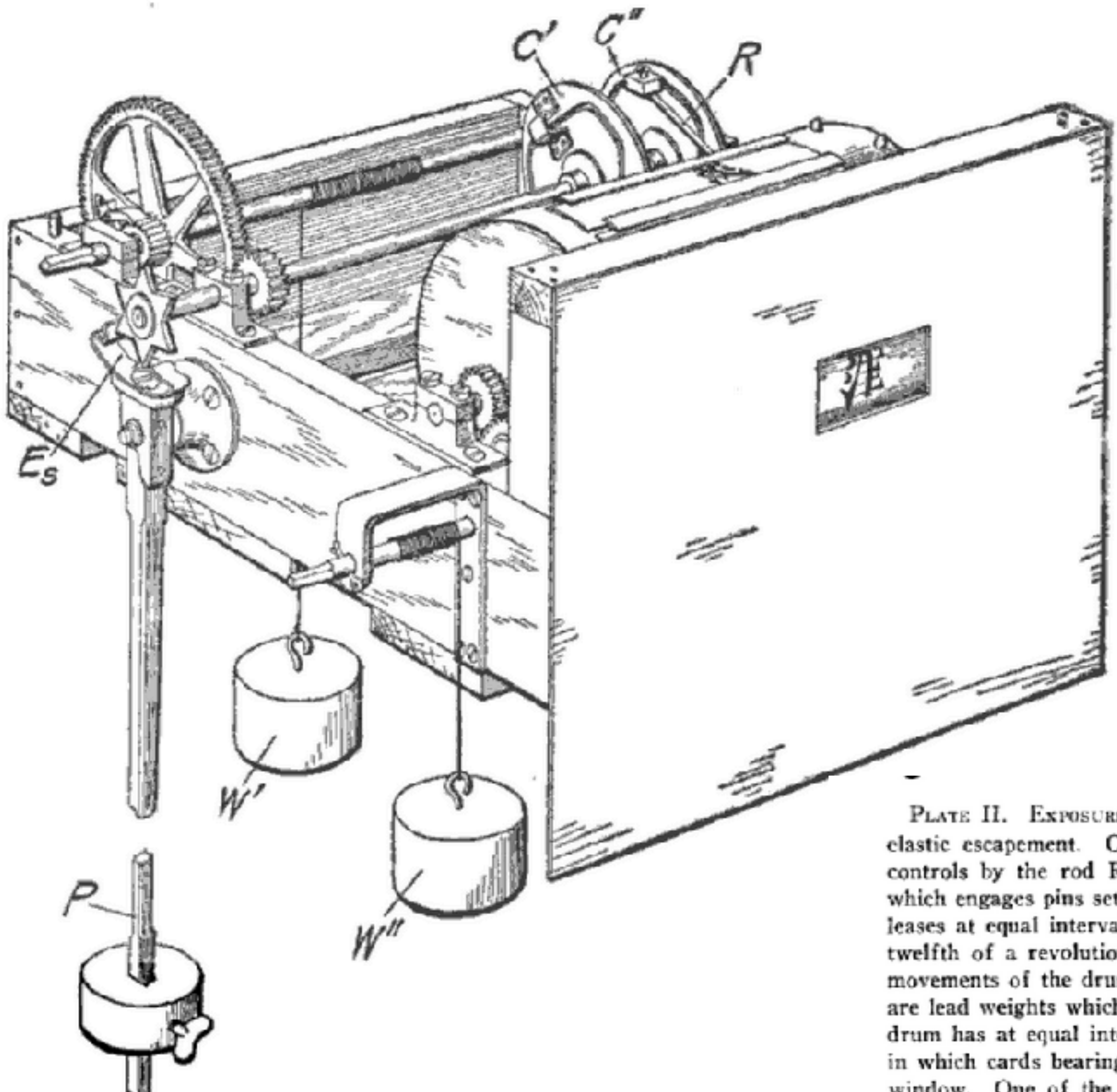


PLATE II. EXPOSURE APPARATUS.—P, pendulum with adjustable bob. Es, elastic escapement. C", cam attached to the escapement shaft. This cam controls by the rod R a secondary escapement (not shown by this view) which engages pins set in the farther end of the drum. This escapement releases at equal intervals of time, permitting the drum instantly to turn one twelfth of a revolution. C' is a second cam which may be used to secure movements of the drum at intervals of time one half as great. W' and W" are lead weights which drive the clockwork and the drum respectively. The drum has at equal intervals on its surface, twelve pairs of shallow grooves in which cards bearing characters are placed for successive exposure at the window. One of the characters is shown at the window as viewed by the subjects in experiments A to I.

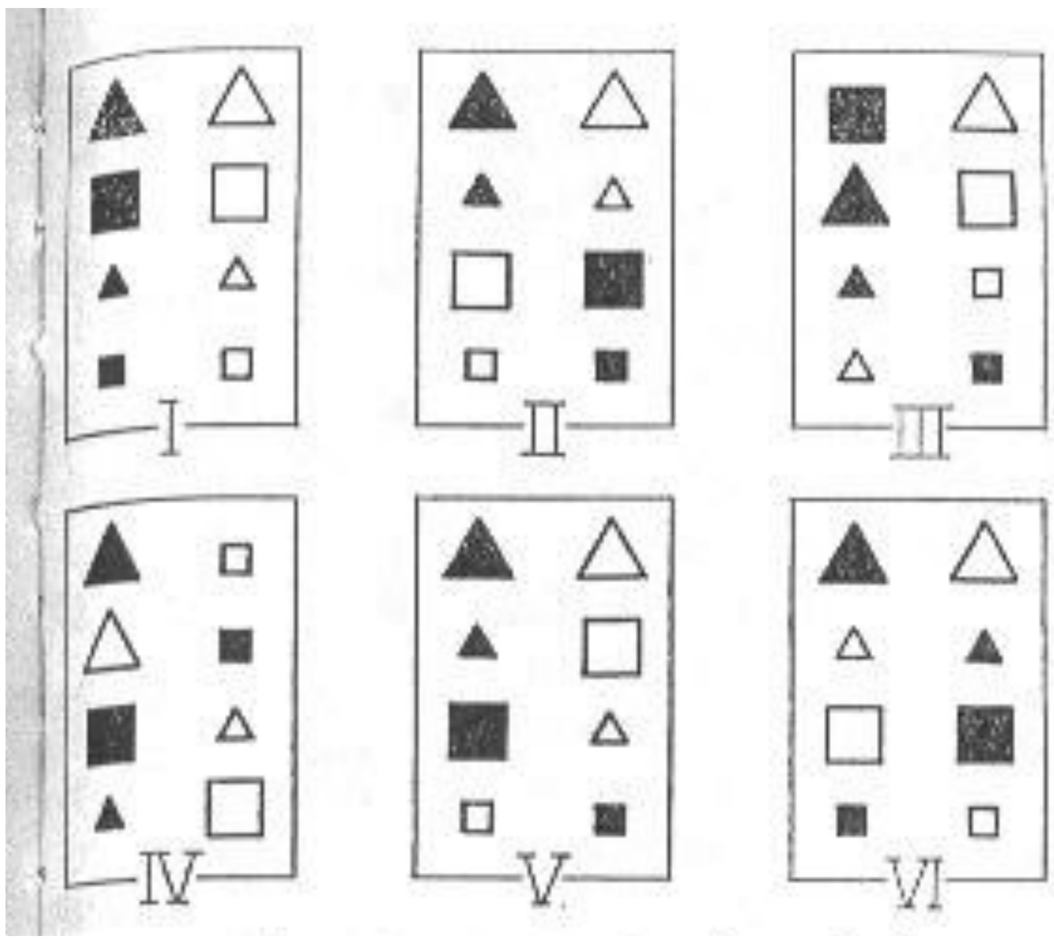
Smoke (1932)

- Seemed to be criticizing the classical view. Wasn't.
- Smoke, after quoting Hull's passage about defining concepts with a common characteristic, writes that "It is our contention that if any concepts have ever been formed in any such fashion, they are very few in number. We confess our inability to think of a single one."
- But his concepts still had clear definitions (just more complex ones):
  - A "dax" is a circle and two dots, one dot inside and one outside the circle
  - A "zif" is three dots, the distance between the two farthest dots being twice the distance between the two nearest dots

What happened next?

- Nothing
- Perhaps due to the dominance of behaviorism, there was very little done in the psychology of concepts for 25 years or so
- Bruner, Goodnow, & Austin (1956) revived the psychology of concepts
  - still well within classical view, but using cognitive approach (people had hypotheses, and learned mental rules)

What did you study in the classical view?



- number of dimensions
- salience of dimensions
- rule type
- stimulus presentation (all at once or sequentially)
- effects of practice
- etc.

- e.g., Shephard, Hovland, and Jenkins (1961) studied the difficulty of learning various rule structures (Type I [easiest] > II > III, IV, V > VI [hardest])

Is the classical view too good to be true?

- Yes
- Next week we will see the downfall of the Classical View.
- For now, let’s talk about the remaining course logistics..

Course agenda

(this, and all material that follows, is on the course website/syllabus)

We have a lecture/discussion format, with a goal of ~30 minutes of discussion per class

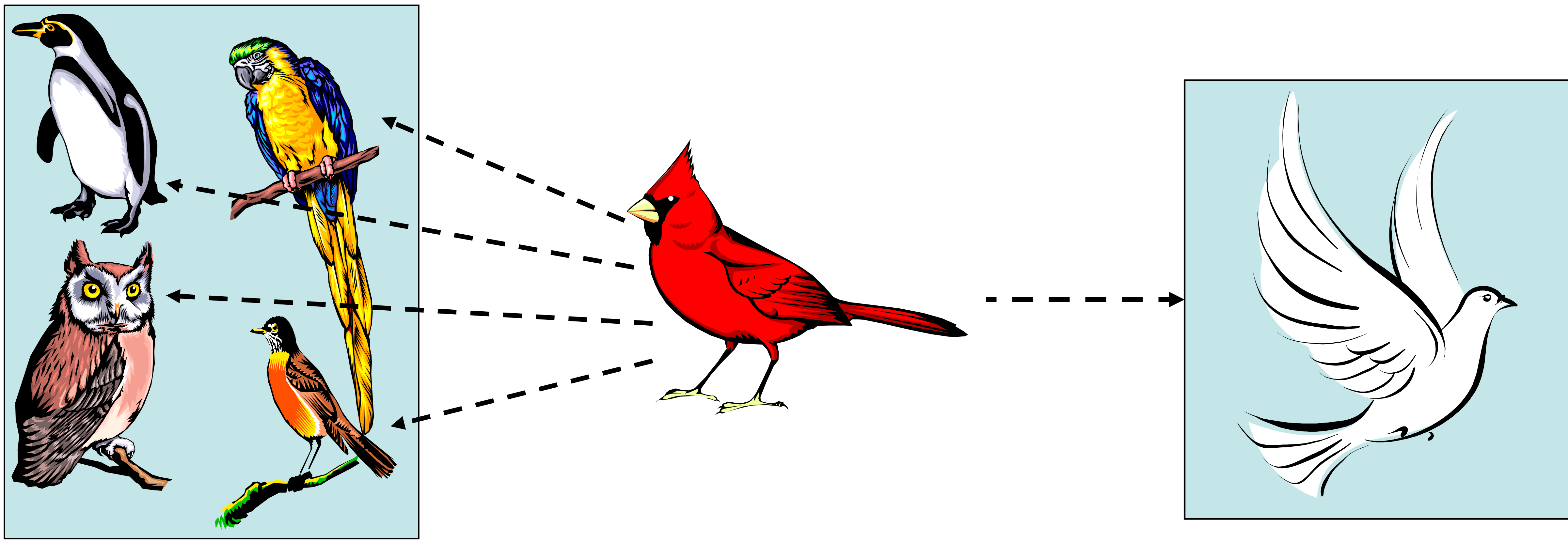
9/11 Introduction; the classical view
9/18 Prototype and exemplar theories
9/25 Concepts as theories and the knowledge view
10/2 Computational models of category learning (part 1)
10/10 (Note special Tuesday time due to Fall recess)
    Computational models of category learning (part 2)
10/16 Computational models of category learning (part 3)
10/23 Computational models of category learning (part 4)
10/30 Taxonomic organization and the basic level
11/6 Category-based induction
11/13 Concepts in infancy
11/20 Conceptual development
11/27 How categories influence perception
12/4 Conceptual combination and exemplar generation
12/11 TBD
Final paper proposal due (Monday 11/13)
Final paper due (Wed 12/13)

(subject to change based on pacing and interest)

Exemplar vs. prototype theories

exemplar theory

prototype theory



Birds
You've Seen

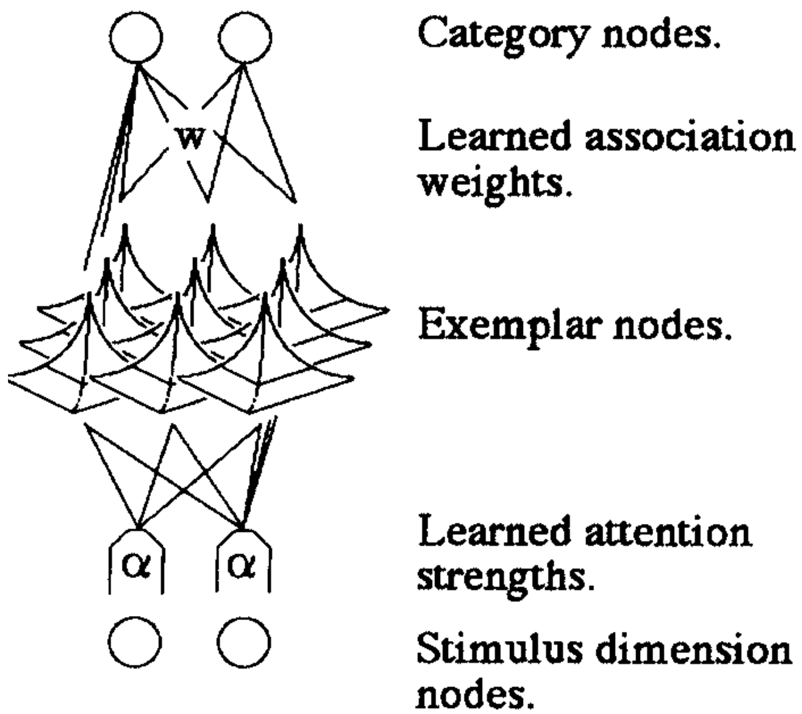
Bird?

Prototypical
Bird

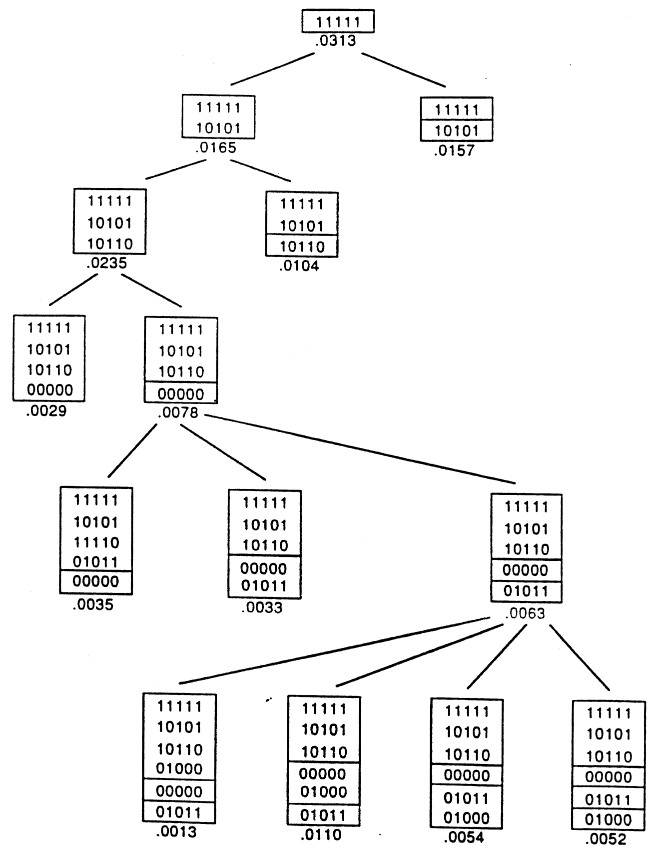
Concepts as theories and the knowledge view

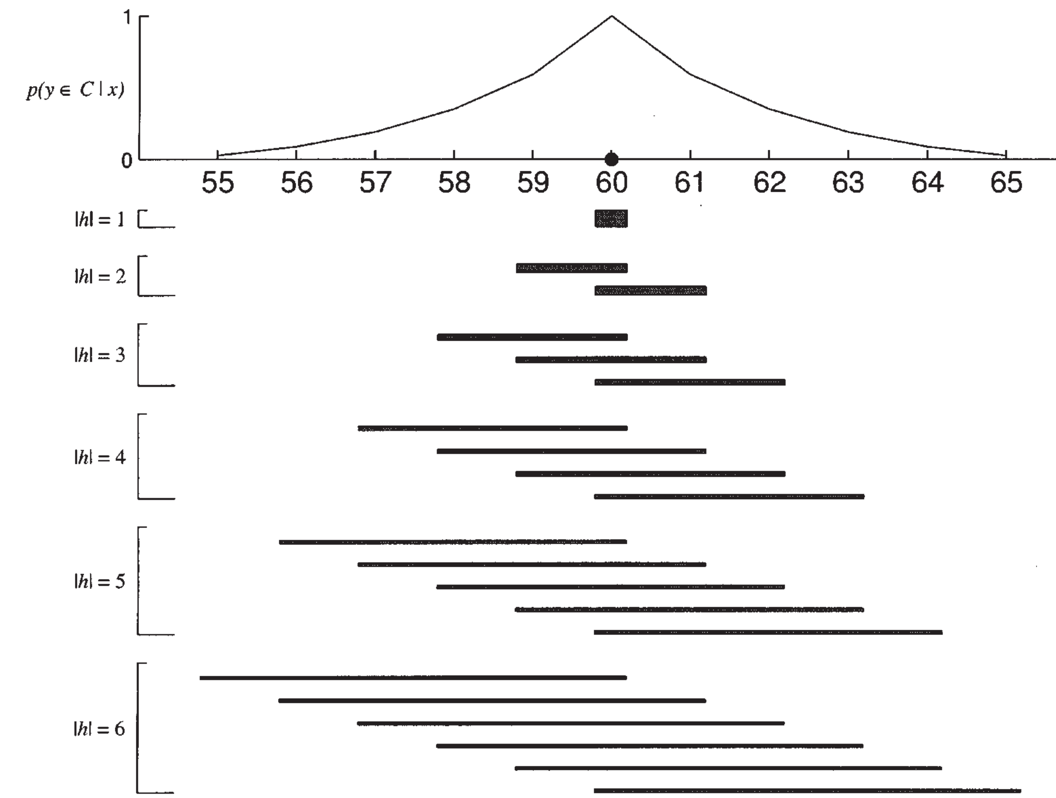
Computational models of category learning

Exemplar / neural net models

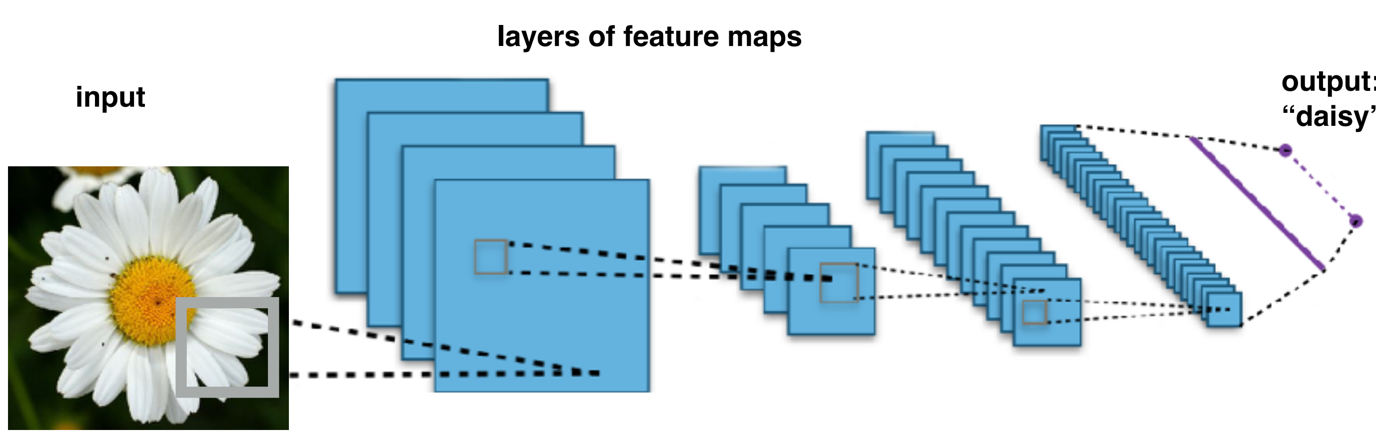


Bayesian concept learning



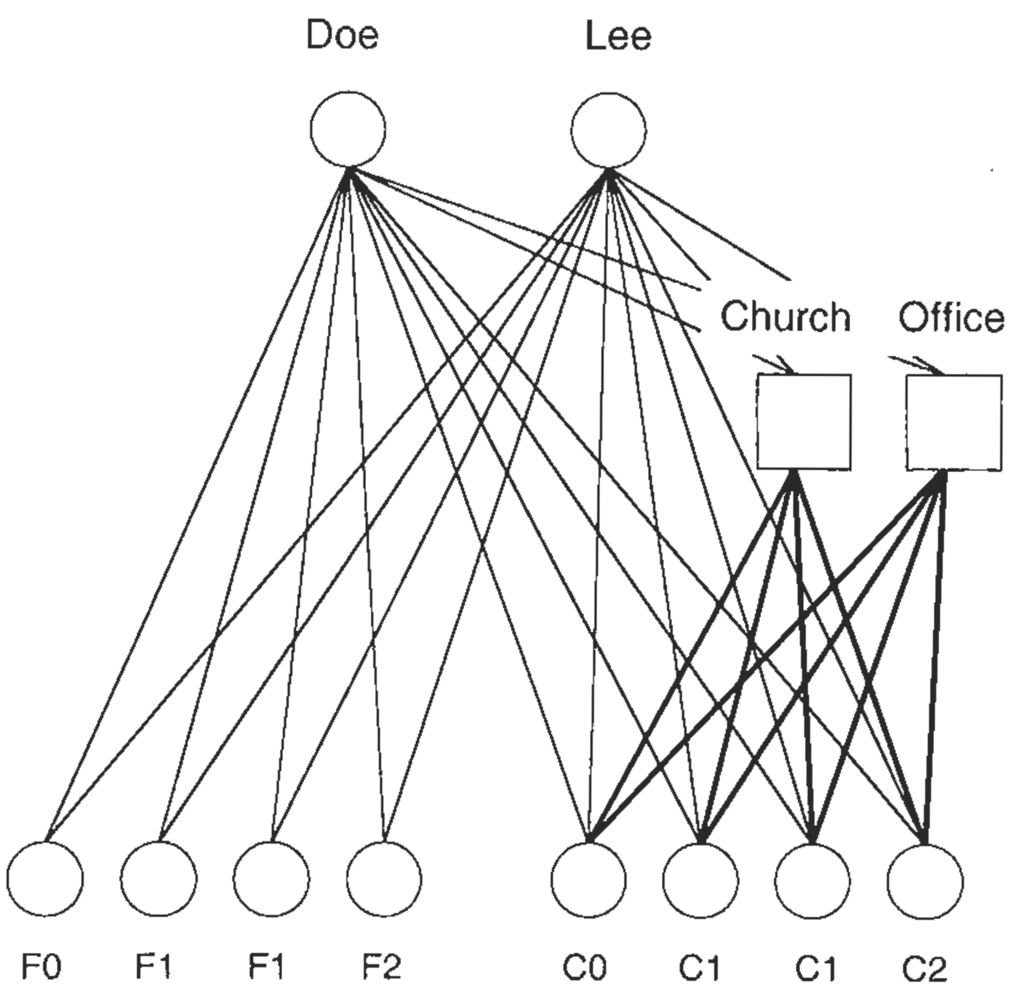


Classification in contemporary AI

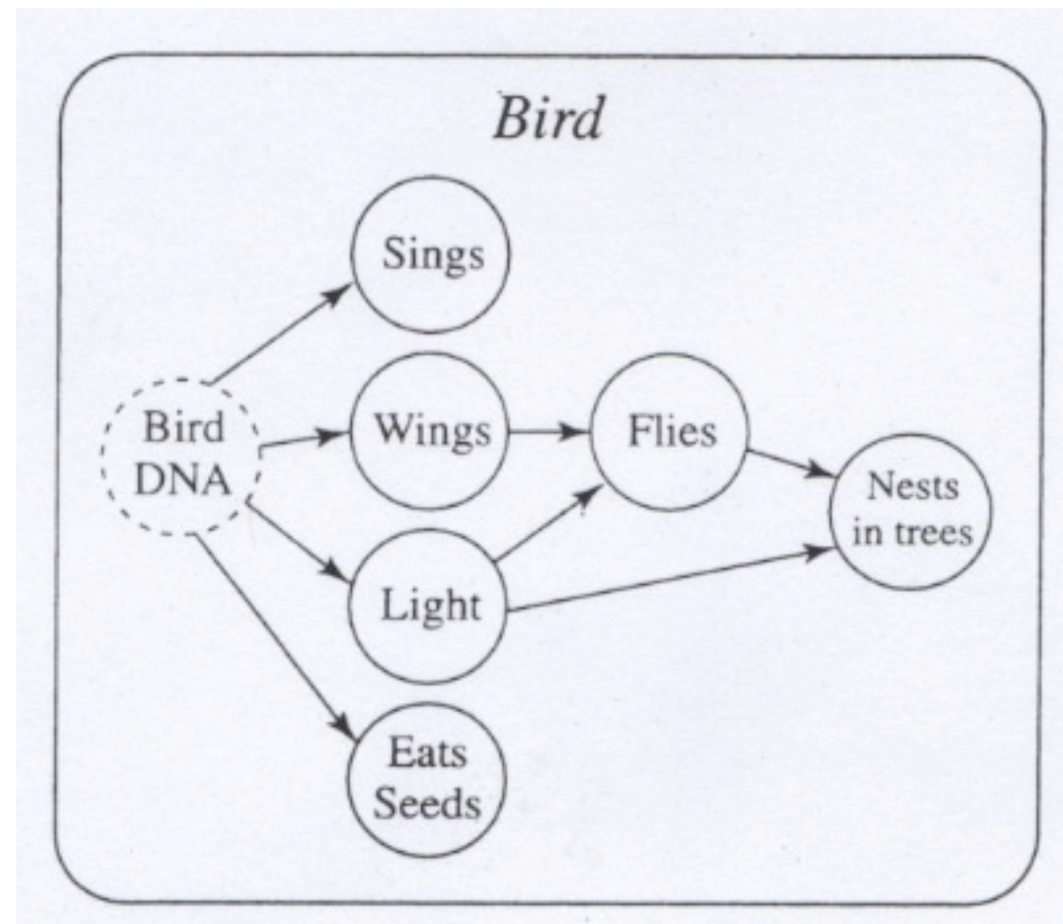


Computational models of category learning

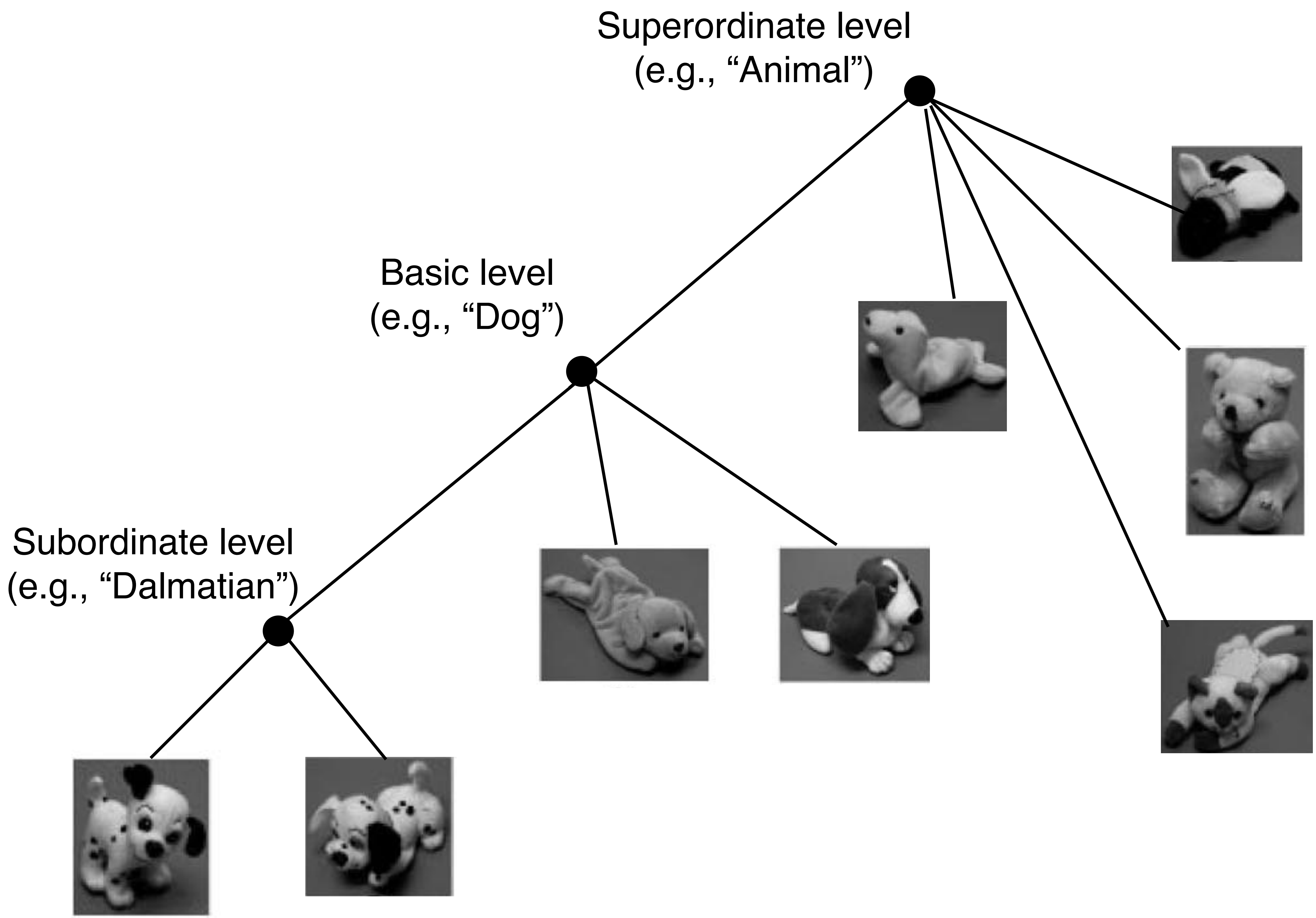
Background knowledge in category learning



Causal models of categorization



Taxonomic organization and the basic level

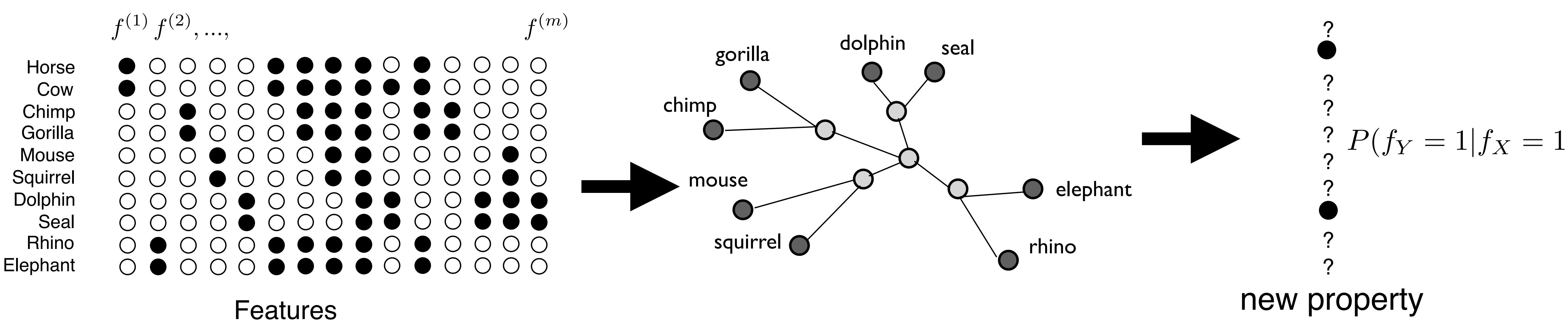


Category-based induction

To make inferences: if you know an object is a "dog", you also know that:
-is a mammal,
-it can bark,
-probably likes squirrels,
-likes to eat doggy chow,
-etc.

Question: "Given that cows and seals have T9 hormones, how likely is it that horses do?"

background knowledge **structure** **inference**



Features for Elephant: 'gray', 'hairless', 'toughskin', 'big', 'bulbous', 'longleg', 'tail', 'chewteeth', 'tusks', 'smelly', 'walks', 'slow', 'strong', 'muscle', 'fourlegs',…

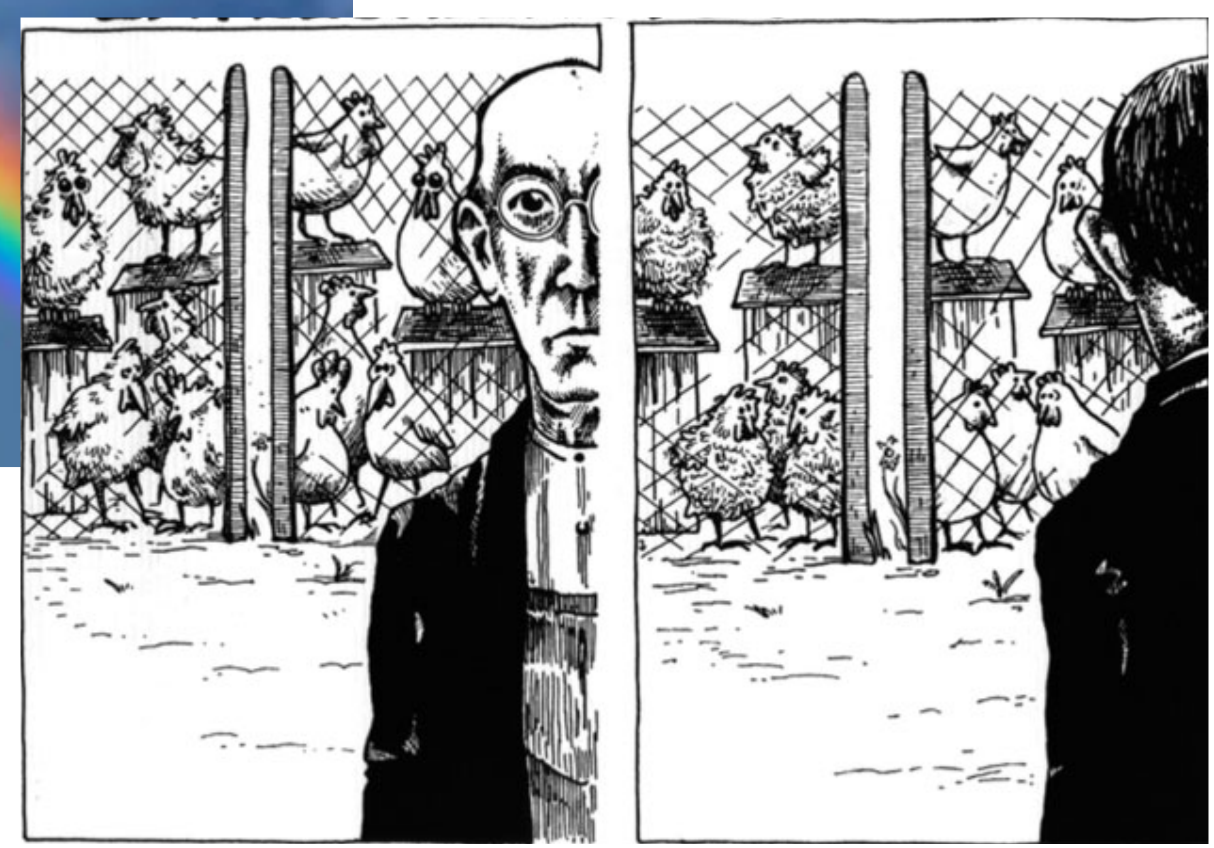
Conceptual development

Figure 1. The animals and vehicles used as stimuli for the global categorization task in Experiment 1.

Here is a “dax”

Which is the other “dax”?

How categories shape perception



Conceptual combination

"pet"

"fish"

? ?

"pet fish"

Be prepared for a lot of reading

- There is a substantial amount of required reading — one book chapter, and 2 or 3 papers per week…
- We will focus on classic empirical and modeling papers

Example from next week:

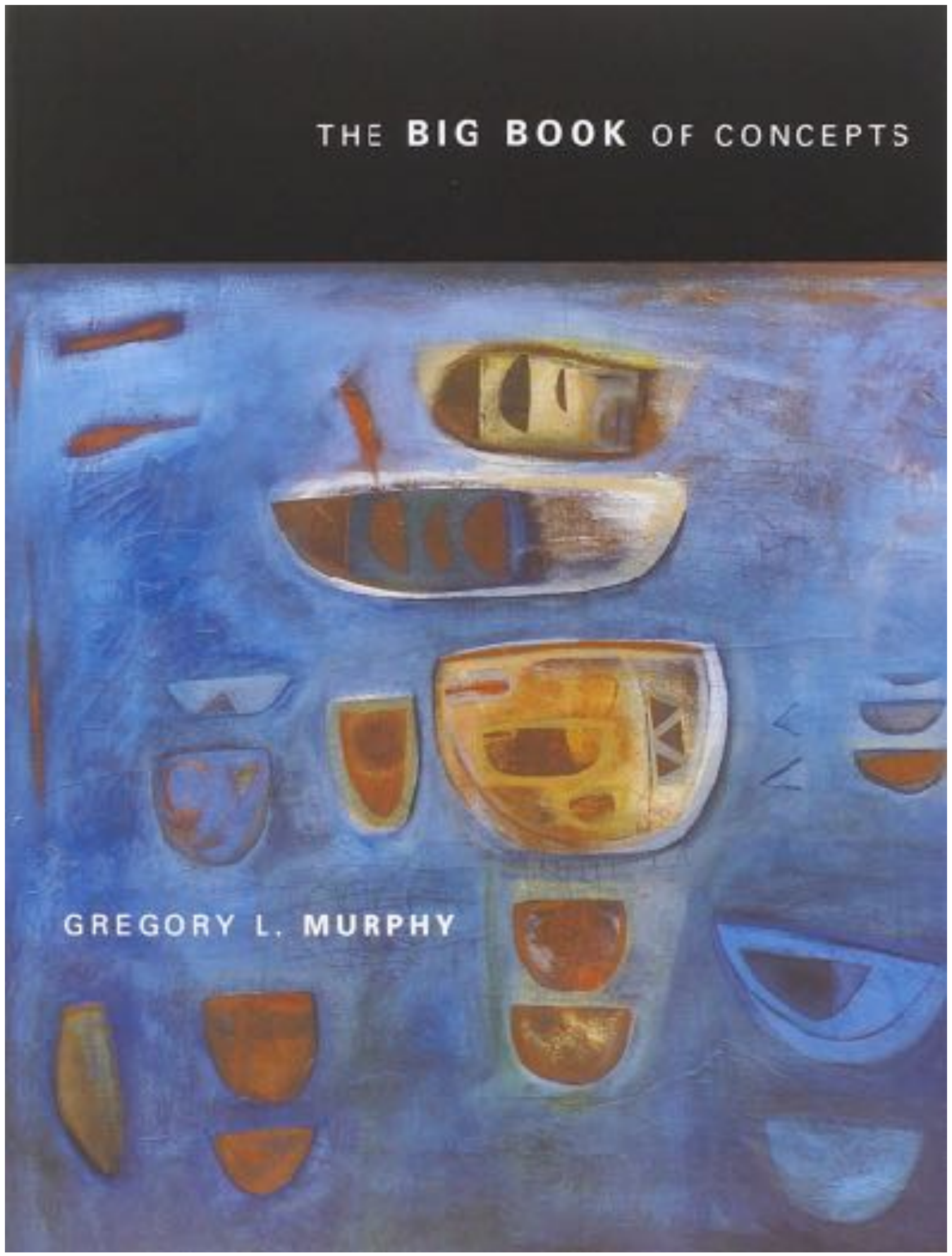
Prototype and exemplar theories

Big Book; Chapter 2 and Chapter 3

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7(4), 573-605.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. Psychological Review, 85, 207-238.

Textbook



Grading

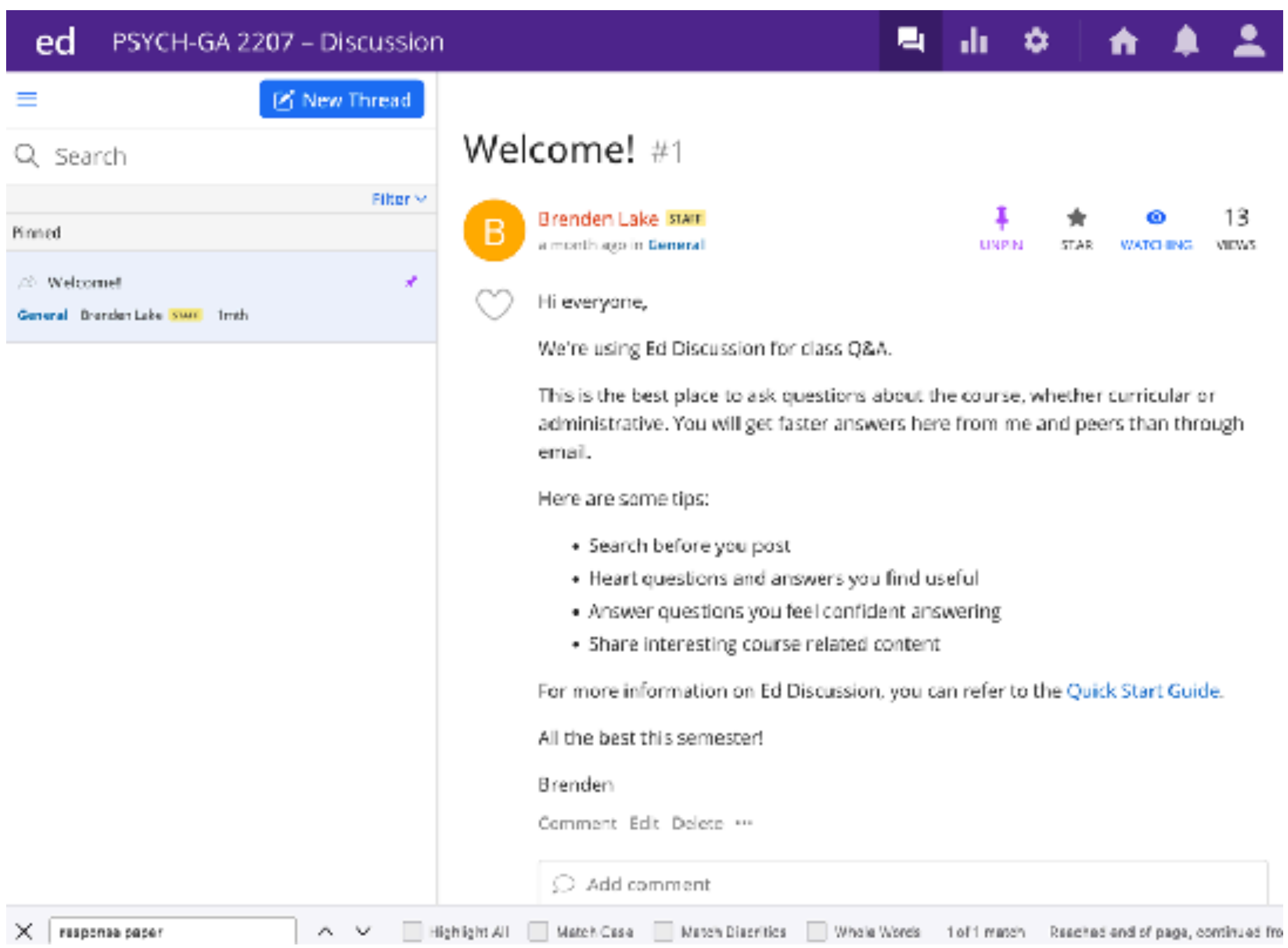
- Final paper (65%)
- Weekly response papers (35%)
- Class participation will be used to decide grades in borderline cases. If you feel sick, you can let me know and please don’t come to class! You can make it up later.

Response papers

- Each week you will write a mini-paper (about 4 paragraphs or about 600 words) in which you will critique the week's readings, discuss an issue raised by it, or propose a new experiment based on it.

- Responses are posted on EdStem (see link on website to sign up)

- Post your response to the class EdStem page before class (the night before would be preferred)



What makes a good response paper?

- There are many ways to write a good response paper, and I would rather leave it up to you than impose a rigid formula.
- Please focus on what issues are most important or interesting and to think about, and what questions are unresolved. Do not give a list of minor questions or flaws.
- Articulate an opinion about the readings, rather than write an exhaustive summary.
- You may skip one weekly response, but any other missed ones will need to be made up.
- Your responses will be graded on a check-plus, check, or check-minus basis, with most responses receiving a check.

Final paper

- Final paper due date is Wed 12/13
- The final paper is written individually (no groups).
- The final paper should address one of the topics covered in the class in more detail. Alternatively, it could investigate a topic that was not covered in class. *Either way, the paper should demonstration your knowledge of the material covered in class.*
- The paper should include a critical review of the literature, along with theoretical conclusions or suggestions for future research. I would expect papers to be about 12 pages long
- If you want to link the paper to your research, that's encouraged.
- Final assignment proposal due on Mon 11/13 (one half page written).

Pre-requisites and expectations

- This course is for graduate students in cognitive science and related fields. All students are expected to have previous coursework in psychology.

- Computational modeling has been central to the study of concepts and categories, and we will cover more computational modeling than in past versions of the class.

    - If you have taken "Computational cognitive modeling” with me, you're in a great position.
    - If you have had linear algebra and statistics as an undergraduate, or MathTools, you will also be in the a good position to understand the modeling details.
    - If you don't have either, don't fret! *Talk to me if you are unsure.*

- Computer programming will not be used in this course, and we will not be implementing models as part of the class. (if you would like to do this on the side, I am happy to provide guidance)

Background survey

- Currently enrolled in what type of program:
  - Psychology Ph.D.? Psychology Masters? Other graduate program? Undergraduate?

- Previous coursework:
  - Cognitive psychology? Cognitive development? Cognitive modeling? Probability, statistics, MathTools?

- Who knows about:
  - Category typicality?
  - Family resemblance?
  - Basic-level categories?
  - Ad hoc categories?
  - Theory theory?
  - ALCOVE model?
  - Backpropagation?
  - Bayes' rule?

What you will come away with…

A deeper understanding of the major topics in the psychology of concepts:

1. Classic empirical findings
2. Main theories of concepts, including the classical view, prototype models, exemplar models, and the knowledge view
3. Influential computational models of concept learning, with a focus on neural networks and Bayesian models
4. Other key topics including taxonomic categories, category-based induction, conceptual development, categorical perception, and conceptual combination

Questions?

Introductions

1) Your name
2) department / degree program
3) Why are you taking this course? What do you hope to learn?