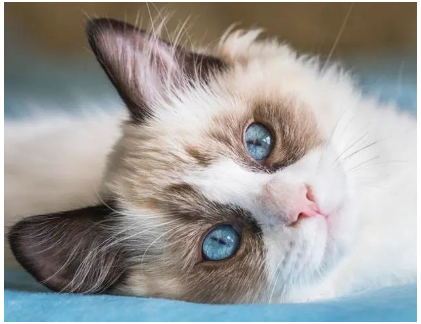
Bayesian Concept Learning

Chelsea Zou

Question: How Do Humans Learn and Generalize Concepts from Specific Examples?

"Cats"



Is this a cat?

Background - Concept Learning

Searching for attributes that can be used to distinguish exemplars from non exemplars of various categories

Physical

Contextual

Background - Concept Learning

Classical Approach: Assigned category if exemplar meets ***definition*** of the category

Prototype Theory (Rosch 1973): Compares exemplar to a ***statistical*** **average** of all known instances

Exemplar Theory (Nosofsky 1991): Storing ***all*** instances in ***memory***

Background - Bayesian Learning Framework

Tenenbaum (1998) first proposed the formalization of **Bayes Theorem** with concept learning in a rectangular guessing task

- **Integration of Prior Knowledge**: Learners combine prior knowledge with new observations.
- **Efficient Decision-Making:** Enables efficient judgments about new information.

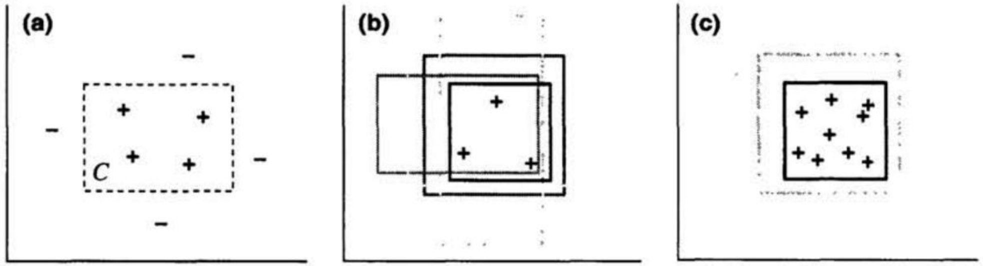


Figure 1: (a) A rectangle concept C. (b-c) The *size principle* in Bayesian concept learning: of the many hypotheses consistent with the observed positive examples, the smallest rapidly become more likely (indicated by darker lines) as more examples are observed.

- Subjects had to guess 2D rectangular concepts based on dots (observations) that represented instances of healthy levels of insulin and cholesterol.
- The number of observations was manipulated during the experiments
- Results: People tend to guess rectangles larger than the examples they saw if ranges of rectangle was larger. Interestingly, if there were more observations, people tend to make smaller sized guesses.

Bayes Theorem → How do humans update their beliefs in light of new information?

- **Prior *p(h)* =** Prior knowledge state
- **Likelihood *p(d|h)* =** New information/observations belonging to a concept
- **Posterior *p(h|d)* =** Updated beliefs about a concept

Aim: How does the number of observations shown affect our ability to update our knowledge and learn concepts?

$$\overset{\text{Posterior probability}}{p(h \mid d)} = \frac{\overset{\text{Likelihood}}{p(d \mid h)}\,\overset{\text{Prior probability}}{p(h)}}{\sum_{h' \in H} p(d \mid h')\,p(h')}$$

Disclaimers

1. **Controversial**: Quite a leap to go from formalisms of probability theory to something as complex as cognition

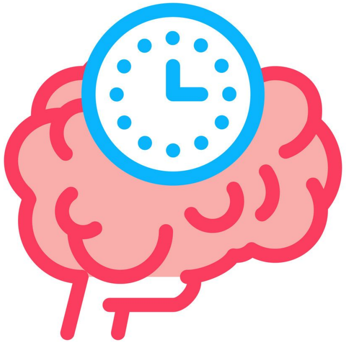
   → could just be a satisfying post-hoc explanation for how cognition works

2. Design of experiment **oversimplifies** something complicated and theoretical
3. Nonetheless, this experiment is still worthy of exploring an interesting analogy

Hypothesis

As the number of positive training observations increase…

1. Accuracy of correctly judging whether a new instance belongs to a concept will **increase**
2. Reaction time will **decrease**



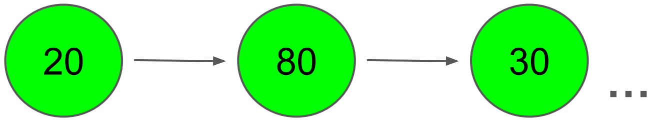
Methods

- Eight participants were recruited
- Four different number concepts were used
  - Even → {2,4,6}
  - Odd → {1,3,5}
  - Multiples of ten → {10,20,30}
  - Negative → {-5,-6,-7}
- **IV:** Number of training observations
  - 1. Sequence of **three** numbers shown
  - 2. Sequence of **seven** numbers shown
- **DV:** Accuracy & response time

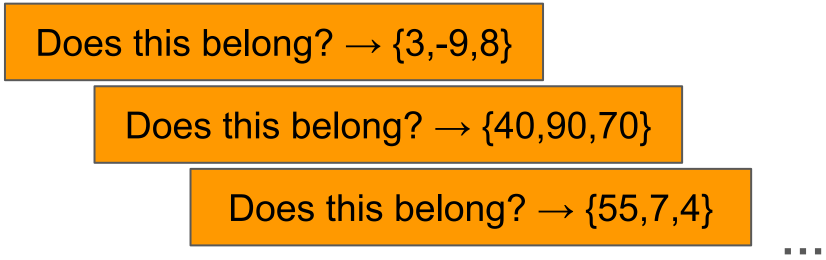
Procedure

Participants experienced four sessions (one for each number concept)

1. For each session → two phases:
2. Phase I (Train): **Three numbers** belonging to concept displayed for two seconds each
3. Phase I (Test): Participants immediately answered six questions sets determining if a group of three numbers belonged or did not belong to the concept.
4. Phase II (Train): **Seven numbers** belonging to concept displayed for two seconds each
5. Phase II (Test): Same as Phase II but with different numbers



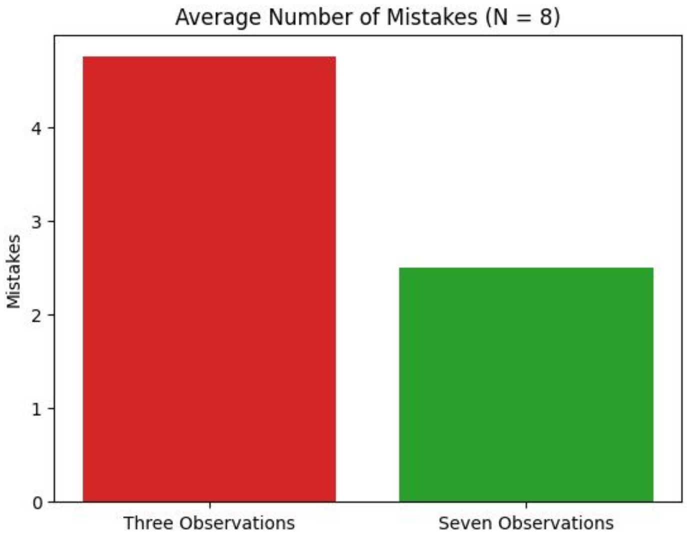
TRAIN

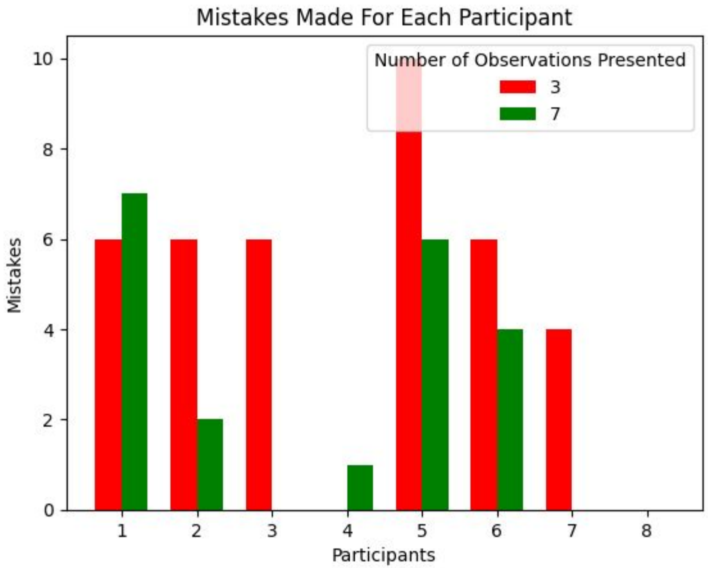


TEST

Results: Accuracy

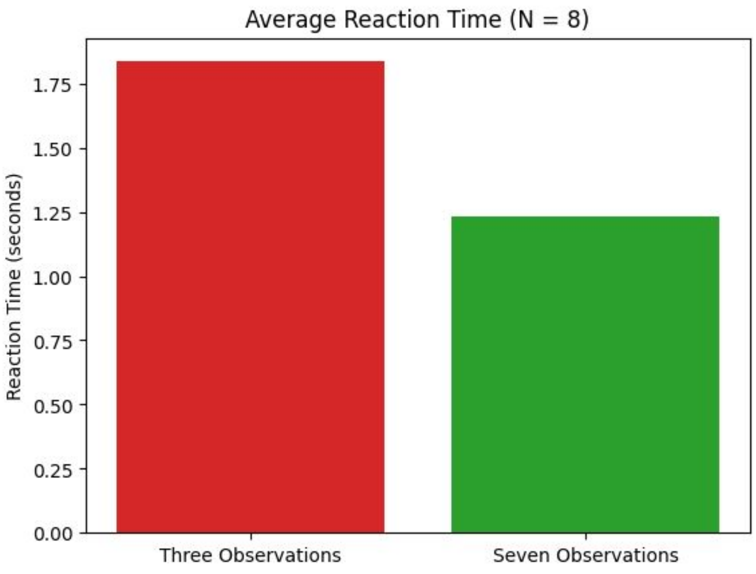
- Average number of mistakes decreased with more observations
- Accuracy improved from 20.8% to 58.3%
- Paired t-tests revealed significance: $t(7) = 2.32$, $p < .05$, $d = 0.57$

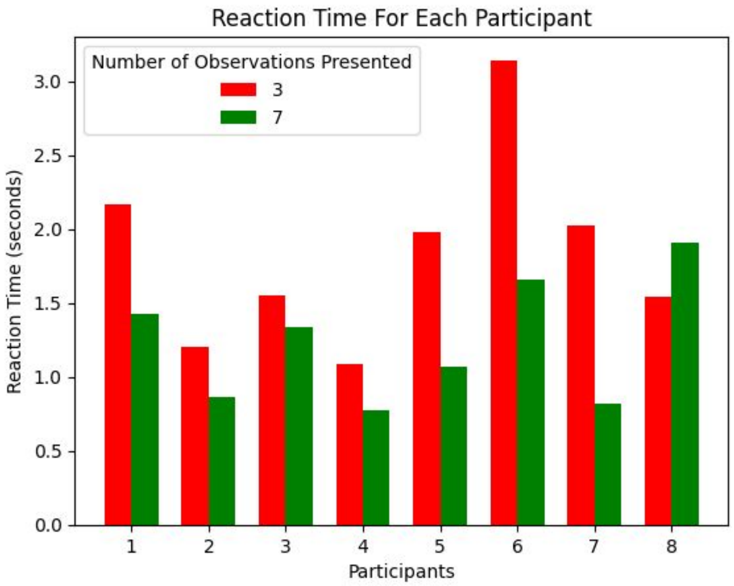




Results: Reaction Time

- Average reaction time decreased with more observations
- Reaction time decreased from 1.83 to 1.23 seconds
- Paired t-tests revealed significance: *t(7)* =1.12, $p < .05$, d =0.34



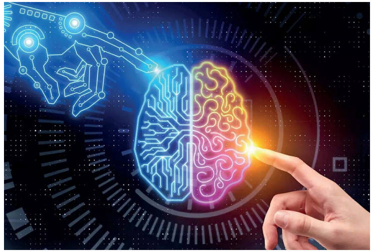


Discussion - Limitations

Besides sample size…

1) **Homogenous** group of participants in terms of age and educational background restricts external validity of the results
2) Because it was designed as a numerical task, **individual differences in mathematical ability** can bias results
3) Because participants did not know the possible options of the four concepts, since number properties **are not mutually exclusive** → chance that participant found some other pattern between concepts that did not necessarily correspond to one of the four, but can be correct as well
4) **Ecological validity**: Using number properties as concepts may undermine the complexity of real world concepts in dynamic settings

Discussion - Future Directions & Applications



Future Questions:

1) **Transferability of Bayesian framework:** To what extent can the Bayesian framework be applied to diverse conceptual domains beyond numerical categories?
2) **Long-term concept retention:** What is the durability of knowledge acquired through Bayesian-informed learning as time passes?

Potential Applications

1) **Educational Settings:** Curricula could be designed to provide increased number of examples to optimize knowledge acquisition
2) **Artificial Intelligence:** Integrating computational Bayesian frameworks into machine learning models may enhance the ability to generalize and adapt to new information

References

Tenenbaum, J. (1998). Bayesian modeling of human concept learning. *Advances in neural information processing systems,* 11.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive science*, 32(1), 108-154.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing Bayesian inference. *Psychonomic bulletin & review*, 17(4).

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. J*ournal of experimental psychology: human perception and performance*, 17(1), 3.

Rosch, E. (1973). Prototype theory. *Cognitive development and the acquisition of language*, 111-144.

Tenenbaum, J. (1999). Rules and similarity in concept learning. *Advances in neural information processing systems*, 12.


Thank You for
Listening
Questions?