# Language Acquisition Computational Intelligence 4/7/05

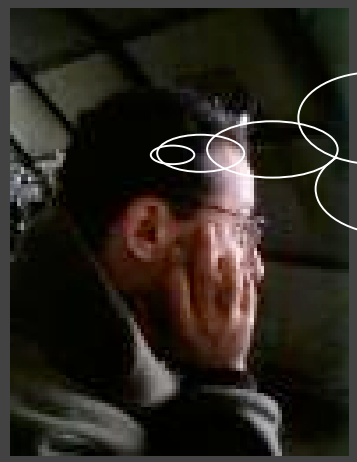
LouAnn Gerken

# Overview

- The language learner's problem space
- A sample of computational capacities of human infants
- Turning the induction problem into a deduction problem

# The Language Learner's Problem Space

# Language vs. Communication

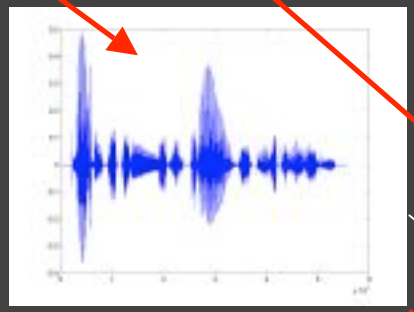


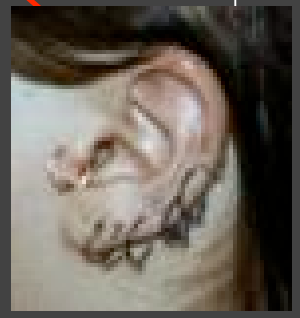
Did it hurt to get all of those holes in your ear?

Interpret structure

Parse sentence structure

Recover words from acoustics





Mental message (coded in some lingua mentis)

Create hierarchical structure and find words to convey message

Ordered sequence of language specific sub-meaning units

# Language vs. Communication

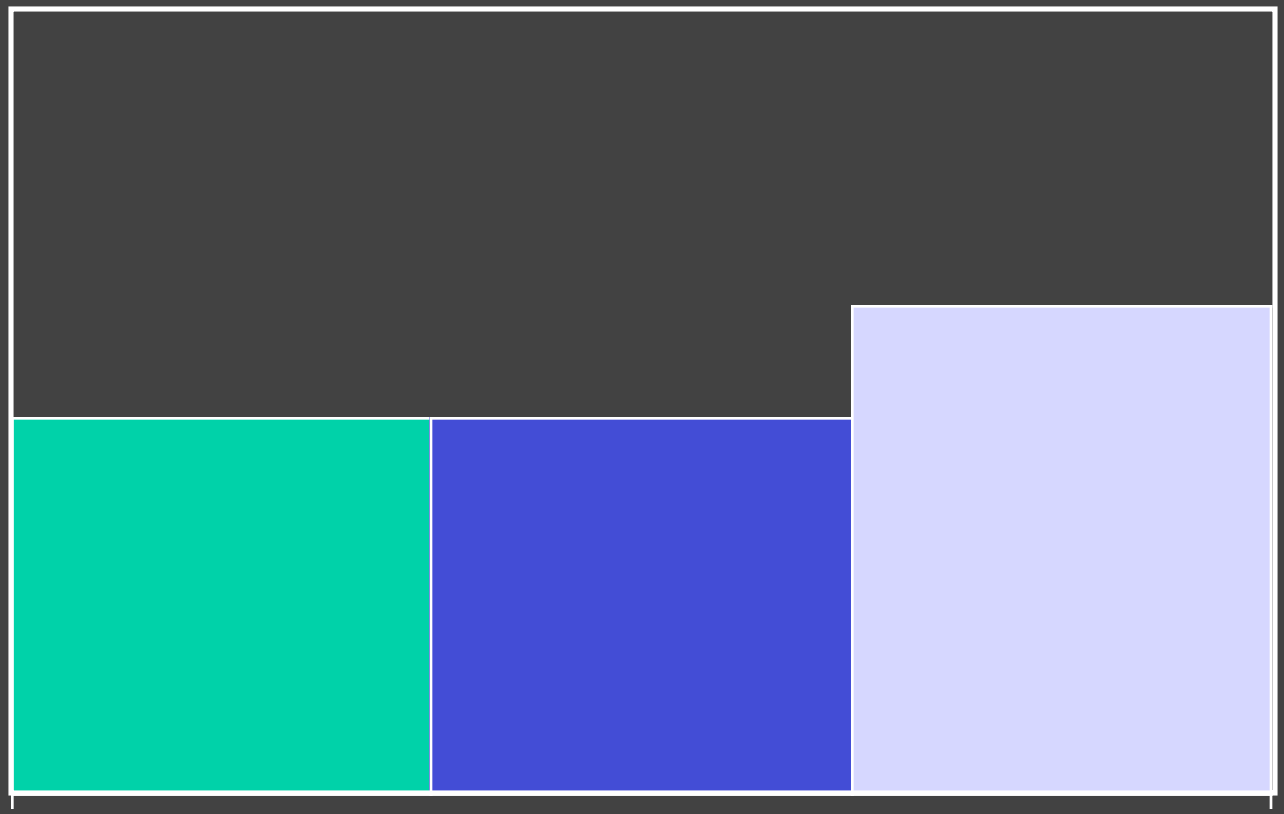
- Language researchers generally reserve the term ‘language’ to refer to a system with ‘duality of patterning’
  - Combinable sub-meaning units
  - Combinable meaning units
- Communication is any system that conveys information from one organism to another.

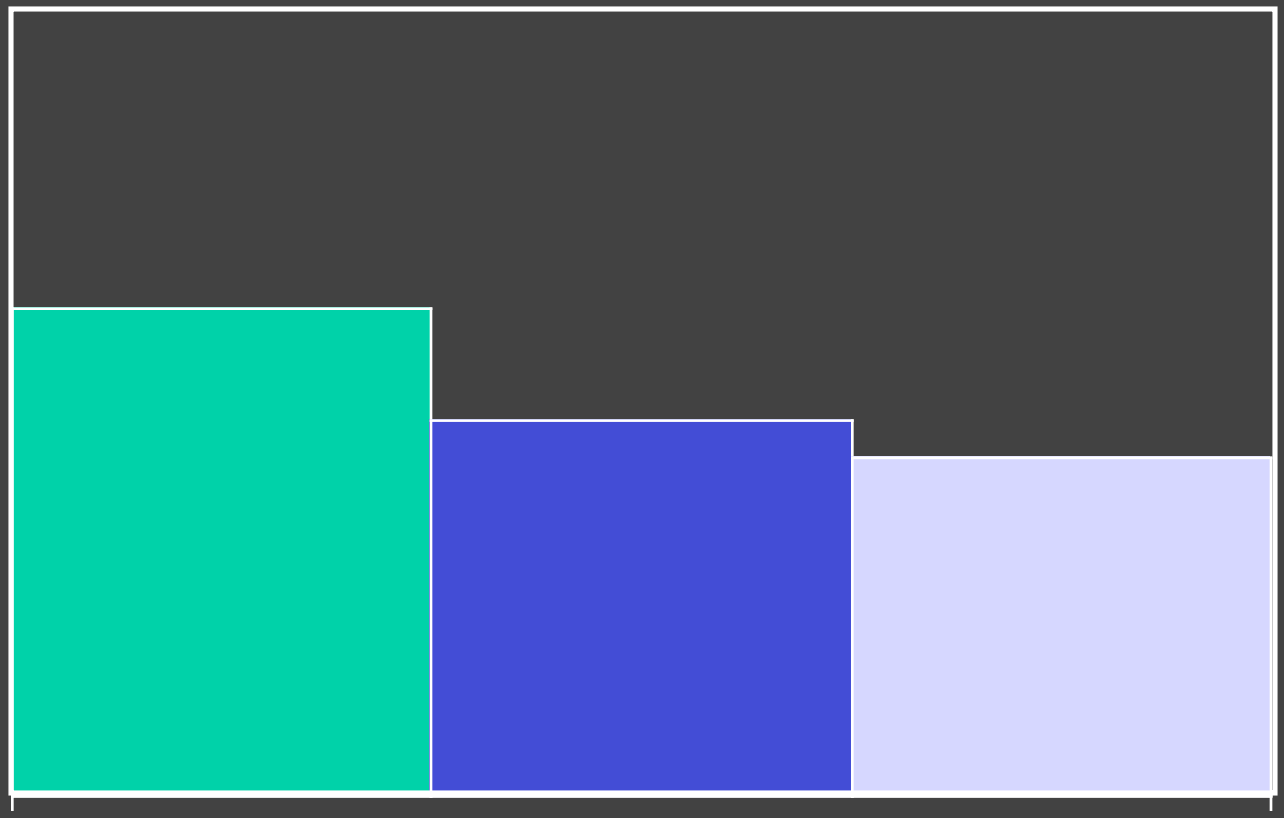
# What problems face a language learner?

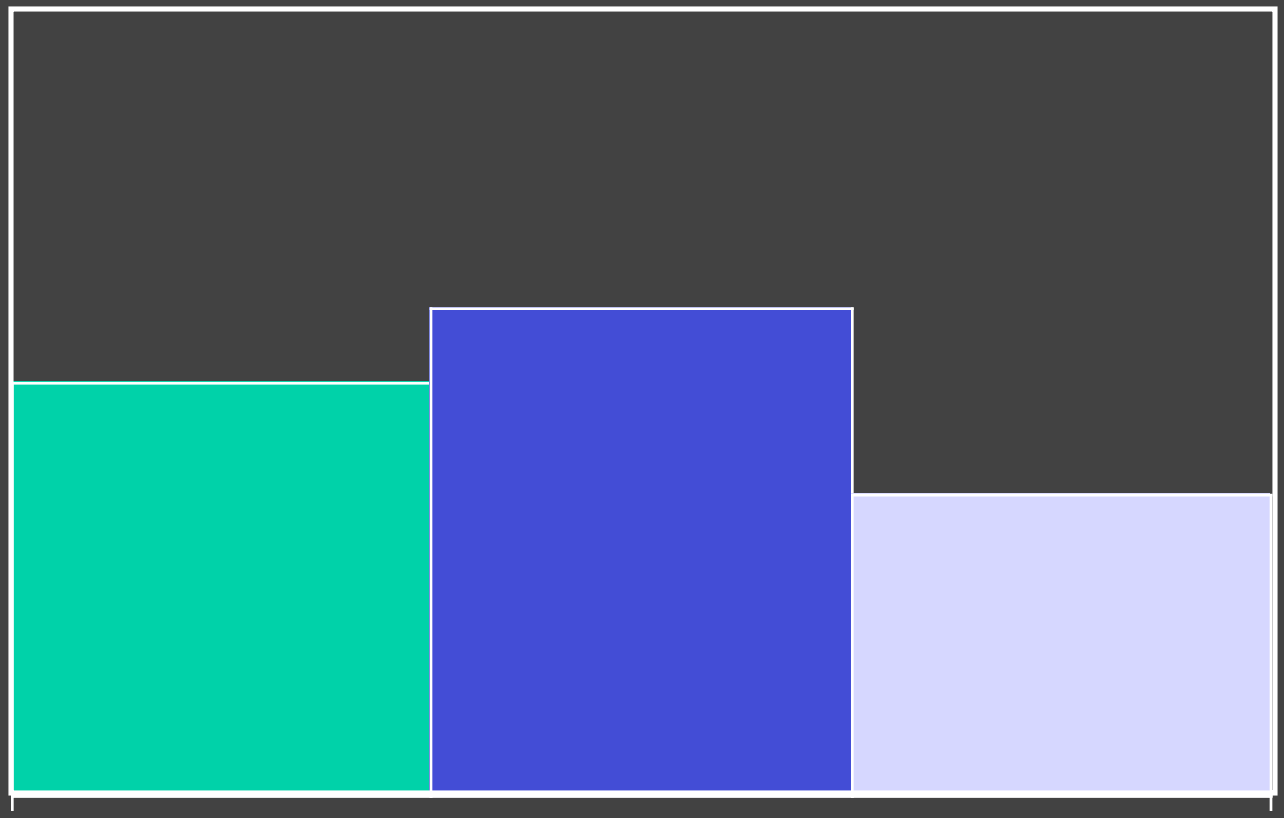
- Induction problem (at 2 levels)
- Poverty of the stimulus (at 2 levels)
- Noisy data (at 2 levels)

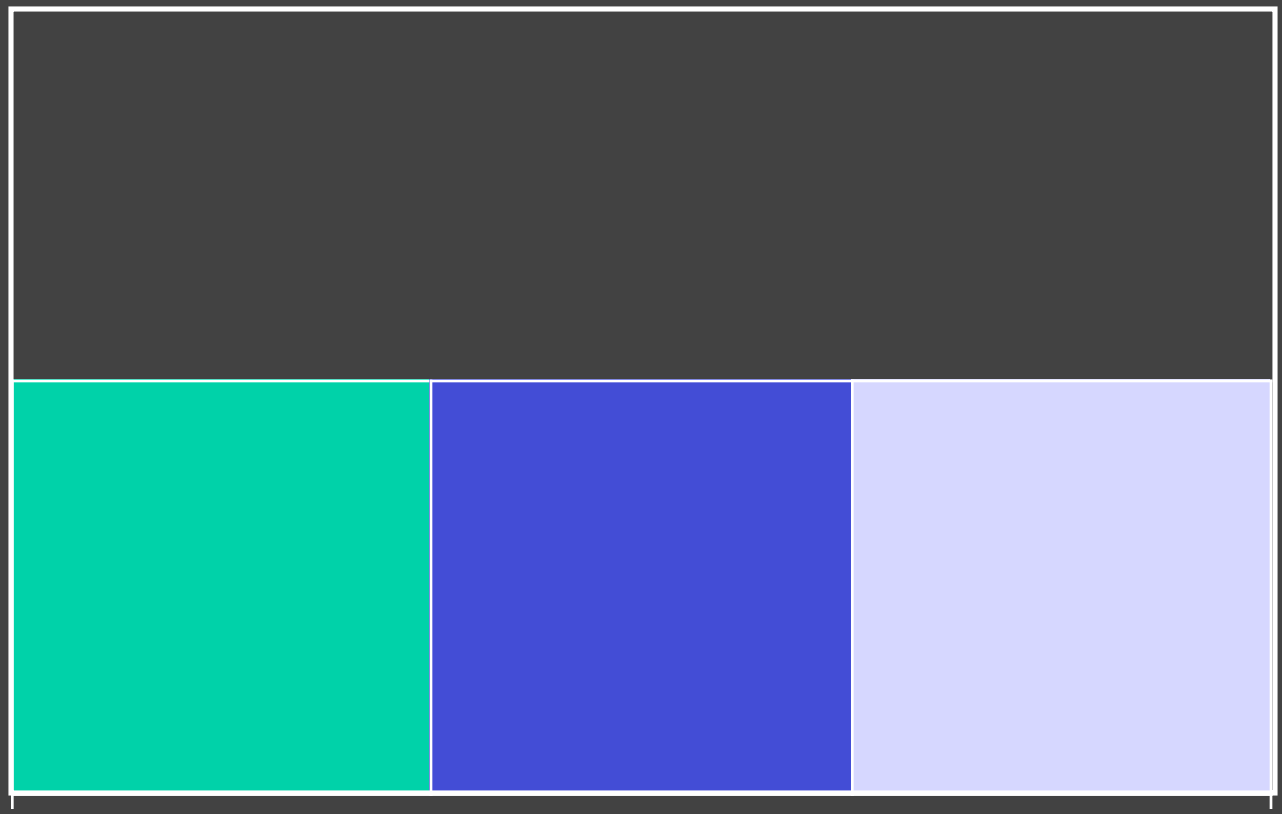
# The Induction Problem

- You're about to see 5 slides of 3 colored bars in frames.
- The first 4 slides exhibit a property that you need to learn.
- Decide whether or not the 5 slide exhibits the property in question.

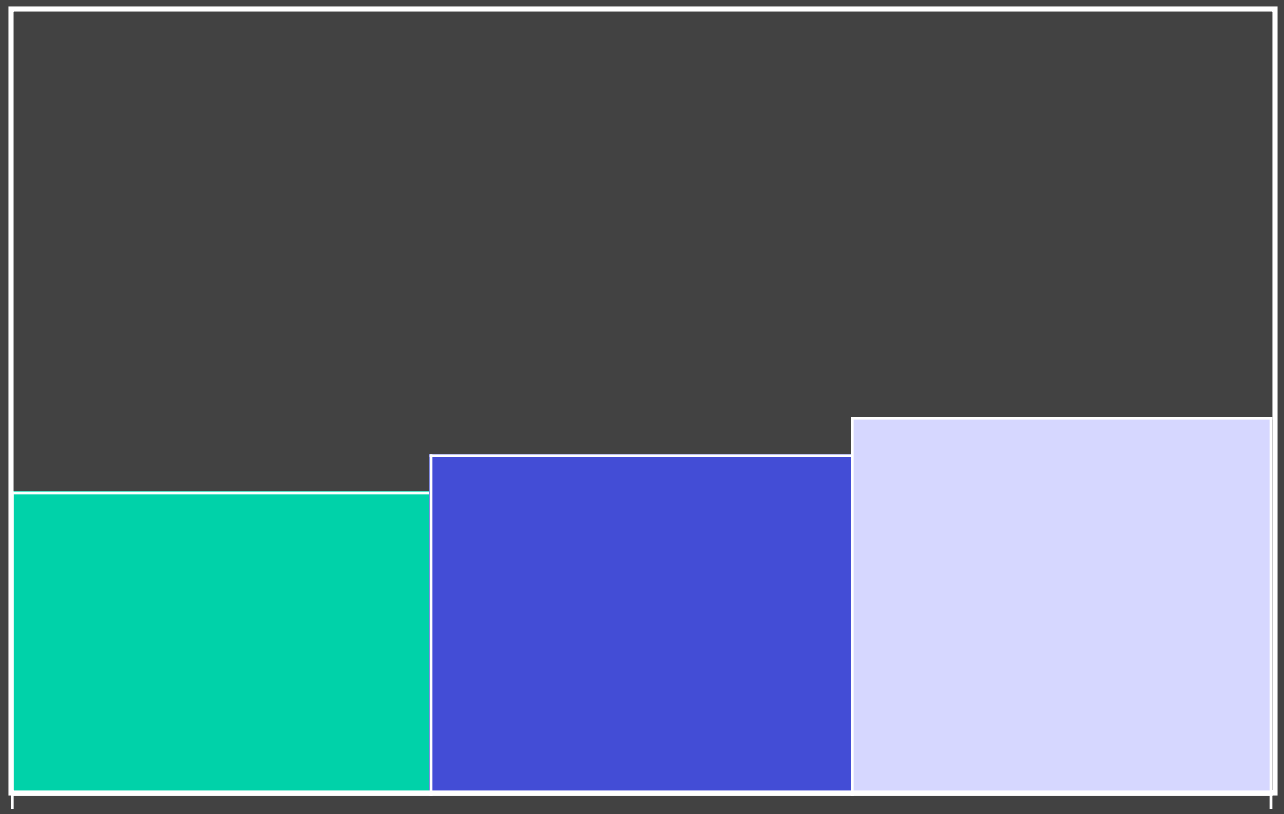








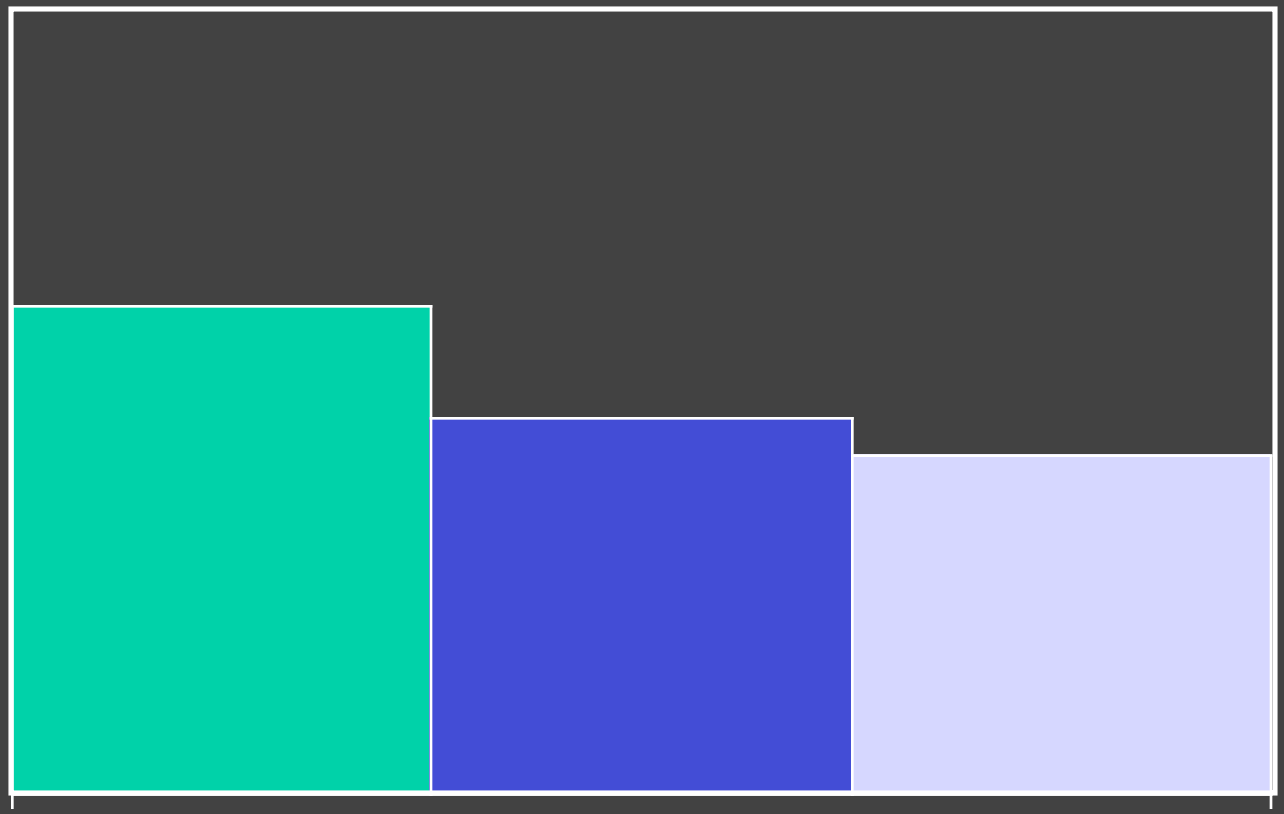
?

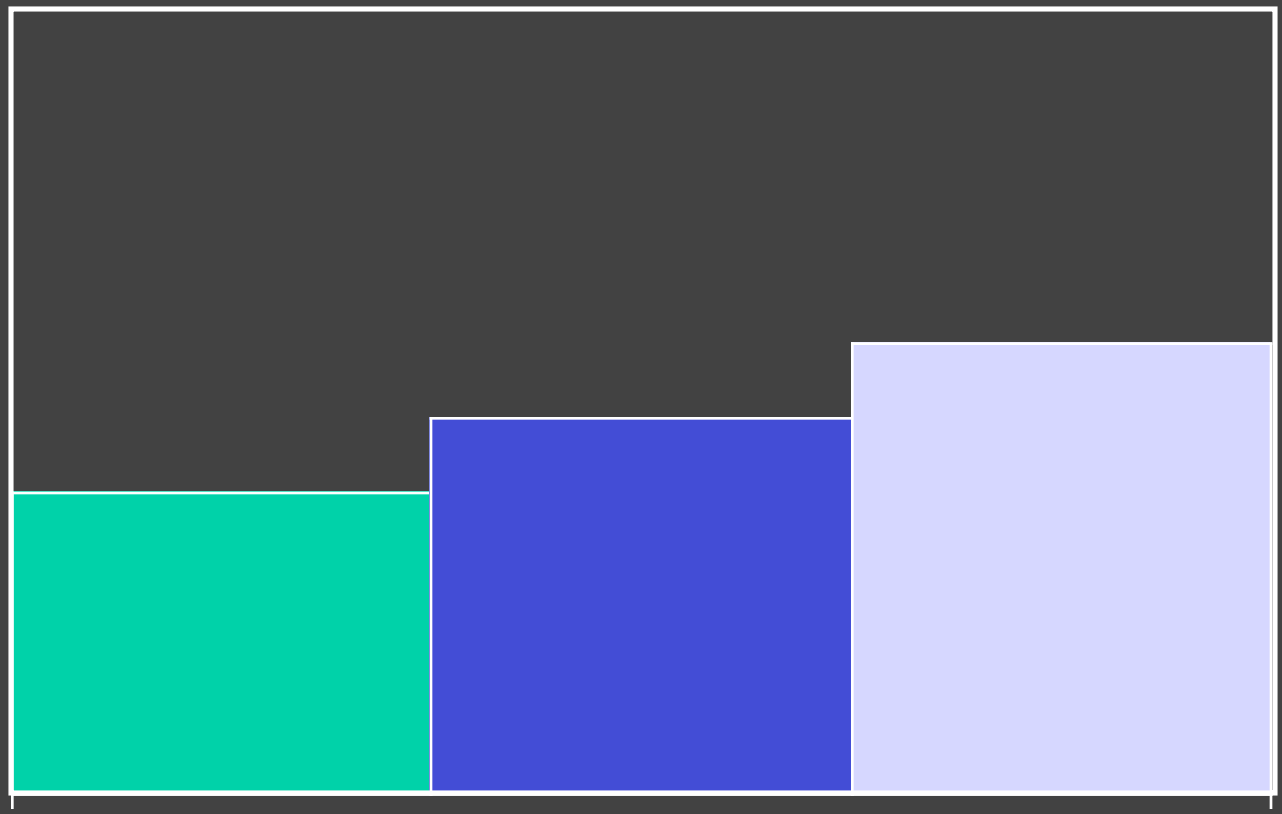


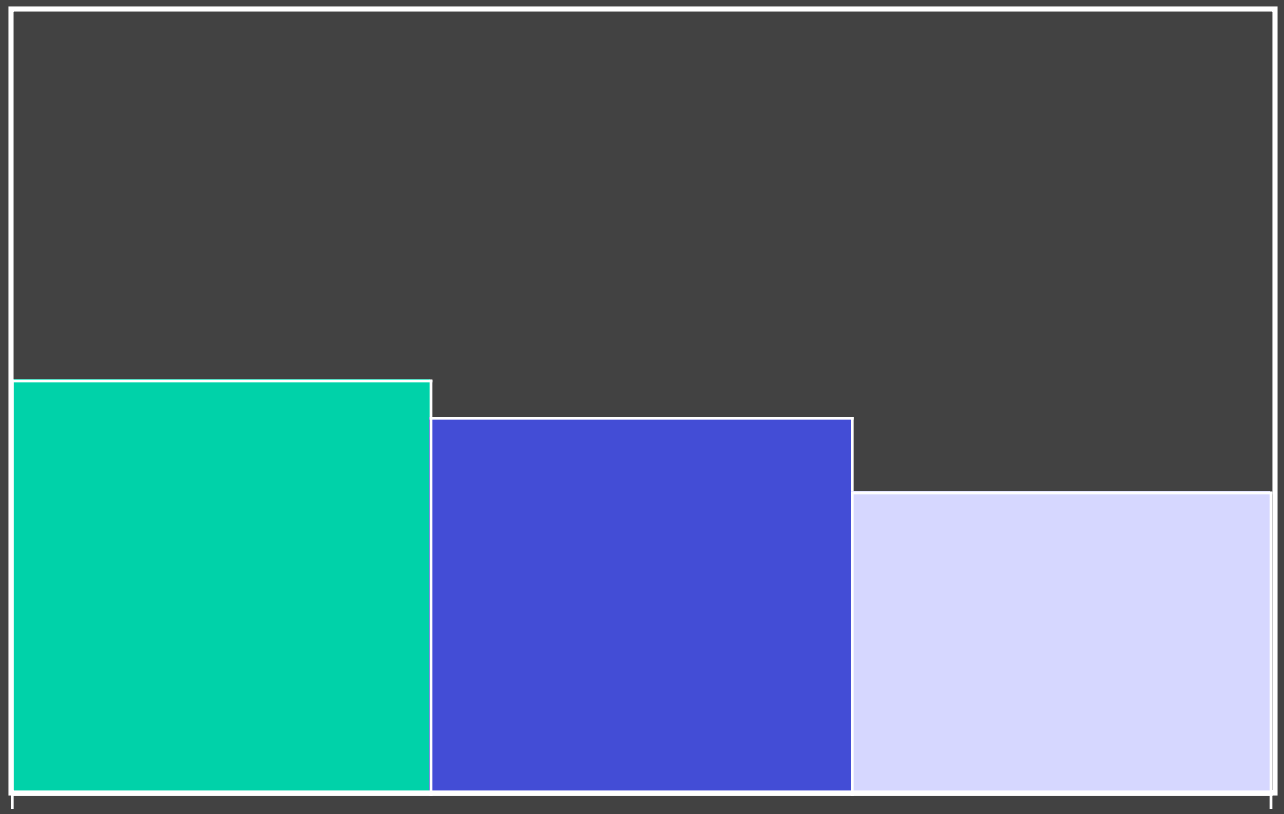
# Answer

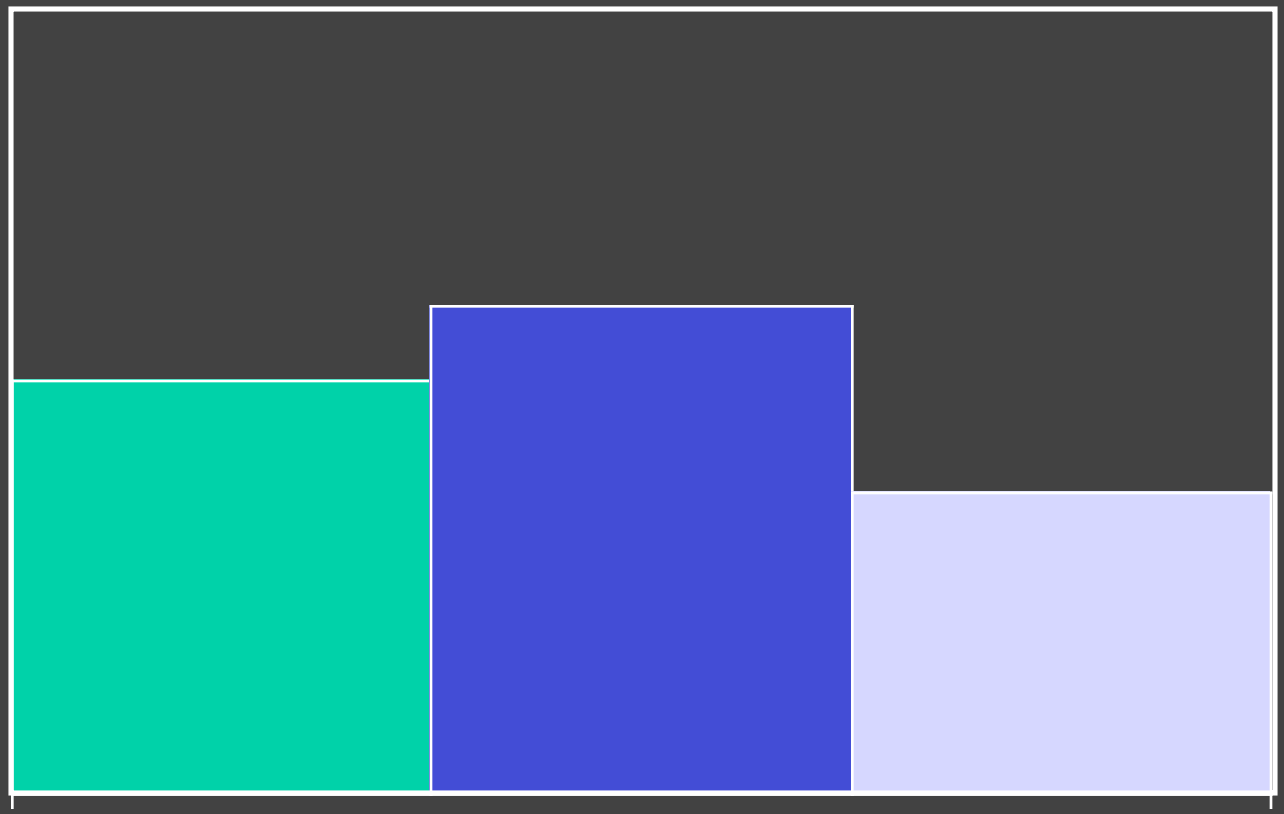
- NO
- The property in question in whether the area covered by the bars is greater than 50% of the area of the rectangle.
- This generalization is more natural for pigeons to learn than for humans.

# Try it again

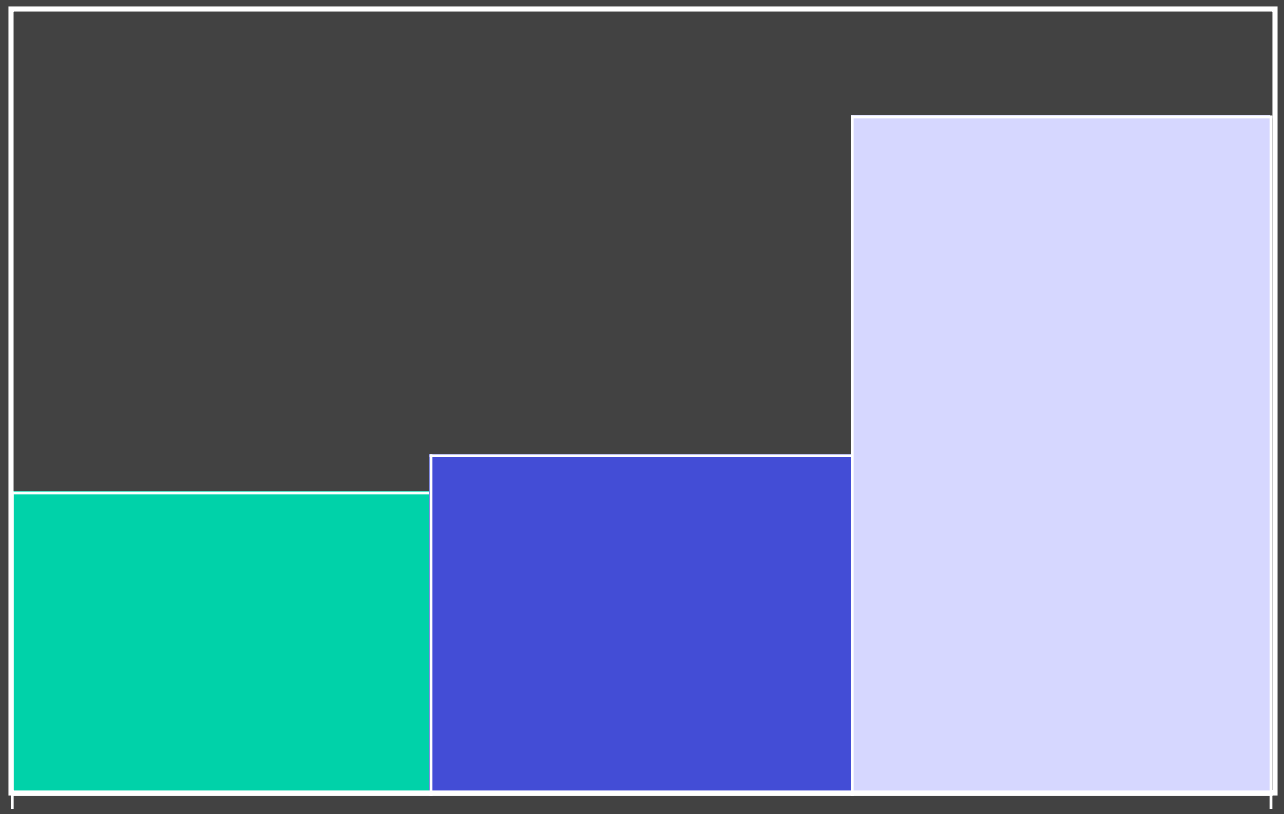








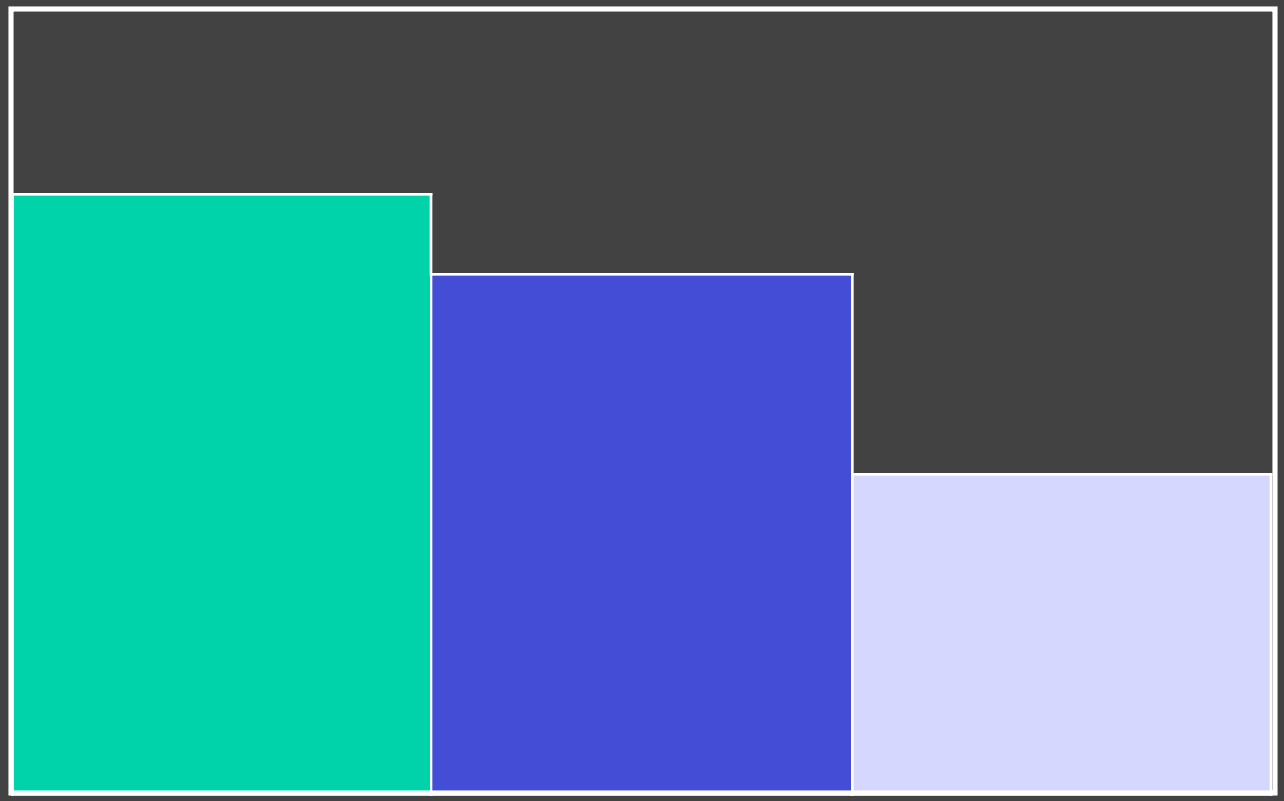
?

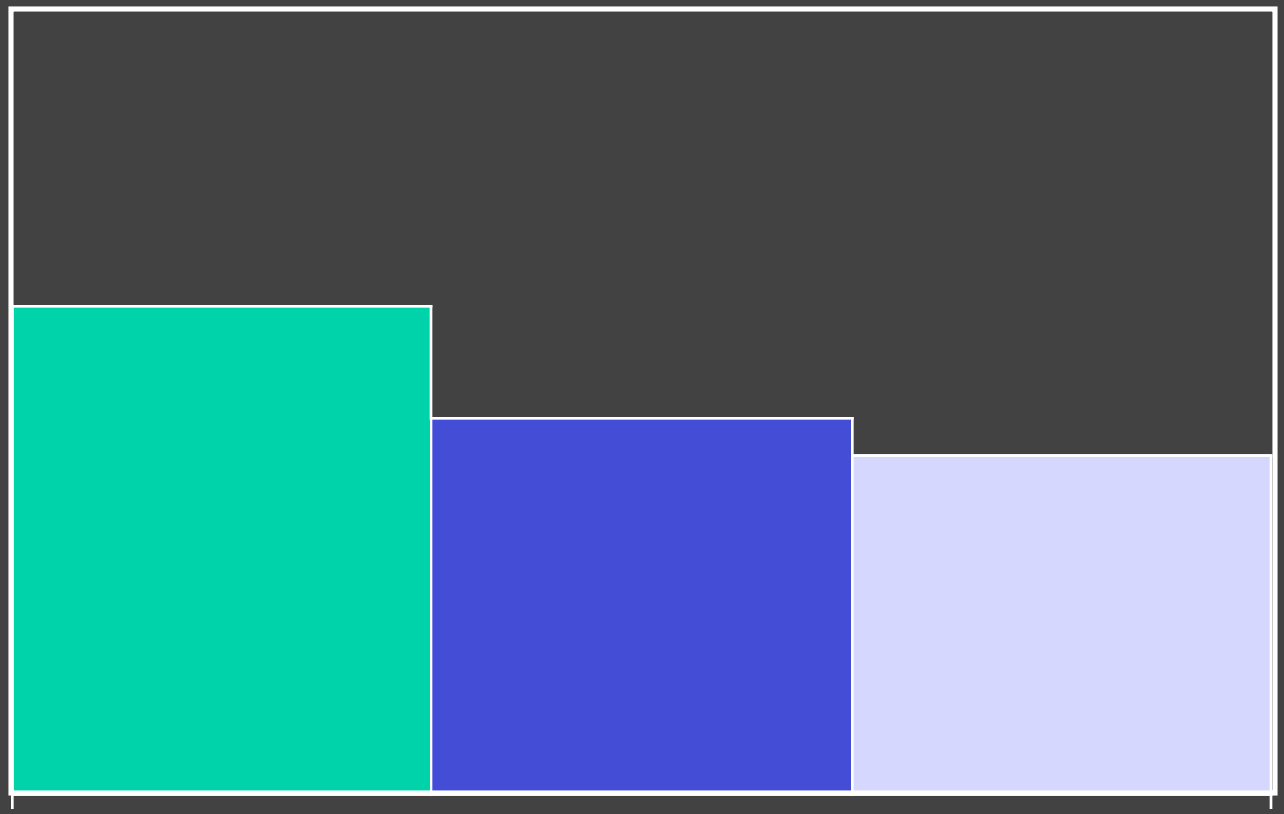


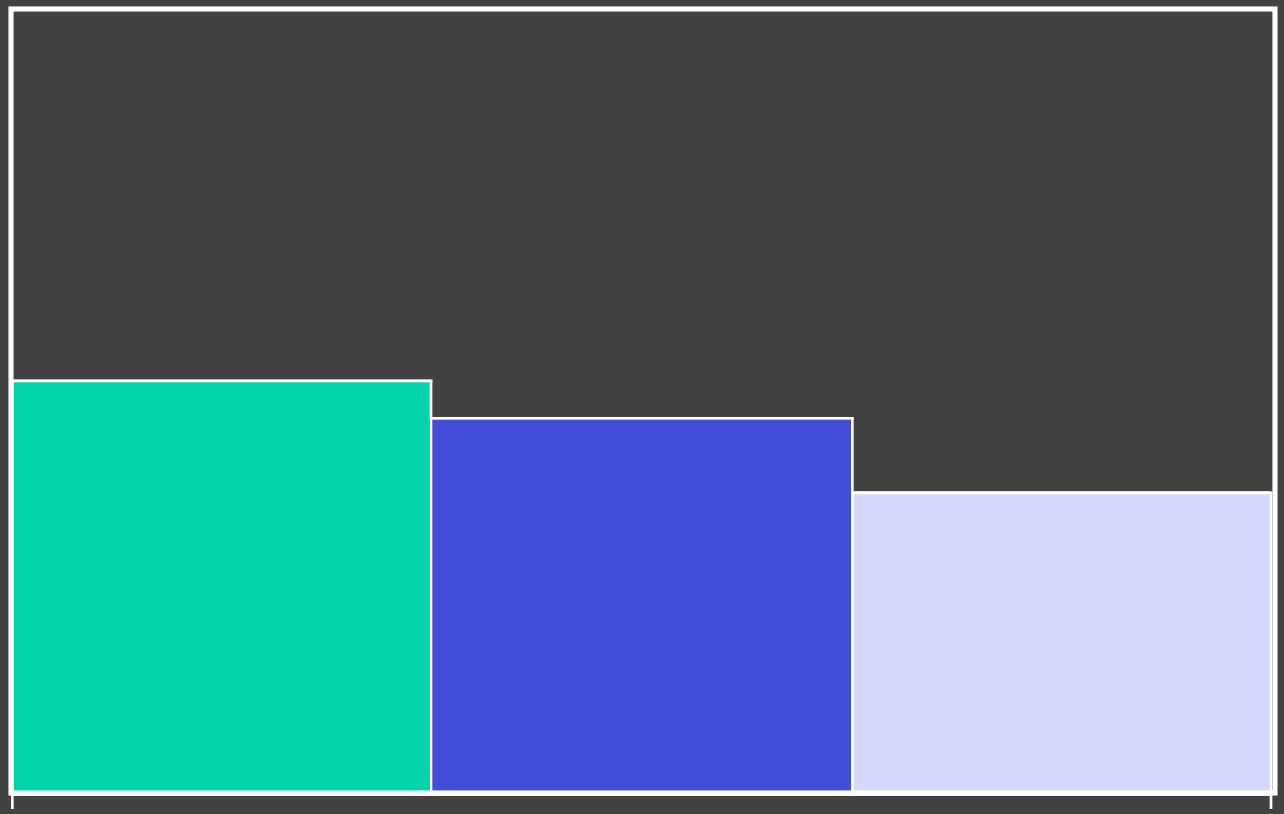
# Answer

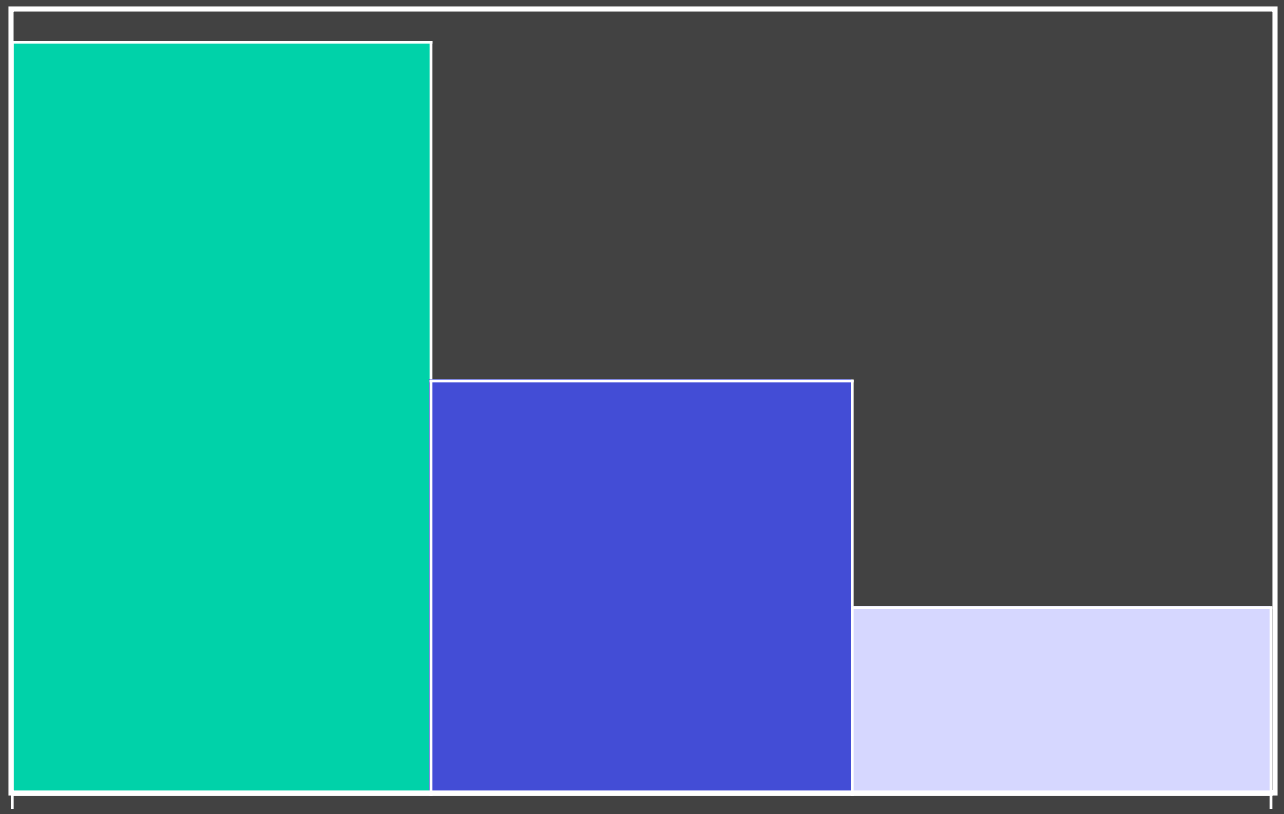
- YES
- Property in question in whether the 3 bars are unequal in height.
- This generalization is more natural for humans than for pigeons.

# Try it again

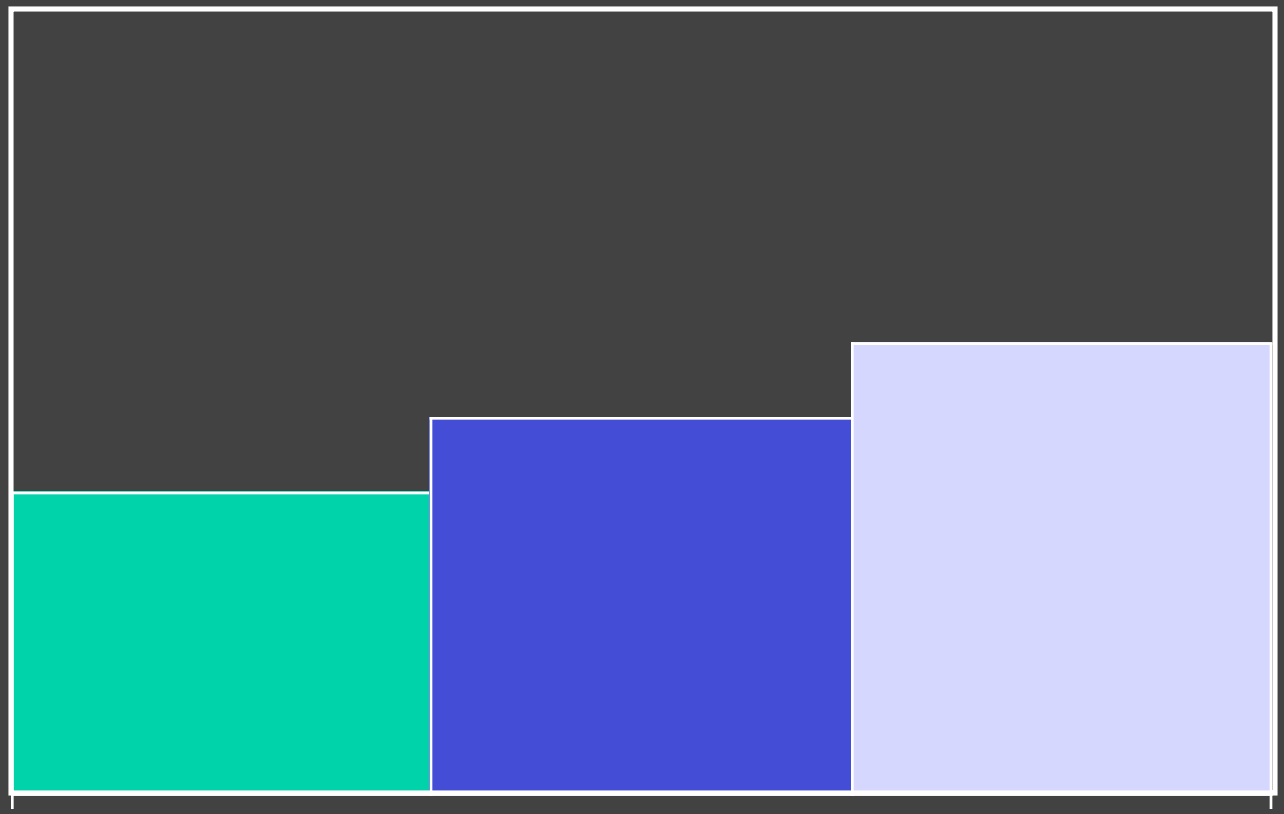








?



# Answer

- YES
- You only saw examples decreasing in height from left to right.  But the generalization was still that the bars only had to be different heights.

# The Induction Problem

- You have to be able to discern the relevant dimension(s)
- Any set of input data potentially allows an infinite number of generalizations.
- If the learner receives a random subset of data, many incorrect generalizations can be ruled out.

# Poverty of the Stimulus

- Generalization in any domain requires the learner to solve the induction problem (as well as the problem of noisy data).
- But some linguists think that there's an additional problem for language learners: Some types of input that could rule out incorrect generalizations is entirely missing during the acquisition process. E.g., …

# Poverty of the Stimulus

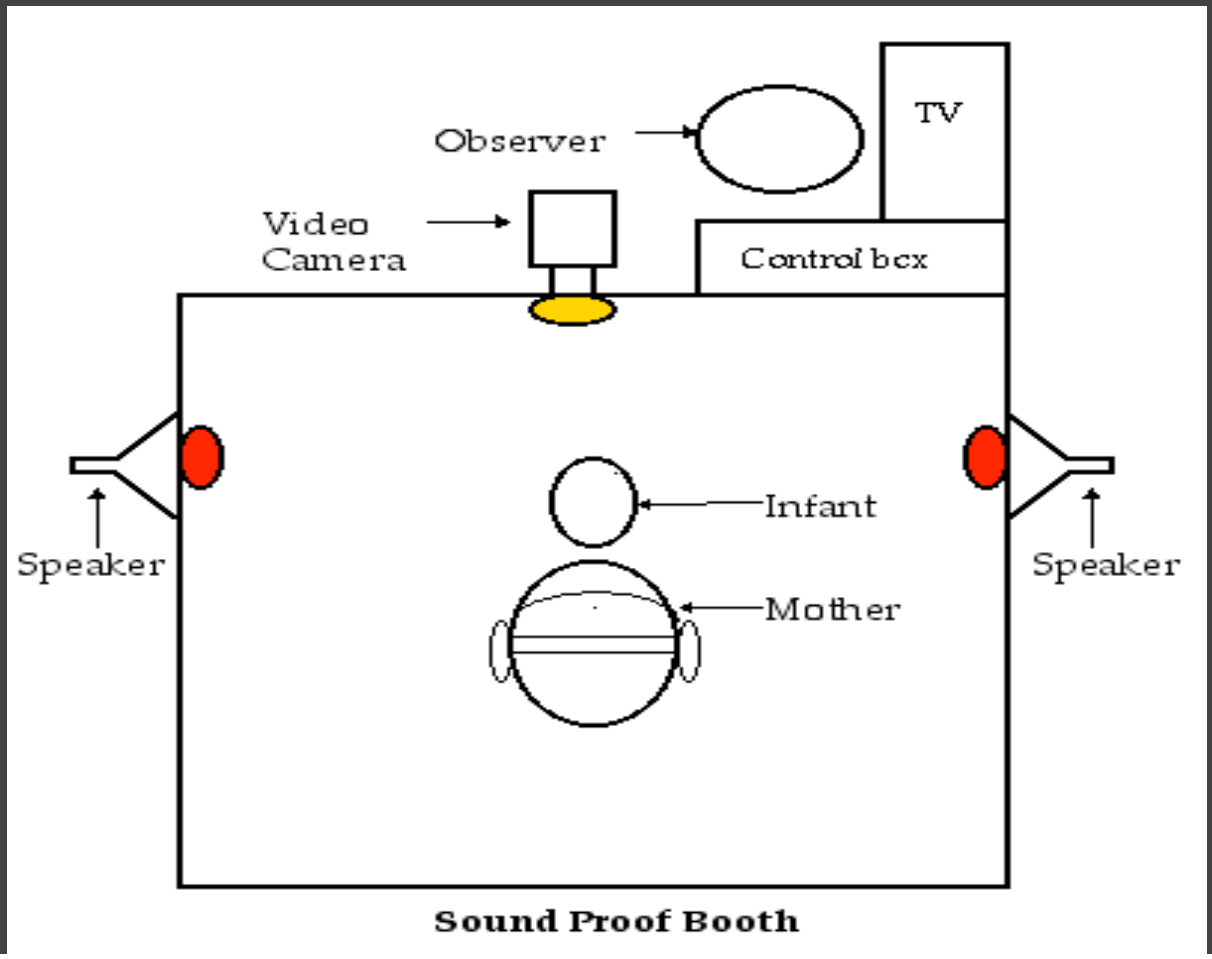
- Statement-question pair with one clause
  - The man is tall.
  - Is the man tall?
- Statement with 2 clauses
  - The man who is tall is nice.
- Possible questions
  - Is the man who tall is nice?
  - Is the man who is tall nice?
- It is argued that children don't hear the correct input before they start to produce the correct 2-clause question.

# Poverty of the Stimulus

- It's my sense that such claims will eventually be seen as incorrect.
- However, if they ARE true, then language learning presents a learning problem above and beyond the induction problem.

# Computational Capacities of Human Infants

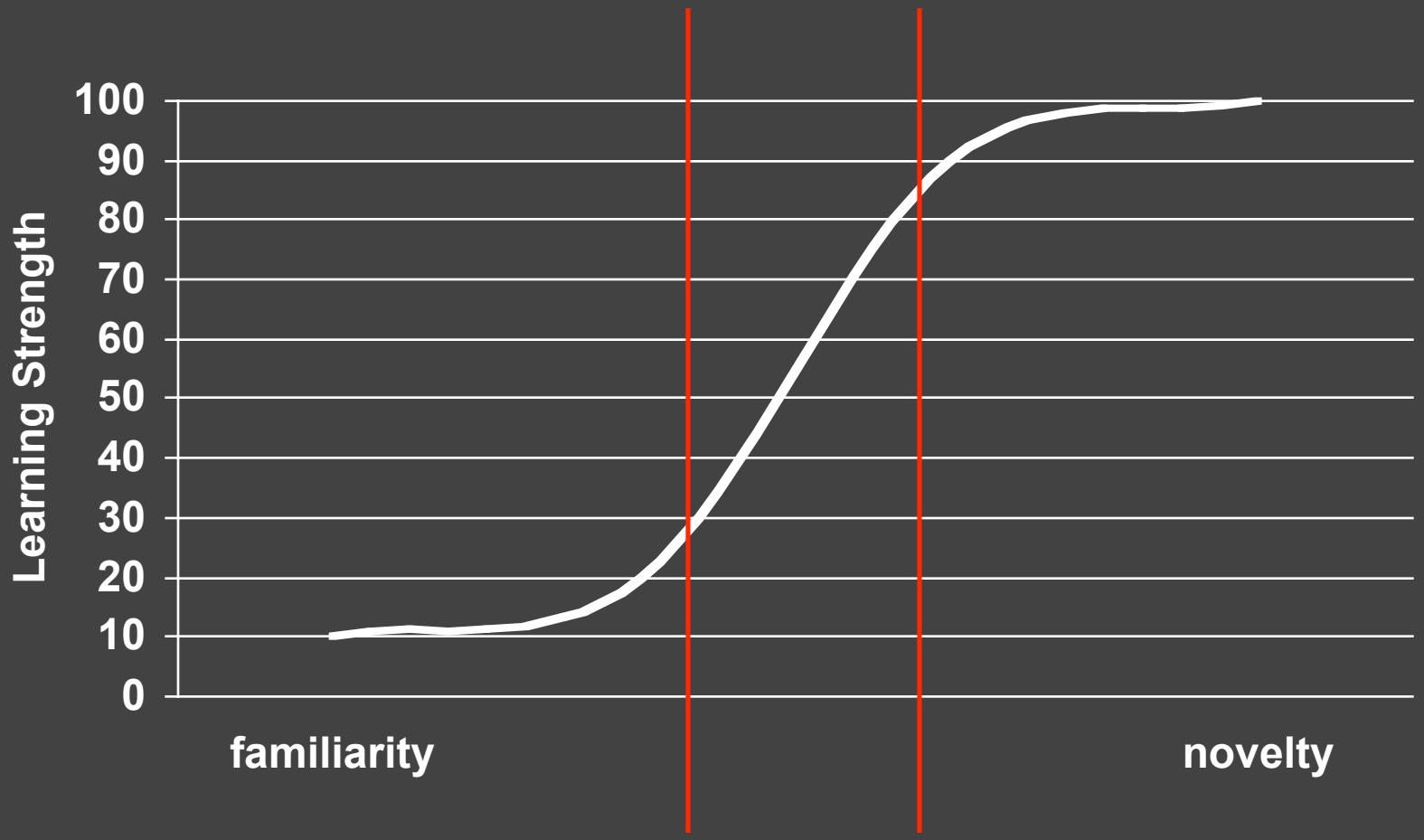
# Headturn Preference Procedure



# Learning in the lab - General procedure

- Infants are exposed for about 2 min. to familiarization stimuli.
- They are then tested on stimuli that are either consistent or inconsistent with the familiarization stimuli.
- The dependent measure is listening time for consistent vs. inconsistent.
- Any significant difference counts

# Novely vs. Familiarity



# Overview of Infant Computation

- Descriptive statistics
- Conditional probabilities
- Patterns of identity
- Input with multiple generalizations
- Categories

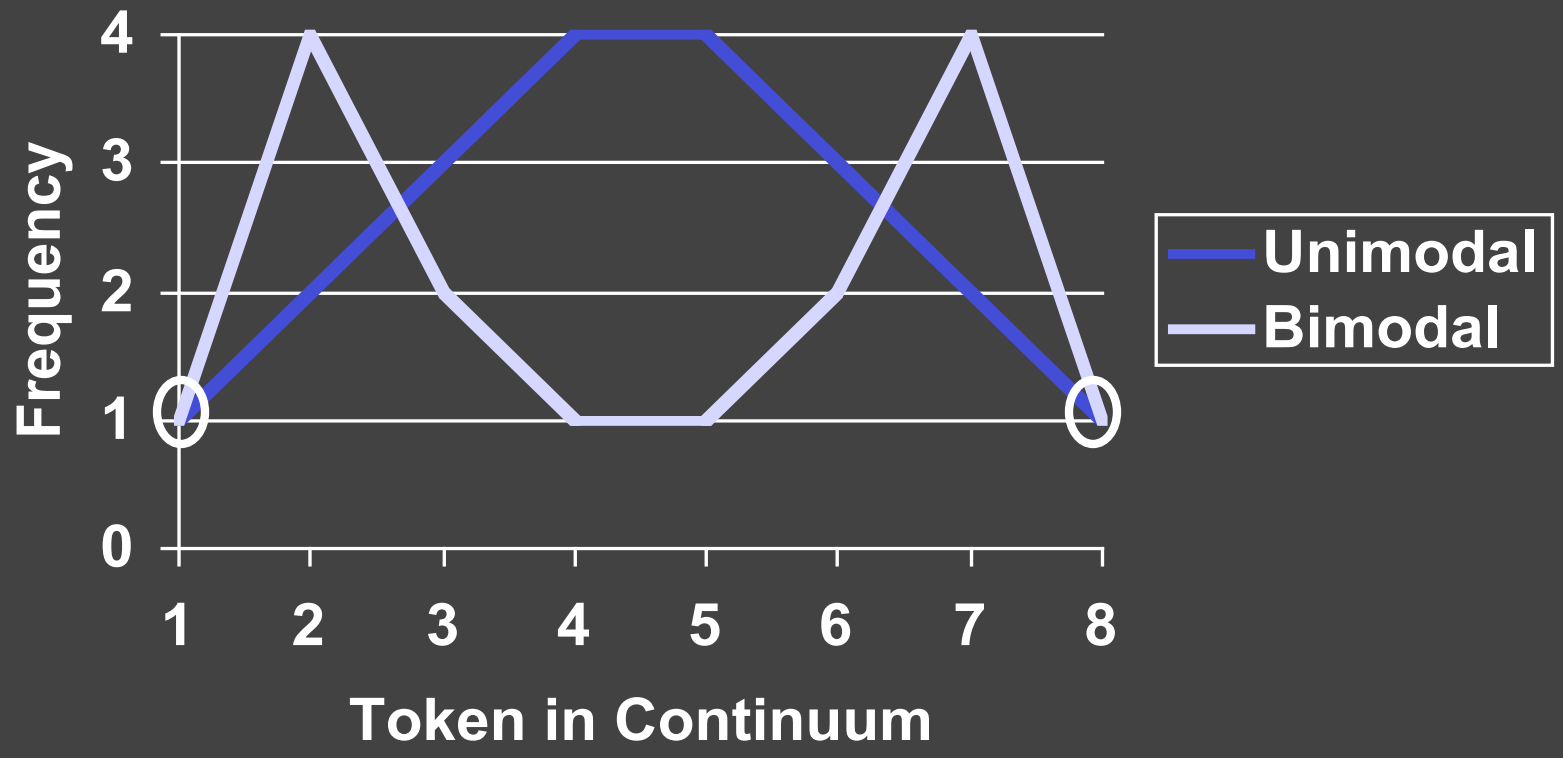
# Descriptive Statistics

Maye, Werker & Gerken, 2001

# Which acoustic differences mark category distinctions?

- The same acoustic differences that are not used to mark meaning distinctions appear allophonically.
- E.g., ‘zip lip’ vs. ‘back lip’ in English vs. Mid-Waghi
- How do infants learn which acoustic differences reflect different categories?

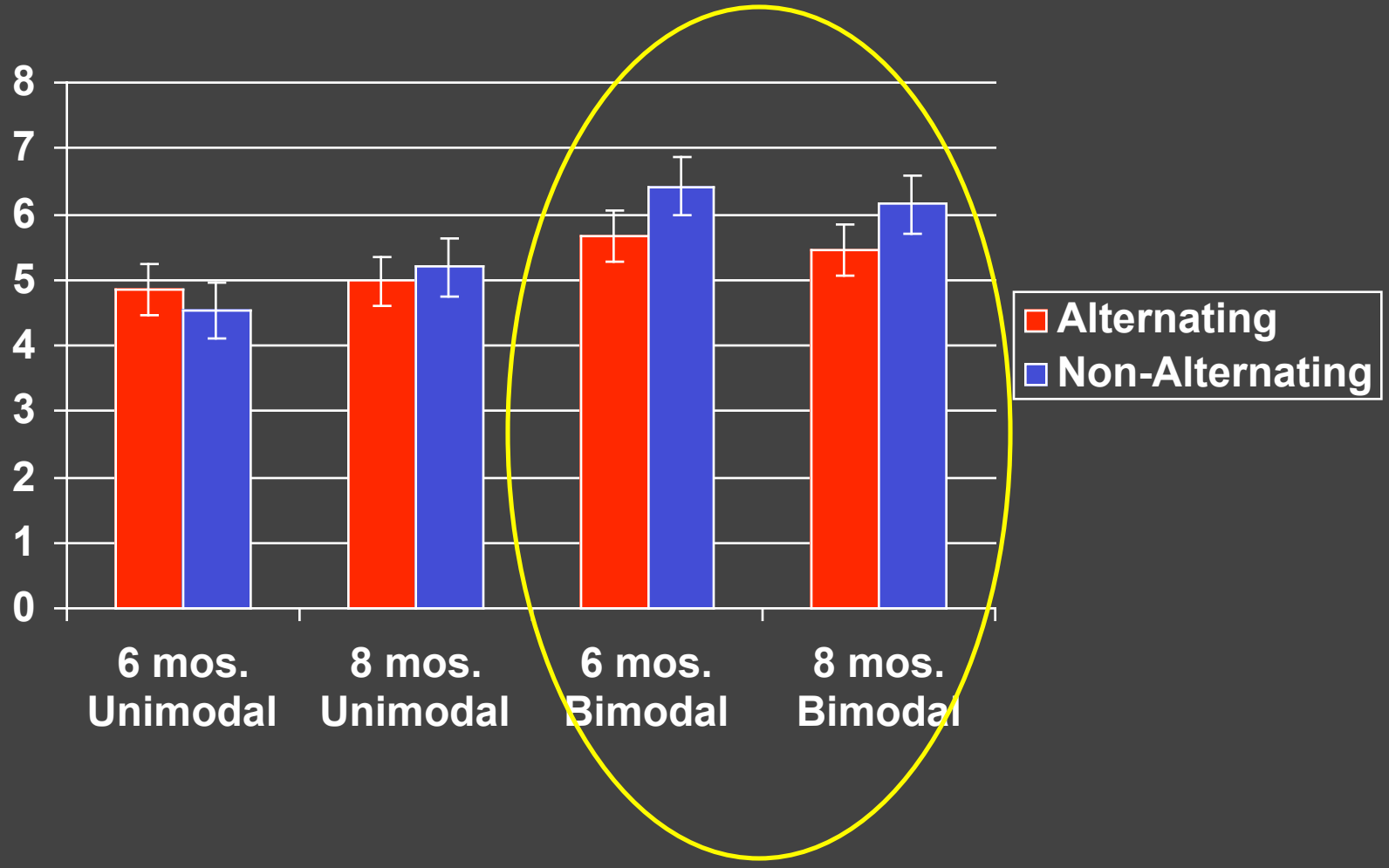
# Maye, Werker & Gerken, 2001 Training Distribution



# Maye et al., 2001

- Infants were familiarized with strings of syllables from the continuum plus fillers.
- Half heard a unimodal and half heard a bimodal distribution from the continuum.
- At test, they either heard one of the end points multiple times in succession of the two end points alternating.

# Maye et al., 2001



# Conditional Probabilities

# Infants' Computational Prowess Saffran et al., 1996

Listen to the following stimulus and try to find the words

- You heard 3 3-syllable words stuck together bidakugolabupadotigolabubidaku
- da follows bi 100% of the time, whereas go follows bu 33% of the time.
- 7.5-month-old infants can use conditional probabilities across adjacent syllables, as shown by the fact that they listen longer to statistical part words like bupado than statistical words like padoti

# Patterns of Identity

# AAB Stimuli from Marcus et al., 1999

| B<br><br>A | di | je | li | we |
|---|---|---|---|---|
| le | leledi | leleje | leleli | lelewe |
| wi | wiwidi | wiwije | wiwili | wiwiwe |
| ji | jijidi | jijije | jijili | jijiwe |
| de | dededi | dedeje | dedeli | dedewe |

# Marcus et al. 1999

- Half of the infants were familiarized with AAB and half on ABA strings.
- Over 12 test trials, all infants heard AAB and ABA strings instantiated in new syllables (*popoga, kokoba* and *pogapo, kobako*)
- Infants listened significantly longer during inconsistent test trials.

# Input with Multiple Generalizations

Gerken (in press)

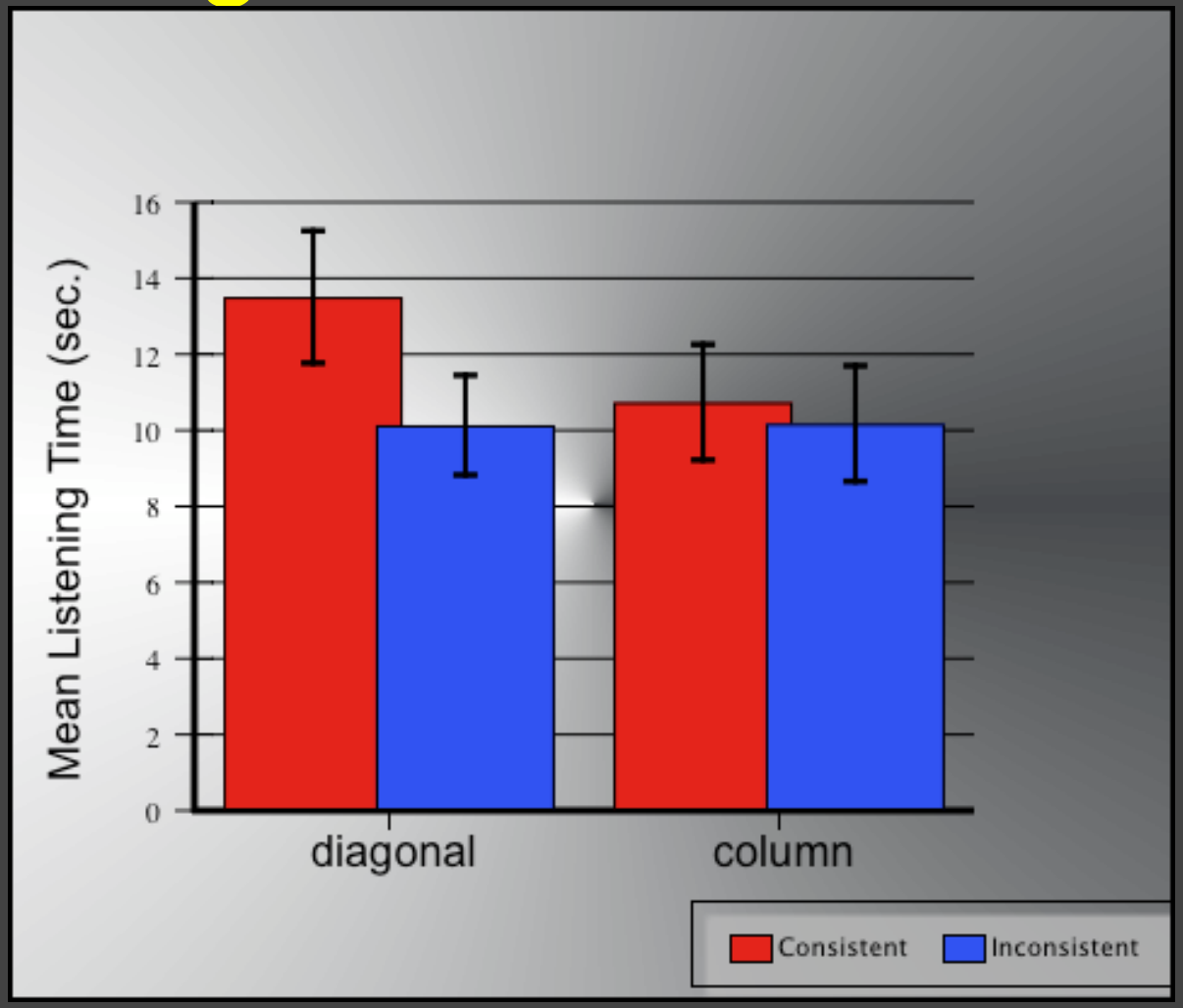
# Different subsets of an input set support different generalizations

| B / A | di | je | li | we |
|---|---|---|---|---|
| le | leledi | leleje | leleli | lelewe |
| wi | wiwidi | wiwije | wiwili | wiwiwe |
| ji | jijidi | jijije | jijili | jijiwe |
| de | dededi | dedeje | dedeli | dedewe |

# Different subsets of an input set support different generalizations

- 2 groups of 9-month-olds were familiarized with synthesized tokens from either diagonal or column of previous table.
- Half in each group heard AAB or ABA strings.
- At test, all infants heard 2 AAB and 2 ABA strings instantiated in new syllables (*popoga, kokoba* and *pogapo, kobako*)

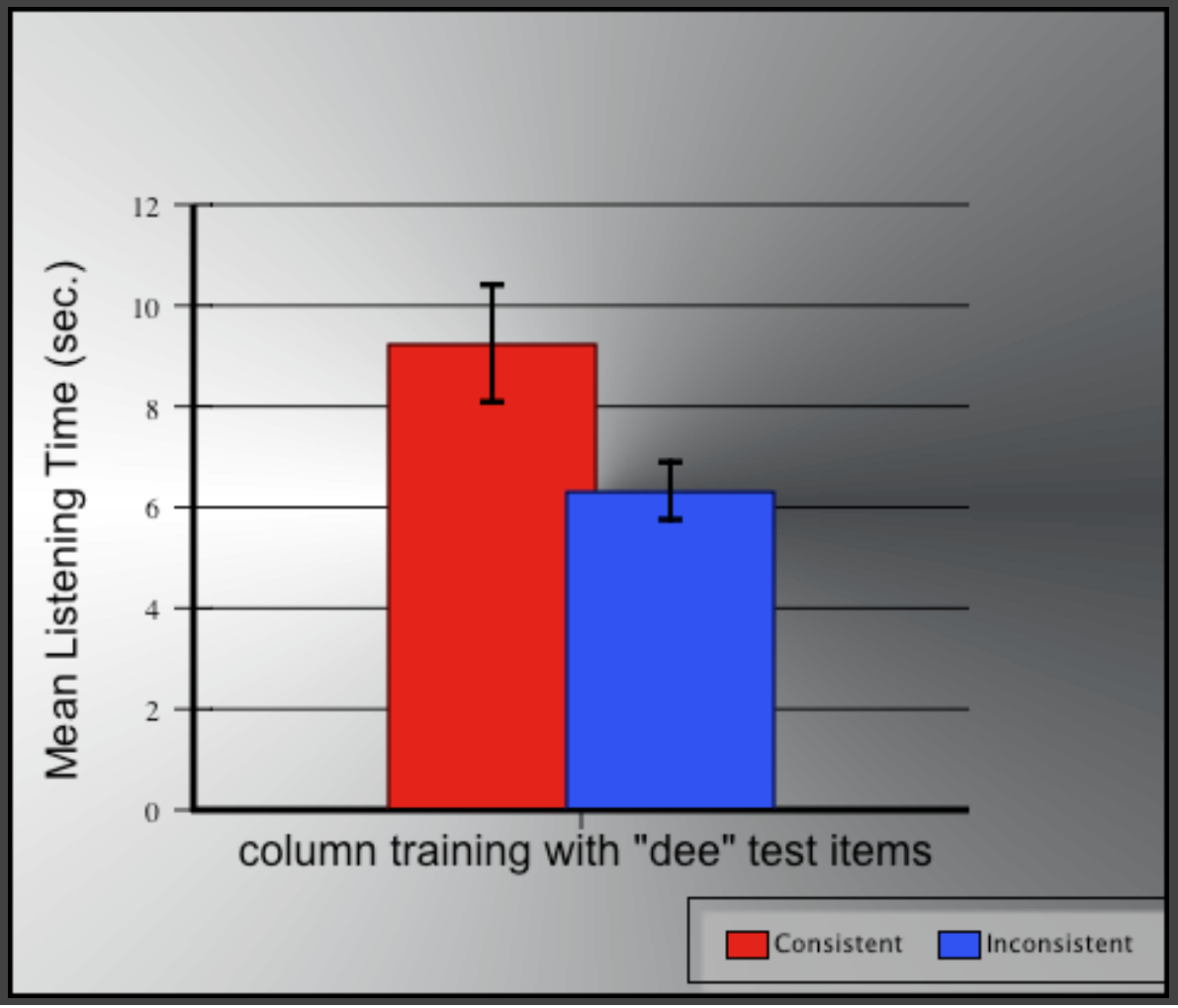
# Generalization to New Strings Diagonal vs. Column



# Did the Column-Exposed Make a Less Abstract Generalization (dee)?

- A third group of infants was familiarized with the column stimuli (AAB or ABA).
- They were then tested on strings containing “dee” *popodi, kokodi* or *podipo, kodiko*)

# Infants Make Less Abstract Generalization from Column



# Categories

(Gerken, Wilson & Lewis, 2005)

# Russian Gender Categories

| | | | | | |
|---|---|---|---|---|---|
| polkoj | rubashkoj | ruchkoj | ??? | knigoj | korovoj |
| polku | rubashku | ruchku | vannu | knigu | ??? |
| zhitelya | uchitelya | stroitelya | ??? | kornya | pisarya |
| zhitel-yem | uchitel-yem | stroiltel-yem | medved-yem | kornyem | ??? |

tested on 6 trials each of

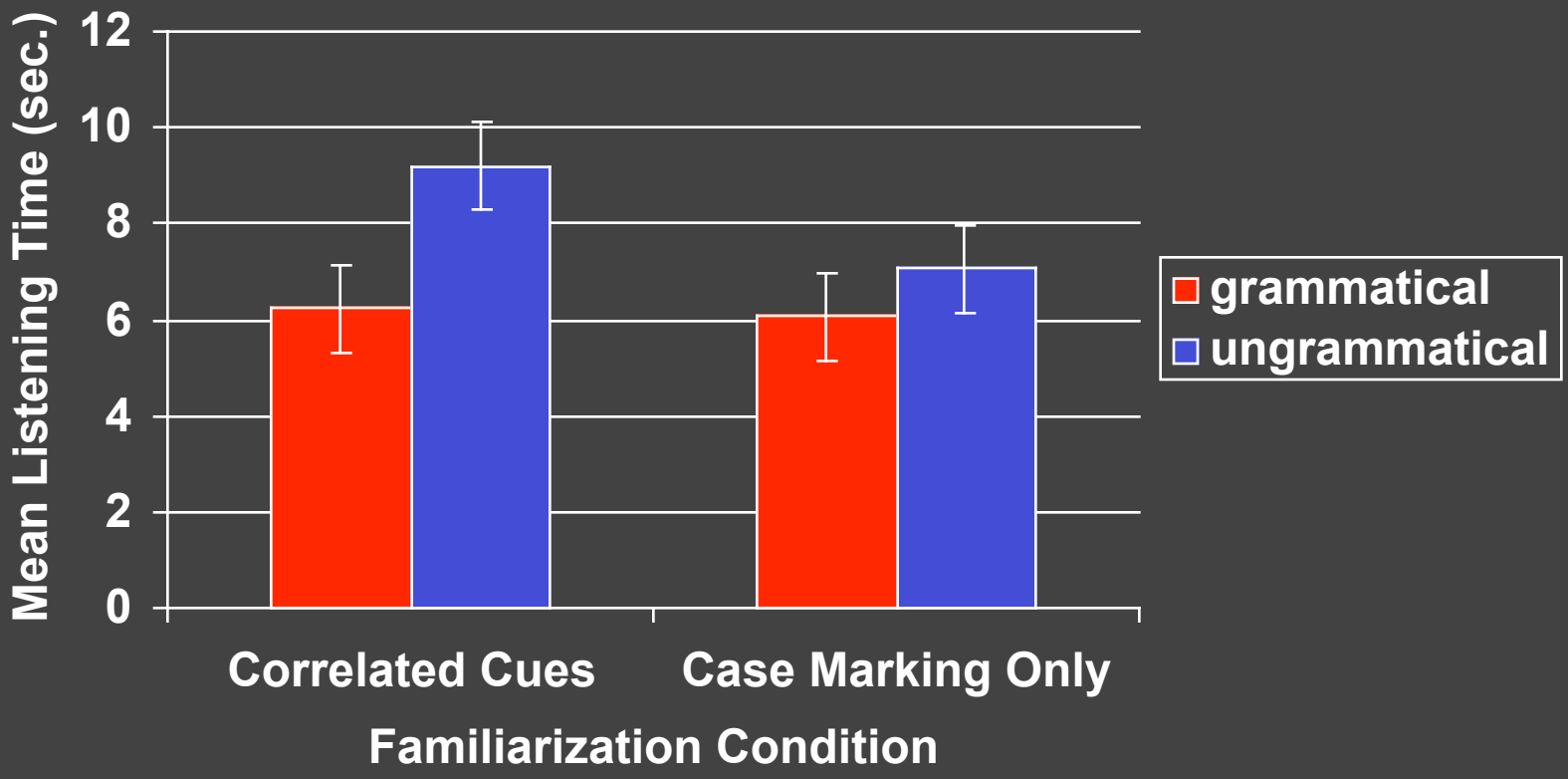
grammatical: vannoj korovu medvedya pisaryem

ungrammatical: vannya korovyem medvedoj pisaru

# Can Infants Learn Russian Gender Categories?

- 2 groups of 17-month-olds
  - Group 1 was familiarized for 2 min. to words in which a subset were double-marked
  - Group 2 received an identical test, but was familiarized with stimuli containing only case ending cues to gender (no words ending in “tel” or “k” were included)

# Correlated Cues for Category Induction

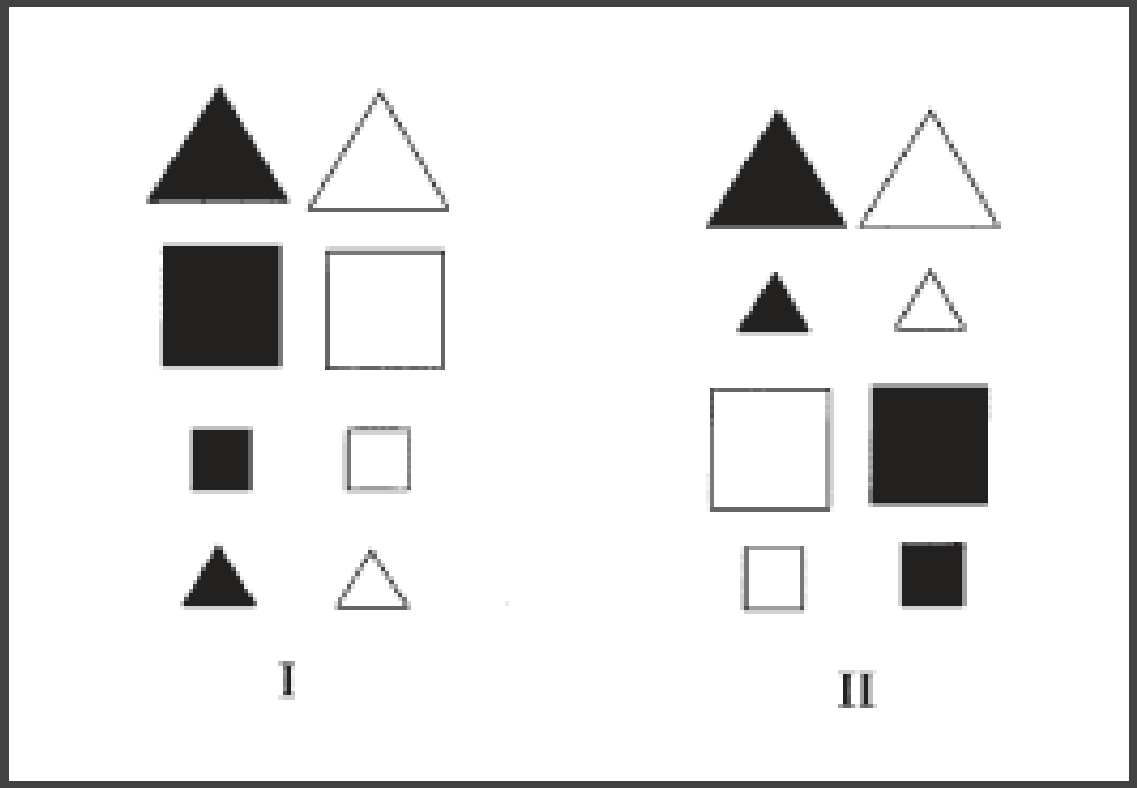


# Turning Induction into Deduction

# Turning Induction into Deduction

- There are 3 approaches to the induction problem for language:
  - ANN's
    - generalize, but do so by pre-coding the relevant dimensions of the problem space
    - produce 'weird' errors
    - don't account for non-linearities in behavior …

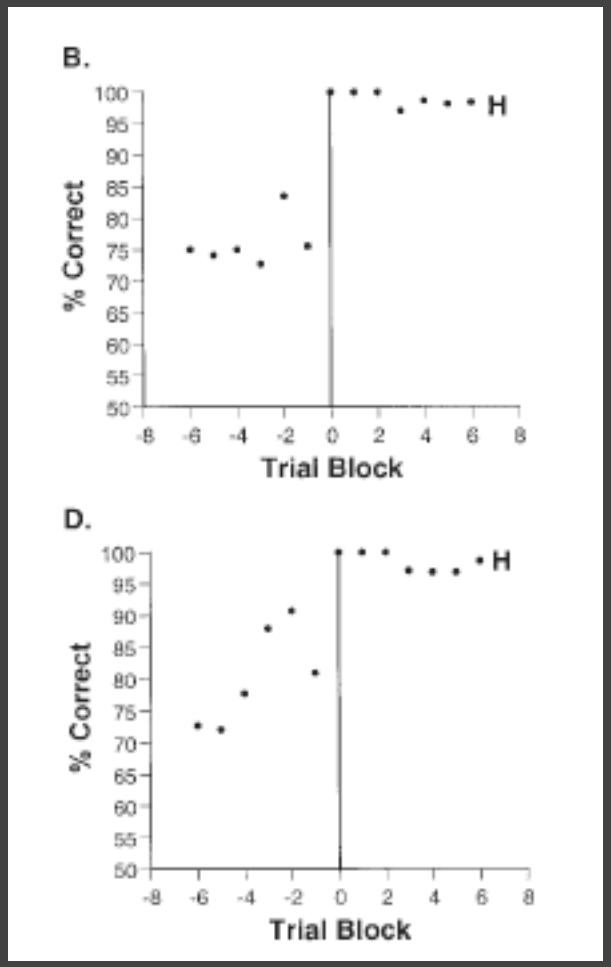
# Abrupt vs. Gradual Learning

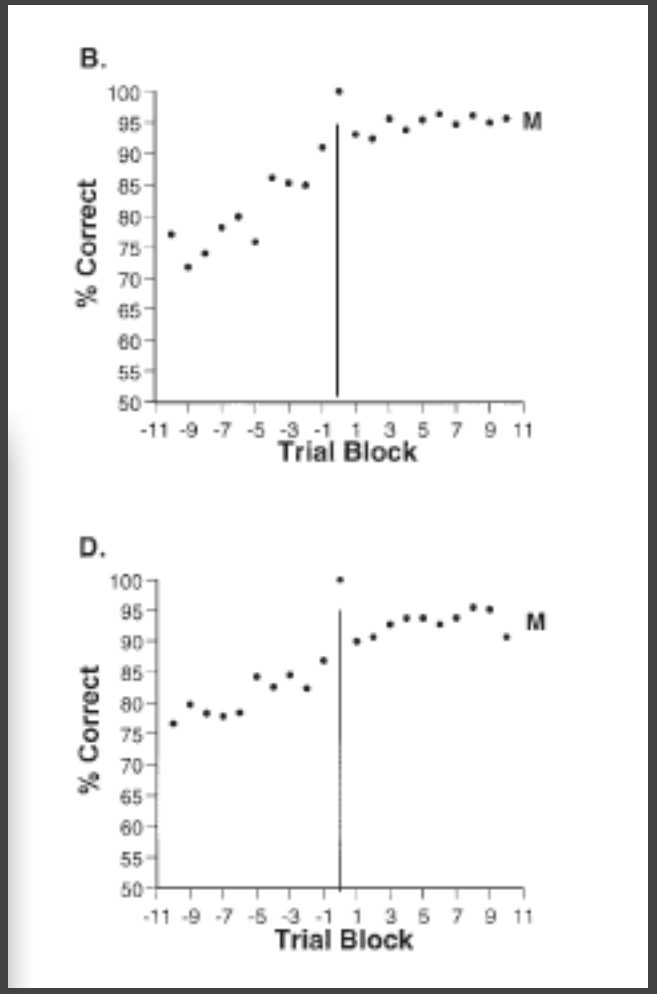



black vs. white

black triangle or white square vs. white triangle or black square

# Abrupt vs. Gradual Learning



# Turning Induction into Deduction

- Hypothesis testing (often Bayesian), like in visual object identification
- Universal Grammar …

# Motivation for UG Approach

- Only humans have it
- Language creation by individual children (and writing system creation)
- Similarities across languages
- Poverty of the stimulus

# The UG Approach

- Languages of the world differ from each other in a limited number of ways (parametric variation).
- Humans are born with all parameter values latently available, and minimal input is needed to set a parameter.
- Early utterances and de novo creation of language reflects default settings.

# Overall Summary

- Language learning entails solving the induction problem at 2 levels.
  - sub-meaning units (phonology)
  - meaning units (syntax)

# Overall Summary

- Language learning also might entail dealing with stimulus poverty, in which some critical information is withheld from young learners.

# Overall Summary

- Infants have many computational skills.

# Overall Summary

- But we don’t know how they might use them to solve the induction problem.

# Overall Summary

- Alternatively, infants might be engaged in deduction. But the empirical evidence for a particular view of deduction (parameter setting) is not very strong.

# Overall Summary

- Research areas:
  - computational approaches to induction
  - more information about the computational skills of infants
  - exploring deductive constraints on learners