

# 9.85 Cognition in Infancy and Early Childhood

## Lecture 3: Theoretical perspectives post-Piaget

# Today

- Red: Gopnik & Meltzoff; Blue: Fodor & Chomsky
  - Red: What (acc. G & M) are the key characteristics of a theory? Blue: “Why does evidence that we are capable of inductive inferences offer support for nativism?”
- Coming up: Critical responses
  - do summarize main findings in a few sentences at the start
  - do suggest alternative explanations for the findings (if any)
  - do suggest alternative interpretations of the findings (if any)
  - do suggest interesting possible extensions or avenues for future research.
  - do not point out reasons (e.g., too few participants, too complicated a task) why the study might not have worked (if they wrote it up -- it worked and the groups are large enough for the differences to be statistically significant).

# Spoiler ...

- Post-Piaget, there is no grand unifying theory of cognitive development ...
- Grand unifying theory of cognitive development will probably require a grand unifying theory of cognition.

# Trends in developmental research post-Piaget

- connectionism (Bates, Elman, McClelland, Munakata)
- information processing approaches (Case, Klahr, Siegler)
- modularity theories (Leslie)
- **core knowledge** (Spelke, Carey)
- **theory theory** (Spelke, Carey, Gelman, Gopnik, Meltzoff, Wellman)
  - also biological, maturational accounts (Diamond, Neville, Thelen)

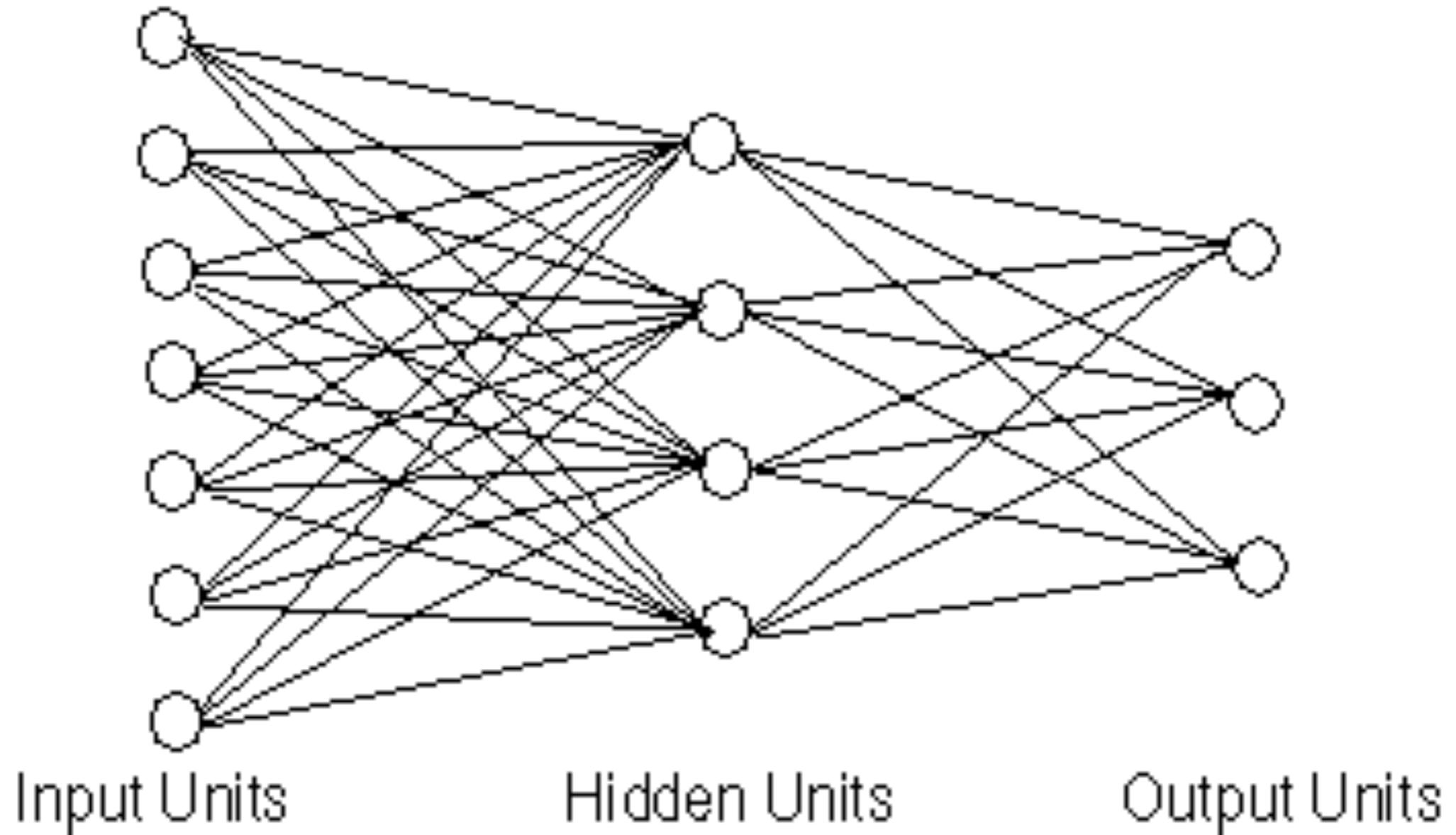
# At issue ...

- Are basic processes central to learning ...
  - domain-general?
  - or domain-specific?
- What is innately specified?
  - Just learning mechanisms?
  - Particular content knowledge?
- What stays constant over development?
  - Just basic learning mechanisms?
  - Particular content knowledge?

# trends post-Piaget

- connectionism (Bates, Elman, McClelland, Munakata)
- information processing approaches (Case, Klahr, Siegler)
- modularity (Leslie)
- core knowledge (Spelke, Carey)
- theory theory (Spelke, Carey, Gelman, Gopnik, Meltzoff, Wellman)
  - also biological, maturational accounts (Diamond, Neville, Thelen)

# Connectionism

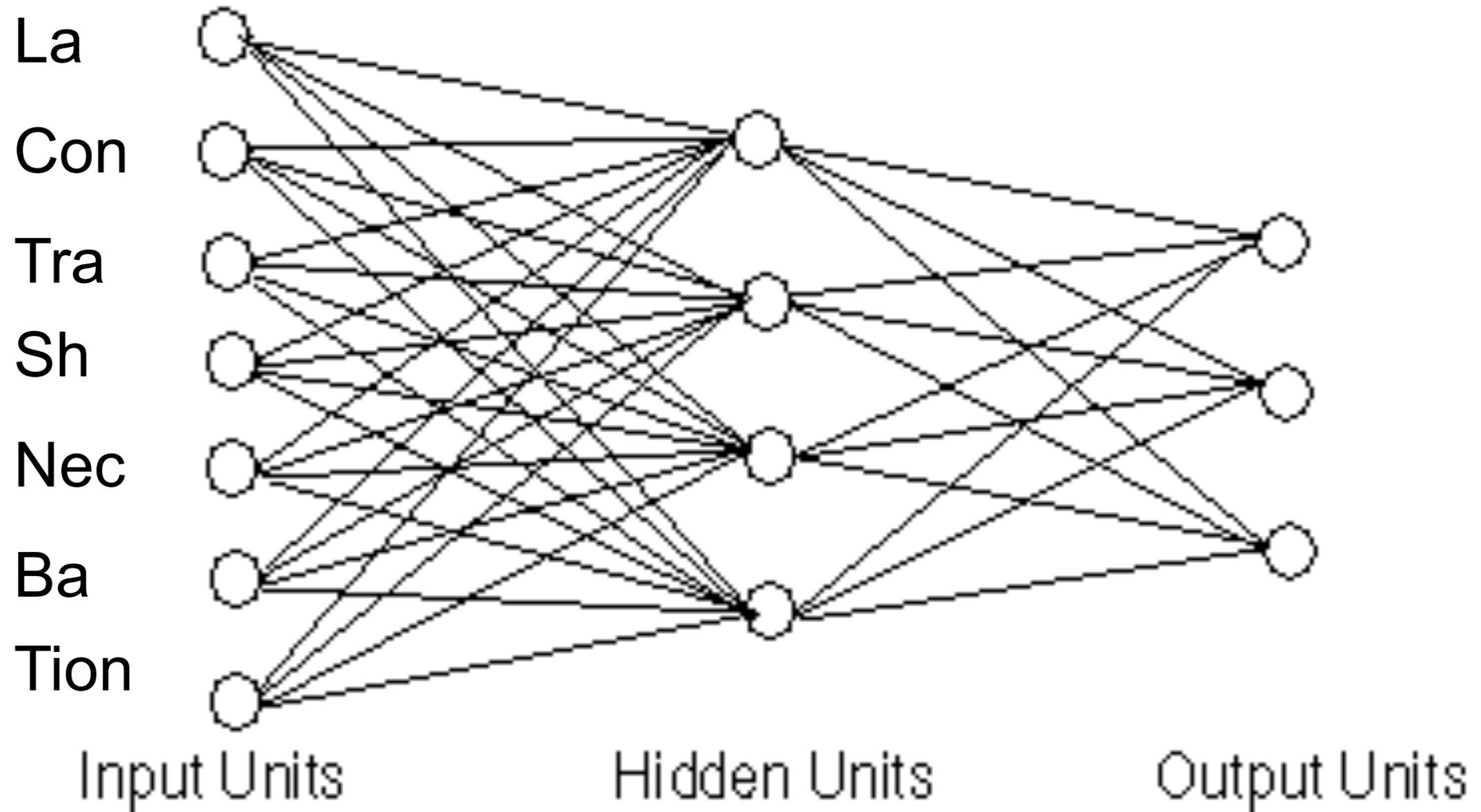


© Metaphysics Research Lab, CSLI, Stanford University. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

# Connectionism

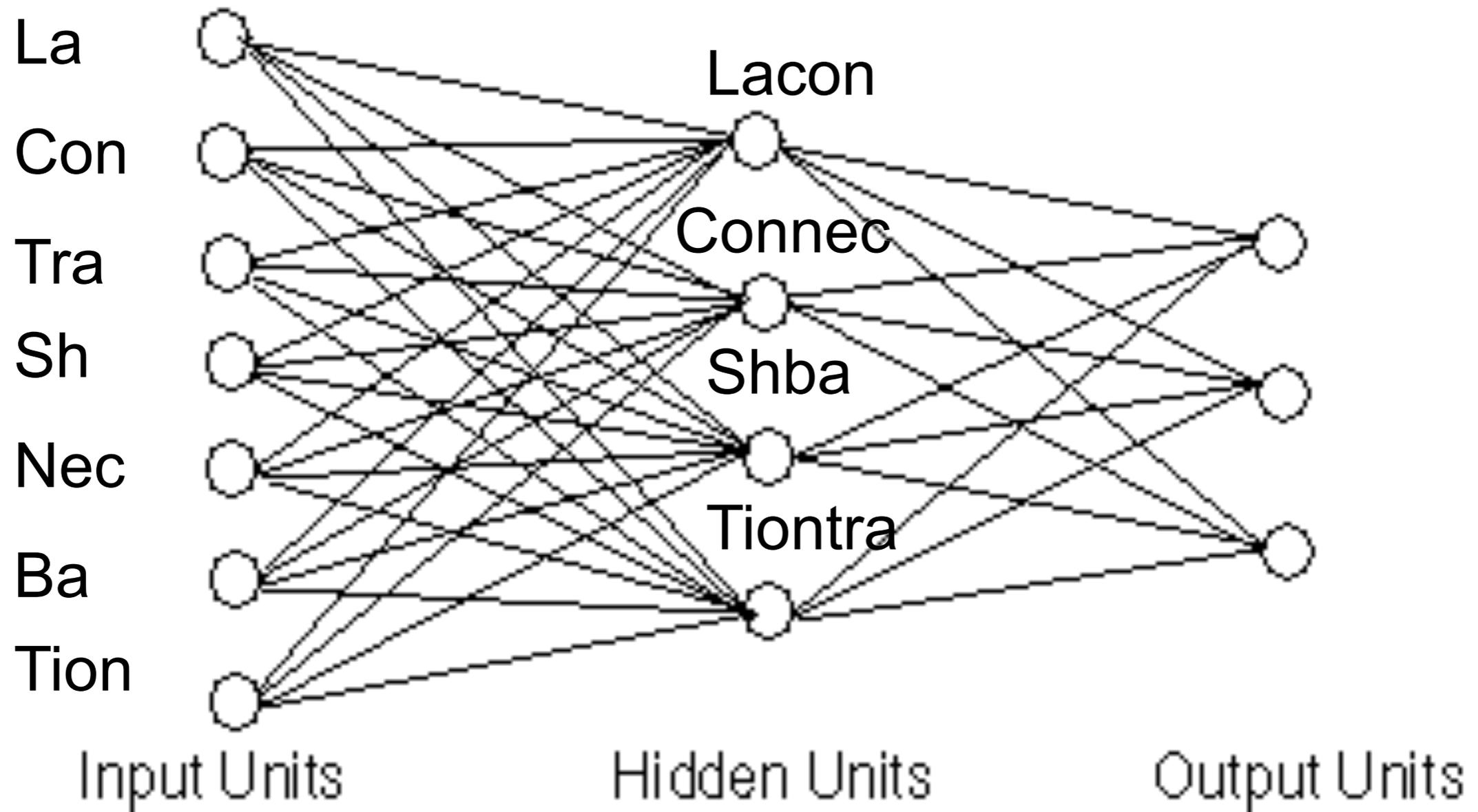
- Input units initially set to random values
- Hidden units can be activated or inhibited by patterns of input
- Patterns of activation are “reinforced” (strengthened) until desired output on the training set is achieved.
- The neural network may then be able to generalize to novel input.

# Connectionism



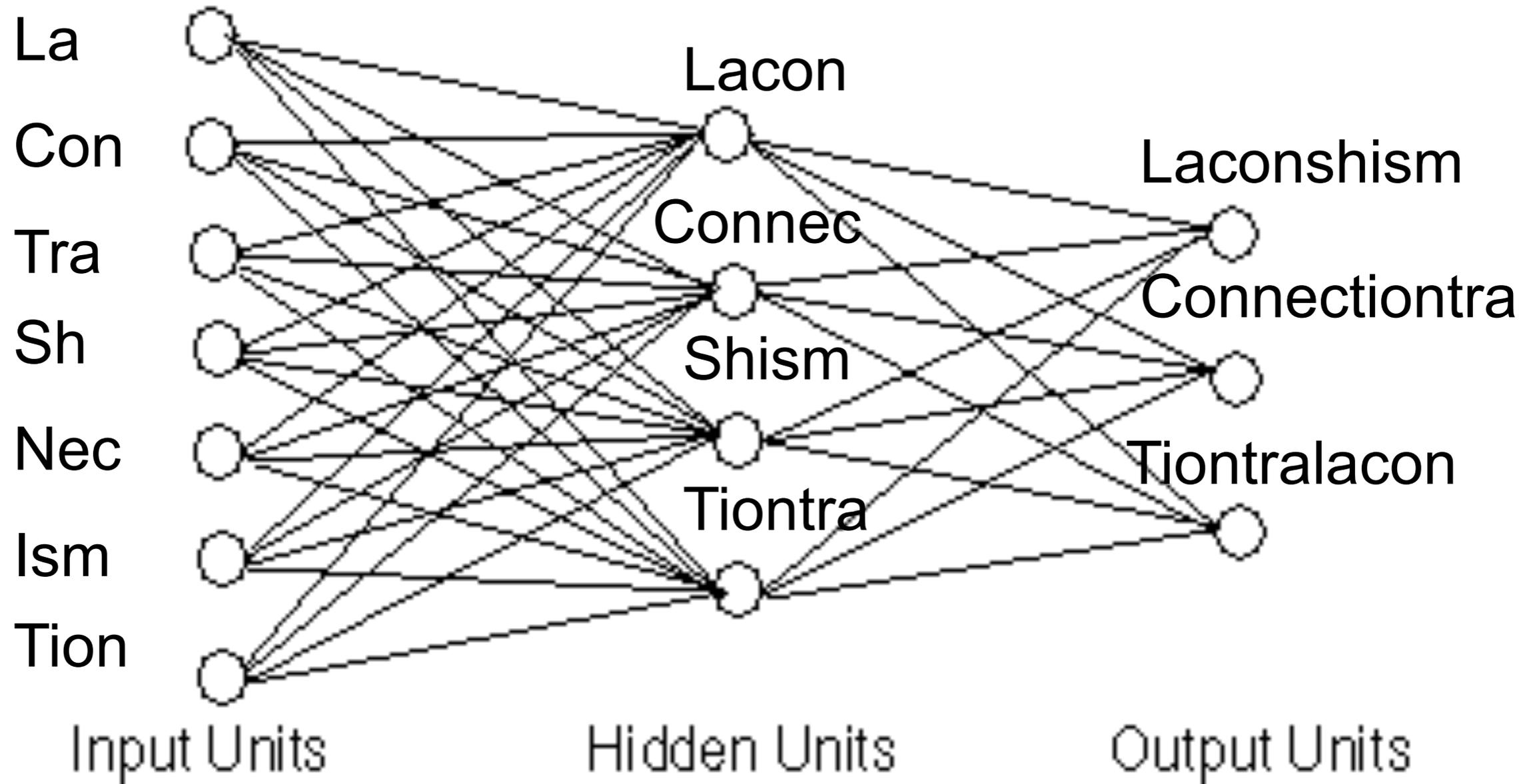
© Metaphysics Research Lab, CSLI, Stanford University. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

# Connectionism



© Metaphysics Research Lab, CSLI, Stanford University. All rights reserved. This content is excluded from our Creative Commons license. For more information, see <http://ocw.mit.edu/fairuse>.

# Connectionism



# Connectionism

- Aim to provide “error-driven, self-organizing, and constructivist learning systems” (Johnson & Munakata, 2005)
- Tries to account for representational change in response to evidence.
- Start with domain-general processes which become domain-specific through learning.

# Connectionism: Virtues

- Self-organizing (that is, it has to compare real output to desired output and adapt but doesn't need a rule for how to get from input to output).
- Accounts for graded responses and sensitivity to subtle statistical regularities (“Connectio-cough” will still trigger “connectionism”)
- “graceful” degradation -- the whole system doesn't fall apart if a part of it is wrong.

# Connectionism: Objections

- Systematicity debate (Fodor, Pylyshyn)
  - Connectionist models learn “John loves Mary” without learning “Mary loves John”
- Doesn’t represent logical relationships. Will need equal amounts of training to learn “John loves Mary” and “John is John”.
- Unlike human language (and presumably the language of thought), it’s not infinitely productive. Learning “John loves Mary and “John is John” doesn’t support: “John who is John, loves Mary who is Mary, who loves John ...” (cannot generate Gertrude Stein)
- Critically, the objection is not that connectionist models cannot do these things (they can be trained to do so) -- it’s that they do not *have* to. That is, nothing about connectionism explains why human cognition has the properties it does.
- Fails to account for one shot learning

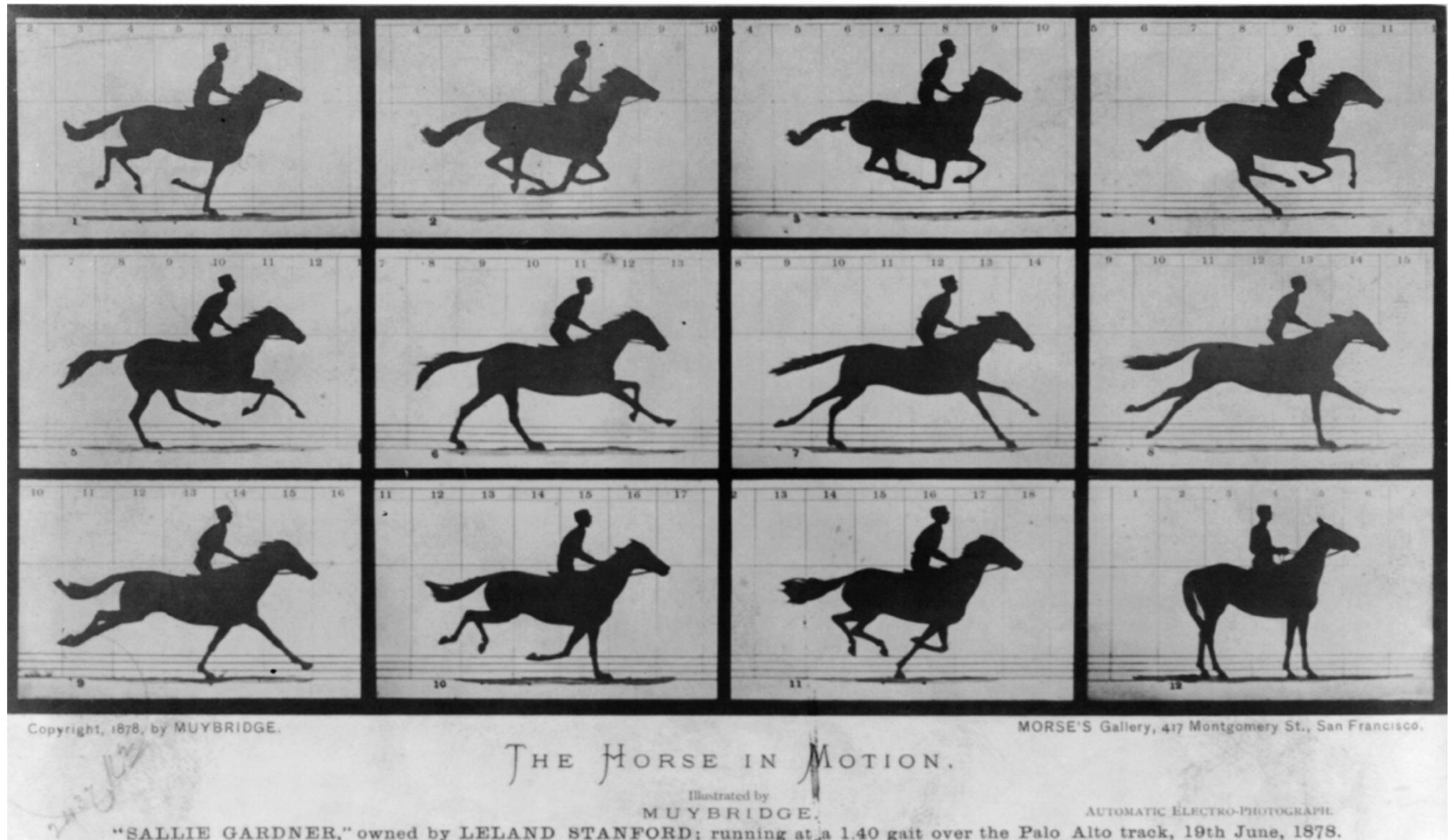
# trends post-Piaget

- connectionism (Bates, Elman, McClelland, Munakata, Rumelhart)
- information processing approaches (Case, Klahr, Siegler)
- modularity (Leslie)
- core knowledge (Spelke, Carey)
- theory theory (Spelke, Carey, Gelman, Gopnik, Meltzoff, Wellman)
  - also biological, maturational accounts (Diamond, Neville, Thelen)

# Information-processing theories

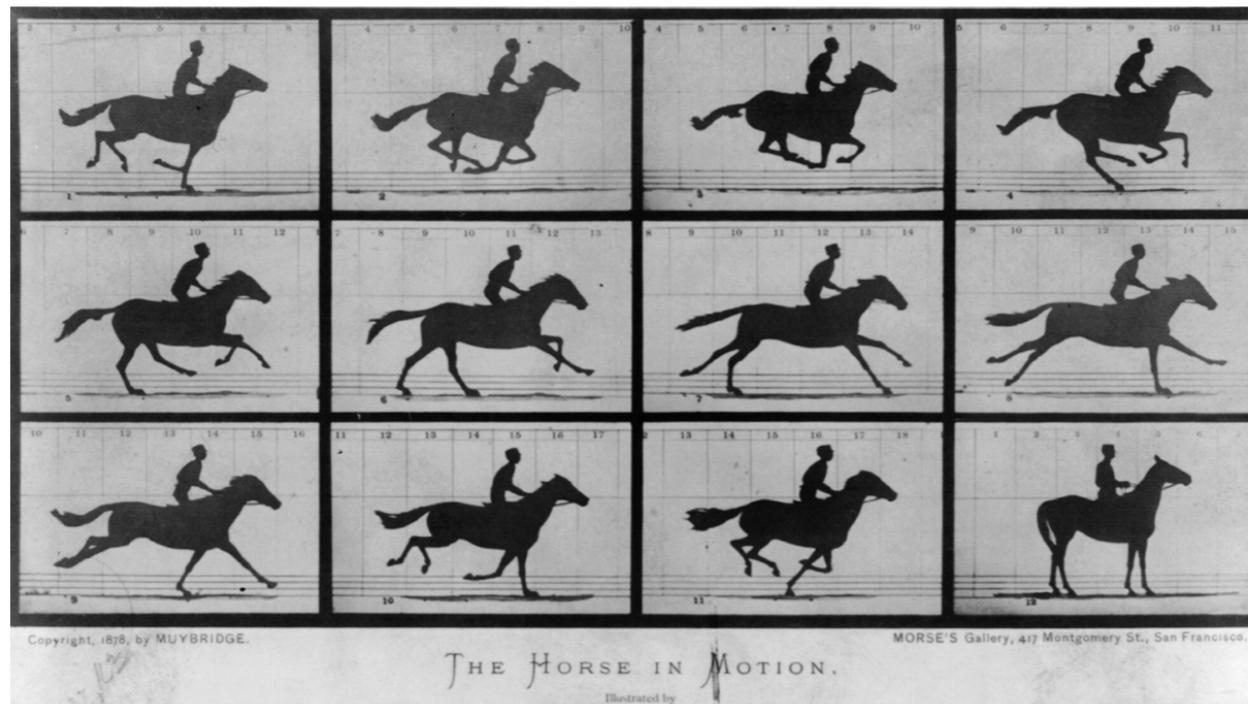
- These researchers often call themselves “Neo-Piagetians” but their focus is on changes in children’s procedures rather than conceptual change per se.
- Focus on cognitive activities (encoding, comparing, storing, attending).
  - Close investigation of processes involved in cognitive tasks
  - Asks what does the system do first, second, third? How much time does each step take?

# Bob Siegler's microgenetic approach



# Bob Siegler's microgenetic approach

- Rate of observation should be relatively dense compared to the rate of change in the phenomenon you are observing.
- A single snapshot won't tell the whole story.



# Information processing - example

- Example
  - Rule 1: only number of weights
  - Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
  - Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
  - Rule 4: compute torque

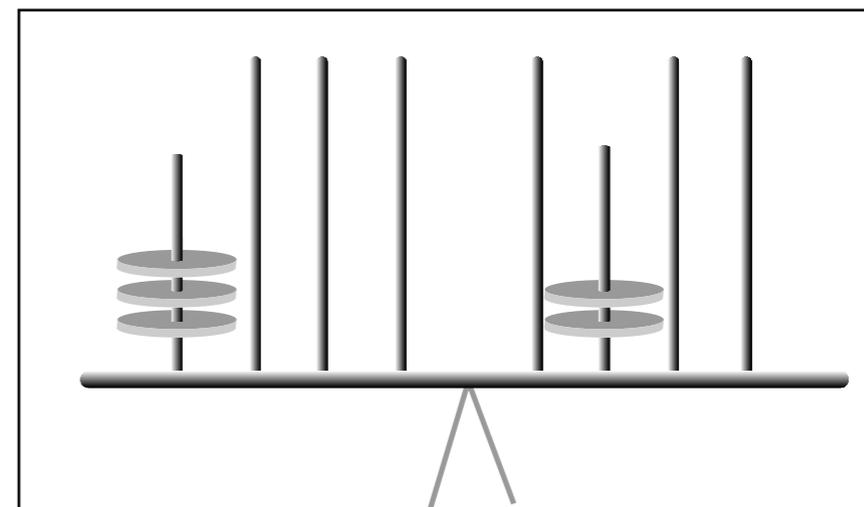


Image by MIT OpenCourseWare.

# Information processing - example

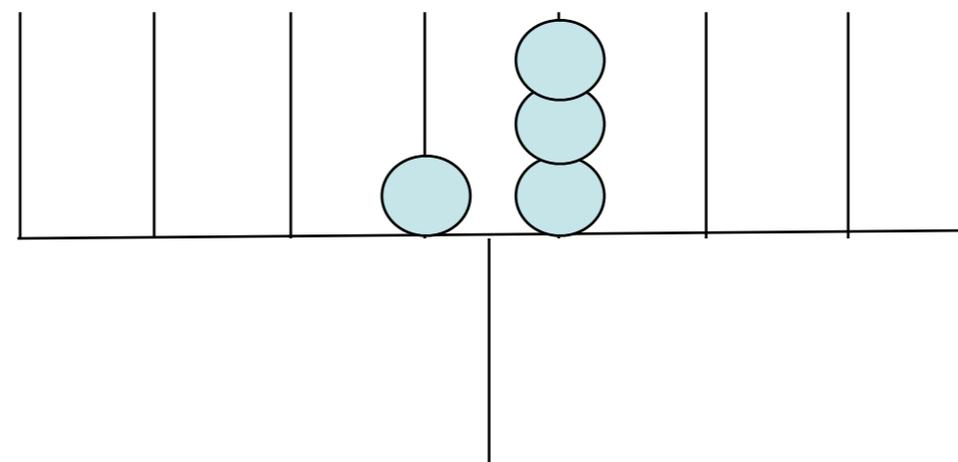
- Example
  - Rule 1: only number of weights
  - Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
  - Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
  - Rule 4: compute torque
- How do you find out what rule children believe?  
By the errors they make ...

# Information processing - example

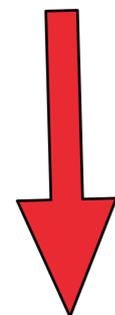
- Example

- Rule 1: only number of weights
- Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
- Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
- Rule 4: compute torque

- If only weight differs ...



All children



# Information processing - example

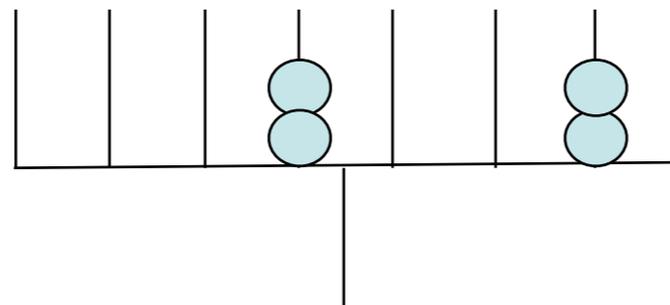
- Example

- Rule 1: only number of weights
- Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
- Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
- Rule 4: compute torque

- If only distance differs ...

Children with rules 2-4

Children with rule 1



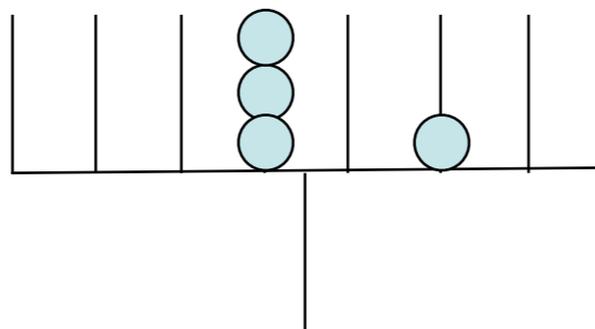
# Information processing - example

- Example

- Rule 1: only number of weights
- Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
- Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
- Rule 4: compute torque

- If one side has more weight (enough to go down) but the other has more distance ...

Children with rules 1,2, and 4

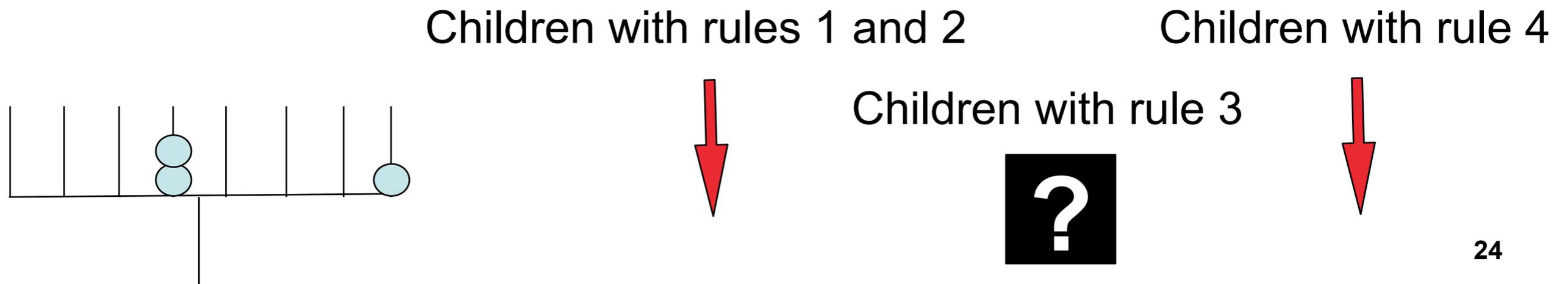


Children with rule 3



# Information processing - example

- Example
  - Rule 1: only number of weights
  - Rule 2: if two sides have equal weight, consider distance from fulcrum as well.
  - Rule 3: consider both weight and distance (but don't know what to do if one side has more weight and the other is further out).
  - Rule 4: compute torque
- If one side has more distance (enough to go down) but the other has more weight ...



# Information processing -- error analysis

1. If only weight differs (All rules predict correct answer)
2. If only distance differs (Rule 1 predicts balance; 2-4 predict correctly)
3. If one side has more weight (enough to go down) and other side more distance. (Rules 1, 2, and 4 predict correctly, Rule 3 predicts at chance)
4. If one side has more weight and other side more distance (enough to goes down) (Rule 1, 2 predict opposite side; Rule 3 predicts at chance, Rule 4 predicts correctly)

# Methods used in information processing approaches

- Microgenetic approaches
- Error analyses
- Tests of memory, recall
- Reaction times
  - assume more steps of information processing take more time. (e.g., children mentally rotate objects more slowly --  $7^\circ/\text{msec}$  than adults  $4^\circ/\text{msec}$ )
- Eye movements
  - as index of attentional control (e.g., older children scan more systematically than younger children)

# Information processing

- Virtues:
  - Detailed analyses of strategies
  - Detailed analyses of change
  - Attention to information processing constraints is critical for determine the kind of evidence the mind can access -- and the kind of evidence it will tend to access

# Information processing

- Critiques:
  - Blurs competence v. performance distinction
    - maybe children have conceptual knowledge but memory, motor limitations, ability to deploy multiple strategies, etc. mask it.
    - “Information-processing changes” might result from conceptual ones (e.g., memory improves because you have more knowledge about a domain).
  - Explains local phenomenon in great detail (balance beams; adding) but no overarching account.
  - Pre-requisites to cognitive changes, not explanations of it. “Buying a telescope doesn’t make you Galileo” (Gopnik & Meltzoff, 1997)
  - Feature of cognition, not theory of it.

# trends post-Piaget

- connectionism (Bates, Elman, McClelland, Munakata)
- information processing approaches (Case, Klahr, Siegler)
- modularity (Leslie)
- core knowledge (Spelke, Carey)
- theory theory (Spelke, Carey, Gelman, Gopnik, Meltzoff, Wellman)
  - also biological, maturational accounts (Diamond, Neville, Thelen)

# Background

- From the beginning of Western philosophy: two competing traditions.
- **Rationalism** -- Some knowledge is innate. (Plato)
- **Empiricism** -- "Nothing is in the intellect which was not first in the senses." (Aristotle)

# Modularity, core knowledge and theory theories of cognitive development

- ... are three different attempts to explain the gap between experience and our understanding of it
- What gap? Can't we just perceive things as they are?
- Removing the veil of self-evidence ...

# What gap? Try this multiple choice quiz

- A speaker of Quinian points to this and says “Gavagi.” What does Gavagai mean?



Image: Wikimedia. Ann Harrison. CC BY-SA.

A. A rabbit    B. Rabbit plus grass    C. lunch

# What gap? Try this multiple choice quiz

- Complete the sentence: “The giraffe has a very long ...”

A. neck

B. temper

C. flight to Kenya

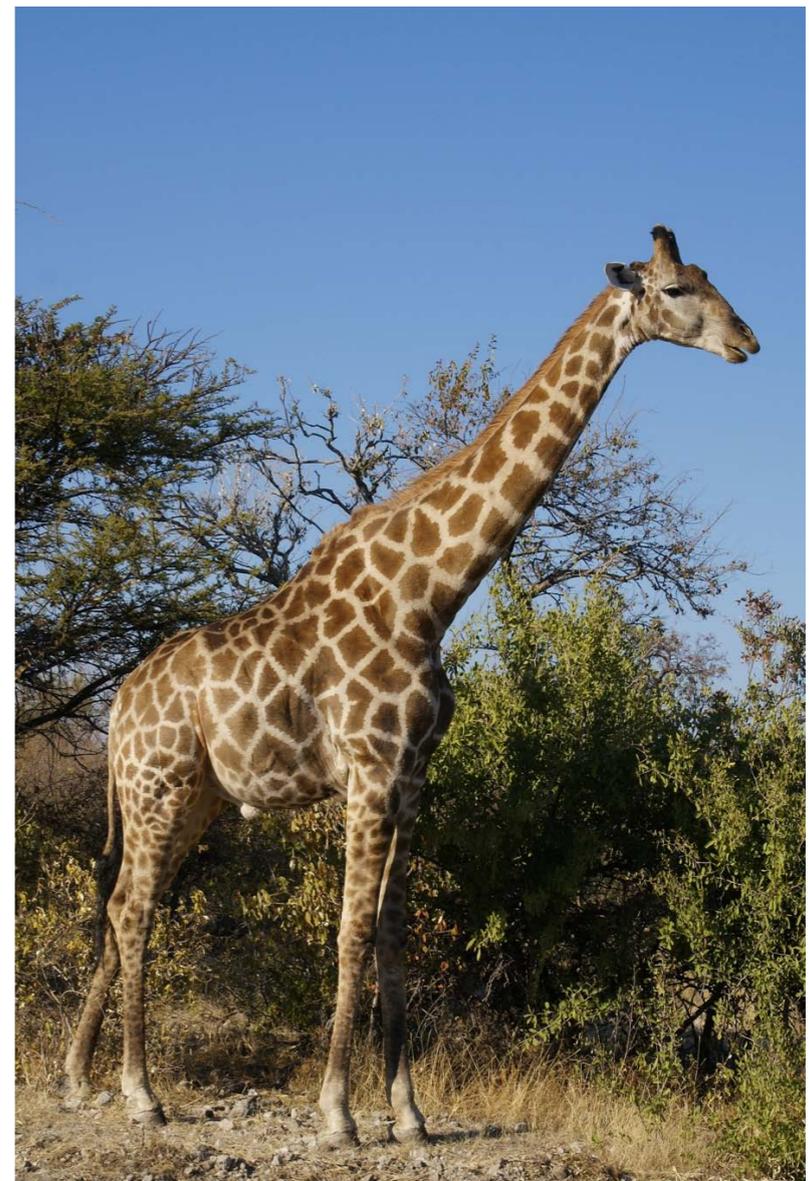
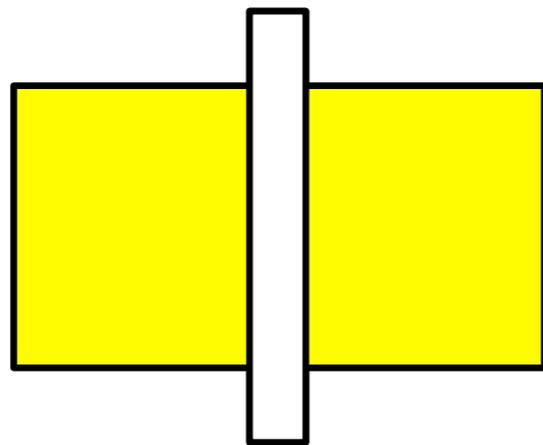


Image: Wikimedia. Hans Hillewaert. CC BY-SA.

# What gap? Try this multiple choice quiz

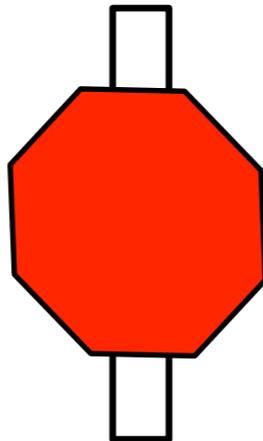
- What's behind the rectangle?



A.



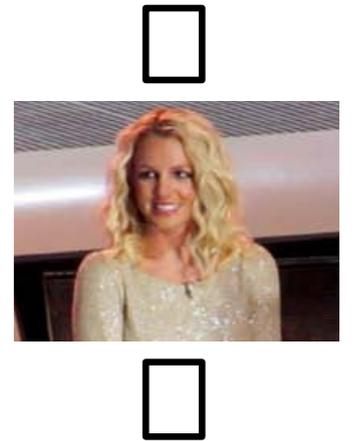
B.



C.



D.



# Why is learning a hard problem?

- Because most problems of learning are massively underconstrained by the data -- this is the problem of **induction**.
- Unless something is deductively true, evidence is always open to many logically possible construals.
- Indeed, we may take identical statistical evidence and interpret it in quite different ways.

# Broad Institute

- Scientists came up with a technique for changing the gender of mice. They injected 100 mice with glucosin before impregnation. 70% of the newborns were male.

# Broad Institute

- On a scale of 1-7 where 7 is “yes I’d take a look at the paper” and 1 is “no I wouldn’t”, how likely are you to seriously consider the possibility that glucosin might affect mice gender?

# Institute of Noetic Sciences

- ESP experts came up with a technique for changing coin flips. They flipped a coin 100 times while chanting “heads”. 70% of the coin flips came up heads.

# Institute of Noetic Sciences

- On a scale of 1-7 where 7 is “yes I’d take a look at the paper” and 1 is “no I wouldn’t”, how likely are you to seriously consider the possibility that chanting “heads” might affect coin flips?

# The problem of induction

- Problems of induction arise all over this building: in vision, in motor planning, in language learning, in causal reasoning.
- Yet despite logically innumerable possibilities, we converge on the same conclusions most of the time.
- Thus there must be constraints on learning.
- Debate between **modularity**, **core knowledge** and **theory theory** approaches is over the nature of the constraints.

# Innate, domain-specific, modular, core knowledge, oh my

- Everyone thinks some aspects of cognition are innate -- but what?
  - Innate domain-general learning abilities
  - Innate cognitive modules
  - Innate domain-specific knowledge

# Innate?

- Not necessarily 'born with it'
- Not necessarily 'experience independent' (cortical columns in cats)
- Inter-defined with learning: what is innate does not require learning.

# Domain-specific

- What's a domain?
- “A body of knowledge that identifies and interprets a class of phenomena assumed to share certain properties and to be of a distinct and general type. A domain functions as a stable response to a set of recurring and complex problems faced by the organism.” (Hirschfeld & Gelman, 1994)

# Domain-specific?

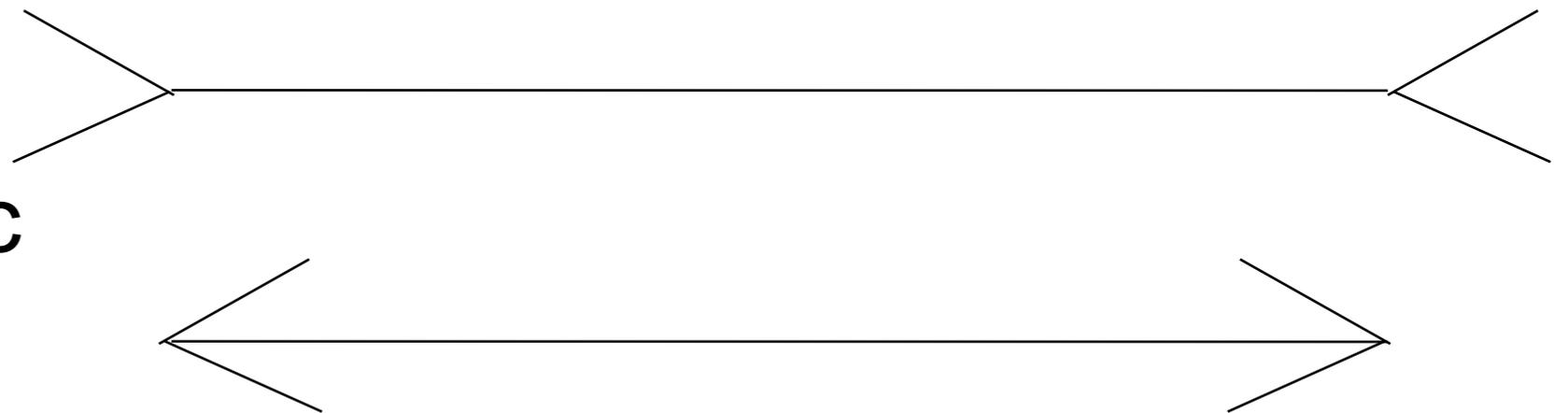
- Physics, psychology, biology
- Language, vision, spatial relations
- Chess?
  
- For practical purposes, believing that cognition is domain specific means:
- believing that there are distinct ways of acquiring and organizing knowledge that may reflect real differences in the structure of the input (e.g., the external world).

# Modular?

Fodor (1983) popularized the idea and specified that modules are:

- innate
- domain-specific
- fast
- **encapsulated**

- Argued for a modular account of vision, language processing, etc.



# What are the candidates for modularity?

- Fodor argued specifically for distinguishing perceptual (peripheral) modularity (things like language and vision) and cognitive (central) modularity.
- However, many developmental researchers have argued for the plausibility of modular approaches to central (e.g., higher-order cognitive) processing.

# Example:Leslie: core architecture for the cognition of agency

Properties of agents	Processing device	Level of understanding
mechanical	ToBY	“agents and objects”
actional	ToMM1	“agents and action”
cognitive	ToMM2	“agents and attitudes”

# Example: Leslie: core architecture for the cognition of agency

- Fodor's response? "Modularity gone mad"

# Critiques of modularity

- Maturation theory; largely independent of interaction with environment
- No explanation of why these modules should come online when they do.

# Characteristics of core knowledge

From Susan Carey slide

- Distinct systems of domain-specific mental representations with conceptual content (beyond sense data)
- Acquisition supported by innate domain-specific learning mechanisms (perceptual input analyzers)
- Evolutionary ancient (often)
- Continuous throughout development

# Core knowledge hypothesis

- “We argue that human reasoning is guided by a collection of innate domain-specific systems of knowledge. Each system is characterized by a set of core principles that define the entities covered by the domain and support reasoning about those entities ...

# Core knowledge hypothesis

- ... Learning on this view consists of an enrichment of the core principles, plus their entrenchment, along with the entrenchment of the ontology they determine.” (Carey & Spelke, 1994)

# What are the candidates for core knowledge?

- Knowledge of objects
- Knowledge of agents
- Knowledge of number
- Knowledge of space
- Moral knowledge?
  - We'll talk about the evidence for core knowledge in each of these domains through the rest of this class ...

# Theory theory

- What's a theory and why think children have theories?

# What theory theory does not mean

- That babies and children have explicit, awareness of their beliefs (although they might).
- That babies and children engage in deliberate hypothesis testing (although they might).
- That babies and children wear white lab coats.

# What are characteristics of a theory?

- Structural features
  - Abstract (goes beyond the evidence)
  - Ontologically committed (categories are defined by the theory)
- Functional features
  - Causal (supports prediction, intervention, explanation, and counterfactual claims).
- Dynamic features
  - Defeasible/revisable with evidence
  - Affect the interpretation of evidence.

# What are characteristics of a theory?

- **Structural features**
  - Abstract (goes beyond the evidence)
  - Ontologically committed (categories are defined by the theory)
- **Functional features**
  - Causal (supports prediction, intervention, explanation, and counterfactual claims).
- **Dynamic features**
  - **Defeasible/revisable with evidence**
  - **Affect the interpretation of evidence.**

# Theory theory

- “Place some children in a universe that is radically different from our own . . . If they come up with representations that are an accurate account of our universe, modularity is right. If they come up with representations that are an accurate account of their universe, the theory theory is right.” (Gopnik & Meltzoff, 1997).

# So we might think children had theories if ...

- Their beliefs about the world were abstract, causal, and coherent
- Those beliefs changed with evidence ...
- Their beliefs affected the interpretation of evidence.
  
- Full disclosure -- most of my own work comes from this tradition

# Fine distinction ...

- **Modularity-nativism**: some higher-order cognitive representations are generated by innate modules and are indefeasible (Leslie).
- **Core knowledge nativism**: some higher-order cognitive representations are innate and indefeasible but not encapsulated (knowledge of space, objects, and number). (Carey/Spelke)
- **Starting-state nativism (Theory theory)**: “child is innately endowed with a particular set of representations and rules operating on these representations . . . (but) such innate structures (are) defeasible; any part of them could be, and indeed will be, altered with new evidence.” (Gopnik & Meltzoff, 1997).

# Summary (rough guide -- don't take too literally ...)

- What types of knowledge are innately specified?
  - Just learning mechanisms (connectionism)
  - Particular content knowledge (modularity; core knowledge; theory theory)
- Are basic processes central to learning ...
  - domain-general (connectionism, theory theory, core knowledge)
  - or domain-specific (core knowledge)
  - no such thing! (Fodor's modularity)
- What stays constant over development?
  - Just basic learning mechanisms (connectionism, theory theory)
  - Particular content knowledge (modularity; core knowledge)

MIT OpenCourseWare  
<http://ocw.mit.edu>

9.85 Infant and Early Childhood Cognition  
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.