

RSA, ambiguity, and relevance

a summary of our progress

Gregory Scontras

Alya Achimova

Christian Stegemann

Martin Butz

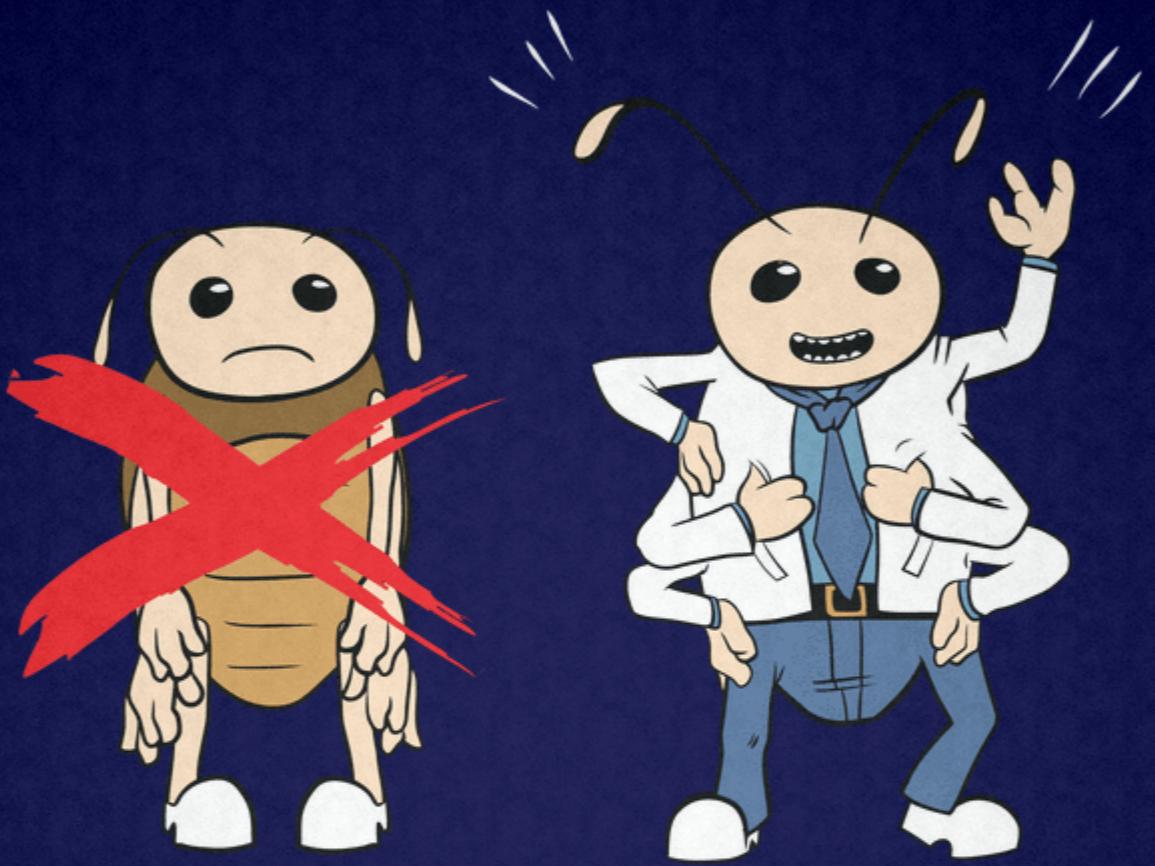
A stylized illustration of a man from the chest up. He is wearing a dark grey bowler hat with a black band. The word "ambiguity" is written in white capital letters on the front of the hat. He has a large, bright red question mark where his head would be. He is wearing a white dress shirt with a black tie and a dark grey suit jacket. A large, solid red circle is pinned to the center of his collar.

ambiguity

ambiguity

a design flaw of language?

(Grice, 1975; Chomsky, 2002)



**It's not a bug
It's a feature!**

efficiency



**It's not a bug
It's a feature!**

(Piantadosi et al., 2012)

but maybe ambiguity adds more than just efficiency...

ambiguity ≈ information gaps

MIND THE GAP

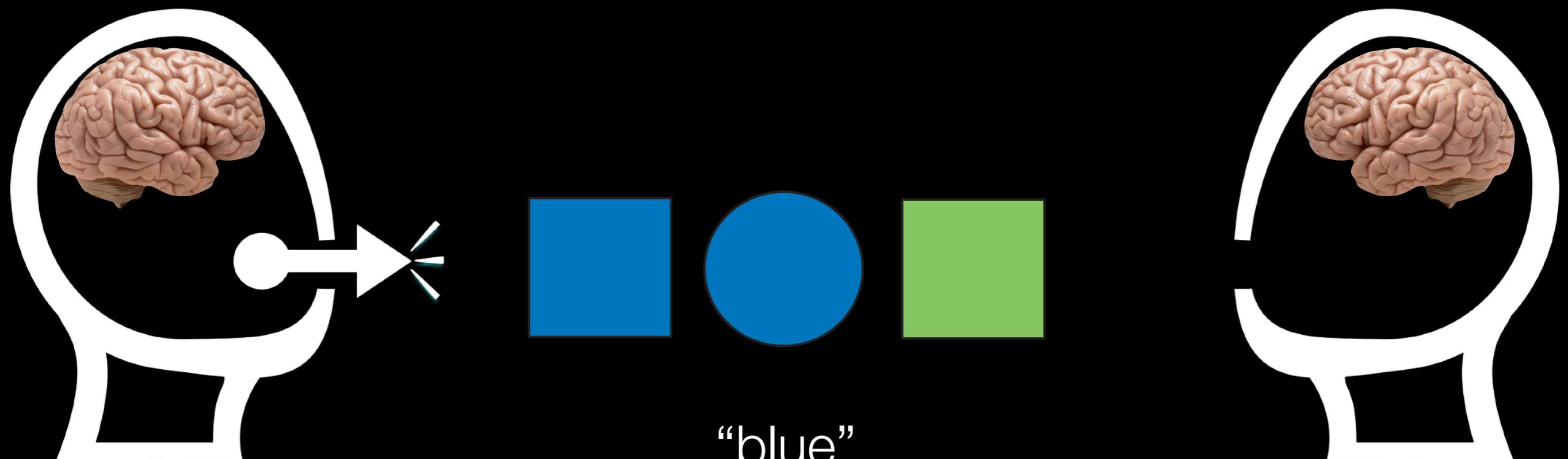
ambiguity

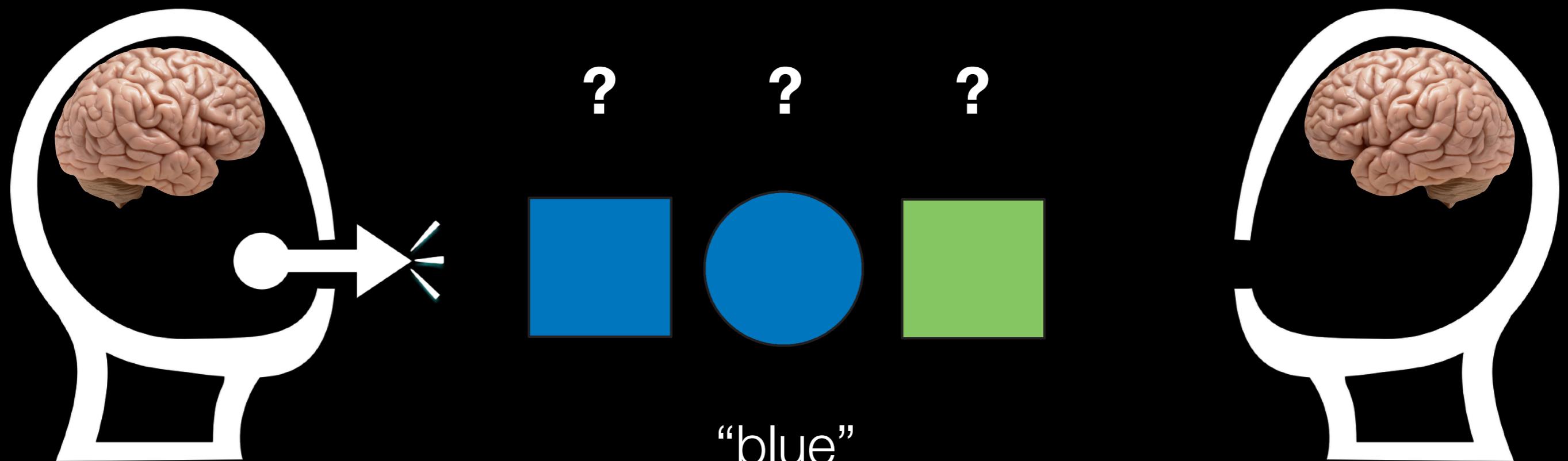
≈

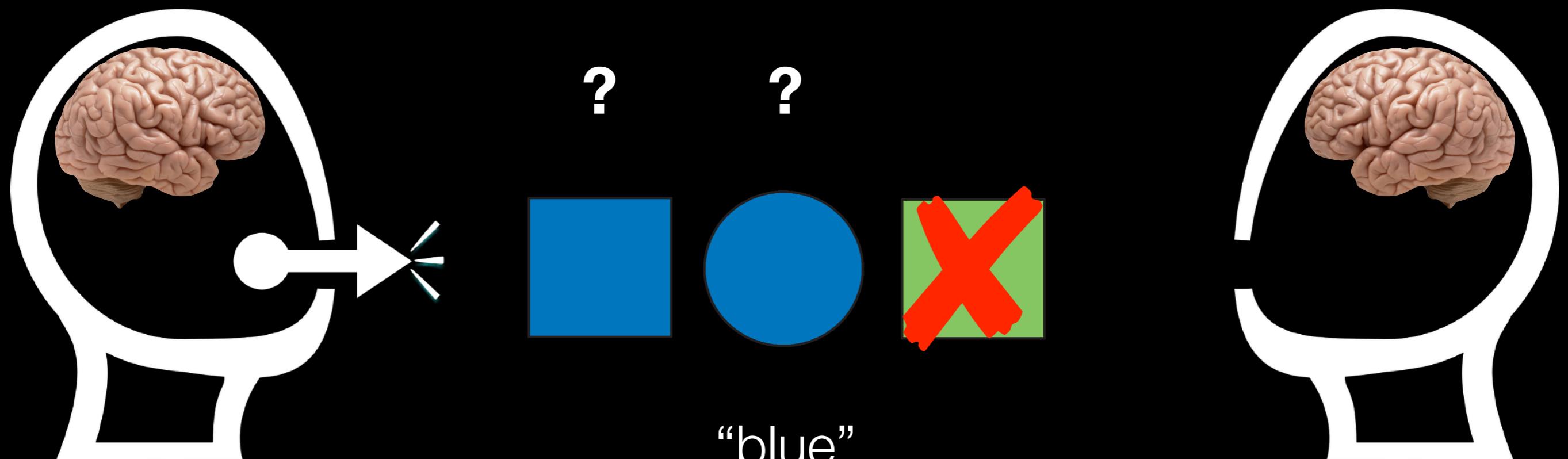
information gaps

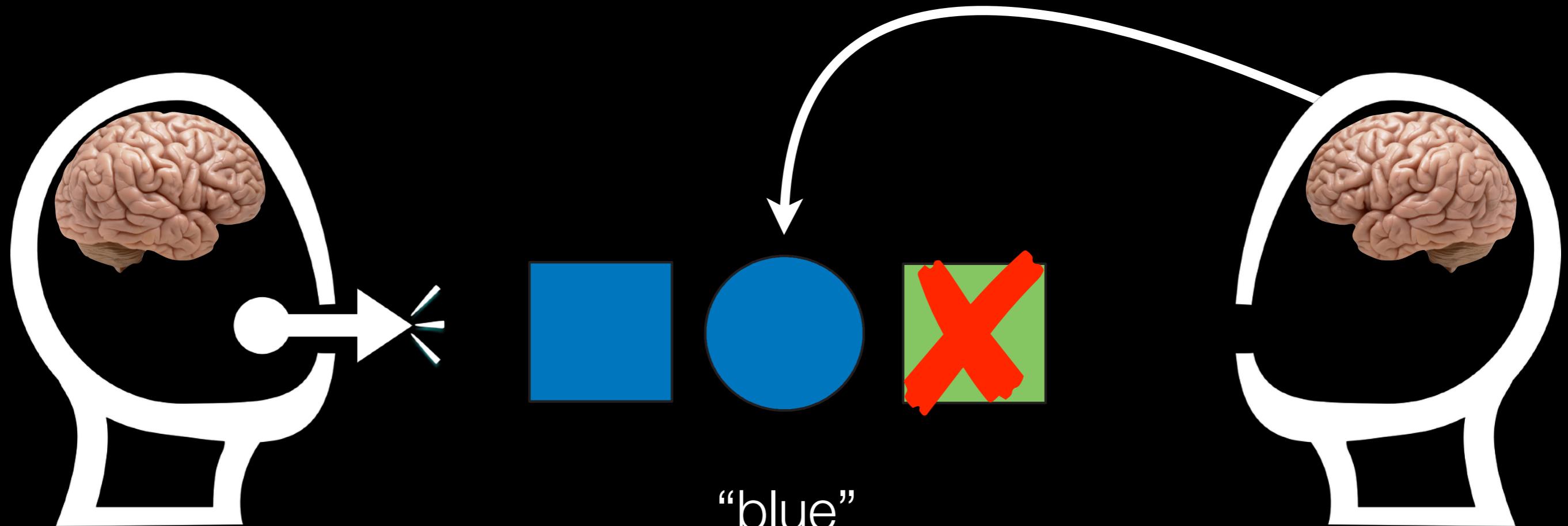




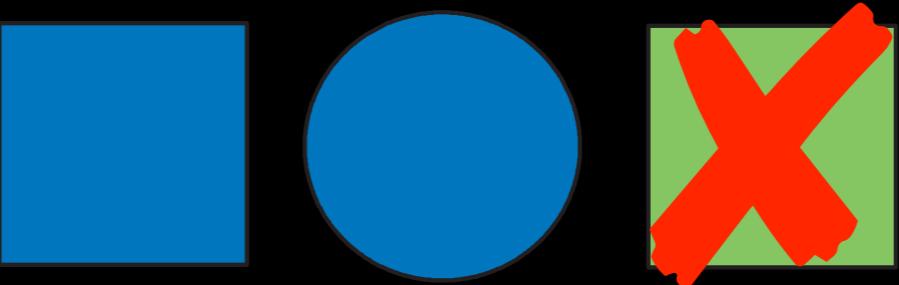








a preference
for circles!



“blue”



R S A

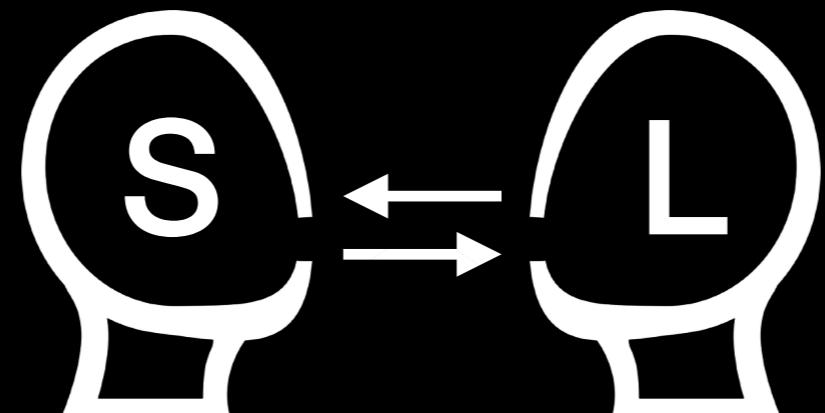
rational

speech

act

R S A

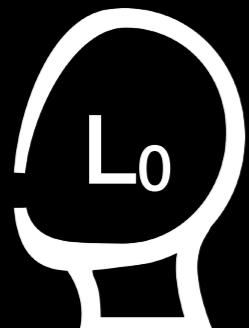
speaker observes state, chooses utterance
listener hears utterance, infers state



speaker and listener coordinate:
utterance + interpretation that maximizes the probability of
correctly resolving the Question-Under-Discussion

(Frank & Goodman, 2012; Goodman & Frank, 2016)

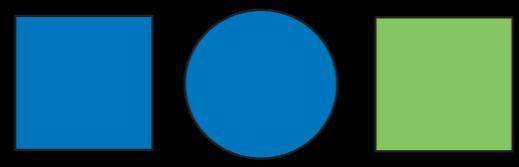
literal listener



$$P_{L_0}(s \mid u) \propto [[u]](s)$$

resolves the QUD (infers the state) given utterance u

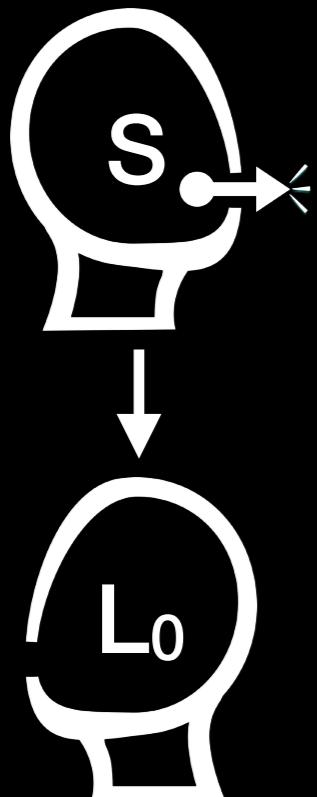
uniform distribution over possible states given utterance



? ? ?

“blue”

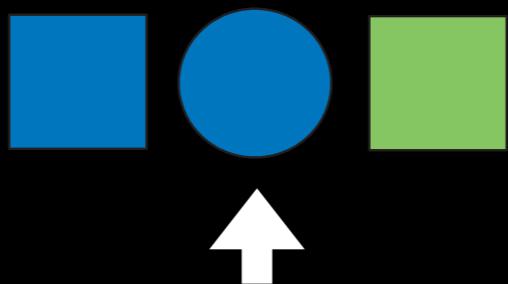
pragmatic speaker



$$P_{S_1}(u \mid s) \propto \exp(\alpha \cdot [\log(L_0(s \mid u)) - C(u)])$$

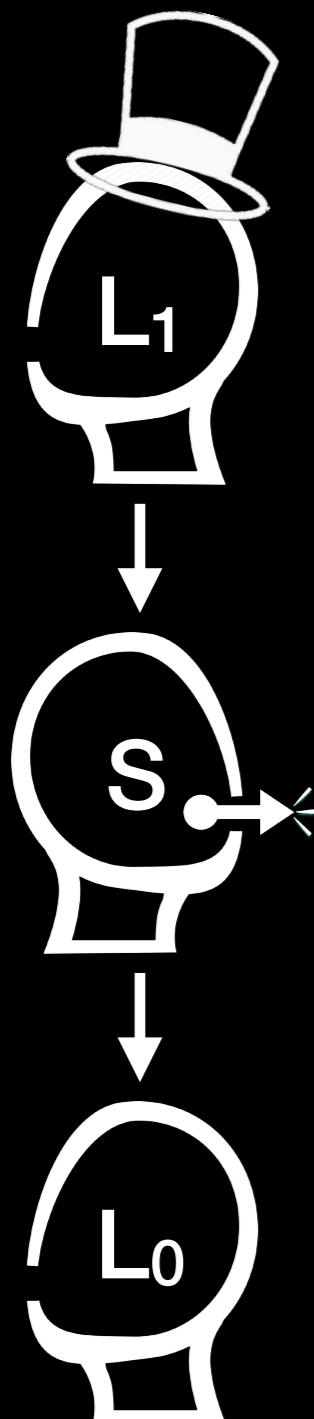
$$P_{L_0}(s \mid u) \propto \llbracket u \rrbracket(s)$$

chooses an utterance that best resolves the QUD
distribution over possible utterances given observed state

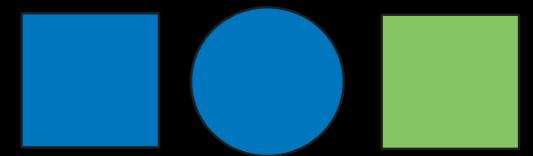


“blue” “circle”

pragmatic listener



$$P_{L_1}(s \mid u) \propto P_{S_1}(u \mid s) \cdot P(s)$$



? ? ?

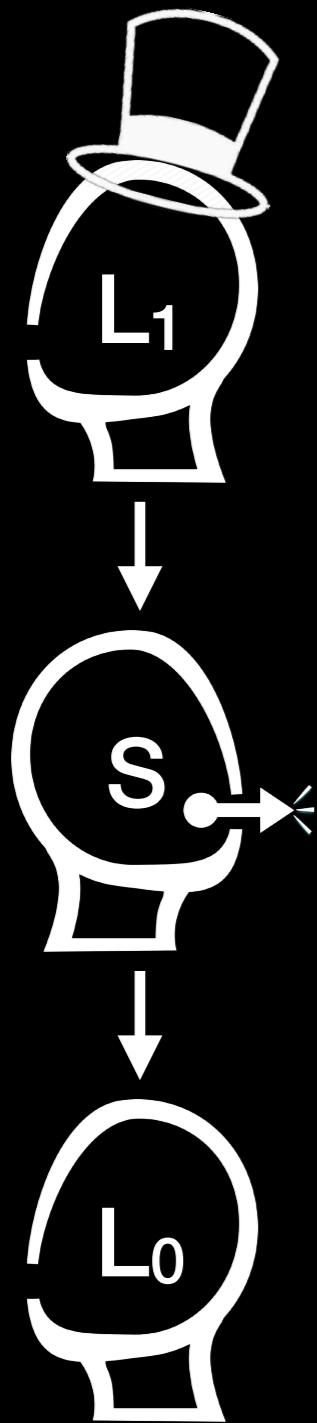
“blue”

$$P_{S_1}(u \mid s) \propto \exp(\alpha \cdot [\log(L_0(s \mid u)) - C(u)])$$

$$P_{L_0}(s \mid u) \propto \llbracket u \rrbracket(s)$$

updates beliefs about state given utterance

vanilla RSA model

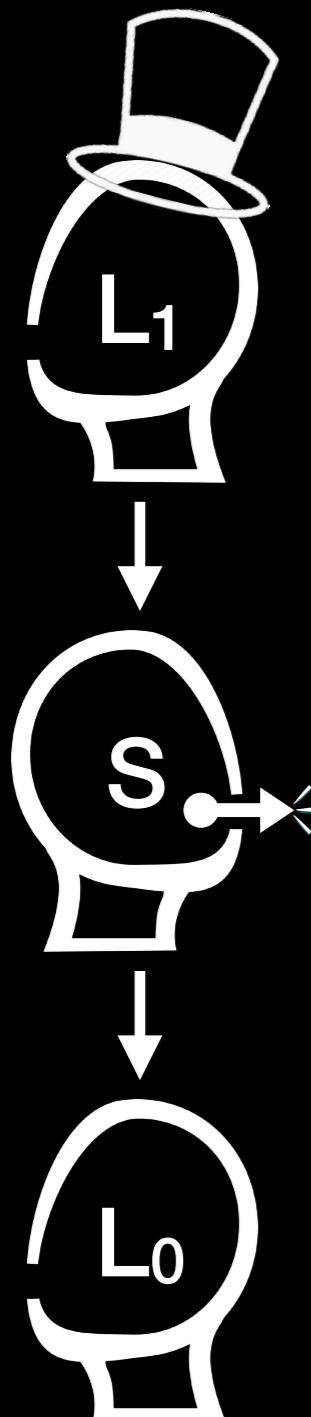


$$P_{L_1}(s \mid u) \propto P_{S_1}(u \mid s) \cdot P(s)$$

$$P_{S_1}(u \mid s) \propto \exp(\alpha \cdot [\log(L_0(s \mid u)) - C(u)])$$

$$P_{L_0}(s \mid u) \propto \llbracket u \rrbracket(s)$$

introducing preferences



$$P_{L_1}(s \mid u, \mathbf{f}) \propto P_{S_1}(u \mid s) \cdot P(s \mid \mathbf{f})$$

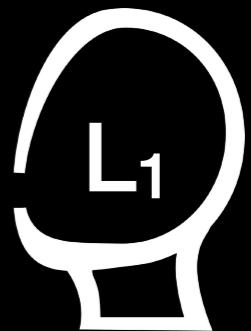
$$P_{S_1}(u \mid s) \propto \exp(\alpha \cdot [\log(L_0(s \mid u)) - C(u)])$$

$$P_{L_0}(s \mid u) \propto \llbracket u \rrbracket(s)$$

inferring preferences



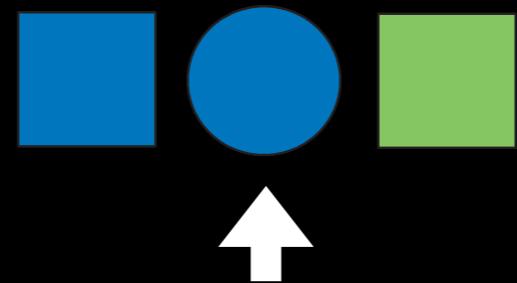
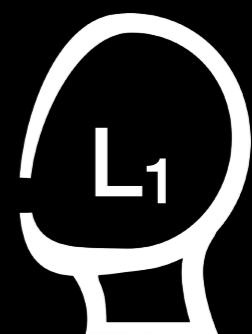
$$P_{L_1}(s \mid u, \mathbf{f}) \propto P_{S_1}(u \mid s) \cdot P(s \mid \mathbf{f})$$



inferring preferences



$$P_{L_1}(s \mid u, \mathbf{f}) \propto P_{S_1}(u \mid s) \cdot P(s \mid \mathbf{f})$$

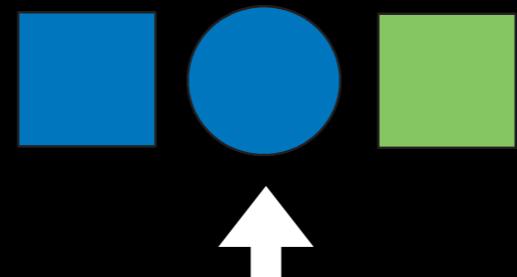
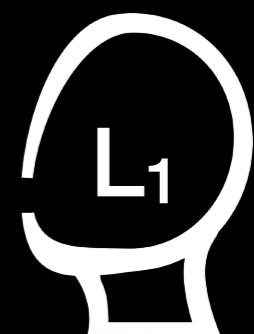


“blue”

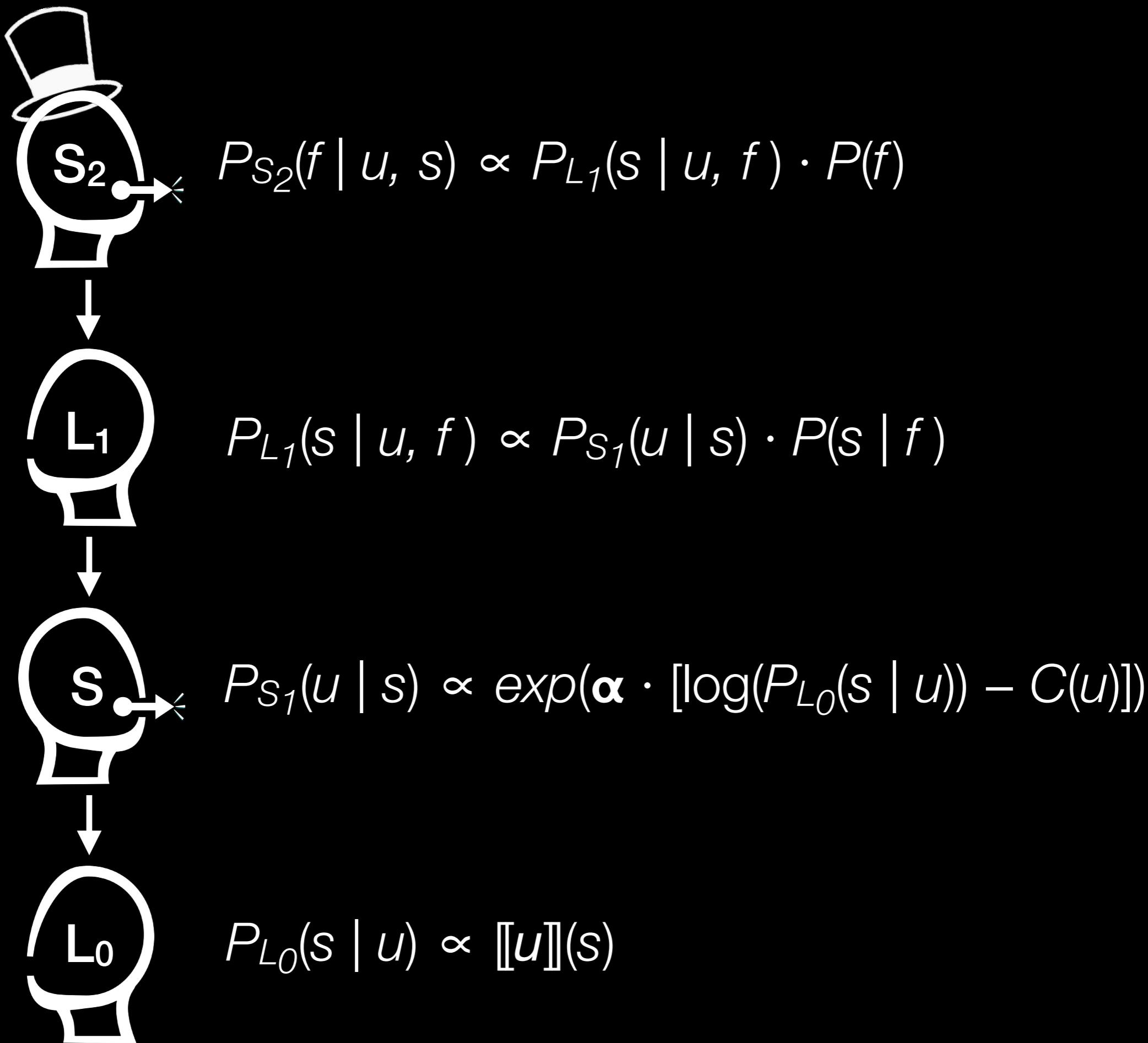
inferring preferences



$$P_{L_1}(s \mid u, \mathbf{f}) \propto P_{S_1}(u \mid s) \cdot P(s \mid \mathbf{f})$$



“blue”



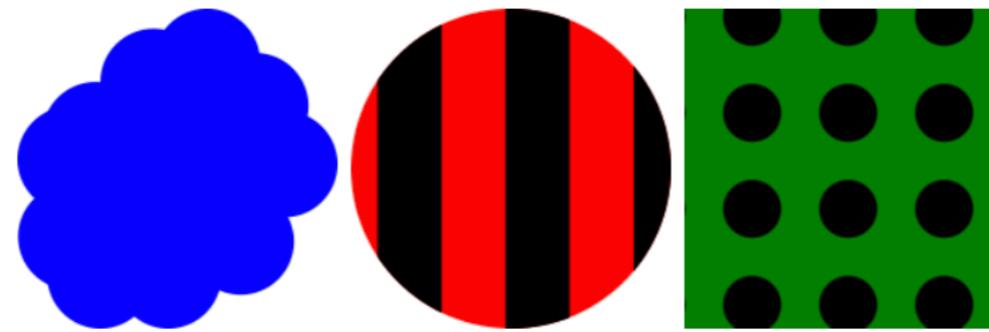
experiment 1

inferring preferences

Progress:

Instructions

You will see a series of scenarios with two people and three objects. For example, a scenario might feature Sue and Bill and the three objects below:



Sue asks Bill to choose one of the objects. In doing so, she is allowed to mention one (and only one) of the features of the target object (i.e., its color, its shape, or its pattern). However, Bill might have a preference for certain colors (blue, red, or green), shapes (clouds, circles, or squares), or patterns (solid, striped, or polka-dotted).

You will observe Sue's utterance and Bill's choice. Your task is to determine the preferences Bill might have had in making the choice.

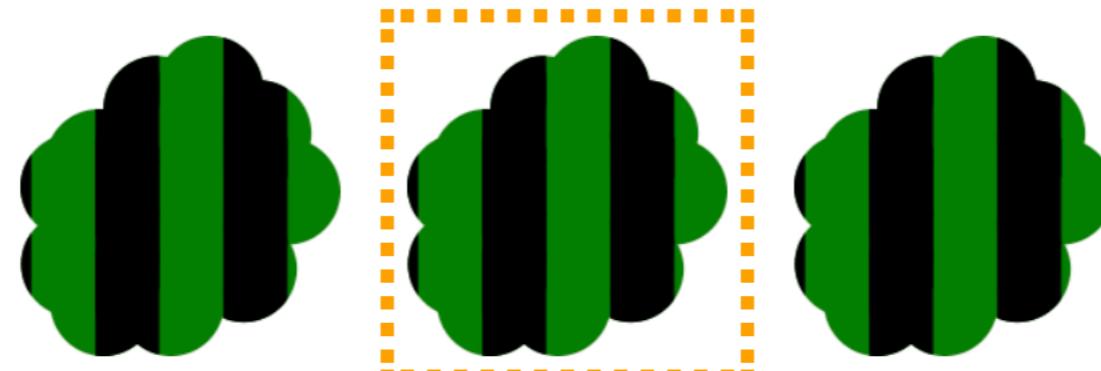
[Continue](#)

Progress: 

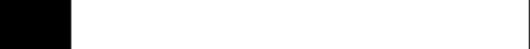
Training

You might see a scenario like the following:

Suppose John wants to signal an object in the following scene to Mary.
John says "**cloud**" and Mary chooses the outlined object:



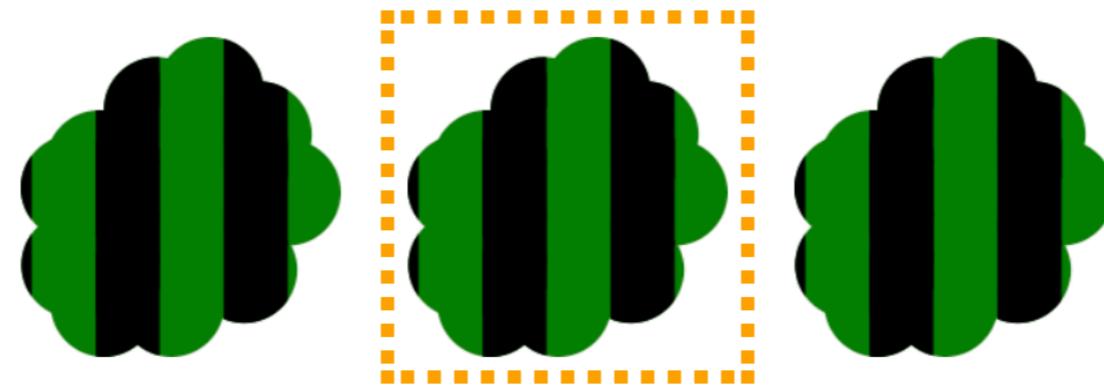
[Next](#)

Progress: 

Training

You might see a scenario like the following:

Suppose John wants to signal an object in the following scene to Mary.
John says "**cloud**" and Mary chooses the outlined object:



You will then be asked whether Mary has any preferences:

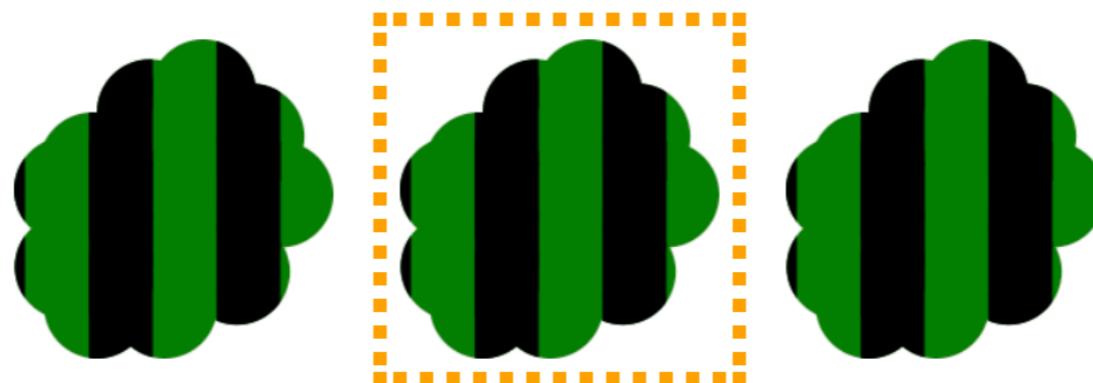
Based on this choice, do you think Mary has a preference
for certain types of objects?

[Next](#)

You might see a scenario like the following:

Suppose John wants to signal an object in the following scene to Mary.

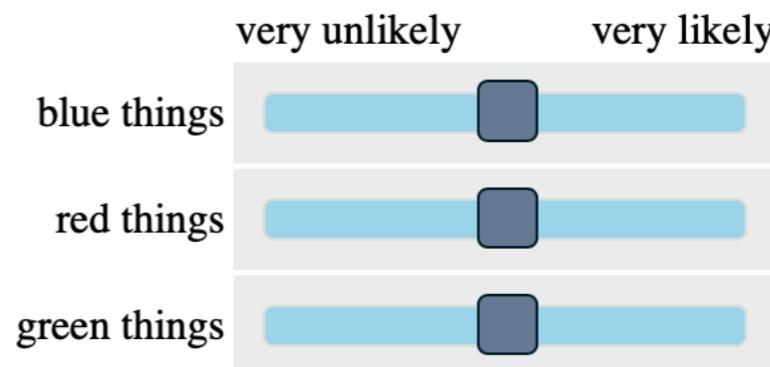
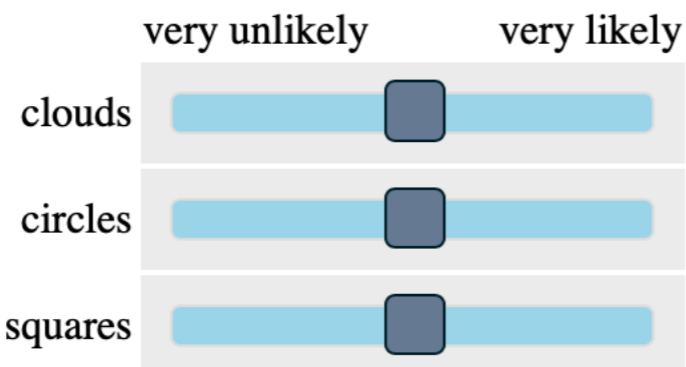
John says "cloud" and Mary chooses the outlined object:



You will then be asked whether Mary has any preferences:

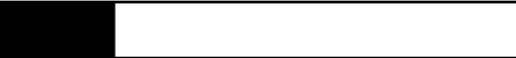
Based on this choice, do you think Mary has a preference
for certain types of objects?

Given that Mary had to choose from three objects of the same type, Mary's choice does not give you any information about her preferences. In this case you should adjust all the sliders to equal values (ideally centered), indicating that Mary's preferences are unknown:



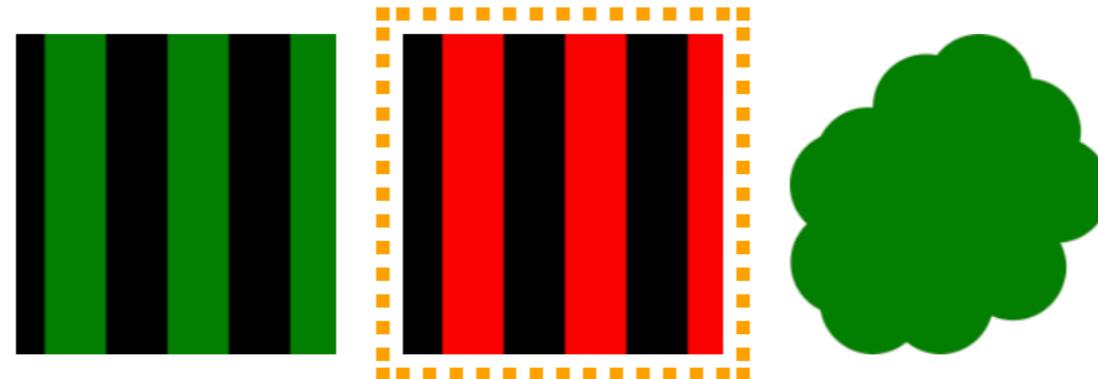
Click "Continue" to begin the experiment.

[Continue](#)

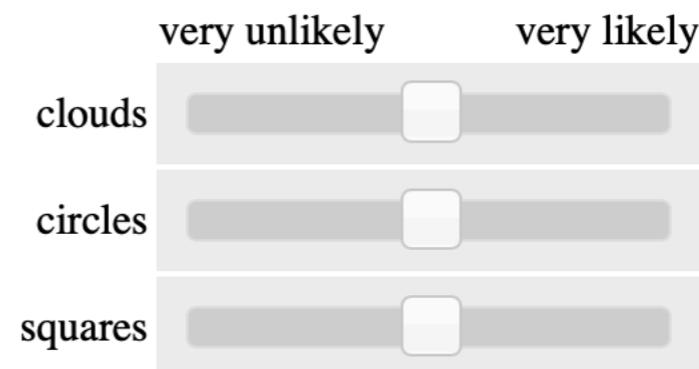
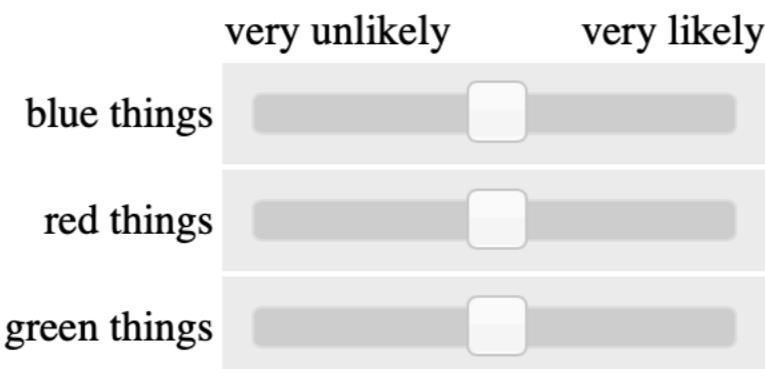
Progress: 

Suppose Joseph wants to signal an object in the following scene to Savannah.

Joseph says "**striped**" and Savannah chooses the outlined object:



Based on this choice, do you think Savannah has a preference
for certain types of objects?



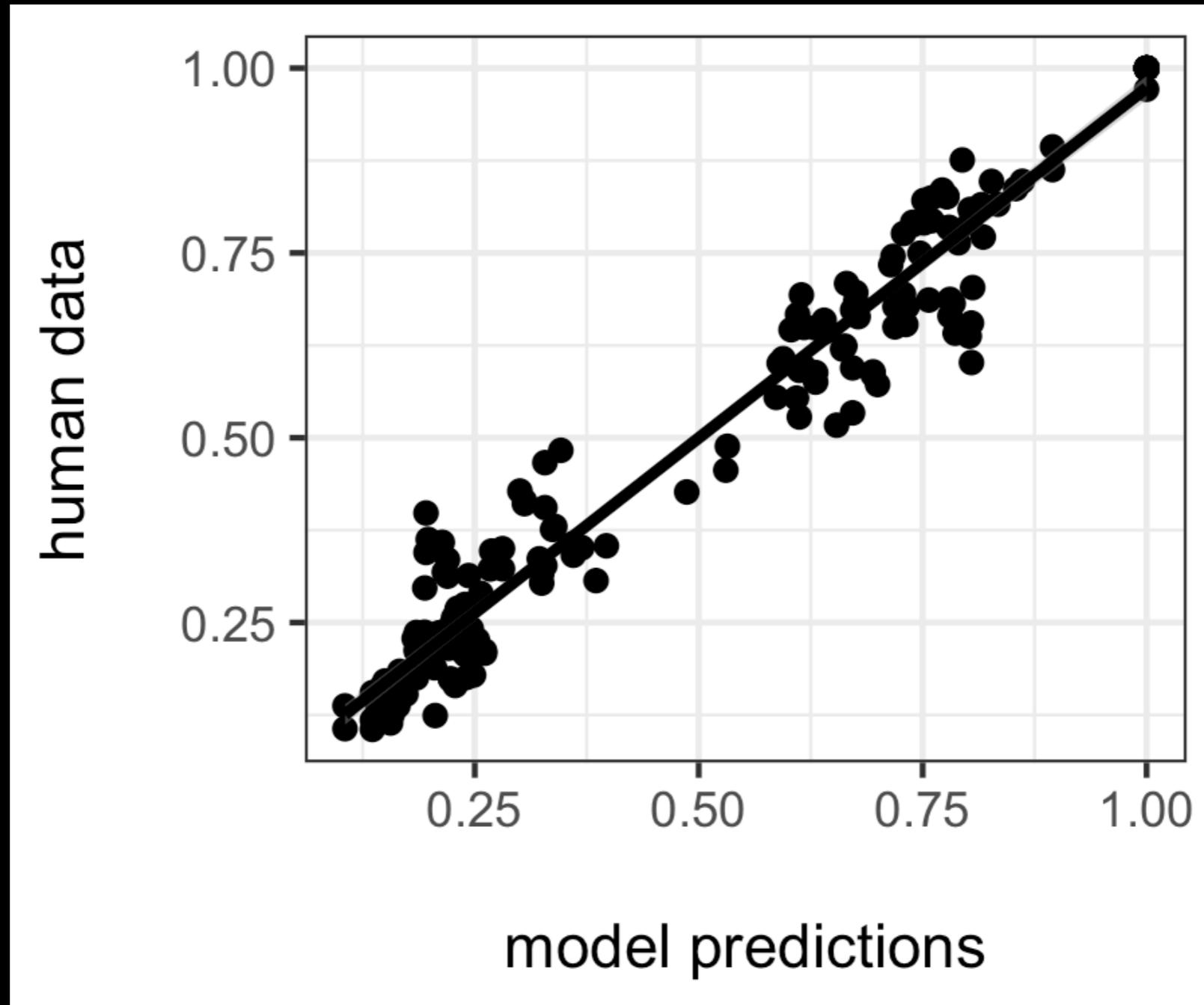
[Continue](#)

15 trials

results

SOMETHING ABOUT MODEL FITTING

results



$r^2 = 0.96$, 95% CI [0.94, 0.97]

the added informativity of ambiguous language

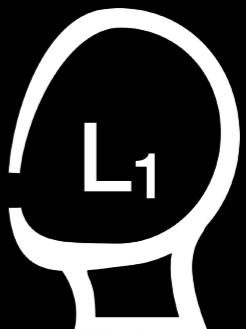
we learn about private knowledge by observing
how our listeners resolve ambiguity

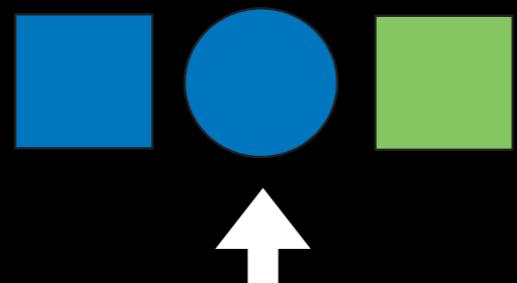
**are speakers able to use this knowledge
to strategically choose their utterances?**

the added informativity of ambiguous language

we learn about private knowledge by observing
how our listeners resolve ambiguity

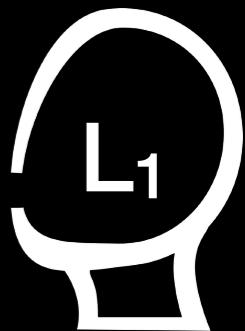
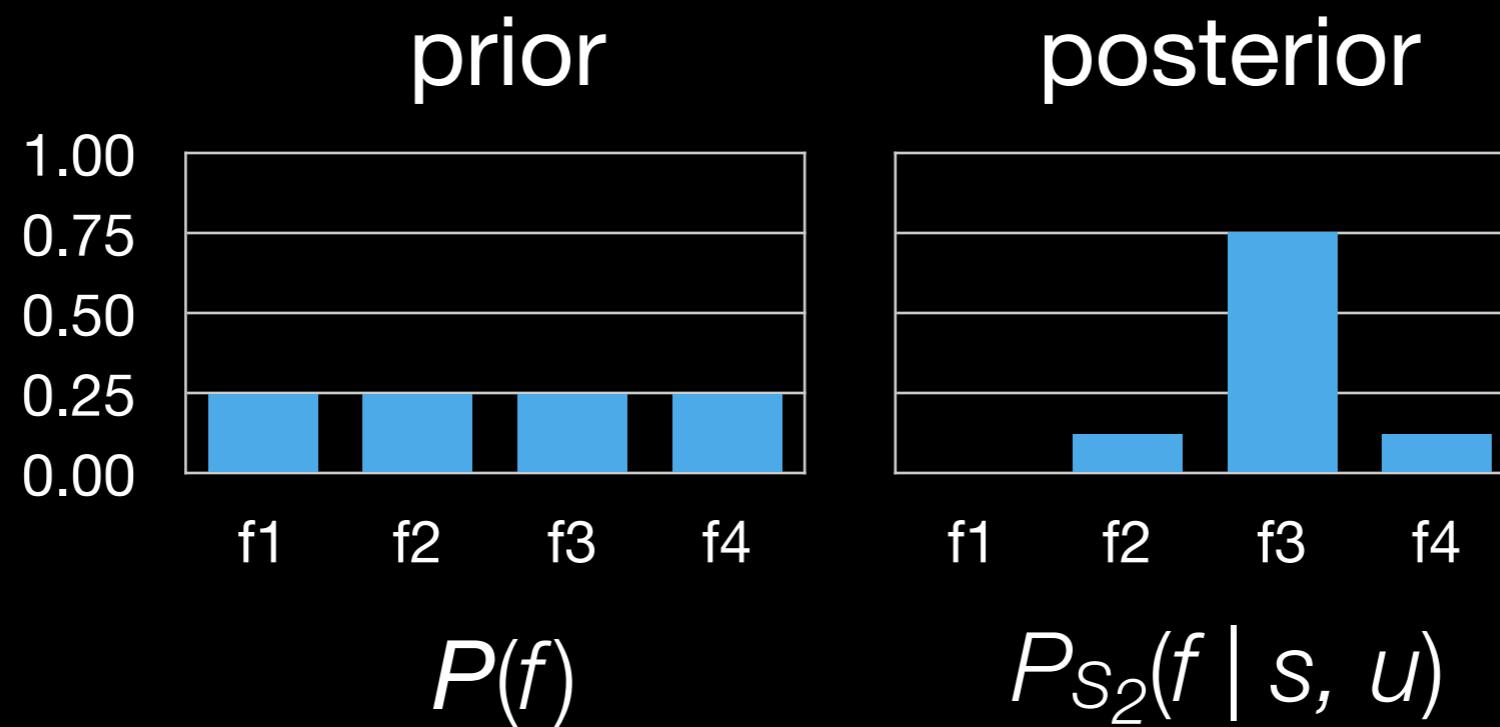
inferring preferences


$$P_{S_2}(f \mid s, u) \propto P_{L_1}(s \mid u, f) \cdot P(f)$$

$$\downarrow \quad \downarrow$$
$$P_{L_1}(s \mid u, \mathbf{f}) \propto P_{S_1}(u \mid s) \cdot P(s \mid \mathbf{f})$$



“blue”

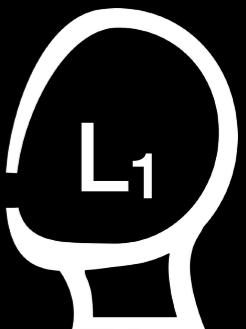
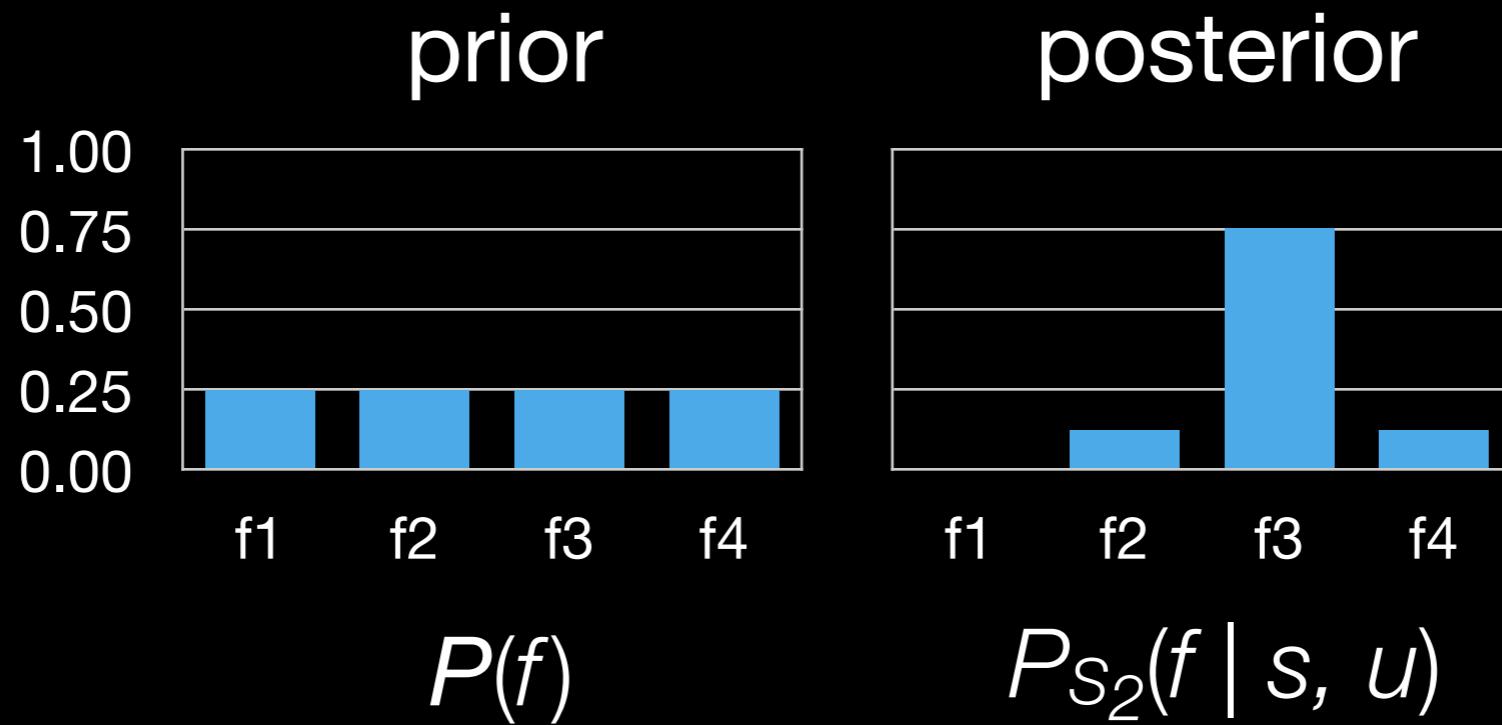
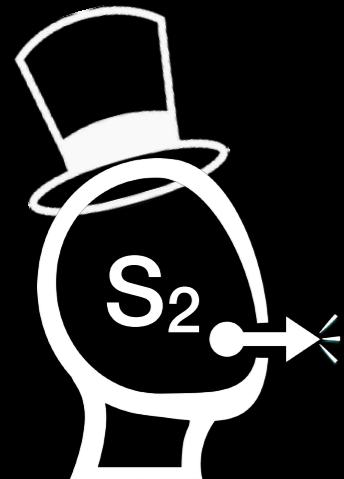
choosing utterances



useful utterances maximize the difference

choosing utterances

$$P_b(u) \propto \sum_{s: \llbracket u \rrbracket(s)} \lambda \cdot \text{KL}(P(f), P_{S_2}(f \mid s, u)) - C(u)$$



useful utterances maximize the difference

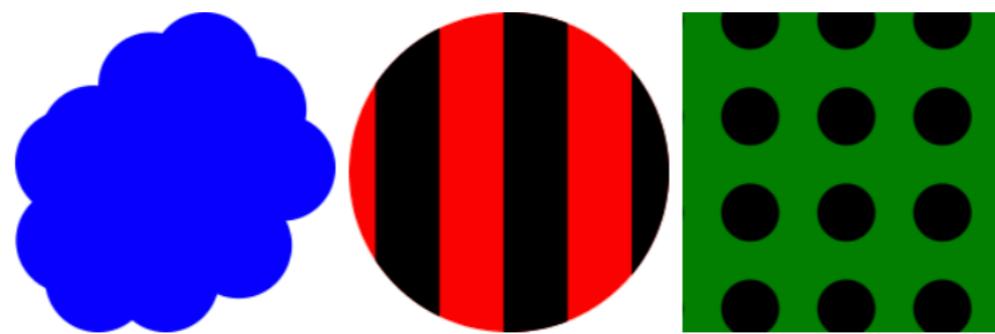
experiment 2

choosing utterances

Progress:

Instructions

You will see a series of scenarios with two people and three objects. For example, a scenario might feature Mary and John and the three objects below:



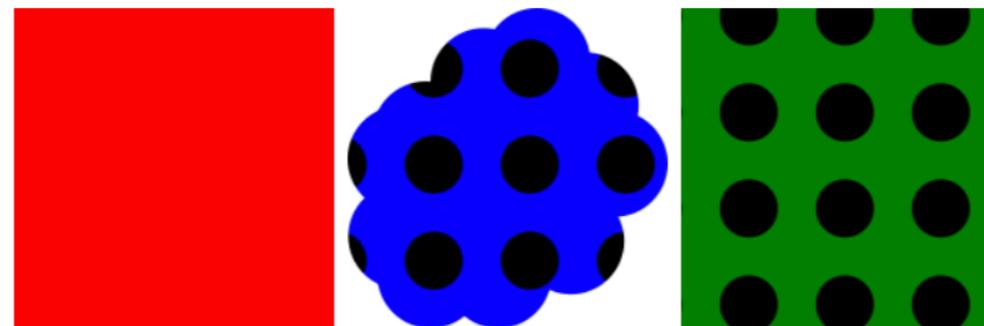
Mary has to choose a single word to communicate one of the objects to John. John is supposed to choose an object that matches the description, and Mary gets to observe his choice. However, John might have preferences for certain colors (blue, red, or green), shapes (clouds, circles, or squares), or patterns (solid, striped, or polka-dotted). These preferences would be reflected in John's object choice.

Your task is to help Mary choose the utterance to signal one of the objects to John. Try to pick an utterance that will reveal John's color, shape, or pattern preferences!

[Continue](#)

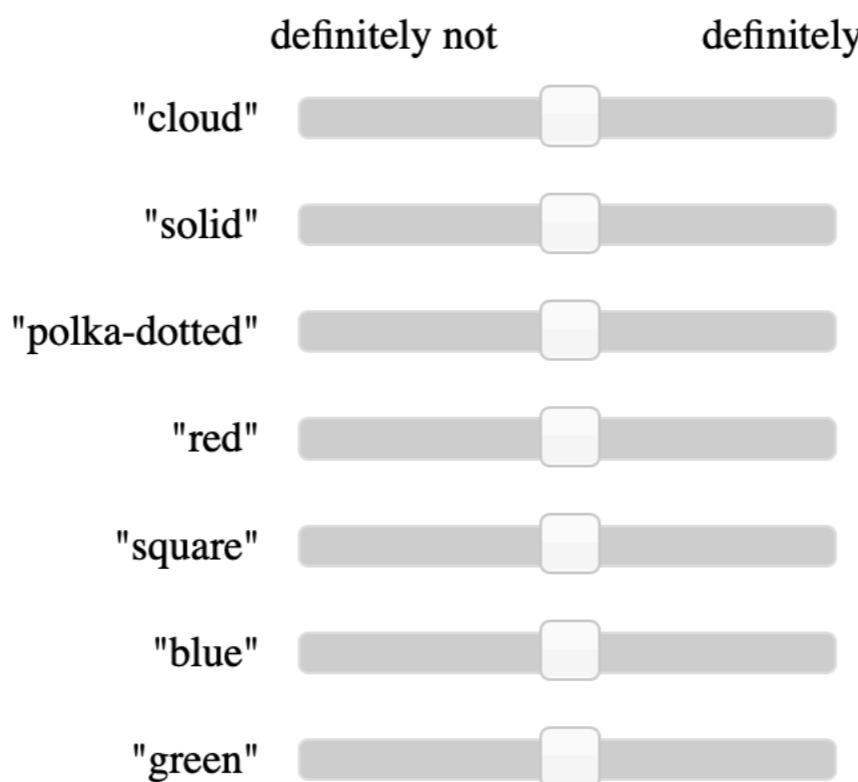
Progress: 

Suppose Devin wants to learn about Mark's preferences in the following scenario:



Devin can choose a single utterance and then watch Mark select an object.

What should Devin say?



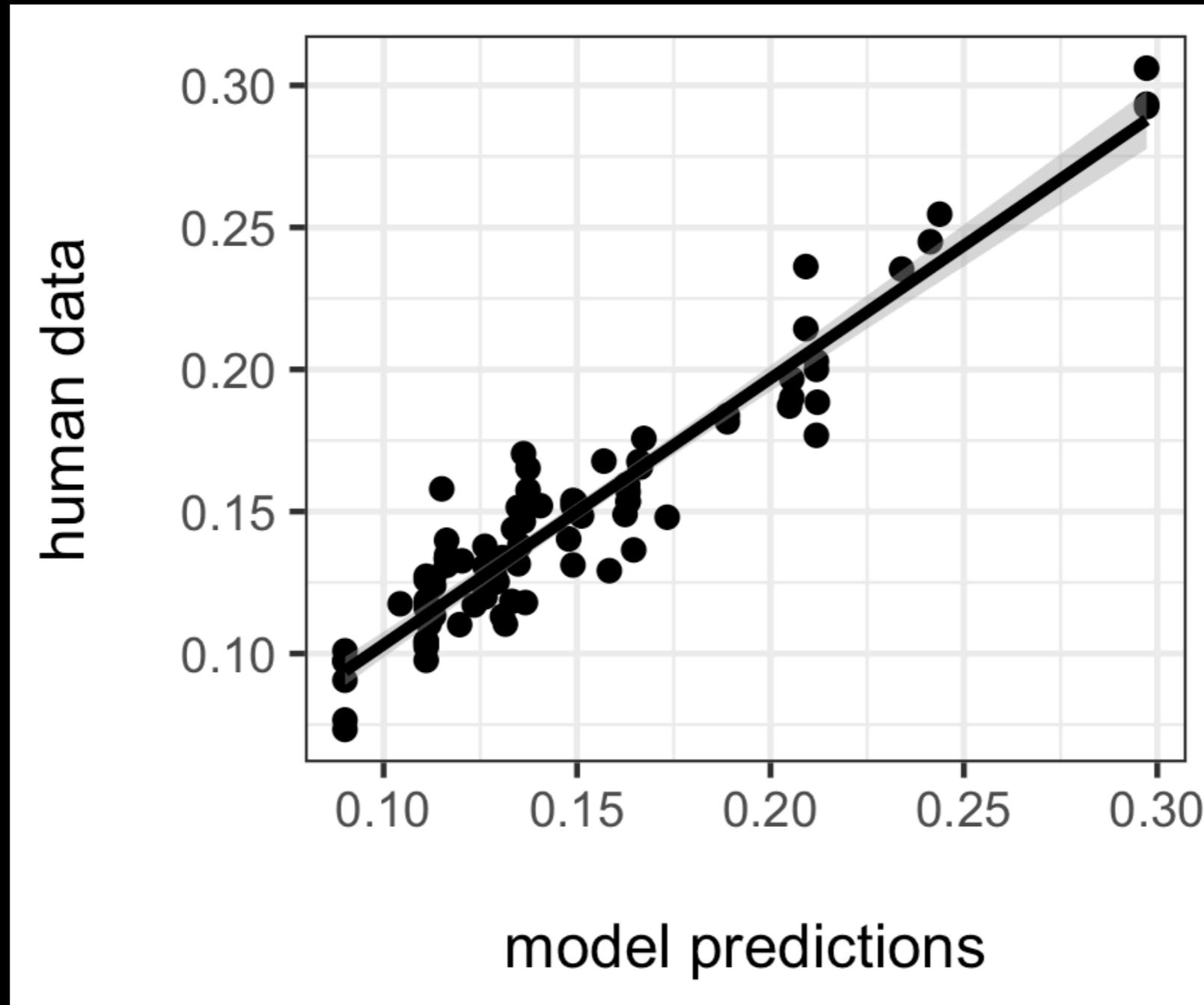
[Continue](#)

15 trials

results

SOMETHING ABOUT MODEL FITTING

results



$r^2 = 0.91$, 95% CI [0.84, 0.95]

discussion

thank you!

