The Rational Speech-Act Model

Leon Bergen

Introduction

Semantics and pragmatics aim to provide theories of how people understand the intended meaning of utterances.

Semantics: a theory of the conventionalized parts of utterance meaning.

Pragmatics: a theory of how people use conventionalized meanings to recover the speaker's intended meaning.

Game-theoretic pragmatics

Game theory provides a natural set of tools for formalizing Gricean reasoning.

Rationality and cooperativeness can be formalized as simple conditions on how the agents make decisions:

- Agents have a utility function which represents their preferences, and a probability distribution which represents their beliefs about the state of the world.

Game-theoretic pragmatics

Game theory provides a natural set of tools for formalizing Gricean reasoning.

Rationality and cooperativeness can be formalized as simple conditions on how the agents make decisions:

- Agents have a utility function which represents their preferences, and a probability distribution which represents their beliefs about the state of the world.
- Rationality is represented by the assumption that agents take actions which maximize their expected utility.

Game-theoretic pragmatics

Game theory provides a natural set of tools for formalizing Gricean reasoning.

Rationality and cooperativeness can be formalized as simple conditions on how the agents make decisions:

- Agents have a utility function which represents their preferences, and a probability distribution which represents their beliefs about the state of the world.
- Rationality is represented by the assumption that agents take actions which maximize their expected utility.
- Cooperativeness is a condition on the speaker's utility function: the speaker receives higher utility when the listener's beliefs match their own.

Game-theoretic pragmatics

Game theory provides a natural set of tools for formalizing Gricean reasoning.

The speaker's and listener's reasoning about each other can be formalized by assuming that each has a model of the other as a rational, cooperative agent:

- The listener interprets an utterance by considering which intended meanings would have made the speaker most likely to choose this utterance.

Game-theoretic pragmatics

Game theory provides a natural set of tools for formalizing Gricean reasoning.

The speaker's and listener's reasoning about each other can be formalized by assuming that each has a model of the other as a rational, cooperative agent:

- The listener interprets an utterance by considering which intended meanings would have made the speaker most likely to choose this utterance.
- The speaker chooses the utterance that is most likely to be interpreted correctly by the listener.

Iterated response models

Obvious problem with this idea: vicious circularity in the definitions of the speaker and listener models. The speaker and listener are being defined in terms of each other.

Two possible solutions:

- Use a game-theoretic equilibrium concept which picks out fixed-points of the speaker-listener recursion.
- Make the setup well-defined by giving the recursion a base-case. Iterated response models (Franke, 2009; Jager, 2013; Frank & Goodman, 2012) fall under this approach.

Rational speech act model (RSA)

L_0 - literal listener

Rational speech act model (RSA)

S_1 - communicates with L_0

$\downarrow$

L_0 - literal listener

Rational speech act model (RSA)

L_1 - reasons about S_1

S_1 - communicates with L_0

L_0 - literal listener

Rational speech act model (RSA)

...

$\downarrow$

S_2

$\downarrow$

L_1 - reasons about S_1

$\downarrow$

S_1 - communicates with L_0

$\downarrow$

L_0 - literal listener

Rational speech acts model (RSA)

Notation:

- u = utterance
- w = possible world
- S(u|w) = probability that the speaker will choose utterance u given the goal of communicating world w.
- L(w|u) = probability that the listener assigns to world w after hearing utterance u.
- $\mathcal{L}$ = lexicon, s.t. $\mathcal{L}(u,w) = \begin{cases} 0 & \text{if } w \notin [\![u]\!] \\ 1 & \text{if } w \in [\![u]\!] \end{cases}$
- P(w) = prior probability of w
- c(u) = cost of utterance u

RSA: Literal listener

Literal listener: $$L_0(w|u,\mathcal{L}) = \frac{\mathcal{L}(u,w)P(w)}{\sum_{w'} \mathcal{L}(u,w')P(w')}$$
$$\propto \mathcal{L}(u,w)P(w)$$

This listener interprets an utterance by filtering out any worlds that are literally incompatible with this utterance.

RSA: Strategic speaker

Speaker model: $S_n(u|w) \propto e^{\lambda U_n(u|w)}$

$\lambda > 0$ is the inverse-temperature, which controls the degree of rationality of the speaker.

The utility function is defined by:
$$\begin{aligned} U_n(u|w) &= -\log\left(\frac{1}{L_{n-1}(w|u)}\right) - c(u) \\ &= \log(L_{n-1}(w|u)) - c(u) \end{aligned}$$

The speaker is more likely to choose utterances that will be interpreted correctly by the listener L_{n-1}.

RSA: Strategic listener

Strategic listener: $L_n(w|u) \propto P(w)S_n(u|w)$

This listener interprets an utterance by integrating their prior knowledge with the likelihood that the speaker S_n would choose the utterance given different states of the world.

Rational speech act model (RSA)

…

↓

S_2

↓

L_1 $$L_n(w|u) \propto P(w) S_n(u|w)$$

↓

S_1 $$U_n(u|w) = -\log\left(\frac{1}{L_{n-1}(w|u)}\right) - c(u)$$

↓

L_0 $$L_0(w|u,\mathcal{L}) \propto \mathcal{L}(u,w) P(w)$$

Scalar implicatures

Derivation of scalar implicatures in this model:

Assumptions:

- 2 meanings: *not all*, *all*
- 2 utterances: “some,” “all”
- Both meanings are equally likely
- Both utterances are equally costly (equivalently, both utterances have cost 0)

Scalar implicatures

Derivation of scalar implicatures in this model:

The listener L_0 interprets "all" as *all*, and "some" as uninformative:

$$L_0(all|\text{"all"}) = 1$$
$$L_0(all|\text{"some"}) = 0.5 = L_0(not\ all|\text{"some"})$$

Scalar implicatures

The speaker S_1 will always choose “some” to communicate *not all*, and will prefer “all” to communicate *all*:

$$S_1(\text{“some”}|not\ all) = 1$$

$$\begin{aligned} S_1(\text{“all”}|all) &= \frac{L_0(all|\text{“all”})}{L_0(all|\text{“all”}) + L_0(all|\text{“some”})} \\ &= \frac{1}{1+\frac{1}{2}} \\ &= \frac{2}{3} \end{aligned}$$

Scalar implicatures

After hearing "some," the listener L_1 infers that the speaker would have been more likely to say this given the goal of communicating *not all*:

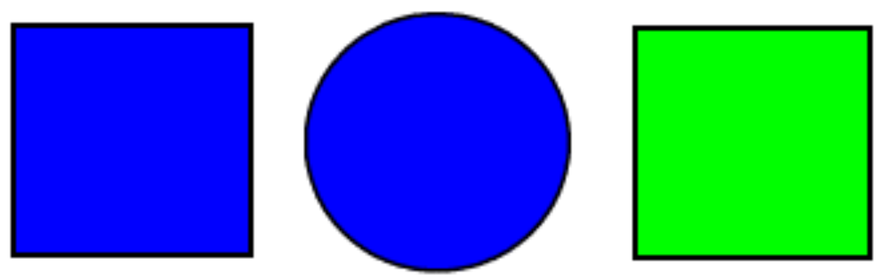
$$
\begin{aligned}
L_1(not\ all|\text{"some"}) &= \frac{P(not\ all)S_1(\text{"some"}|not\ all)}{P(not\ all)S_1(\text{"some"}|not\ all) + P(all)S_1(\text{"some"}|all)} \\
&= \frac{\frac{1}{2}\cdot 1}{\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot\frac{1}{3}} \\
&= \frac{3}{4}
\end{aligned}
$$

Frank & Goodman, 2012

Tested the predictions of the RSA model in a simple domain.

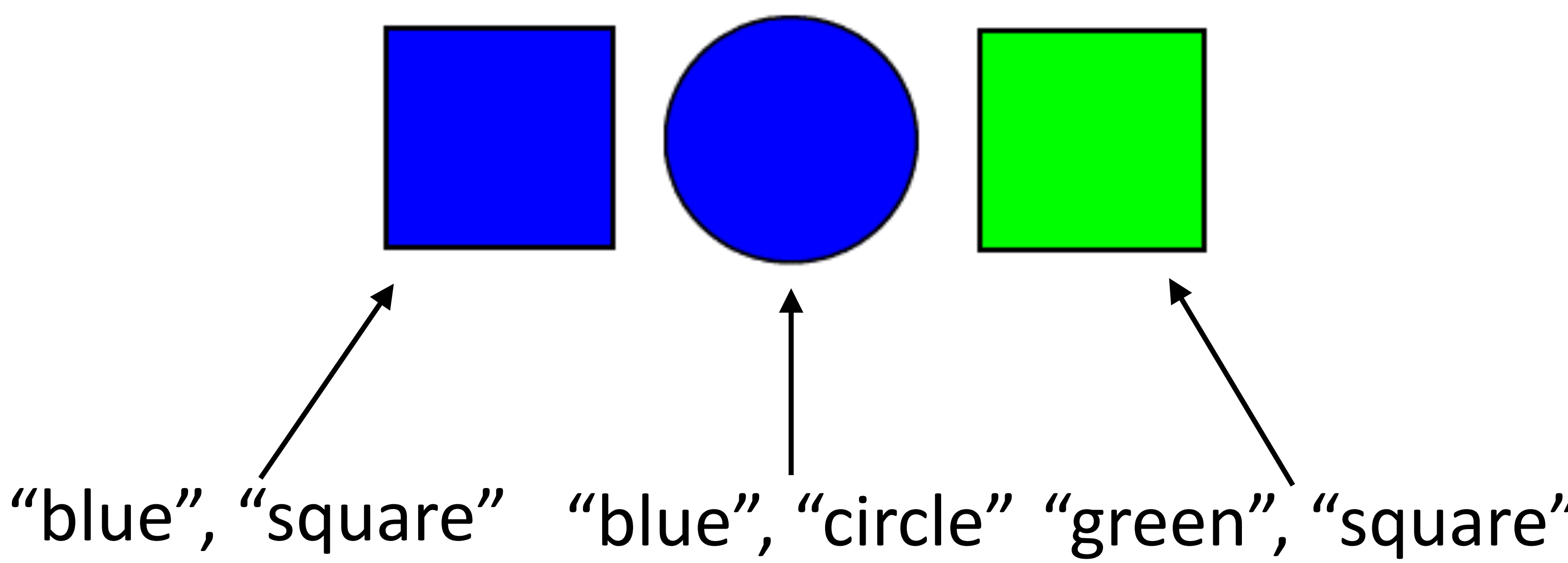
Frank & Goodman, 2012

Objects varied in three dimensions: color, shape, and texture. In every scene, only two of these dimensions were varied.



Frank & Goodman, 2012

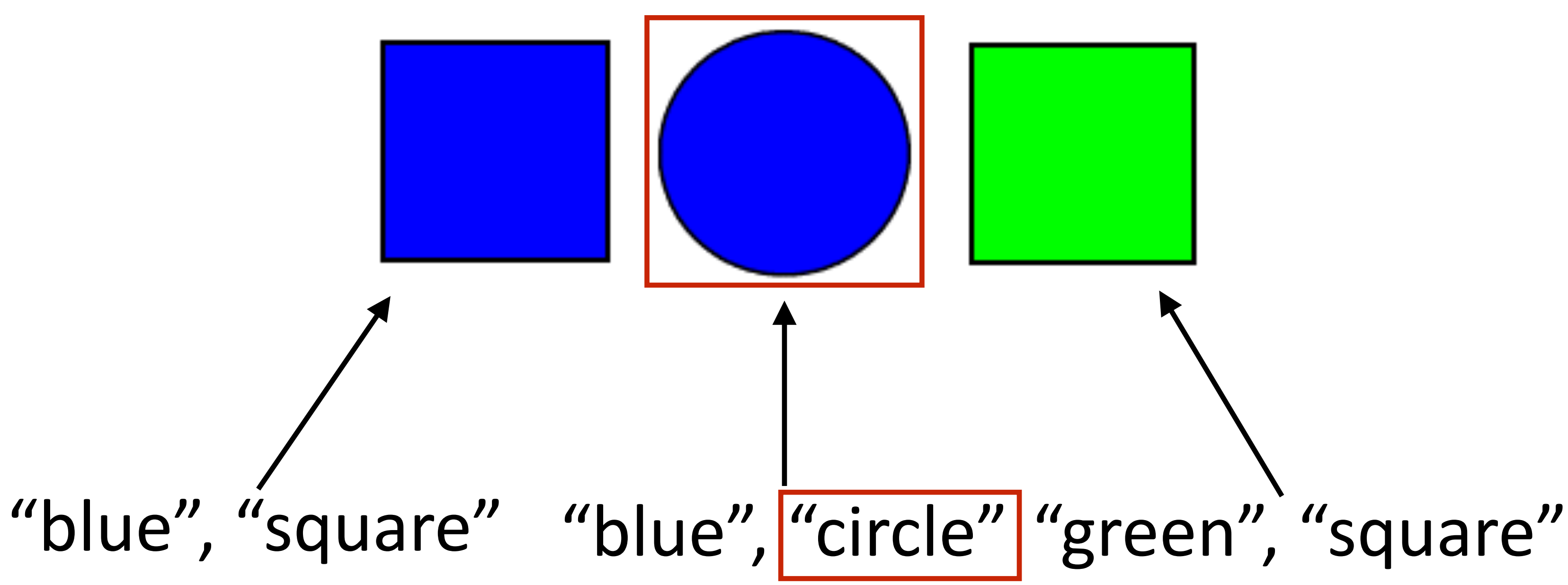
Depending on their features, the objects were either easy or difficult to uniquely describe.



Frank & Goodman, 2012

Depending on their features, the objects were either easy or difficult to uniquely describe.

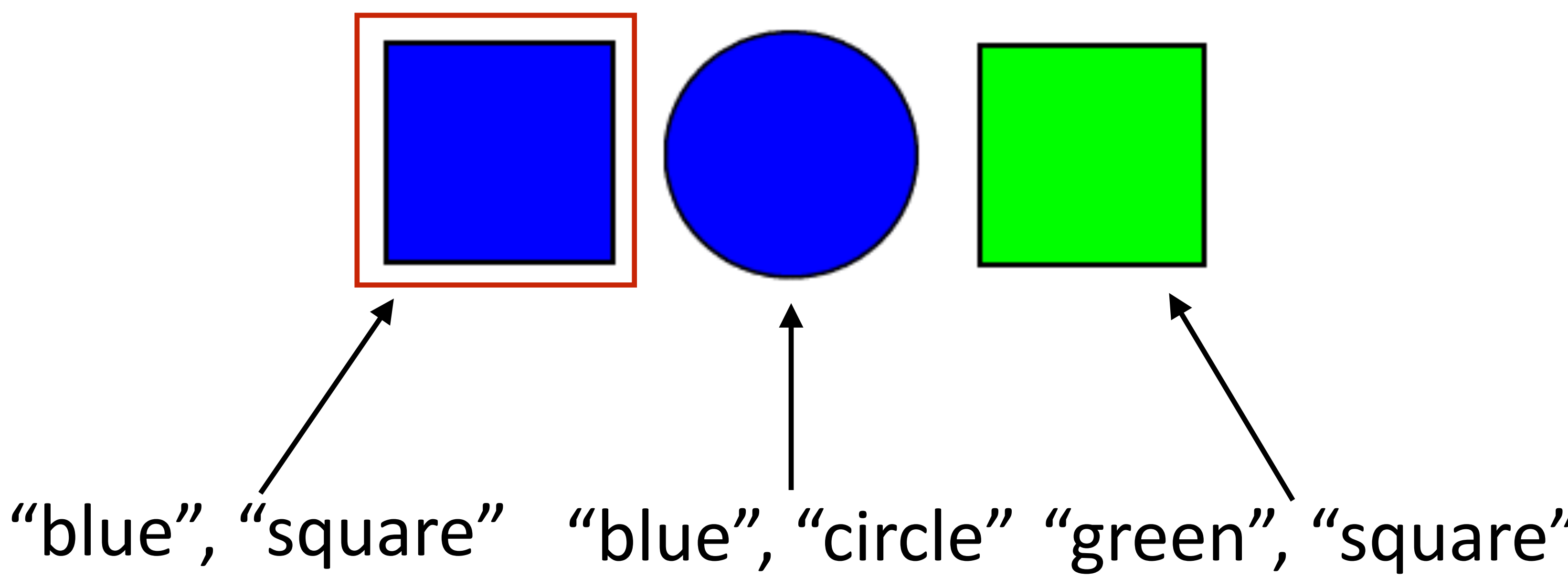
The circle has a simple unique description.



Frank & Goodman, 2012

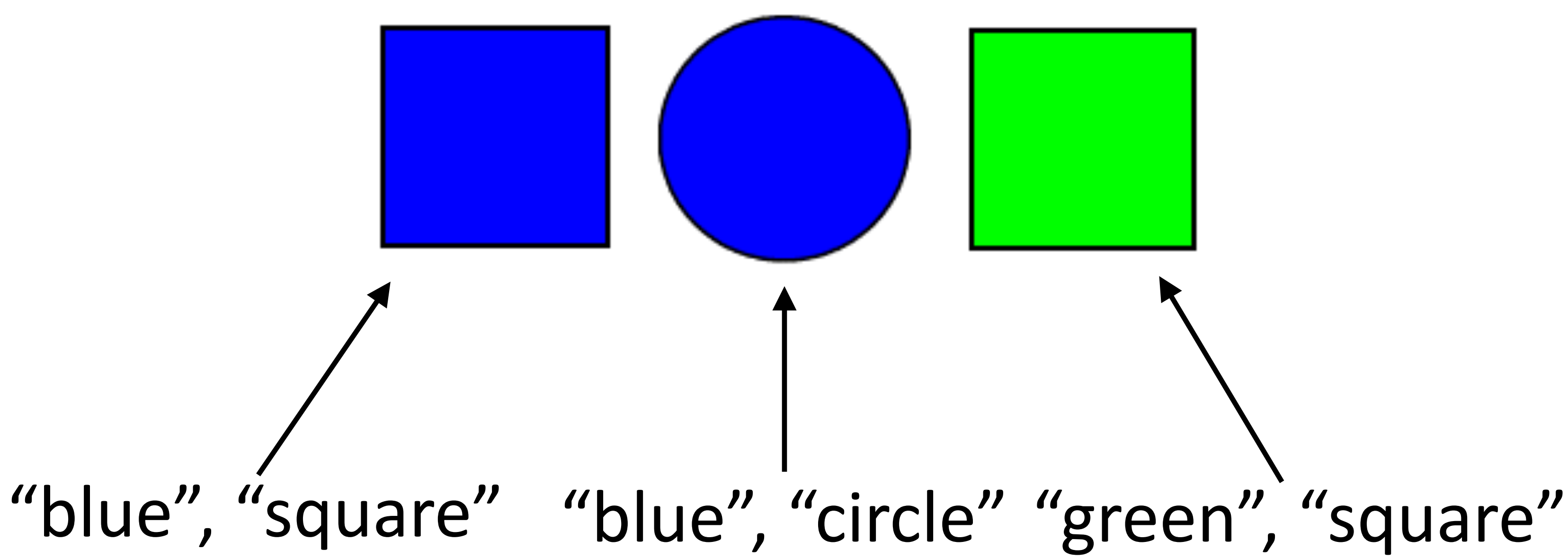
Depending on their features, the objects were either easy or difficult to uniquely describe.

The blue square does not.

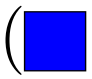


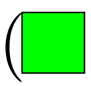
Frank & Goodman, 2012

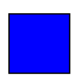
Now suppose someone uses “square” to pick out one of the objects. Which object are they referring to?

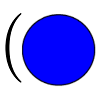


Frank & Goodman, 2012

$L_0($  $|\text{“}square\text{”}) = 0.5$

$L_0($  $|\text{“}square\text{”}) = 0.5$

$L_0($  $|\text{“}blue\text{”}) = 0.5$

$L_0($  $|\text{“}blue\text{”}) = 0.5$

Frank & Goodman, 2012

$$
\begin{aligned}
S_1(\text{“}square\text{”}|{\color{blue}\blacksquare}) &= \frac{L_0({\color{blue}\blacksquare}|\text{“}square\text{”})}{L_0({\color{blue}\blacksquare}|\text{“}square\text{”}) + L_0({\color{blue}\blacksquare}|\text{“}blue\text{”})} \\
&= \frac{0.5}{0.5 + 0.5} \\
&= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
S_1(\text{“}square\text{”}|{\color{green}\blacksquare}) &= \frac{L_0({\color{green}\blacksquare}|\text{“}square\text{”})}{L_0({\color{green}\blacksquare}|\text{“}square\text{”}) + L_0({\color{green}\blacksquare}|\text{“}green\text{”})} \\
&= \frac{0.5}{0.5 + 1} \\
&= \frac{1}{3}
\end{aligned}
$$

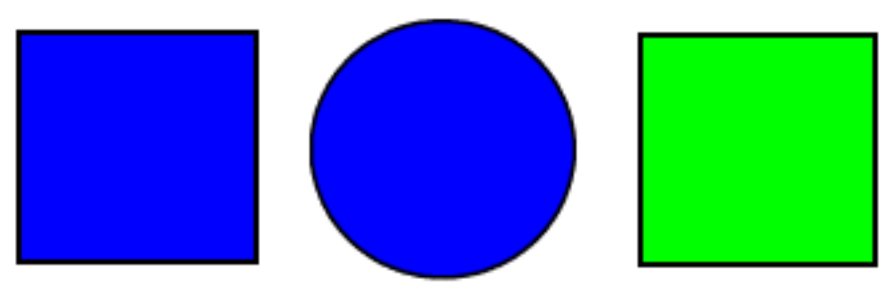
Frank & Goodman, 2012

$$\begin{aligned}
L_1(\textcolor{blue}{\blacksquare} \mid \text{“}square\text{”}) &= \frac{S_1(\text{“}square\text{”} \mid \textcolor{blue}{\blacksquare})}{S_1(\text{“}square\text{”} \mid \textcolor{blue}{\blacksquare}) + S_1(\text{“}square\text{”} \mid \textcolor{green}{\blacksquare})} \\
&= \frac{0.5}{0.5 + \frac{1}{3}} \\
&= \frac{3}{5}
\end{aligned}$$

Frank & Goodman, 2012

Speaker task:

Speaker: Imagine you are talking to someone and you want to refer to the middle object. Which word would you use, “**blue**” or “**circle**”?

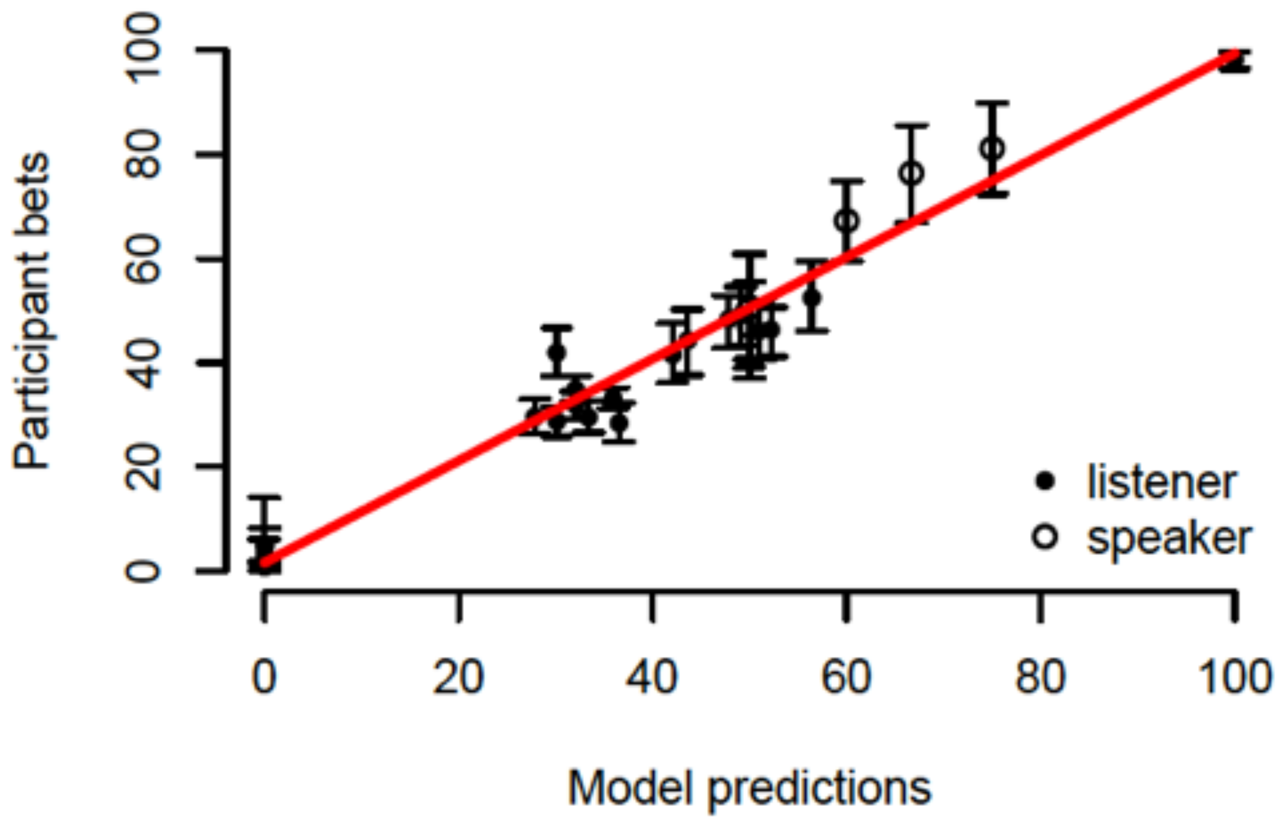


Frank & Goodman, 2012

Listener task (also used to determine prior salience of the objects):

Listener/Salience: Imagine someone is talking to you and uses [the word “**blue**”/a word you don’t know] to refer to one of these objects. Which object are they talking about?

Frank & Goodman, 2012



Hyperbole

Speakers use exaggerated utterances in order to convey information both about the state of the world and about their attitude towards it.

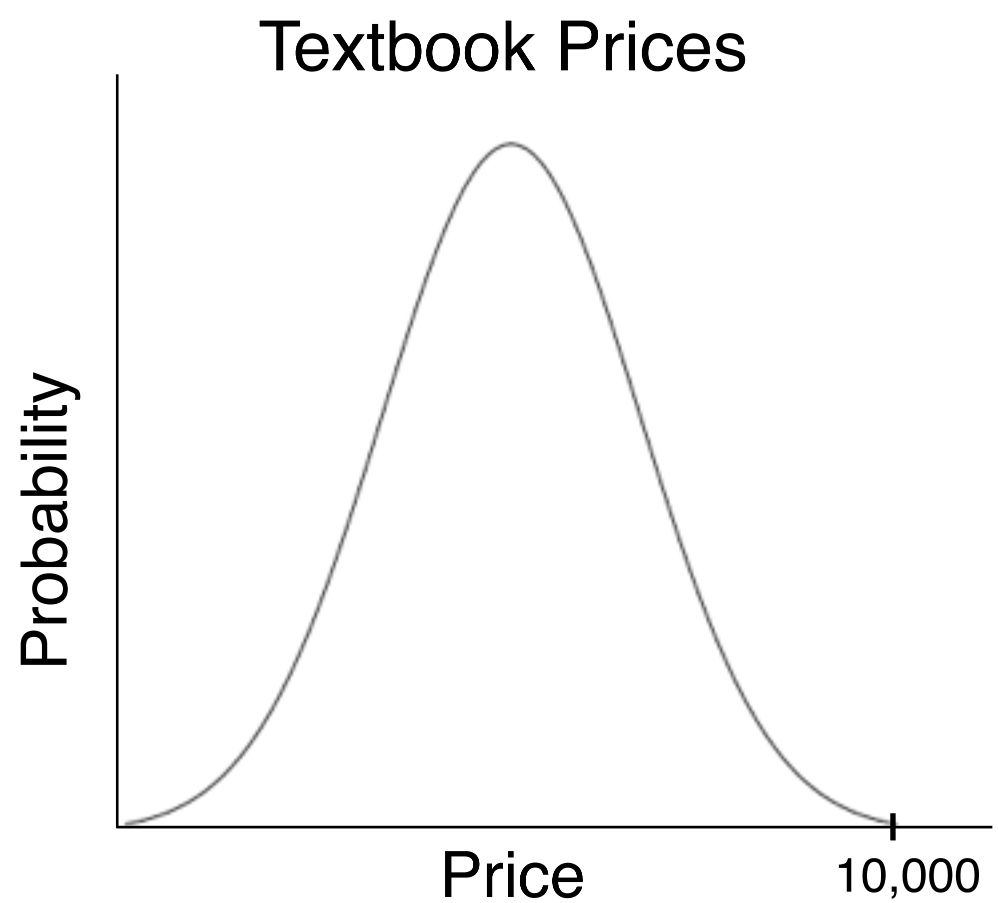
- "That book cost \$10,000"
- Means that the book cost a lot of money.
- Also means that it cost too much (i.e. more than I would have preferred).

(Kao, Wu, Bergen, & Goodman, 2014)

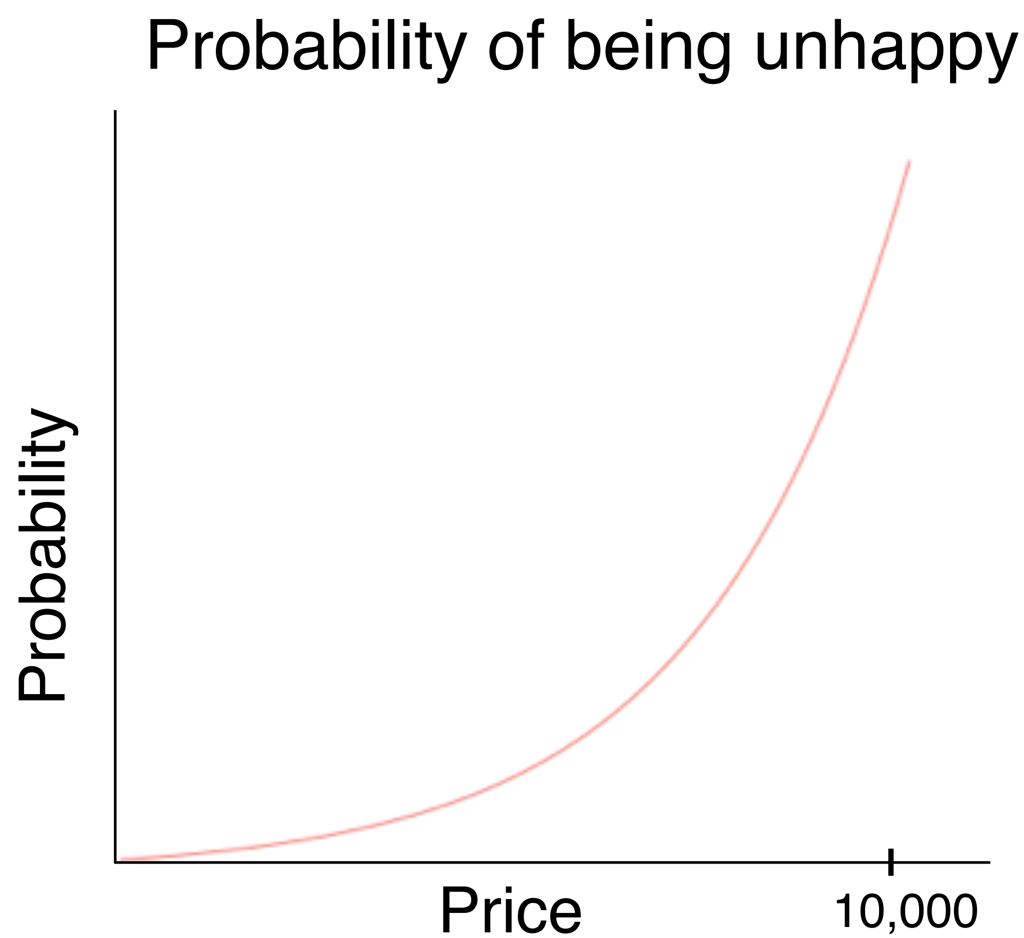
Hyperbole Model: Joint Inference

Idea: Hyperbole depends on the possibility that the speaker does not want to communicate information about the state of the world, but instead wants to communicate only their affect.

Hyperbole Model



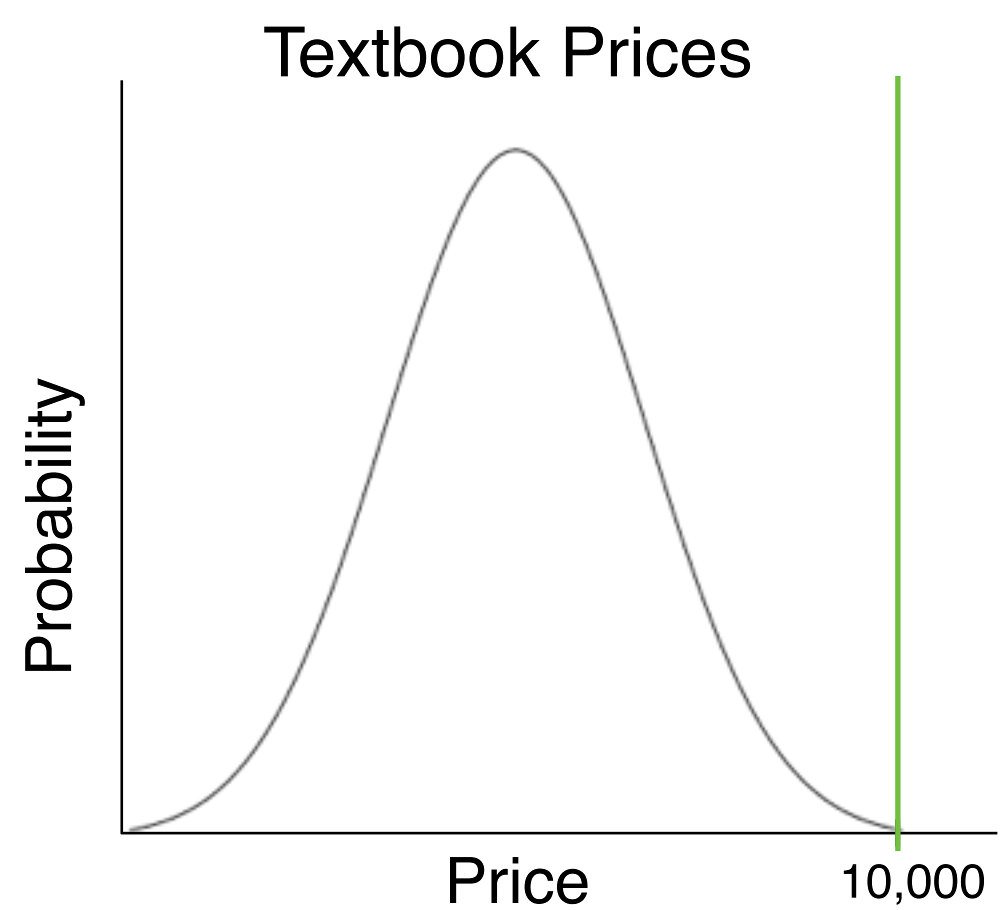
Hyperbole Model



Hyperbole Model

Suppose that the literal listener interprets “The textbook cost $10,000.”

– They infer that the textbook really costs $10,000.

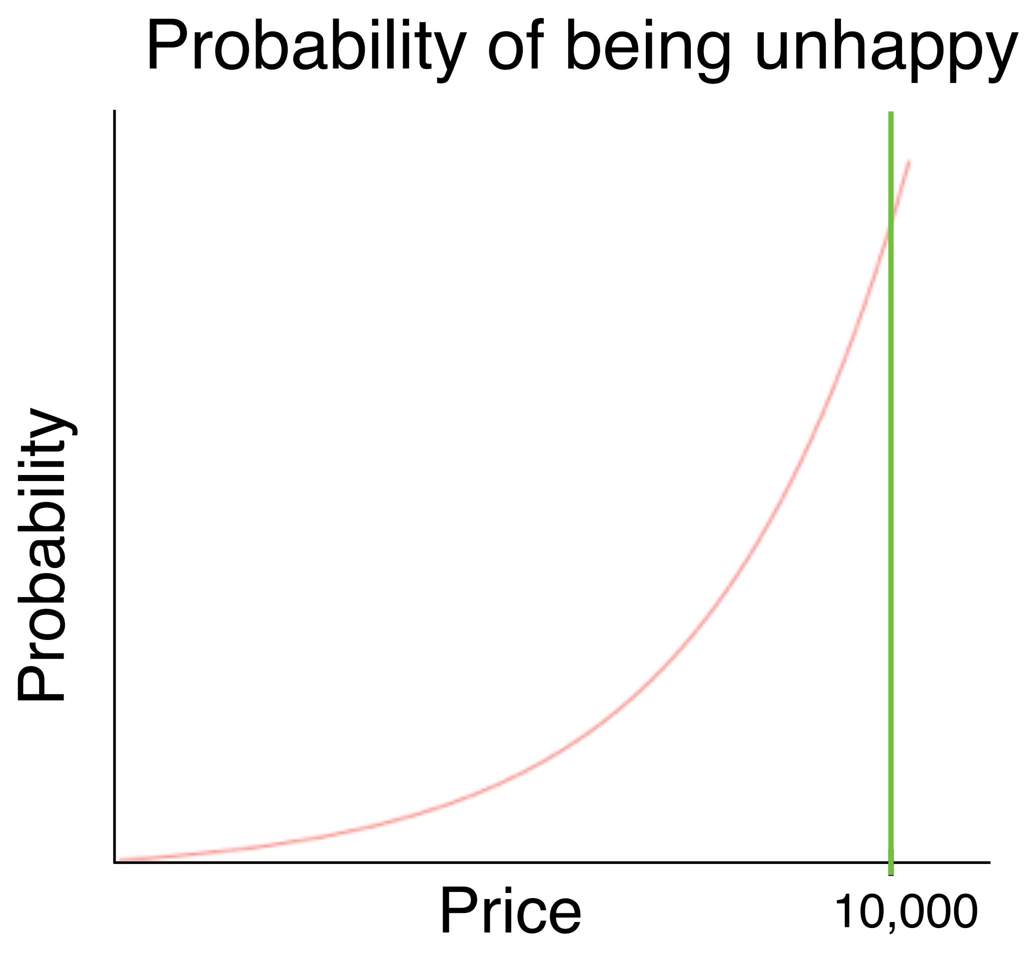


Hyperbole Model

Suppose that the literal listener interprets “The textbook cost $10,000.”

– They infer that the textbook really costs $10,000.

– If a textbook costs $10,000, it was probably too expensive.



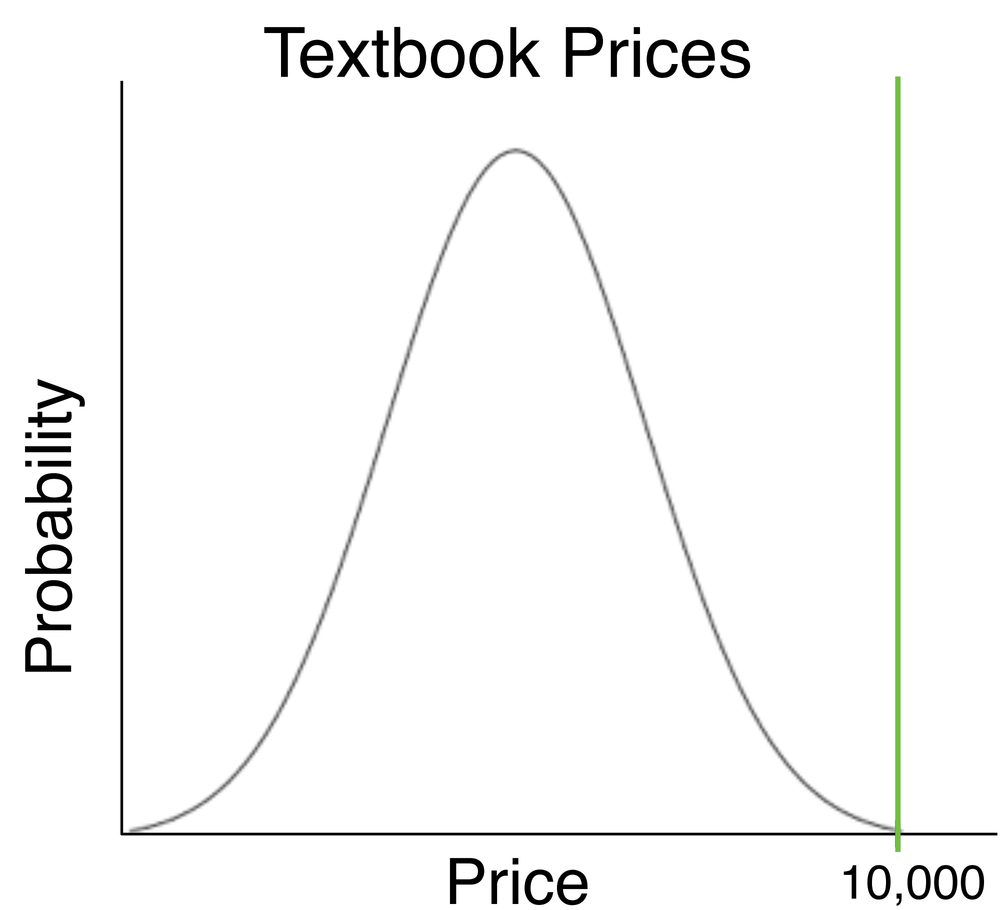
Hyperbole Model

If the speaker's goal is to communicate only that the book was too expensive (and they are unhappy), then saying "The textbook cost $10,000" will do this effectively.

Hyperbole Model

Now suppose that a more sophisticated listener hears "The textbook cost $10,000."

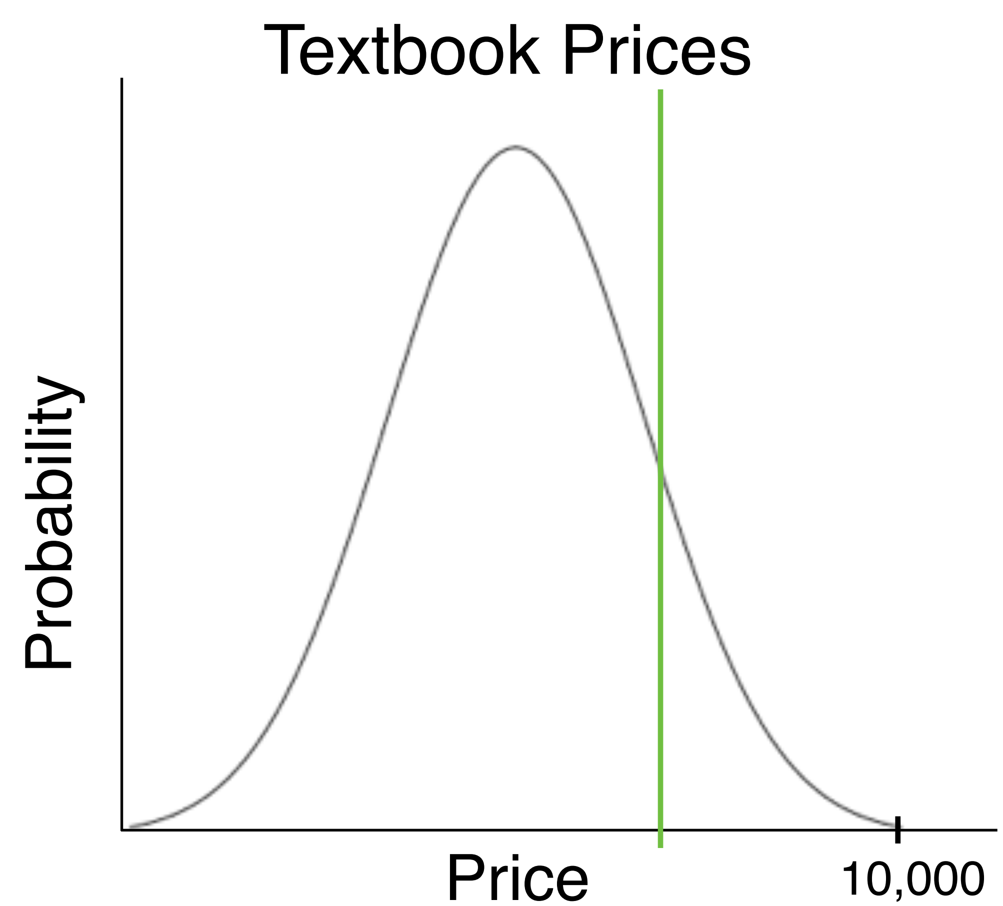
– It is very unlikely that the book actually cost $10,000.



Hyperbole Model

Now suppose that a more sophisticated listener hears "The textbook cost $10,000."

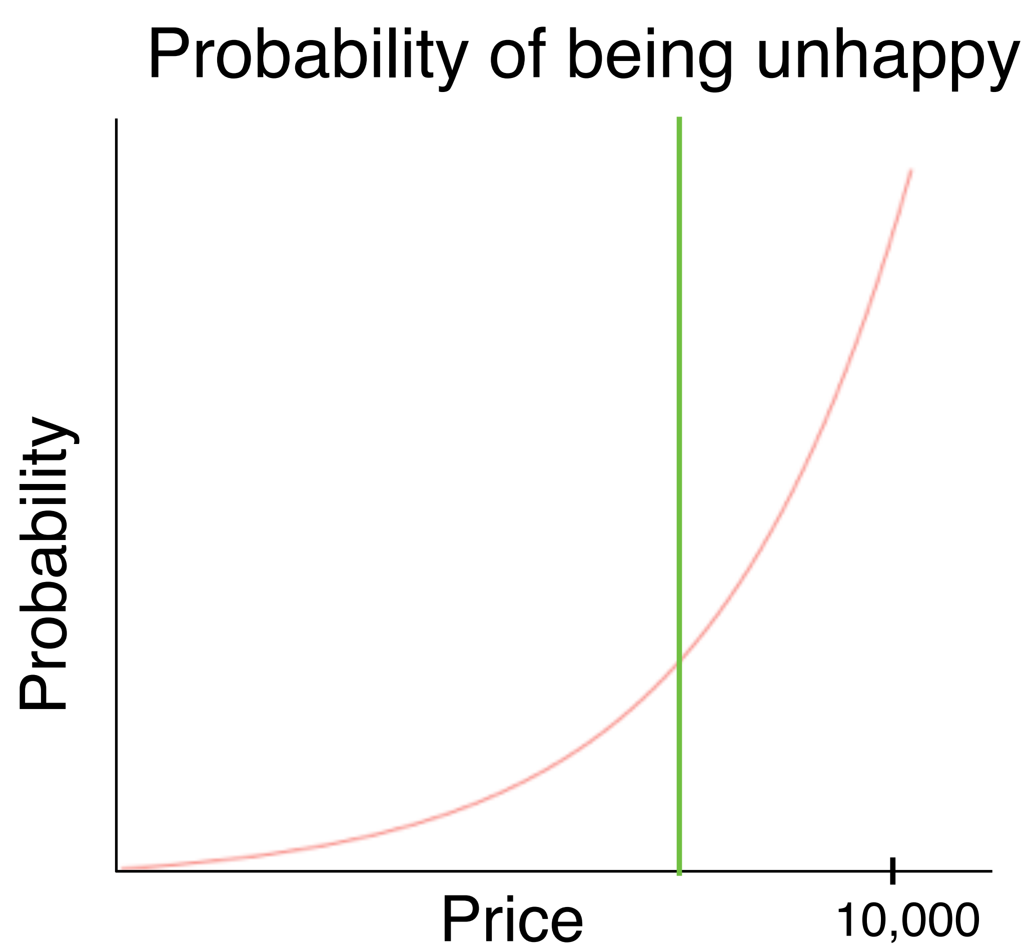
- It is very unlikely that the book actually cost $10,000.
- It is more likely that the textbook was just somewhat expensive, and the speaker wanted to communicate their unhappiness about the price.



Hyperbole Model

Now suppose that a more sophisticated listener hears "The textbook cost $10,000."

- It is very unlikely that the book actually cost $10,000.
- It is more likely that the textbook was just somewhat expensive, and the speaker wanted to communicate their unhappiness about the price.



Hyperbole Model

Speaker model: $S_n(u|s,a,g) \propto e^{U_n(u|s,a,g)}$

- s is the state of the world
- a is the speaker’s affect
- g is the speaker’s goal

Hyperbole Model

The speaker's goal g picks out a particular dimension/set of dimensions. The speaker only cares about the listener's distribution on these dimensions.

Hyperbole Model

The speaker's goal g picks out a particular dimension/set of dimensions. The speaker only cares about the listener's distribution on these dimensions:

$$U_n(u|s,a,g) = \log L_n(g(s,a)|u) - C(u)$$

$$L_n(x|u) = \sum_{s',a'} \delta_{x=g(s',a')} L_n(s',a'|u)$$

This is the probability that the listener assigns to the speaker's goal state.

Hyperbole Model

The speaker's goals vary along two dimensions: what they want to communicate about (state, affect, or both), and how precisely they want to communicate.

Hyperbole Model

The speaker's goals vary along two dimensions: what they want to communicate about (state, affect, or both), and how precisely they want to communicate.

These goals are constructed using two types of projections: one which determines the dimension(s) of interest, and one which determines the desired level of precision:

Dimension(s) of interest:

$$r_s(s,a) = s$$
$$r_a(s,a) = a$$
$$r_{s,a}(s,a) = s, a.$$

Precision:

$$f_e(s) = s$$
$$f_a(s) = \mathrm{Round}(s),$$

Hyperbole Model

The speaker's goals vary along two dimensions: what they want to communicate about (state, affect, or both), and how precisely they want to communicate.

These goals are constructed using two types of projections: one which determines the dimension(s) of interest, and one which determines the desired level of precision:

Dimension(s) of interest:
$$r_s(s,a) = s$$
$$r_a(s,a) = a$$
$$r_{s,a}(s,a) = s, a.$$

Precision:
$$f_e(s) = s$$
$$f_a(s) = \mathrm{Round}(s),$$

Goals can then be defined by composition: $g(s,a) = r(f(s),a)$

Hyperbole Model

The listener L_0 interprets numerical utterances according to their literal meanings:

$$L_0(s, a|u) = \begin{cases} P_A(a|s) & \text{if } s = u \\ 0 & \text{otherwise} \end{cases}$$

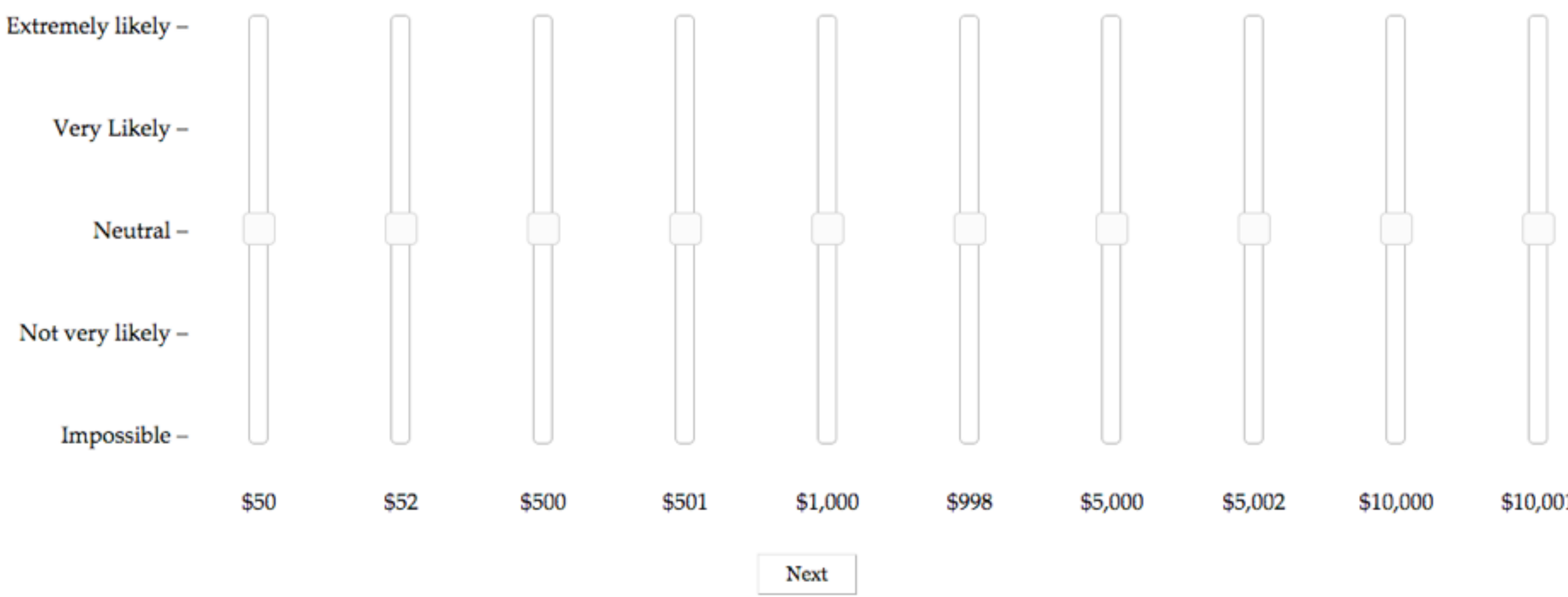
More sophisticated listeners sum over the speaker's possible goals:

$$L_n(s, a|u) \propto \sum_g P_S(s) P_A(a|s) P_G(g) S_{n-1}(u|s, a, g)$$

Experiment: Prior elicitation

Bob bought a new ***electric kettle***.

Please rate how likely it is that the electric kettle cost the following amounts of money.



Experiment: Prior elicitation

Eric bought a new ***electric kettle***.

It cost 10,000 dollars.

How likely is it that Eric thinks the electric kettle was expensive?

Impossible | Not very likely | Neutral | Very Likely | Absolutely certain

Next

Experiment: Exaggeration

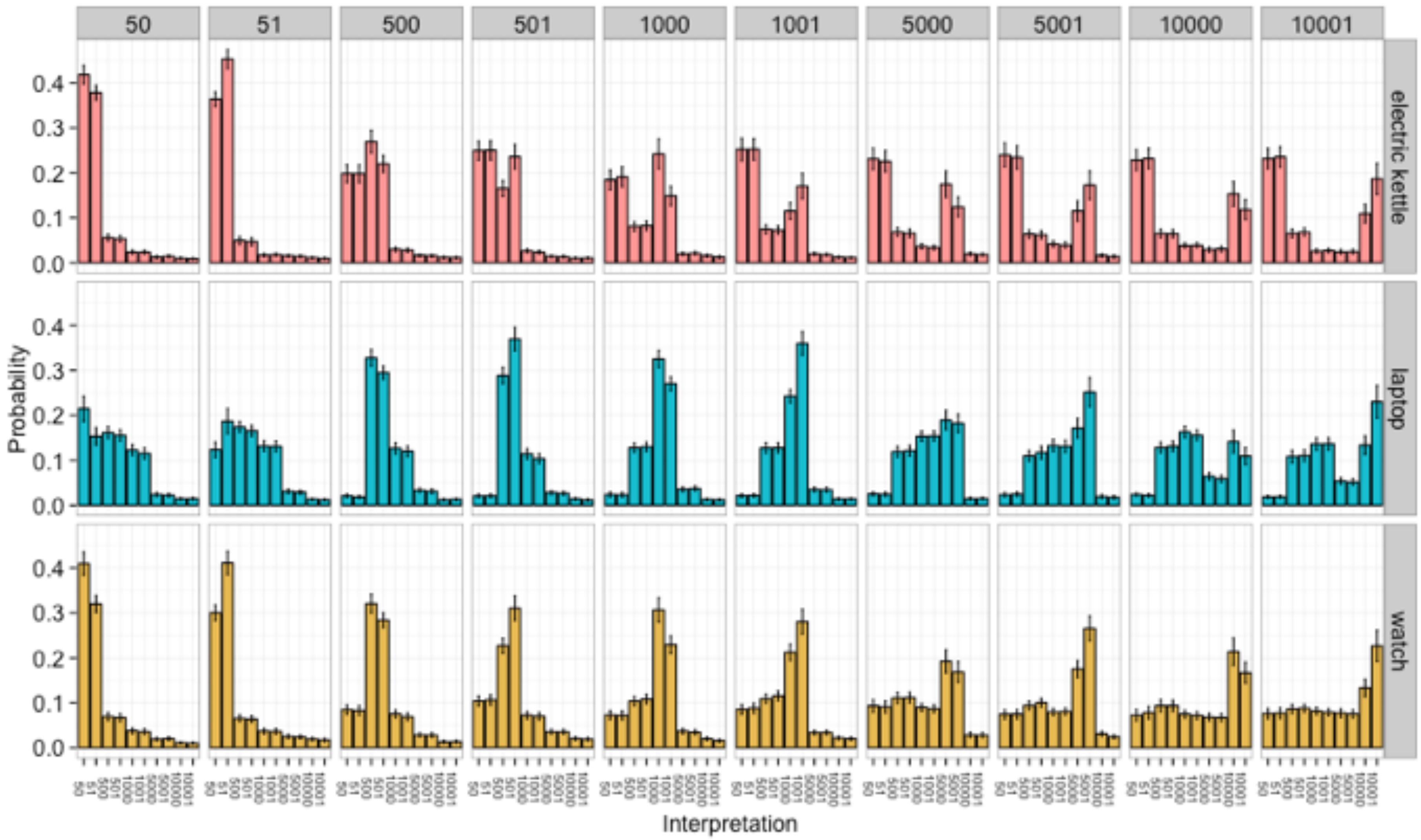
Eric bought a new ***electric kettle***.

A friend asked him, "Was it expensive?"

Eric said, "It cost 10,000 dollars."

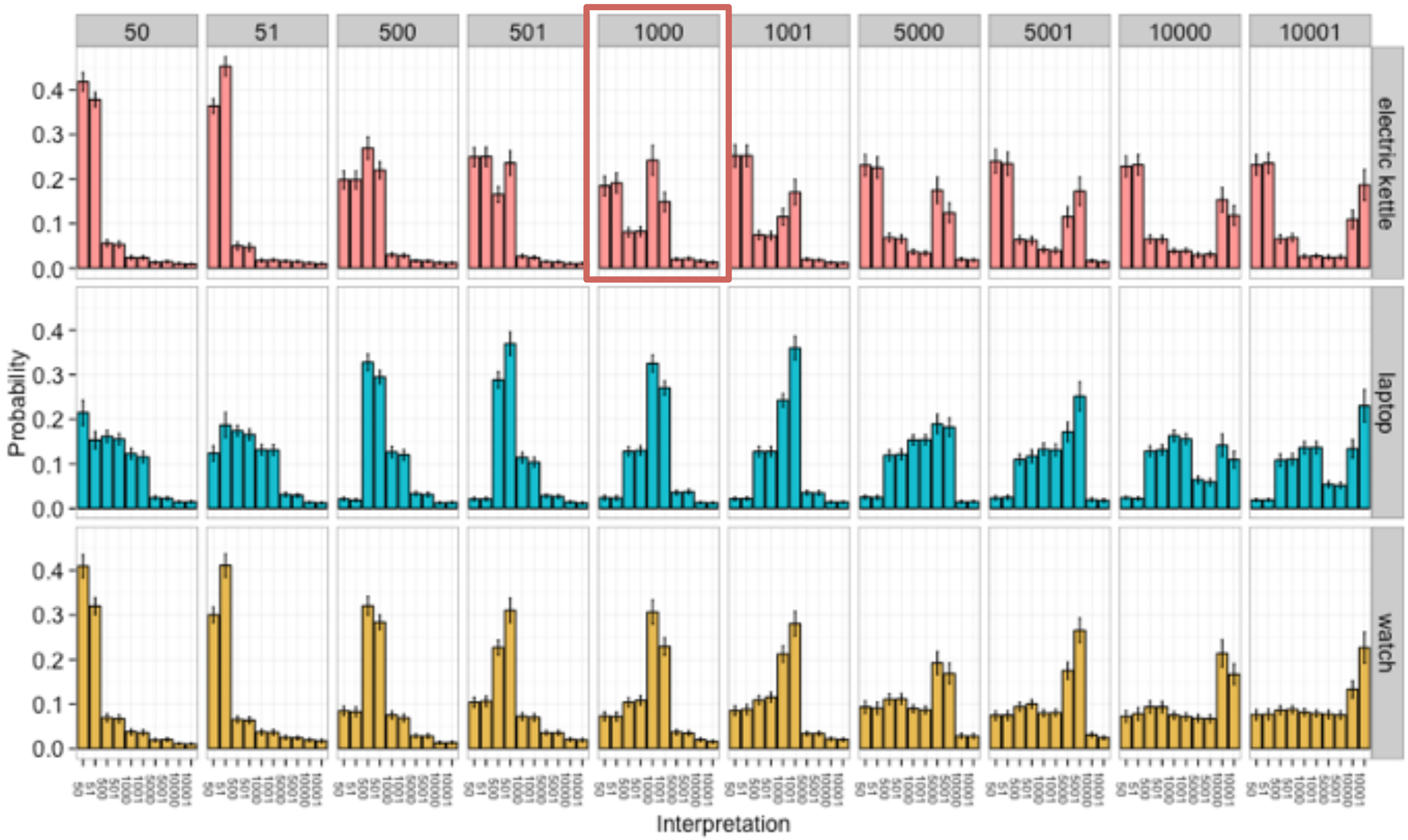
Please rate how likely it is that the electric kettle cost the following amounts of money.

Results



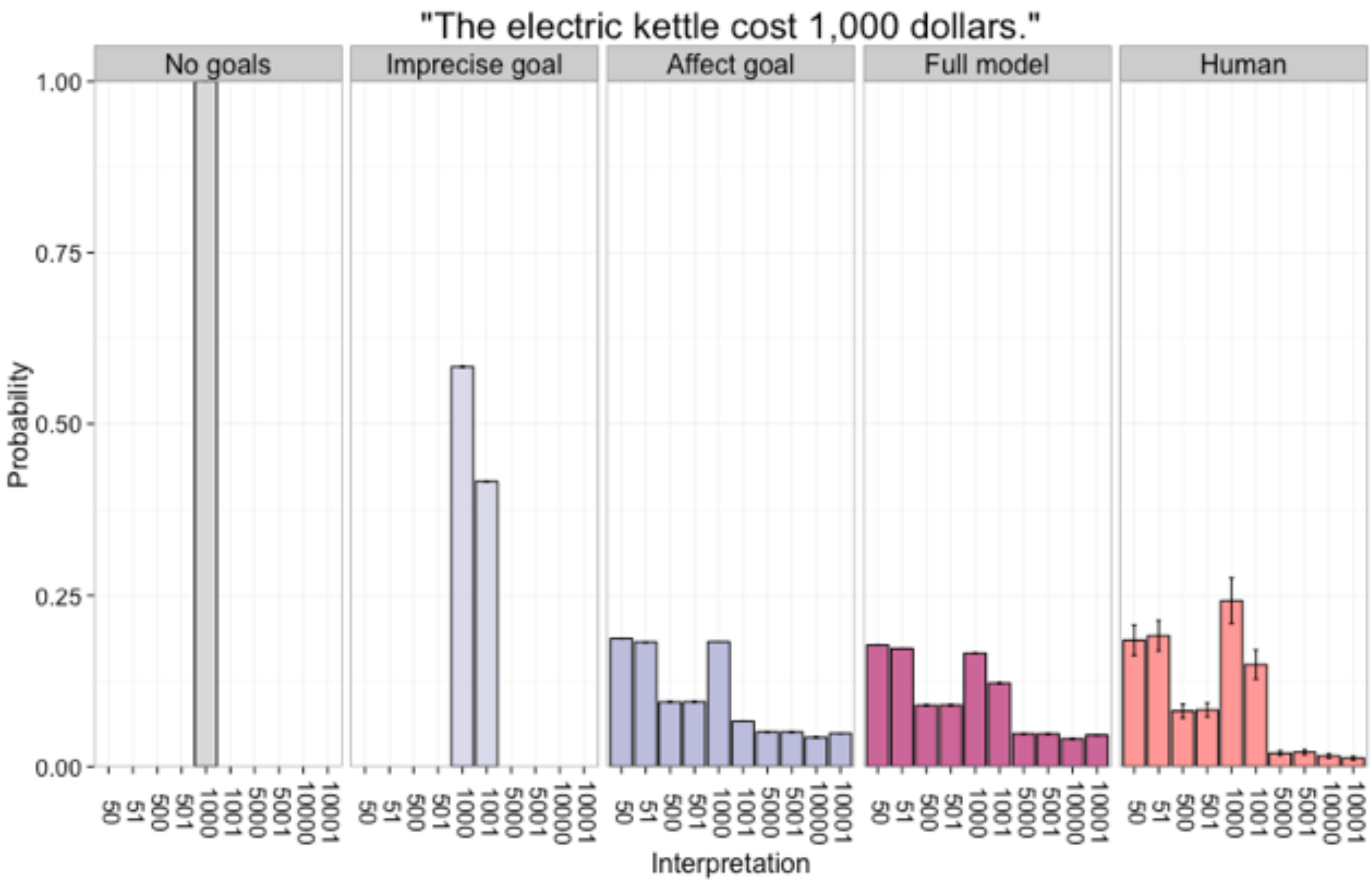
50
51
500
501
1000
1001
5000
5001
10000
10001
electric kettle
laptop
watch
Probability
0.0
0.1
0.2
0.3
0.4
Interpretation

Results



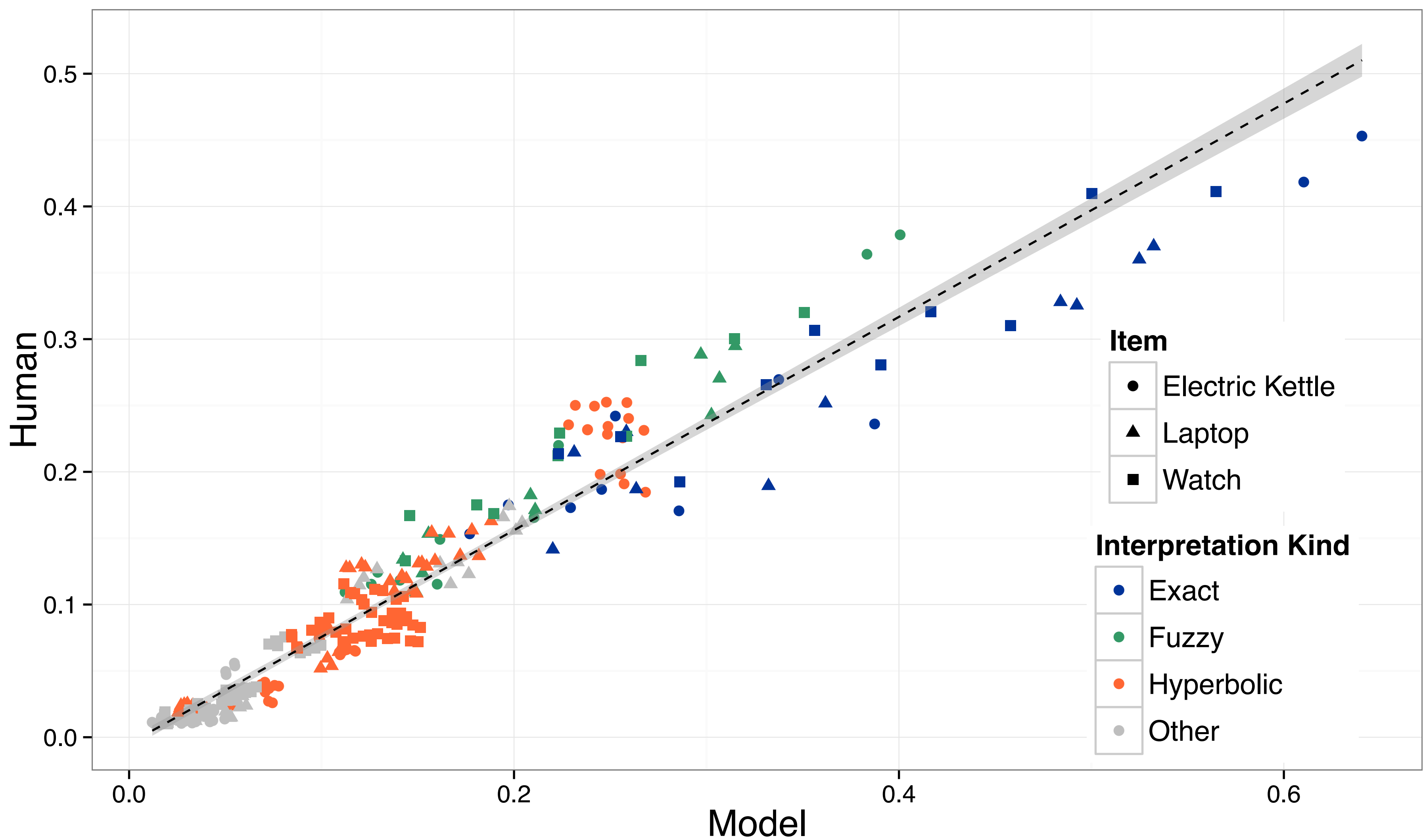
50
51
500
501
1000
1001
5000
5001
10000
10001
electric kettle
laptop
watch
Probability
0.0
0.1
0.2
0.3
0.4
Interpretation

Results



"The electric kettle cost 1,000 dollars."
No goals
Imprecise goal
Affect goal
Full model
Human
Probability
1.00
0.75
0.50
0.25
0.00
50
51
500
501
1000
1001
5000
5001
10000
10001
Interpretation

Results



0.5
0.4
0.3
0.2
0.1
0.0
Human
0.0
0.2
0.4
0.6
Model
Item
Electric Kettle
Laptop
Watch
Interpretation Kind
Exact
Fuzzy
Hyperbolic
Other

Experiment: Affect

David had to buy a new ***electric kettle*** that cost him 501 dollars.

A friend asked him, "Was it expensive?"

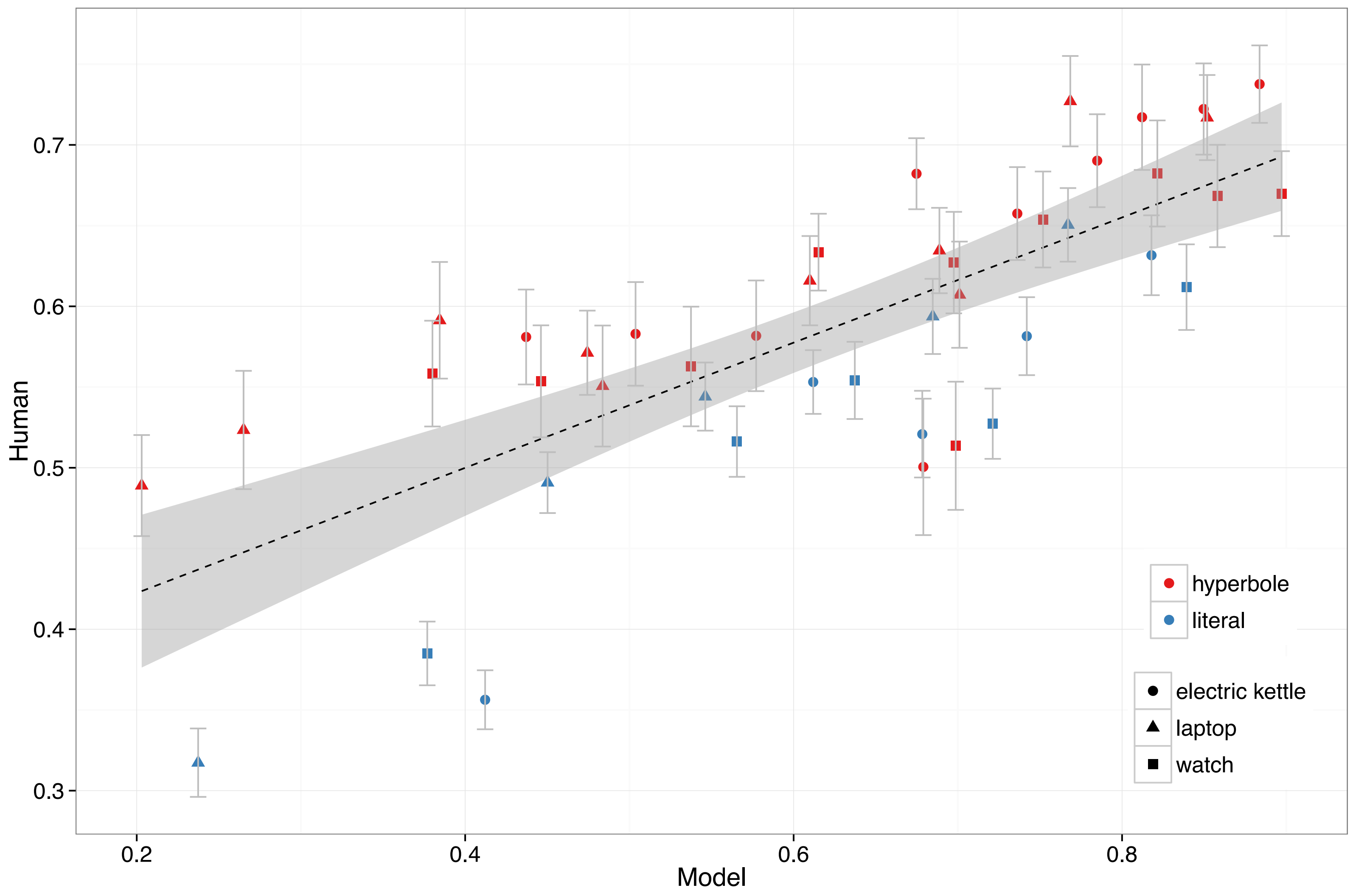
David said, "It cost 10,000 dollars."

How likely is it that David thinks the electric kettle was expensive?

Impossible | Not very likely | Neutral | Very Likely | Absolutely certain

Next

Results



Human
0.7
0.6
0.5
0.4
0.3
0.2
0.4
0.6
0.8
Model
hyperbole
literal
electric kettle
laptop
watch