EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Arguing with quantifiers

Fausto Carcassi, Chris Cummins, Michael Franke

@CarcassiFausto

fausto.carcassi@gmail.com

Helen of Troy

Dante Gabriel Rossetti (1828–1882)

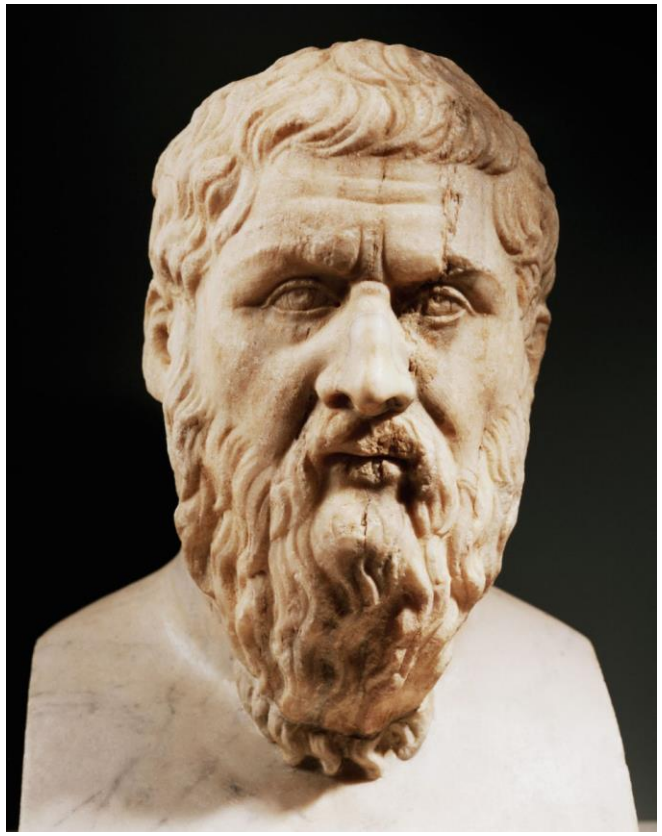
Some say a host of horsemen, others of infantry and others of ships, is the most beautiful thing on the dark earth but I say, it is what you love

Full easy it is to make this understood of one and all: for she that far surpassed all mortals in beauty, Helen her most noble husband

Deserted, and went sailing to Troy, with never a thought for her daughter and dear parents.

— Sappho, fragment 16

Gorgia's *Encomium of Helen*



Gorgias (483–375 BC)

- by the gods
- by physical force
- by love
- by speech

"Speech is a powerful master and achieves the most divine feats with the smallest and least evident body. It can stop fear, relieve pain, create joy, and increase pity"

(Gorgias 31)

Argument and quantities

- They are offering at most 50% off last season’s stock.

- They are offering up to 50% off last season’s stock.

Background story

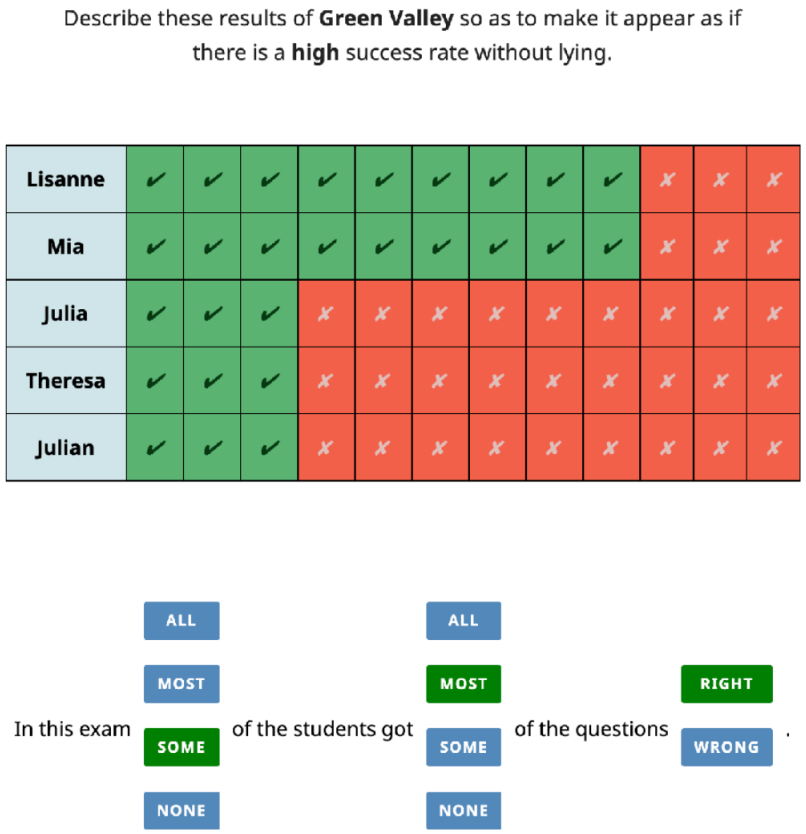
Imagine you have been hired as a marketing consultant for Green Valley High School. Part of your job is to write a report on the results of standardized math exam questions. These results have been published for Green Valley and for your main rival, Riverside High School.

It's important that you don't tell any lies in the report, but you don't have to report objectively on the facts. **Your aim is to make Green Valley sound like a school whose students have a high probability of success on the exam questions, and Riverside sound like a school whose students have a low probability of success.**

Arguing with quantifiers

Experimental design

- N=60 (via Prolific)
- Stimulus: answers array
  - 0, 3, 9, and 12 right answers possible
  - In decreasing order in array
- Task: Complete sentence frame
  - ‘Q1 of the students got Q2 of the questions ADJ’
  - Q: {none, some, most, all}
  - ADJ: {right, wrong}
  - 32 possible utterances
- Two conditions: high vs low framing
  - Within-subjects
- Each participant saw all 20 situations
  - 10 in high cond, 10 in low cond



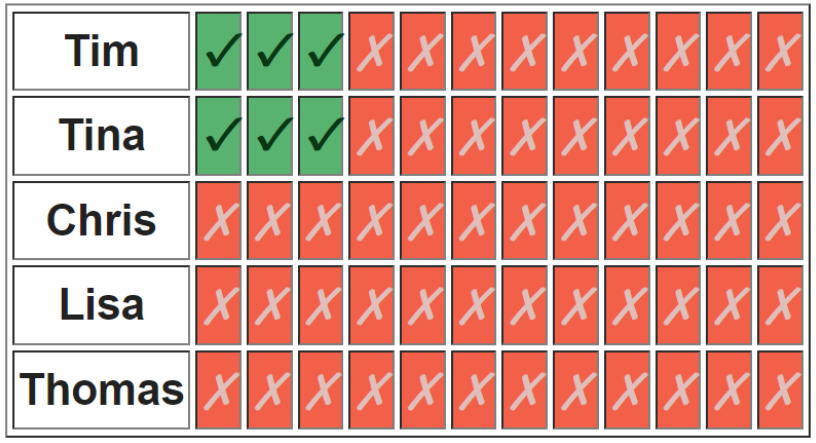
Raw data

Positive condition

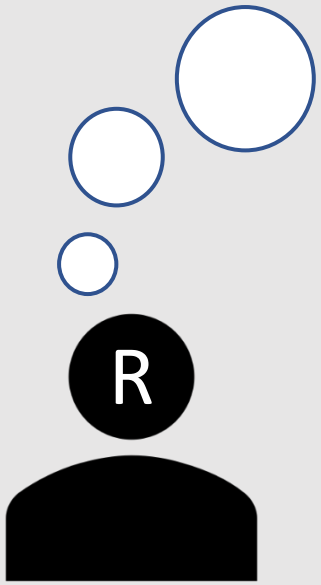
response	obs
some|some|right	3|3|0|0|0
all|some|right	3|3|3|3|3
most|some|right	9|9|3|0|0
some|all|right	12|12|0|0|0
some|all|right	12|12|9|0|0
most|all|right	12|12|12|0|0
most|most|right	12|12|9|3|3
most|all|right	12|12|12|3|3
most|all|right	12|12|12|9|9
all|all|right	12|12|12|12|12

Raw data

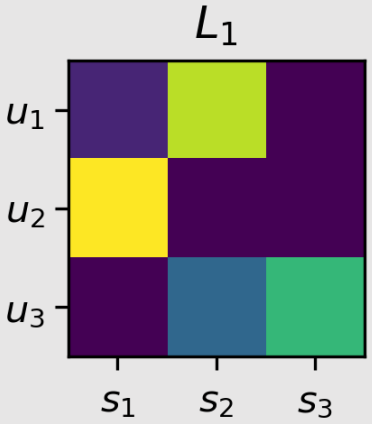
- High condition:
  - 22 ‘some|all|right’
  - 2 ‘most|some|right’
- Low condition:
  - 16 ‘most|most|wrong’
  - 7 ‘some|most|wrong’

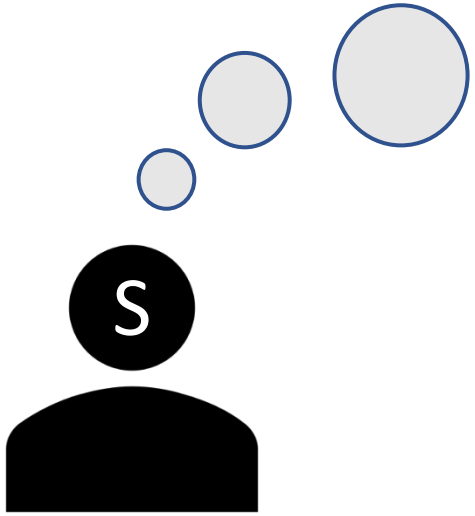


- High condition:
  - 24 ‘some|some|right’
  - 1 ‘none|none|wrong’ (=all|some|wrong)
- Low condition:
  - 16 ‘most|all|wrong’
  - 6 ‘all|most|wrong’

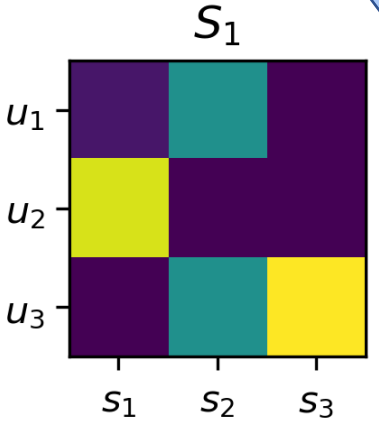


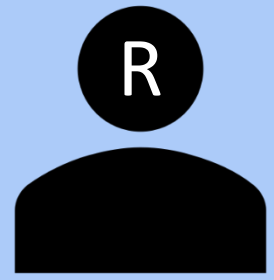
Gets a signal, does Bayesian updating



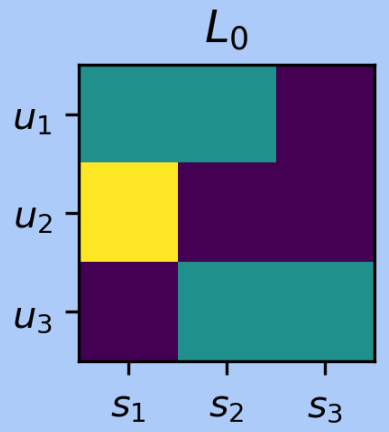


Sees a state, tends to send the message most useful to S





Gets a signal, guesses uniformly among the compatible states.



Frank et al (2017)

Extending the RSA model

$$\text{Util}(s, u) = \log([s \in [\![u]\!]]) + \log(|[\![u]\!]|^{-1}) - \text{cost}(u)$$

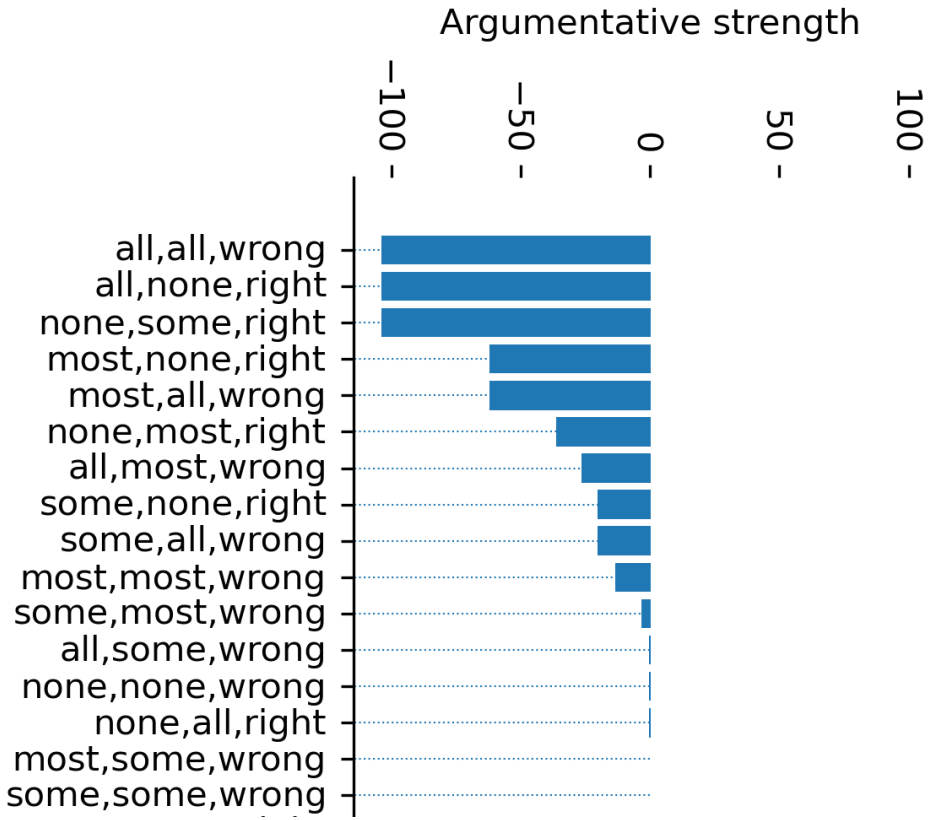
How much more surprising is the data under H_A than H?

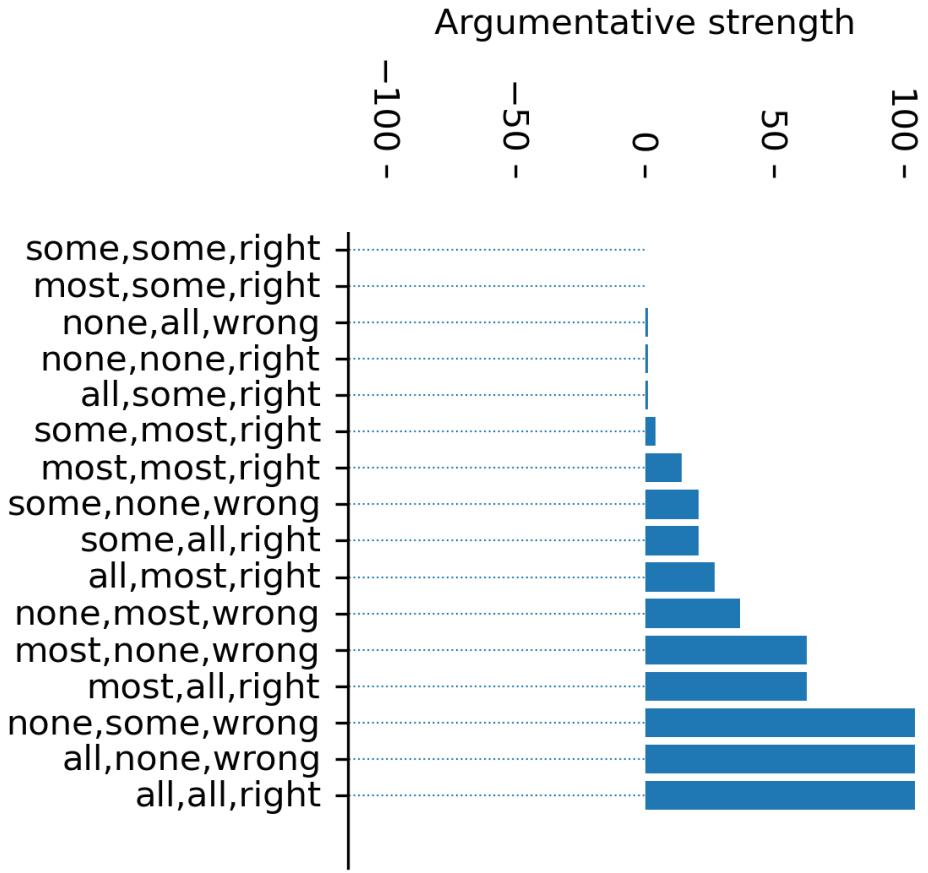
$$\text{argstrength}(u) =$$

$$P_{L_0}(u \mid \gamma) = \sum_o \underbrace{[o \in [\![u]\!]]}_{\text{Truth}} \quad \underbrace{P(o \mid \gamma)}_{\text{Plausibility of scenario}} \quad \underbrace{|\{u \mid o \in [\![u]\!]\}|^{-1}}_{\text{Utterance choice}}$$

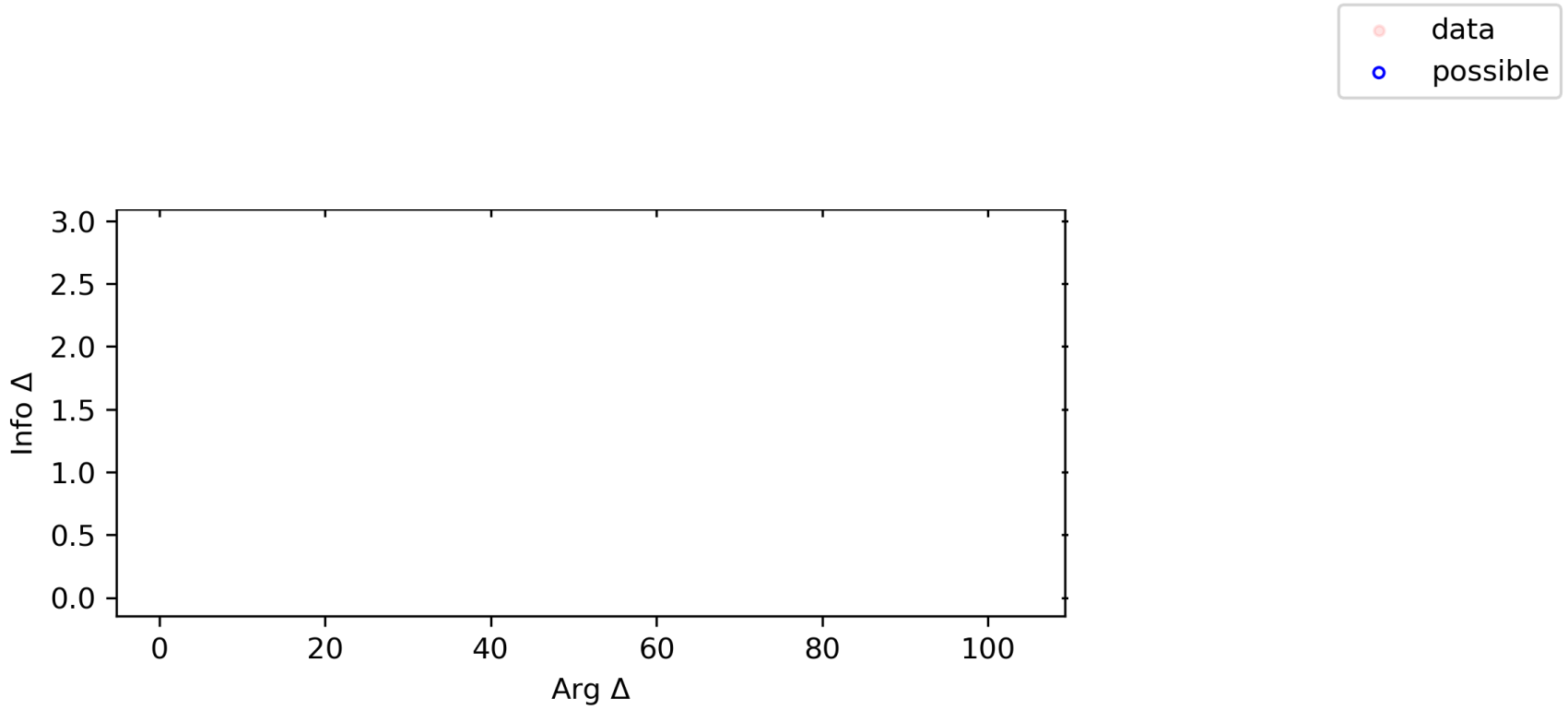
Ducrot (1969, 1973), Anschmore & Ducrot (1983)

Quantifying argumentative strength?



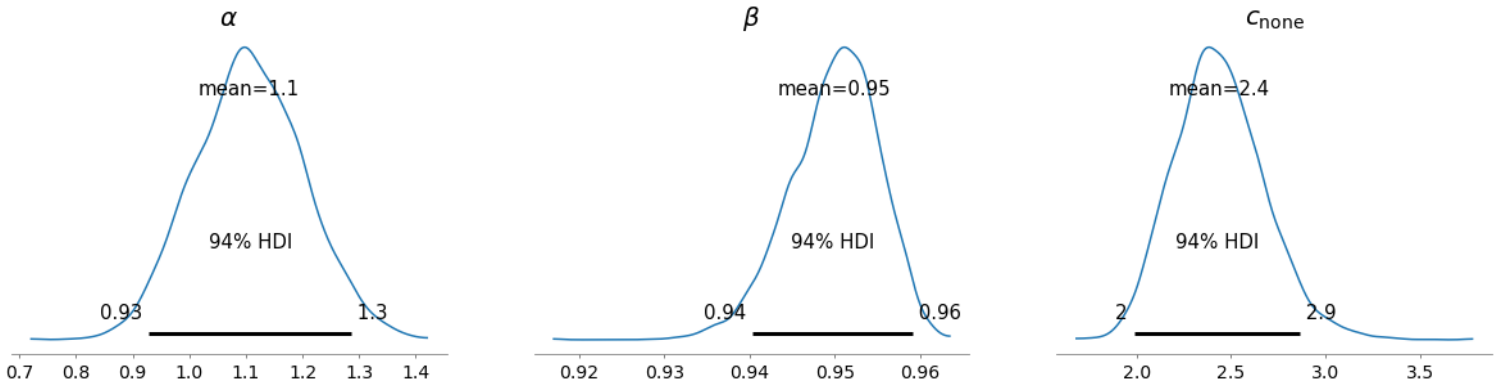


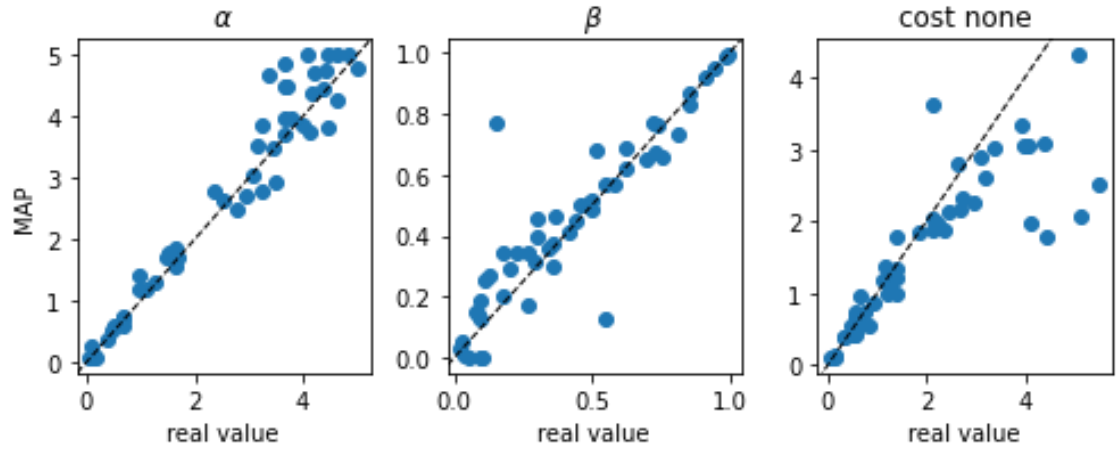
Informativity vs. argumentative strength

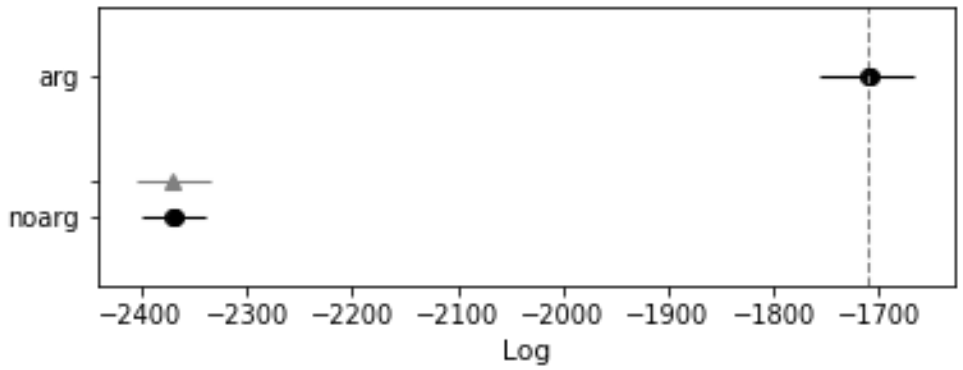


Basic RSA model

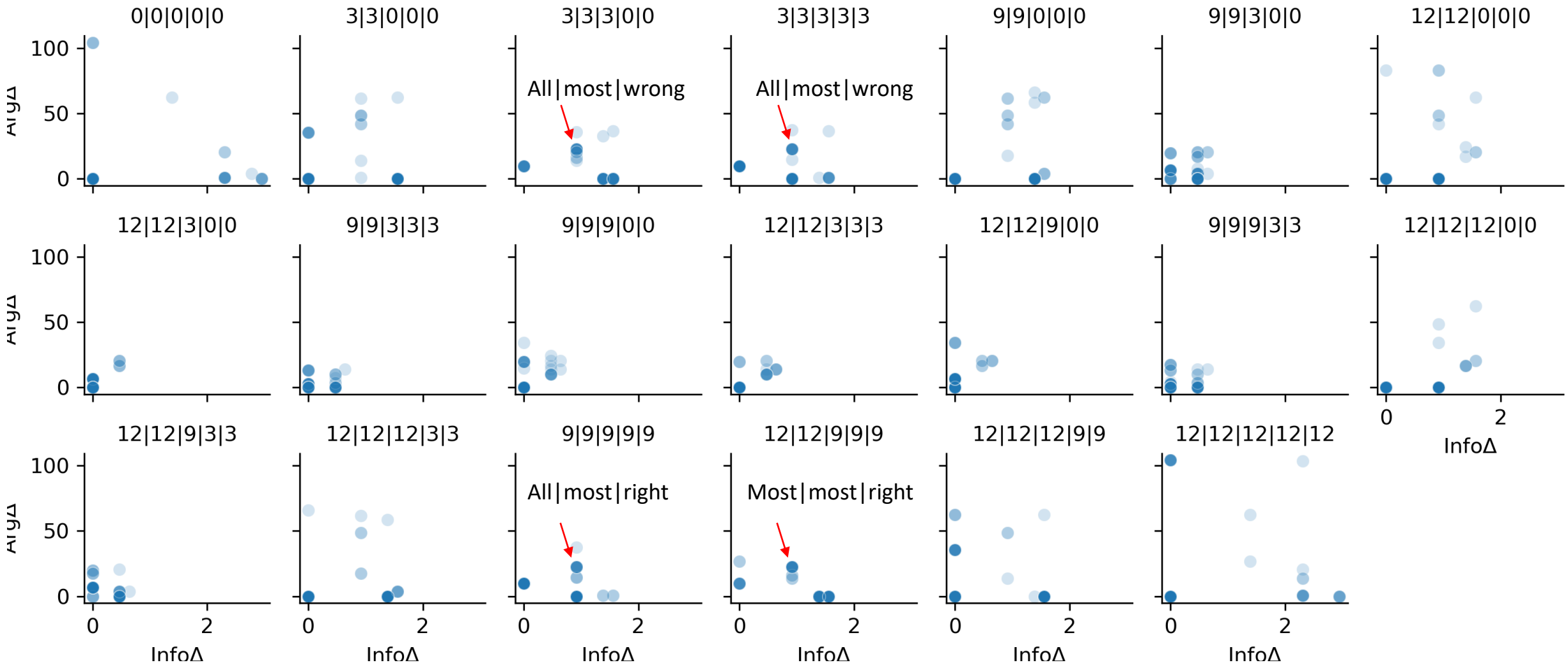
$$\alpha \sim \mathcal{U}(0, 5)$$
$$\beta \sim \mathcal{U}(0, 1)$$
$$c_{\text{None}} \sim \text{Exp}(\lambda = 0.5)$$
$$p_i = RSA(o_i, \alpha, \beta, c_{\text{None}})$$
$$u_i \sim \text{Cat}(p_i)$$





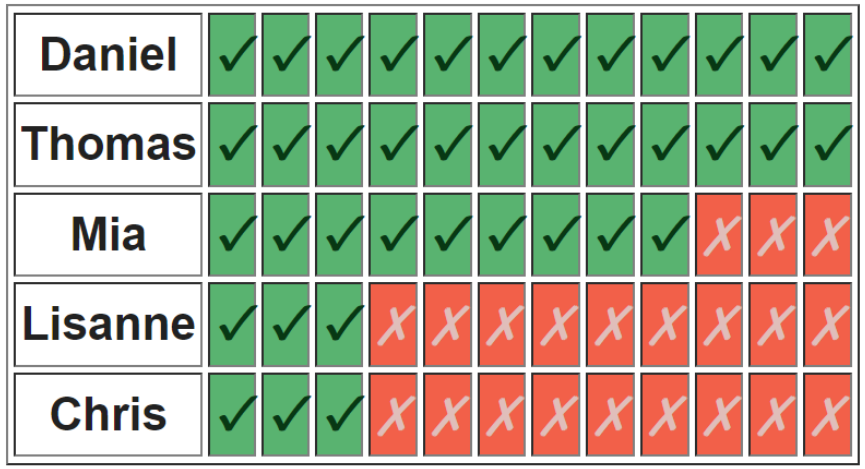


What does the model get wrong?



What does the basic model get wrong?

NOTE: This is **exploratory**!



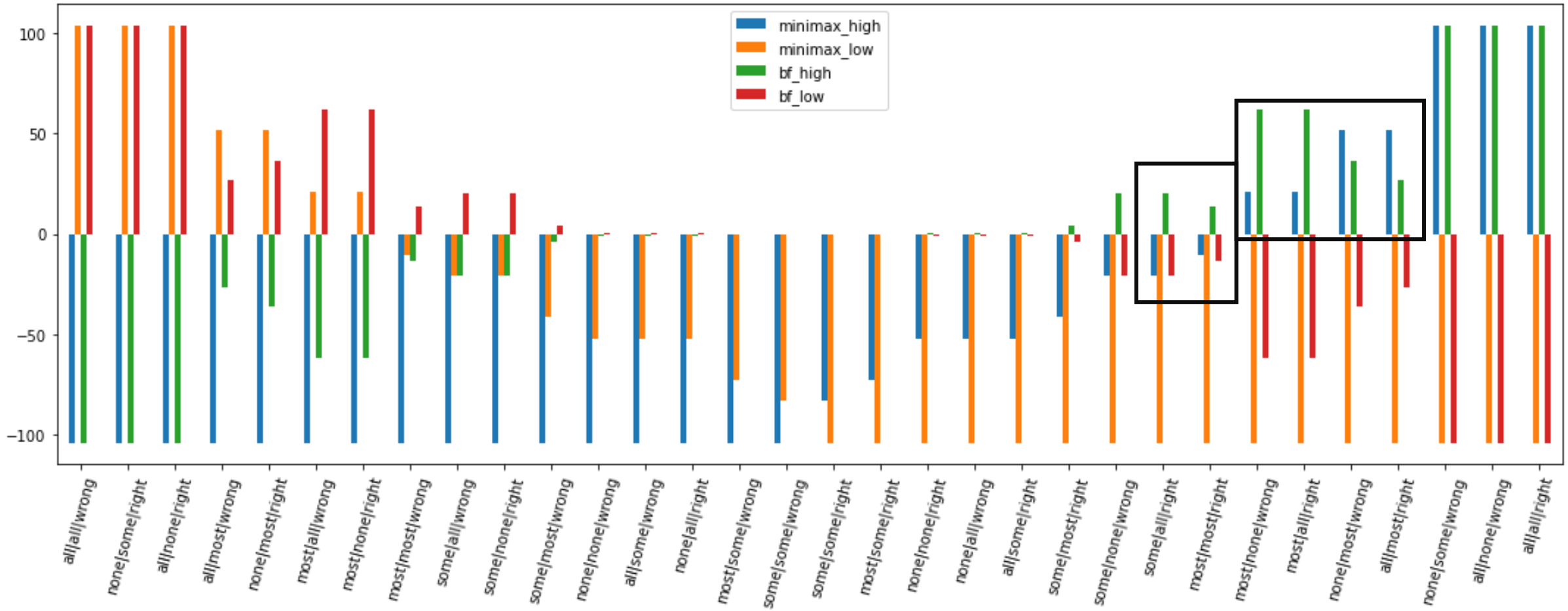
High condition:

20 ‘most|most|right’

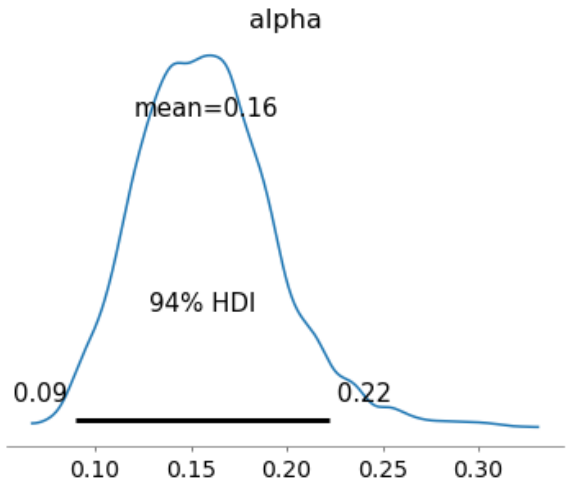
4 ‘some|all|right’ (argbest)

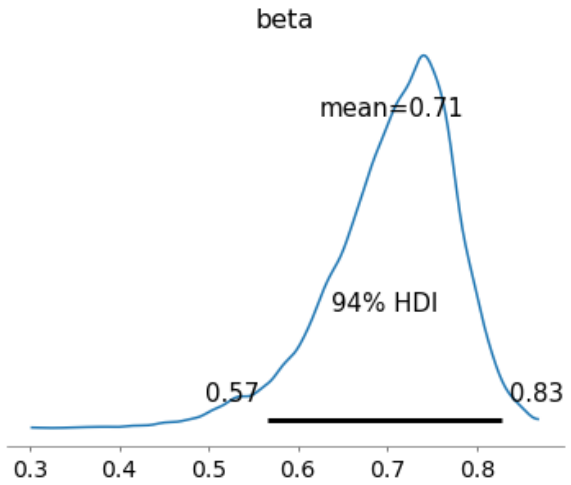
$$\text{maximin-argstrength}(u) = \min_{o} \log \frac{p(u \mid o, \phi_+)}{p(u \mid o, \phi_-)}$$

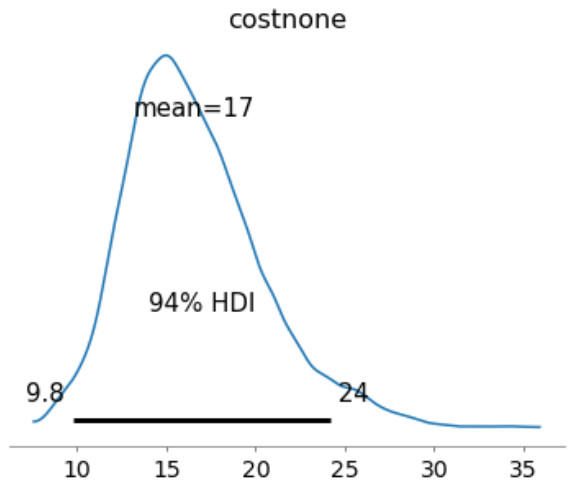
Maximin argstrength

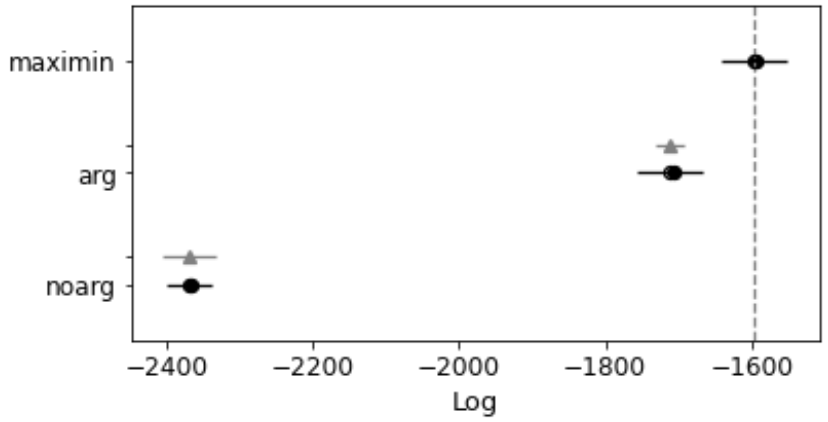


Maximin vs. log lik ratio argstrength

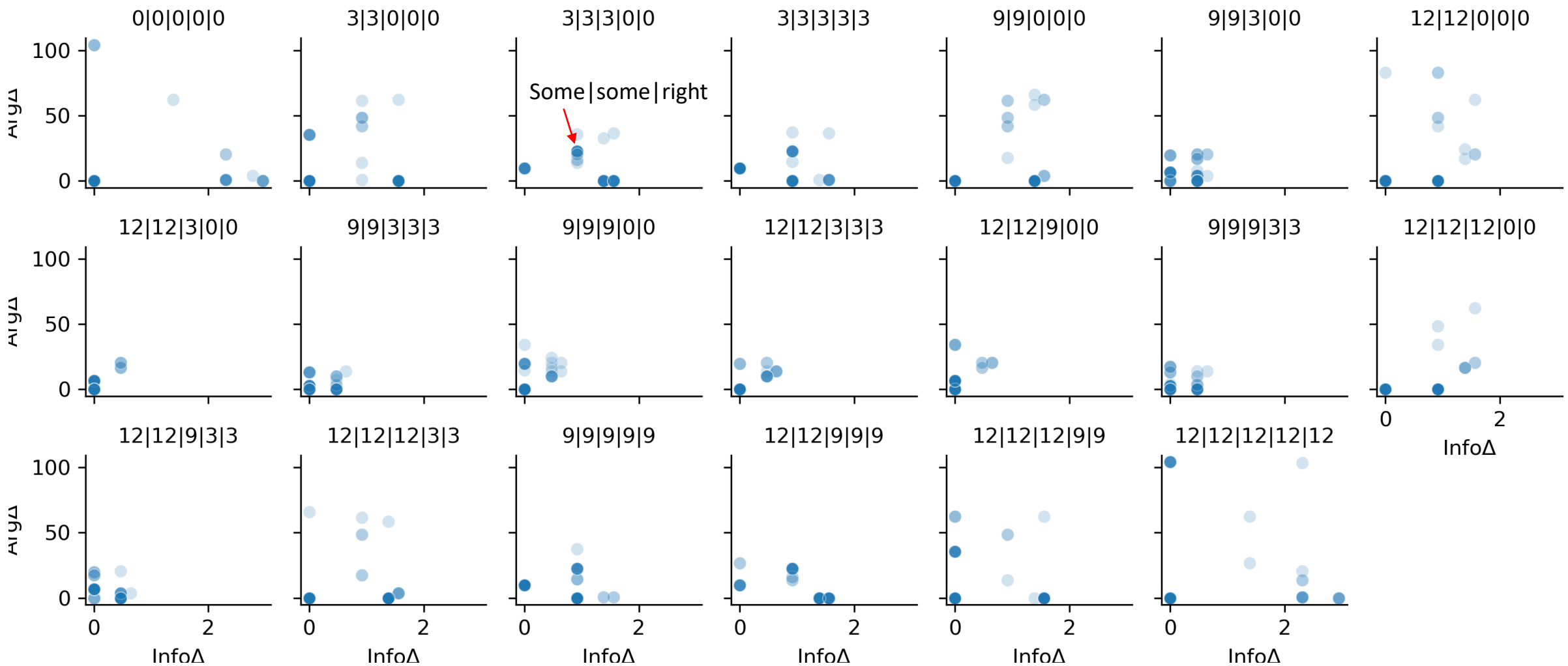








Maximin argstrength



Maximin vs. log lik ratio argstrength

Maximin +

- ‘all|most’ and ‘most|most’ consistently preferred over ‘some|all’.
  - E.g. (12,12,9,0,0), (12,12,9,3,3), (9,9,9,9,9), etc.
- Larger variety of messages for observations that are difficult to argue with
  - E.g. (12,12,12,12,12) in low condition

Maximin -

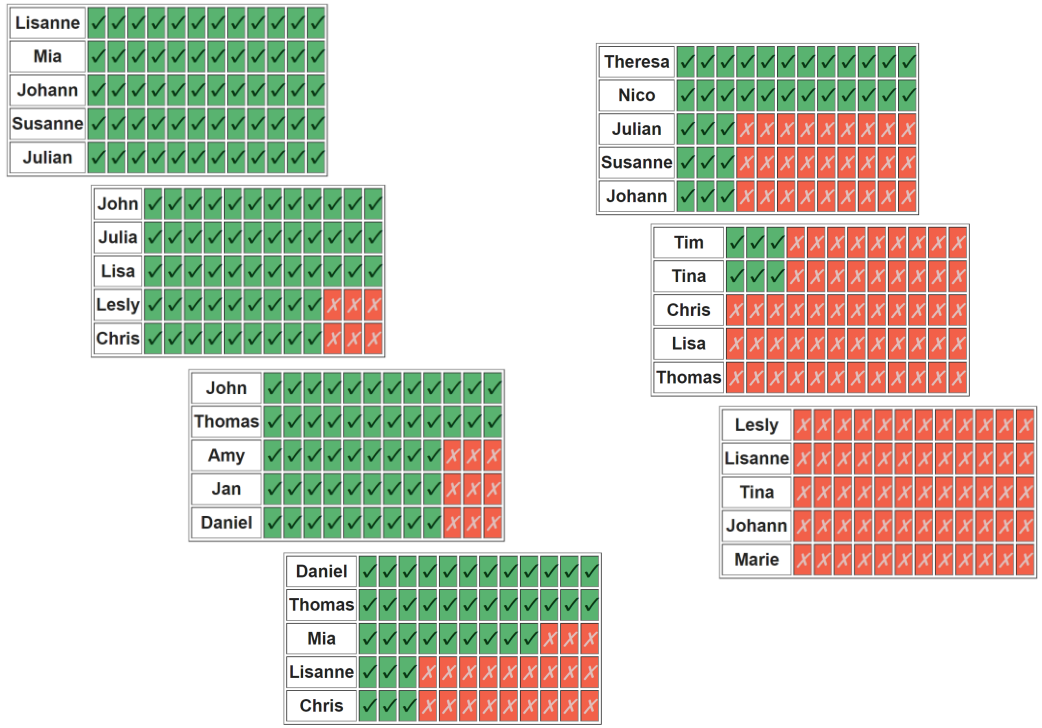
- ‘Weak’ signals can be bad for both argumentative aims (although possibly best *available*)
- (12,12,12,9,9), high condition: ‘most|all|right’ (predicted by loglikratio) preferred over ‘all|most|right’ (argbest for maximin).

Conclusions

- We have asked participants to choose an utterance while optimizing for a certain argumentative aim.
- Participants were able to do this online.
- We presented an extension of the RSA model that captures much of the participants' behaviour. In this picture we model argumentative strength with log likelihood ratio.
- We have further shown that the basic model systematically fails to predict some pattern.
- We have exploratorily proposed an alternative way of calculating argstrength that predicts the observed pattern.
- Much more work to do!
- Thank you!

Hierarchical RSA model

Quantifying argumentative strength?



- Most | most | right
- All| some |right
- All | some |wrong