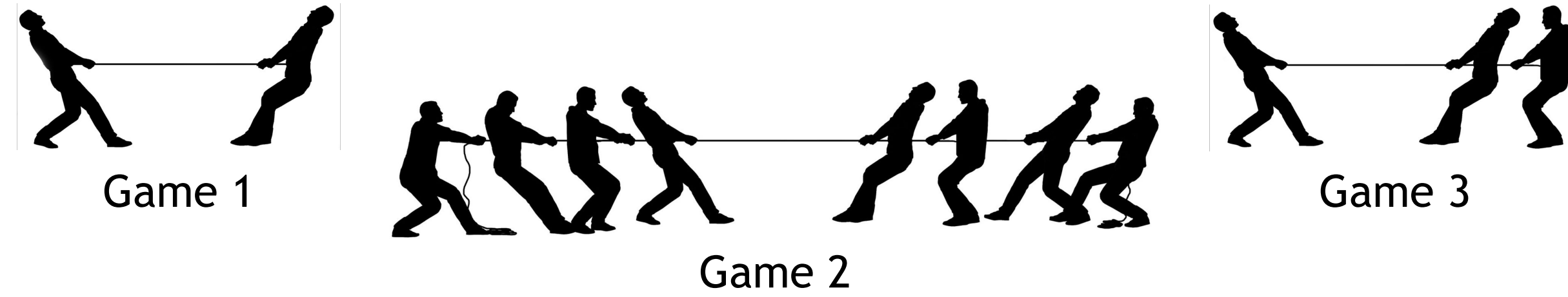




Imagine a Tug of War Tournament



The Probabilistic Language of Thought (PLOT)

Hypothesis 1

Humans reason about the world using concepts

strength, winner, team, player, laziness, pulling

Hypothesis 2

Concepts are compositional

A team is made up of players

Team 1 = Alex + Jim + Sarah

Hypothesis 3

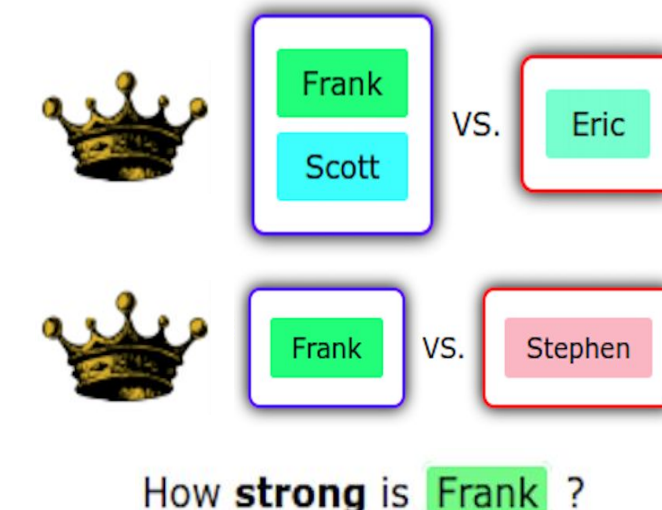
Concepts are probabilistic

People aren't always lazy or not lazy

Maybe they are lazy 30% of the time

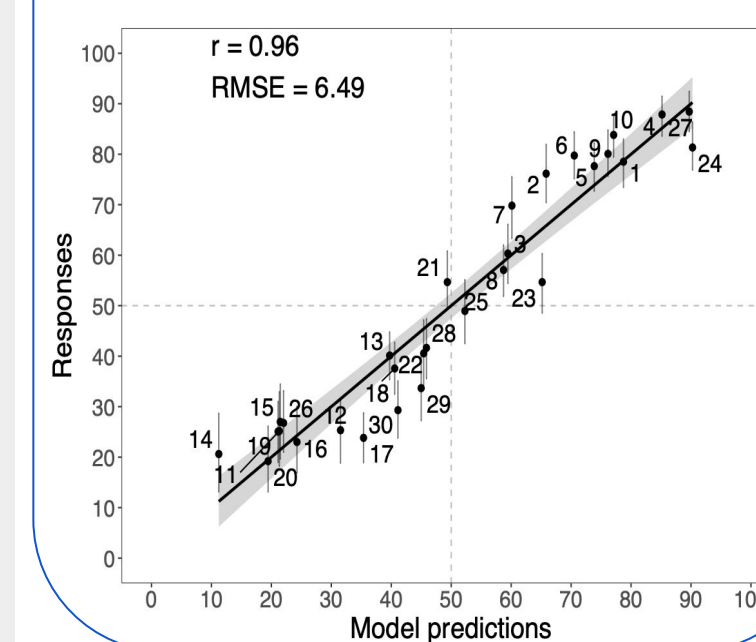
Tug of war PLOT model matches human intuitions strikingly well

Experiment 1 Strength

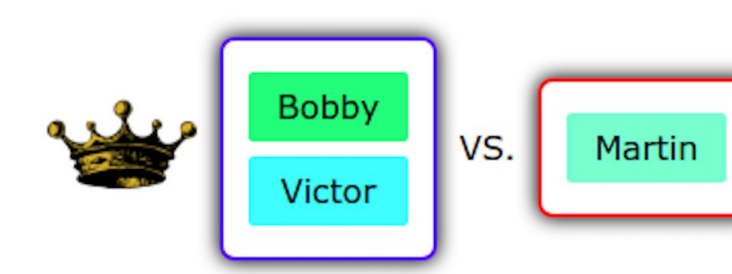


How strong is Frank?

How strong is player X?

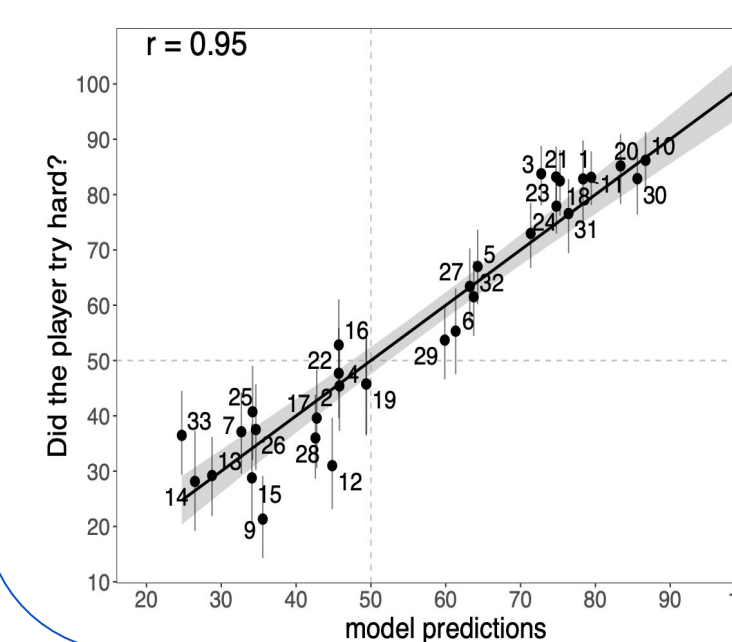


Experiment 2 Laziness

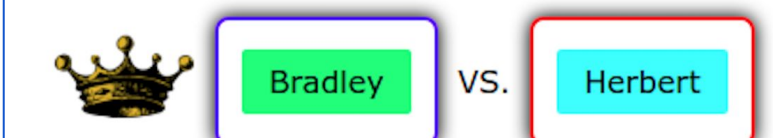


How much effort did Martin put in for Game 1?

How much effort did player X put in?



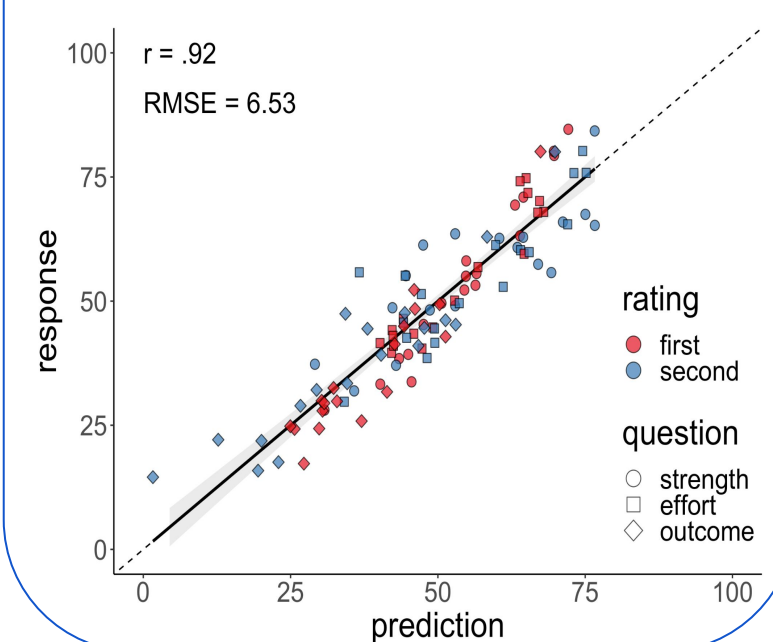
Experiment 3 Multiple Questions & Updated Beliefs



Herbert won 9 out of 10 games in the past.

Who would win in another match?
definitely Bradley definitely Herbert

Updated beliefs

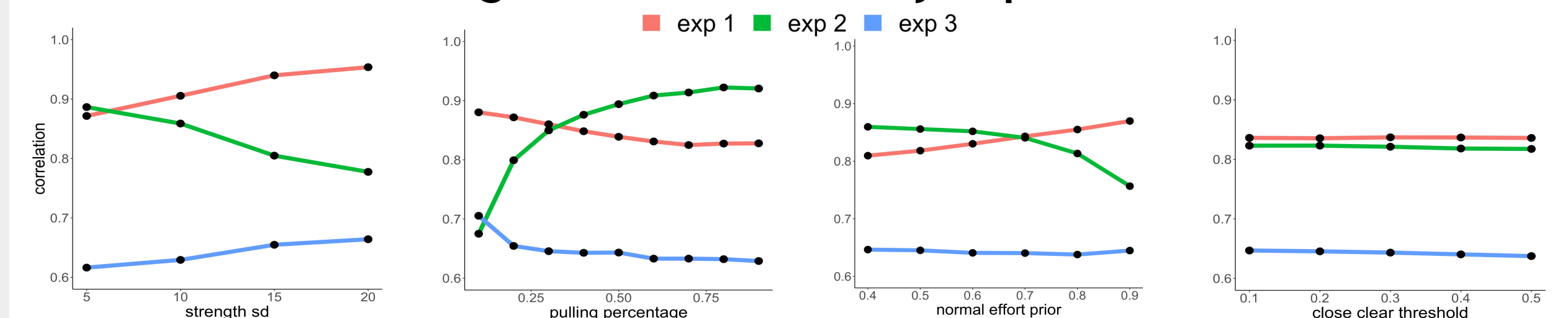


How sensitive is the PLOT model to parameter choices?

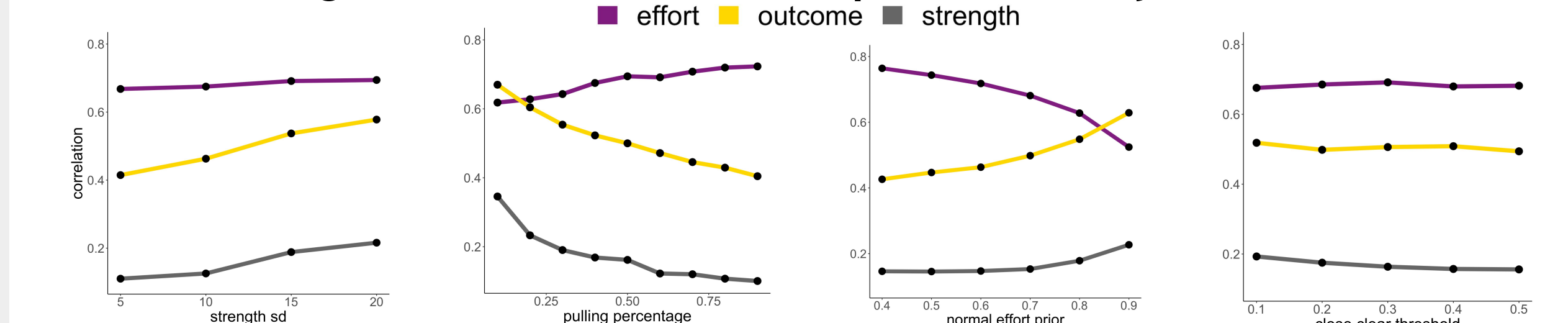
Parameter grid search over four parameters, all other parameters were dependent on one of these four

Parameters	Range
strength standard deviation	[5, 10, 15, 20]
pulling percentage	[.1, .2, .3, .4, .5, .6, .7, .8, .9]
normal effort prior	[.4, .5, .6, .7, .8, .9]
close clear threshold	[.1, .2, .3, .4, .5]

Marginal Distributions by Experiment



Marginal Distributions for Experiment 3 by Question



Alternative Models

Large Language Model (GPT-3)

Is it possible for a model to have the same kind of flexibility and generalizability of human thought just through large amounts of data?

Example 1

...How strong is player 1 on a scale of 0 to 100 based on the evidence below?

In Game 1, player 1 competed against player 2 and player 1 won.

In Game 2, player 2 competed against player 3 and player 2 won.

In Game 3, player 2 competed against player 4 and player 2 won.

Player 1 is stronger than player 2, who is stronger than players 3 and 4. Therefore, player 1 is at least as strong as player 2, and is likely stronger. I would rate player 1 as a 90.

Example 2

...How strong is player 1 on a scale of 0 to 100 based on the evidence below?

In Game 1, player 1 competed against player 2 and player 1 won

In Game 2, player 1 competed against player 3 and player 3 won.

Player 1 is 50% strong.

Player 1 won 2 out of 10 games in the past. Given this new evidence, on a scale of 0 to 100, how strong is player 1?

Player 1 is 80% strong.

Feature Based Models

Can features of the tug of war game accurately predict human judgments?

- # games lost
- # games won
- # players on team
- # players on opposite team
- Did player win game being asked about

Discussions & Future Work

Parameter Grid Search

- Parameters matter to some degree
- There isn't only one parameter setting that works
- But not the case that all of the parameter settings work

PLOT & Alternative Models

- PLOT model works well but need to complete evaluation of alternative models

Tug of War PLOT Model

```
var tugOfWarModel = function() {
  var strength = mem(function(person) {
    return sample(DiscretizedGaussian({mu: strengthMean, sigma: strengthSD})))
  }
  var effort = mem(function(person, match) {
    return categorical({ps: [lazyEffortPrior, normalEffortPrior, extraEffortPrior],
      vs: [lazyPulling, normalPulling, extraPulling]}))
  }
  var pulling = function(person, match) {
    return effort(person, match) * strength(person)
  }
  var totalPulling = function(team, match) {
    return sum(map(function(person) {return pulling(person, match)}, team))
  }
  var deltaTeamStrength = function(team1, team2, match) {
    return totalPulling(team1, match) - totalPulling(team2, match)
  }
  var winner = function(team1, team2, match) {
    return deltaTeamStrength(team1, team2, match) > 0 ? 1 : 2
  }
  var isClose = function(team1, team2, match) {
    return Math.abs(deltaTeamStrength(team1, team2, match)) < closeThreshold
  }
  var isClear = function(team1, team2, match) {
    return Math.abs(deltaTeamStrength(team1, team2, match)) > clearThreshold
  }
}
```

concepts

evidence

query

```
condition(winner(['Bob', 'Mary'], ['Tom'], ['match1']))
condition(winner(['Bob', 'Sue'], ['Tom'], ['match1']))
```

```
return strength('Bob') }
```