Collaborative decision making is grounded in representations of other people's competence and effort

Supplemental Material

Yang Xiang,[1,*] Natalia Vélez,[1] Samuel J. Gershman[1,2,3]

[1]Department of Psychology, Harvard University

[2]Center for Brain Science, Harvard University

[3]Center for Brains, Minds and Machines, MIT

[*]Corresponding author: yyx@g.harvard.edu

August 12, 2022

Supplemental Material Guide

1 A safe joint effort model

Recall that in the joint effort model, the lift outcome is determined by each agent's effort, strength, and the box weight:

$$L = \mathbb{I}[\textstyle\sum_a E_a S_a \geq W], \tag{1}$$

where a indexes agents. To explore the hypothesis that people hedge their predictions, we create a *safe joint effort model*, where we add a hindrance factor H to Eq. 1:

$$\widetilde{L} = \mathbb{I}[\textstyle\sum_a E_a S_a \geq (W + H)]. \tag{2}$$

The hindrance factor H is defined as:

$$H = k \cdot \sum_a (1 - E_a), \tag{3}$$

where $k \geq 0$ is a scaling parameter.

The optimal effort is thus given by:

$$\mathbf{E}^* = \underset{\mathbf{E}}{\operatorname{argmax}}\, R \cdot P(\widetilde{L} = 1 | \mathbf{E}, W) - C(\mathbf{E}), \tag{4}$$

where $\mathbf{E} = [e_1, \ldots, e_n]$ is the vector of efforts for n agents and $C(\mathbf{E})$ is the cost function:

$$C(\mathbf{E}) = \sum_a \alpha_a E_a + \beta G(\mathbf{E}), \tag{5}$$

where $\beta \geq 0$ is a scaling parameter and $G(\mathbf{E})$ is the Gini coefficient, defined as half of the relative mean absolute difference in effort. In the next section, we contrast the predictions of the safe joint effort model (k set to 3.5), the joint effort model, the joint effort model without the Gini coefficient, and behavioral data.

2 Variations of the joint effort model

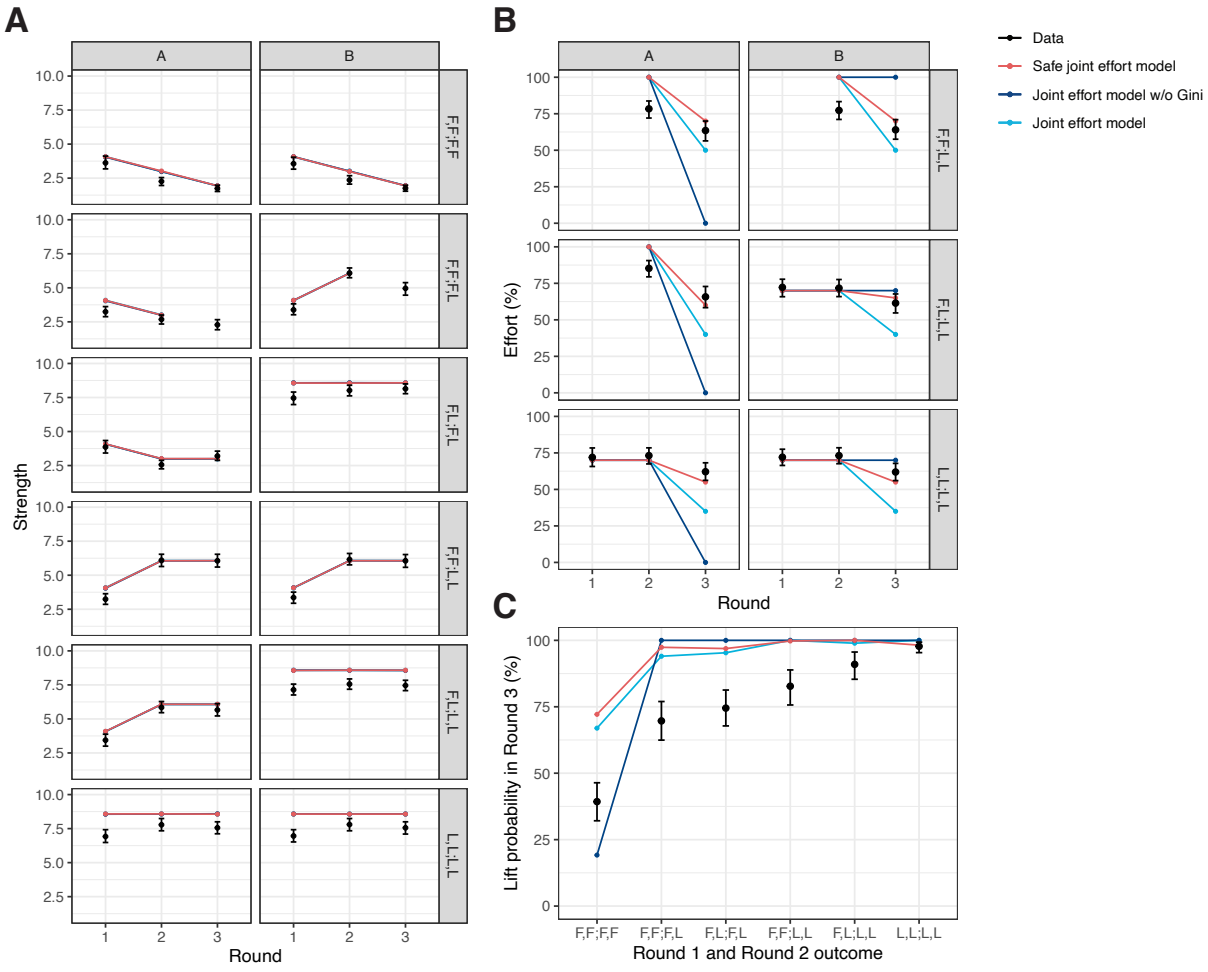


Figure S1: (A) Predictions of contestants' strength in Experiment 1. (B) Predictions of contestants' effort in Experiment 1. Effort judgments were not elicited when the lift outcome was Fail. (C) Predictions of the lift probability in Round 3 of Experiment 1. Model simulations averaged over 10 runs. Error bars indicate bootstrapped 95% confidence intervals.

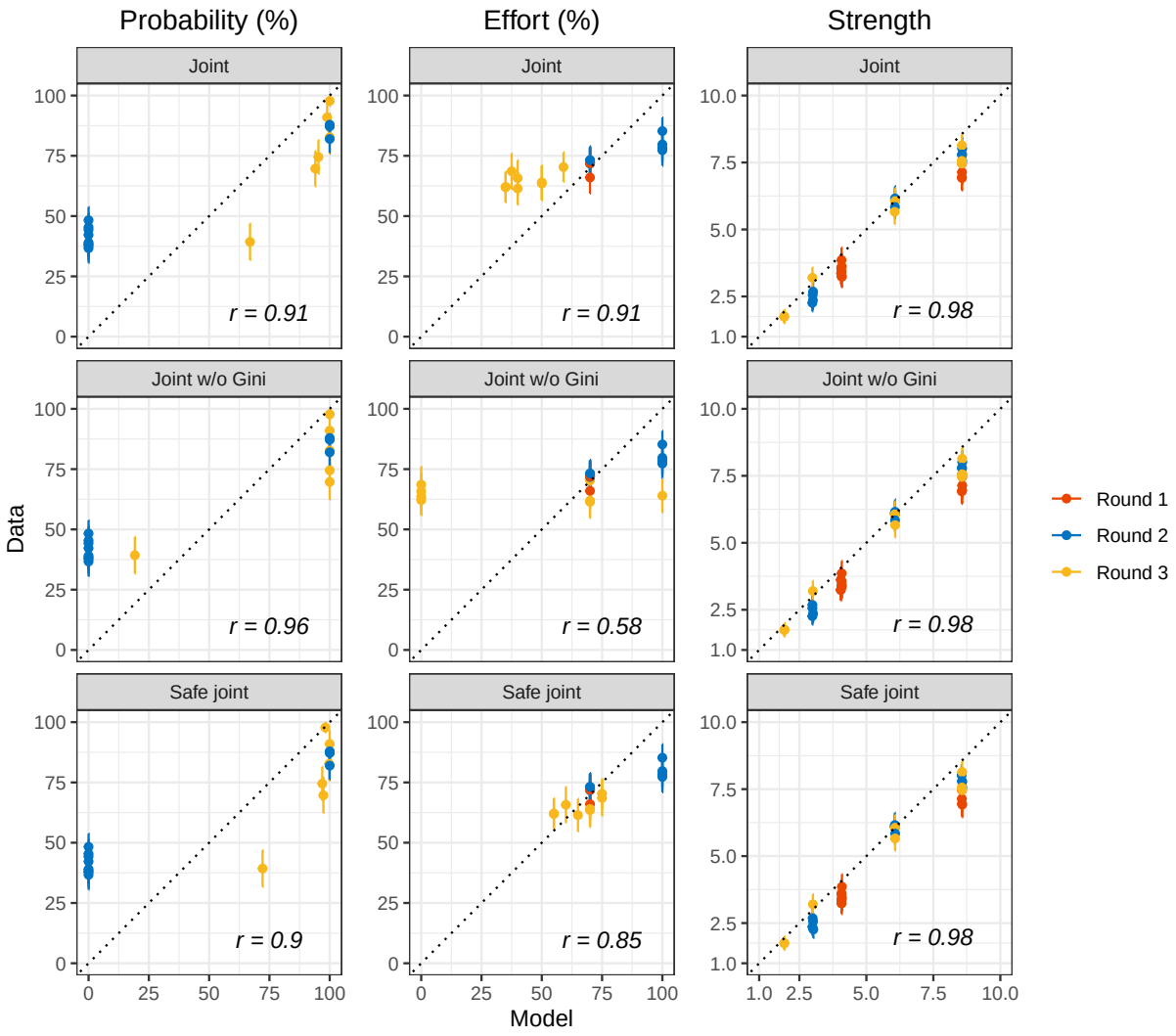


Figure S2: Comparing model predictions to behavioral data across predictions of lift probability, effort, and strength in Experiment 1. Model simulations averaged over 10 runs. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars indicate 95% normal confidence intervals.

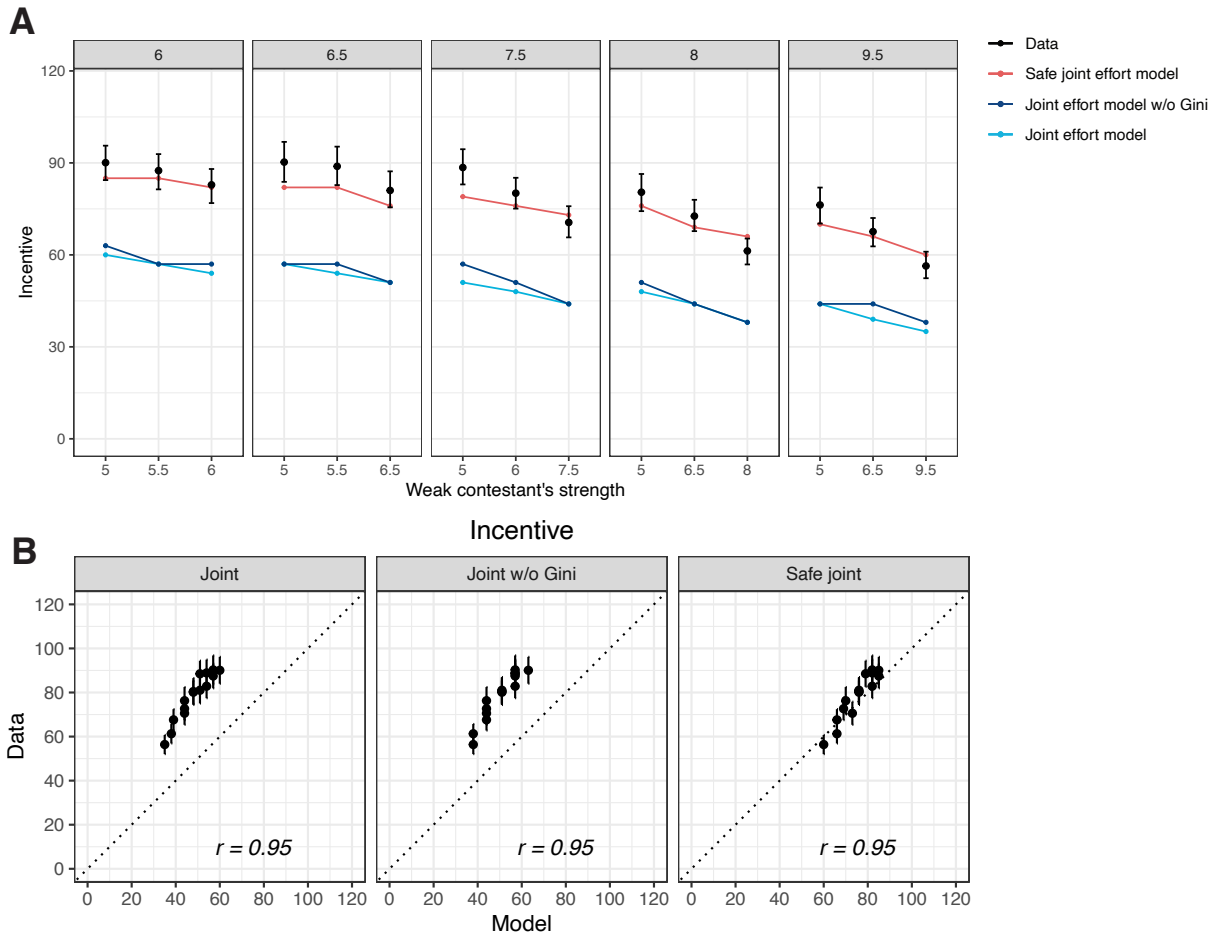


Figure S3: (A) Incentive provided to pairs of contestants in Experiment 2. Each panel corresponds to the strongest contestant's strength in each contest. (B) Comparison between behavioral data and model predictions of incentive in Experiment 2. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars in (A) indicate bootstrapped 95% confidence intervals; error bars in (B) indicate 95% normal confidence intervals.

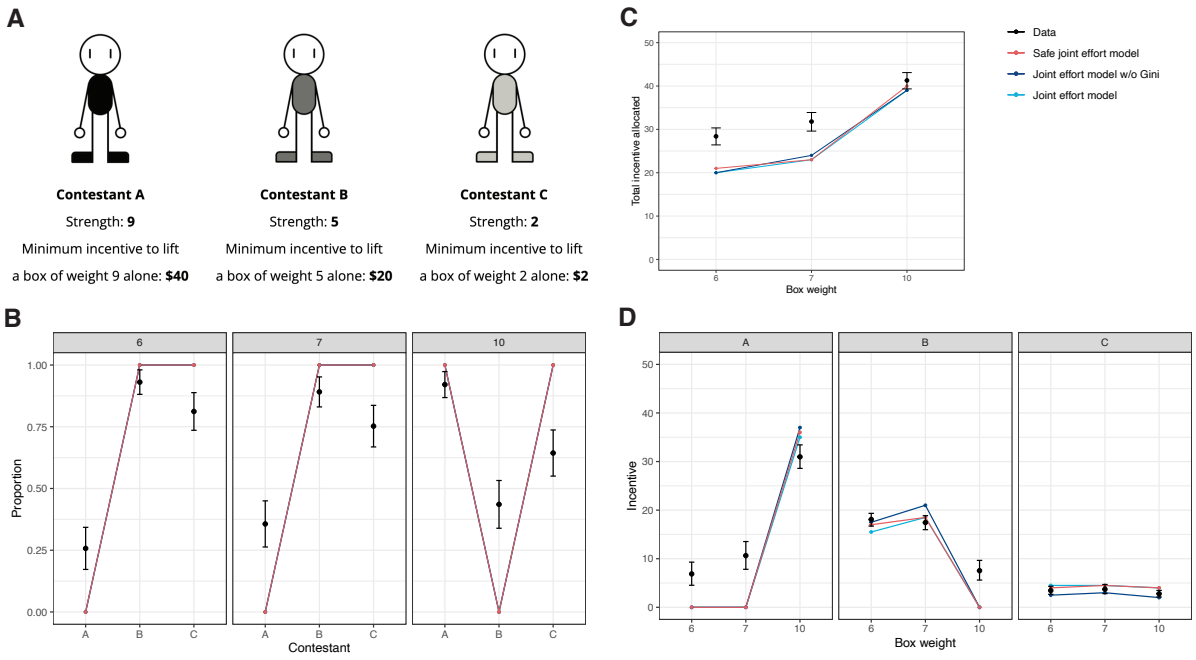


Figure S4: (A) Contestants in Experiment 3. (B) Proportion of times each contestant was selected for each contest in Experiment 3. Contests are identified by their box weights, which appear above each plot. (C) Total incentive provided to contestants in Experiment 3. (D) Incentive provided to each contestant in Experiment 3. Each plot corresponds to an individual contestant. Error bars in (B) indicate 95% confidence intervals of proportions; error bars in (C) and (D) indicate bootstrapped 95% confidence intervals.

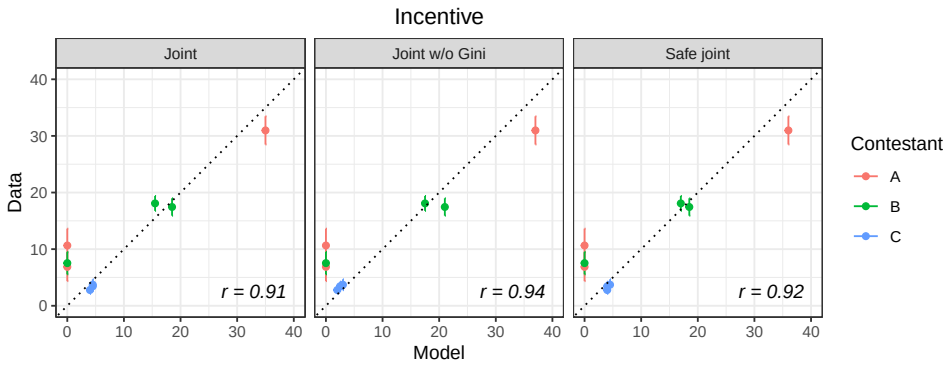


Figure S5: Comparison between behavioral data and model predictions of incentive in Experiment 3. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars indicate 95% normal confidence intervals.

3 Including Gini coefficient in all models

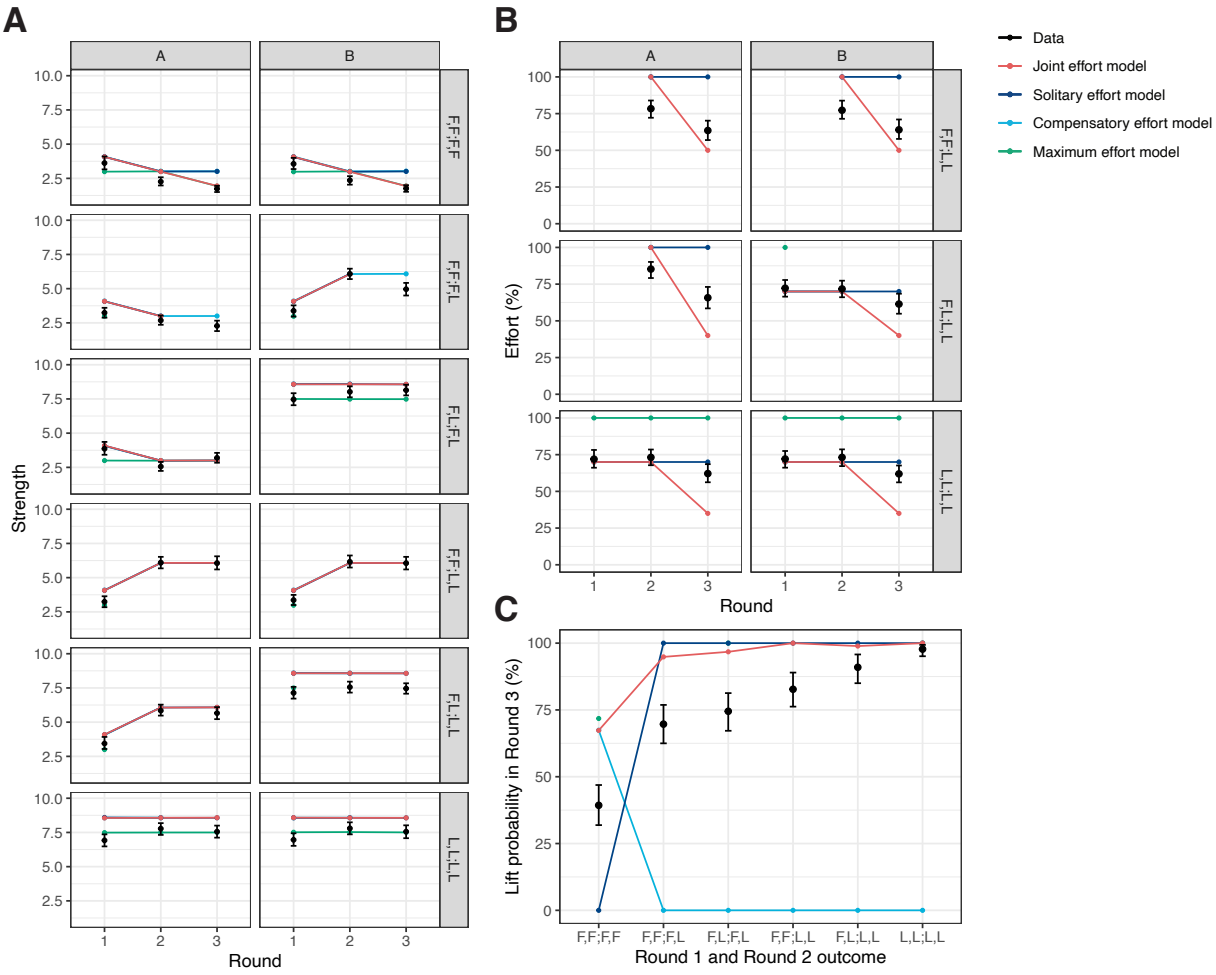


Figure S6: (A) Predictions of contestants' strength in Experiment 1. (B) Predictions of contestants' effort in Experiment 1. Effort judgments were not elicited when the lift outcome was Fail. (C) Predictions of the lift probability in Round 3 of Experiment 1. Model simulations averaged over 10 runs. Error bars indicate bootstrapped 95% confidence intervals.

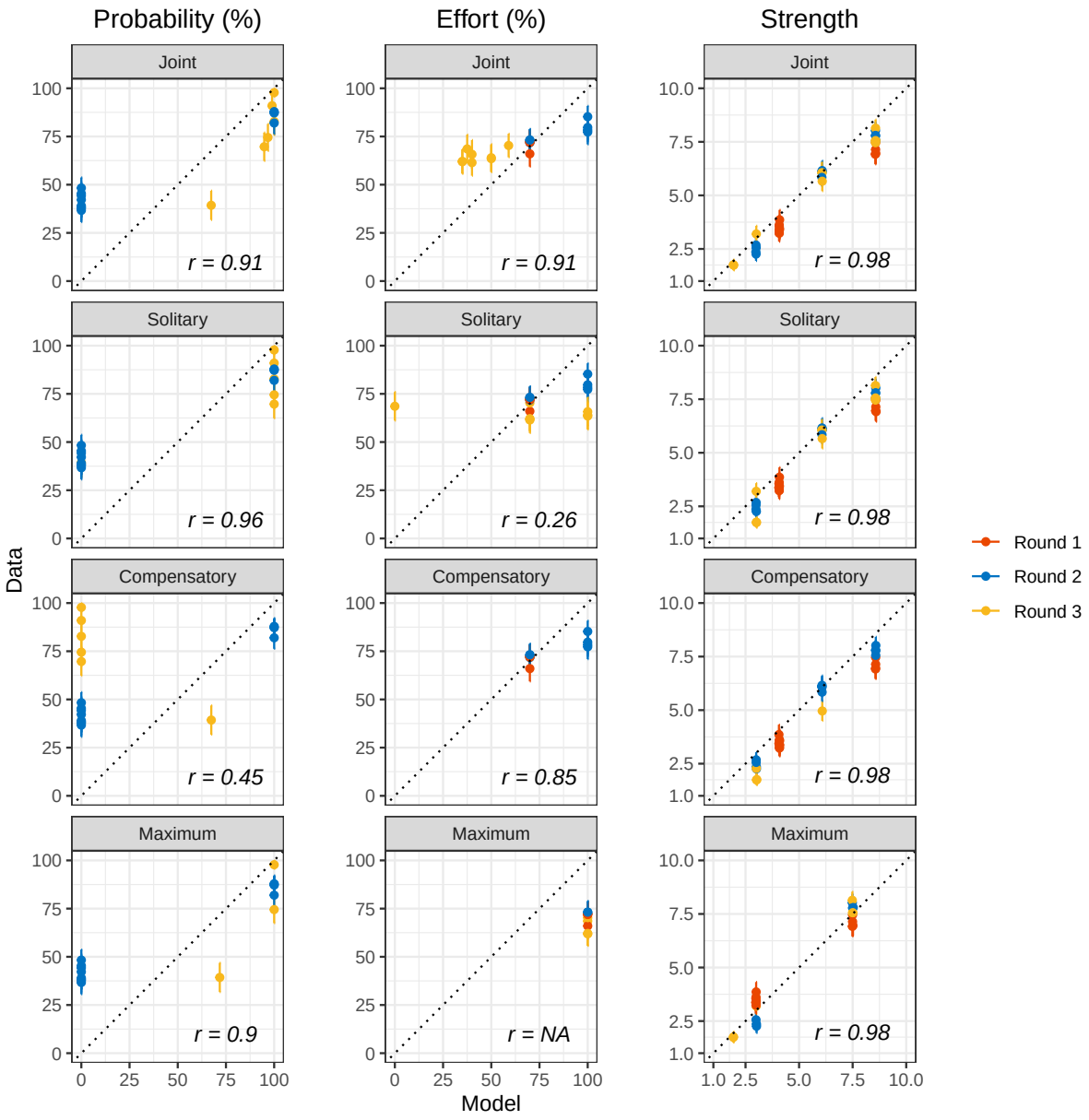


Figure S7: Comparing model predictions to behavioral data across predictions of lift probability, effort, and strength in Experiment 1. Model simulations averaged over 10 runs. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars indicate 95% normal confidence intervals.

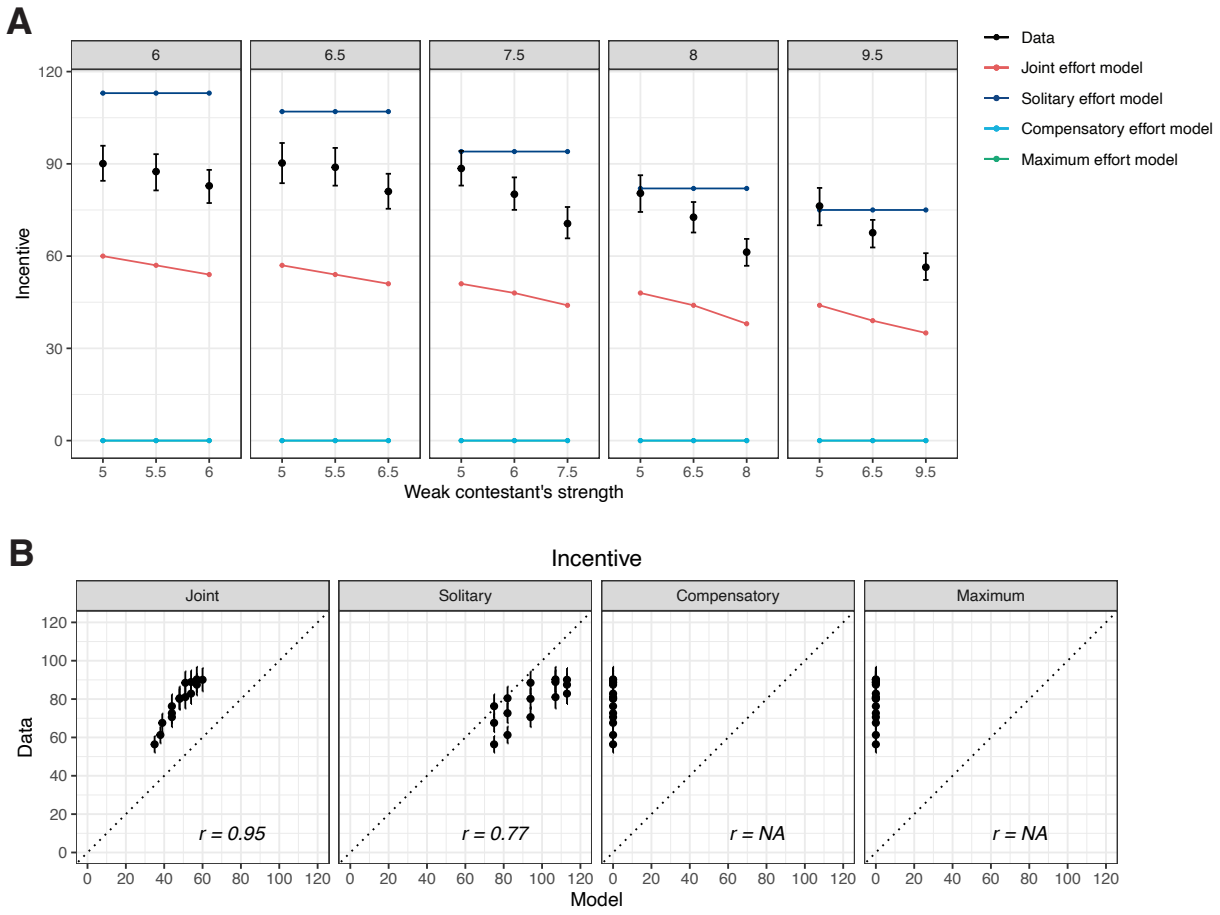


Figure S8: (A) Incentive provided to pairs of contestants in Experiment 2. Each panel corresponds to the strongest contestant's strength in each contest. (B) Comparison between behavioral data and model predictions of incentive in Experiment 2. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars in (A) indicate bootstrapped 95% confidence intervals; error bars in (B) indicate 95% normal confidence intervals.

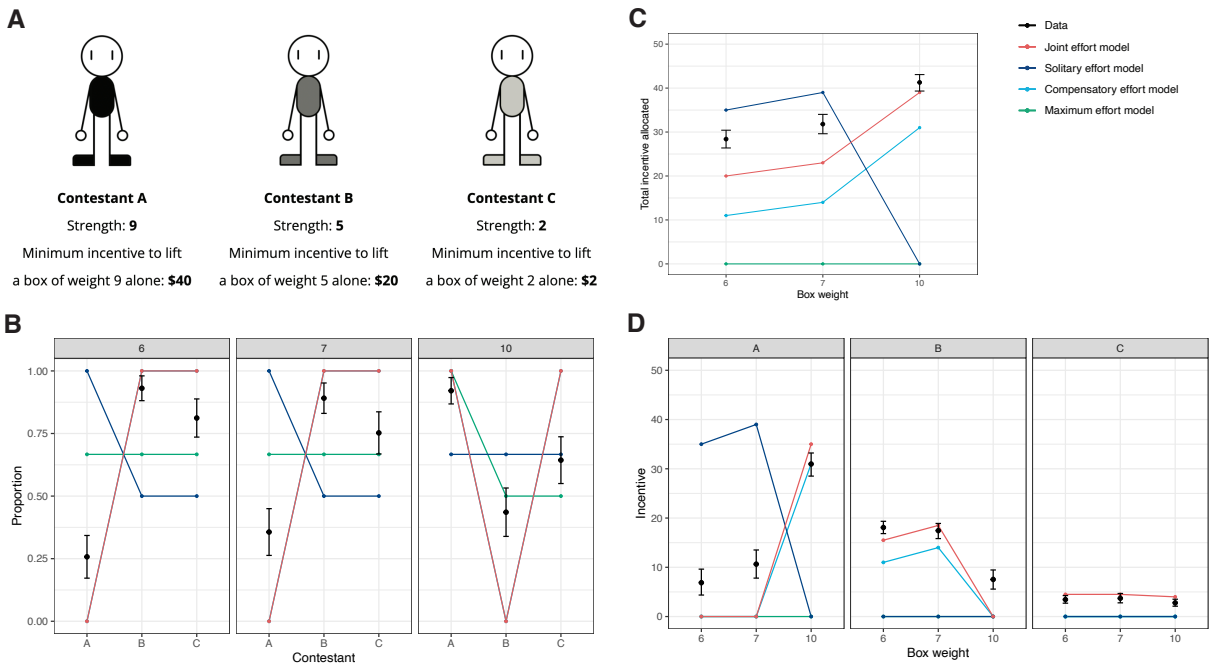


Figure S9: (A) Contestants in Experiment 3. (B) Proportion of times each contestant was selected for each contest in Experiment 3. Contests are identified by their box weights, which appear above each plot. (C) Total incentive provided to contestants in Experiment 3. (D) Incentive provided to each contestant in Experiment 3. Each plot corresponds to an individual contestant. Error bars in (B) indicate 95% confidence intervals of proportions; error bars in (C) and (D) indicate bootstrapped 95% confidence intervals. Predictions of the compensatory effort model in (B) is covered by the joint effort model. This is because when the agent chooses not to act, the penalty for unequal effort allocation is very high, such that the compensatory effort model predicts that an agent would act because they need to avoid the high penalty, rather than a result of recursive reasoning. Decreasing the β value will decrease the penalty, therefore resulting in a similar prediction as in the main text. Predictions of the compensatory effort model in (D) is covered by the solitary effort model (predicting \$0 for Contestant C across all contests).

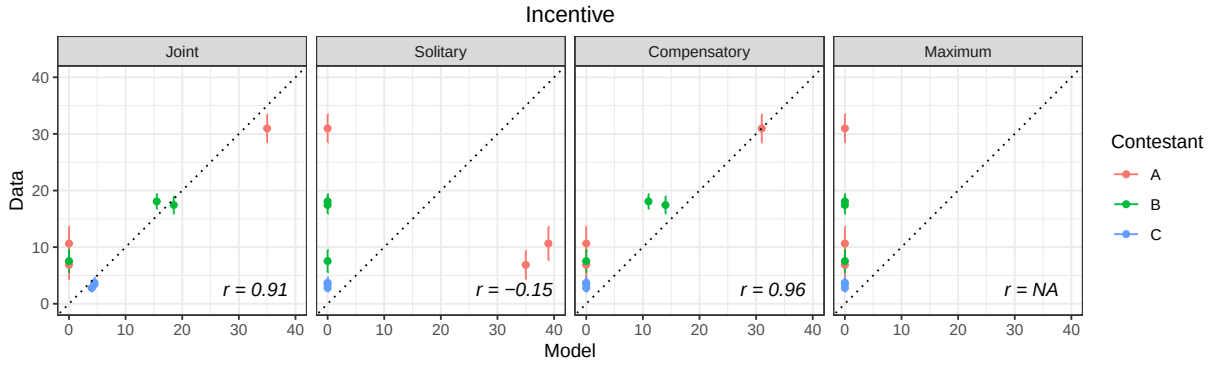


Figure S10: Comparison between behavioral data and model predictions of incentive in Experiment 3. Pearson correlation coefficient shown at the bottom right of each subplot. Error bars indicate 95% normal confidence intervals.

4 Experiment instructions

4.1 Experiment 1

Welcome!

In this experiment, you will watch a game show called *Lift That Box!* and answer some questions.

The game show consists of different contests between different pairs of contestants.

In each contest, the contestants will be given **three attempts** to lift a box.

We will show you whether they succeeded in lifting the box after each attempt.

In the **final round** of each contest, the two contestants will try to lift the box **together**.

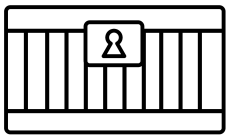
Page Break

The weight of the box is always the same, across all rounds and contests.

To give you an idea of how heavy the box is:
Suppose we rated the weights of objects on a scale of 1 to 10, where 1 is a weight so light that virtually anyone can lift it, and 10 is a weight so heavy that only the strongest humans can lift it.

We loaded the box so that it would measure a 5 on this scale.

In other words: An average person (strength 5) should be able to lift the box if they put in an all-out, 100% effort, and an extremely strong person (strength 10) should be able to lift the box with only a moderate, 50% effort.



Page Break

Contestants can win **$10** or **$20** for lifting the box, as indicated on the box.

In the final round, if they lift the box together, **each of them gets the prize money** shown on the box.

The prizes change from round to round, but the box doesn't change.

Page Break

You will predict how likely the contestants are to lift the box, observe the actual outcome, and answer questions regarding their strength and effort.

When making your guesses, you will see a table of all the previous outcomes.

By the end of the game, we will randomly pick a round to calculate your bonus.

If the outcome on that round is Lift, then your bonus will be the probability of lifting that you guessed.

If the outcome on that round is Fail, then your bonus will be 1 minus the probability of lifting that you guessed.

Contestant	I	J
Round 1 ($10)	Fail	Lift
Round 2 ($20)	Lift	Lift
Round 3 ($20 for each)	?	

Page Break

Keep in mind that some people are stronger and some are weaker.

Even two contestants with the same strength may have different outcomes (some people are lazier and need higher reward to do the same thing!).

The contestants know very well their own strength and the other contestant's strength.

They also know how heavy the box is.

Page Break

Also remember that **the box is always the same**.

Different pairs of contestants will come in to lift the box; each pair will play the game for three consecutive rounds.

If the contestant does not feel the prize is worth their effort, they might not lift the box even if they are strong enough to do so.

Note that effort is costly; If the contestant can lift the box with 50% effort, they would not exert more effort than needed.

4.2 Experiment 2

Welcome!

In this experiment, you will watch a game show called *Lift That Box!* and answer some questions.

The game show consists of different contests between different teams of contestants.

We will change all of the contestants at the end of each contest.

In each round of the contest, two contestants will be given an attempt to lift a box together.

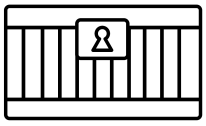
Page Break

The weight of the box is always the same, across all rounds and contests.

To give you an idea of how heavy the box is:
Suppose we rated the weights of objects on a scale of 1 to 10, where 1 is a weight so light that virtually anyone can lift it, and 10 is a weight so heavy that only the strongest humans can lift it.

We loaded the box so that it would measure a 5 on this scale.

In other words: An average person (strength 5) should be able to lift the box if they put in an all-out, 100% effort, and an extremely strong person (strength 10) should be able to lift the box with only a moderate, 50% effort.



Page Break

Contestants can win a certain amount of prize money as an incentive for lifting the box.

We will tell you the minimum incentive that each contestant would accept to lift the box alone.

Your job is to decide how much incentive you would like to provide each pair of contestants to lift the box together.

If they lift the box together, **each contestant gets the incentive** you provide.

For example, if you provide an incentive of $50, then each contestant will get $50 if they lift the box together.

No matter how the prizes change, the box is always the same.

Page Break

For every successful lift, you will get a bonus of $0.4.

For every dollar of incentive you give, $0.002 will be deducted from your bonus payment, regardless of whether the result is Lift or Fail.

Suppose you provide $50 for each contestant to lift a box.

If they succeed, then your bonus payment increases by $0.4-50x$0.002 = $0.3.

If they fail, you lose 50x$0.002 = $0.1.

Your total bonus payment can go up to $6, and will also not be less than $0.

You will see whether the contestants lift the box or not with the incentive you gave by the end of each contest.

Page Break

Keep in mind that some people are stronger and some are weaker.

The box is always the same, its weight equivalent to strength of 5.

The contestants know very well their own strength and the other contestant's strength.

They also know how heavy the box is.

Page Break

Also remember that the more effort a contestant needs to lift the box, the higher the incentive they will require.

In other words, stronger contestants can lift the box with less effort, therefore they require lower incentives.

Note that **effort is costly**; If the contestant does not feel the prize is worth their effort, they might not lift the box even if they are strong enough

4.3 Experiment 3

Welcome!

In this experiment, you will watch a game show called *Lift That Box!* and answer some questions.

The game show consists of different contests between different teams of contestants.

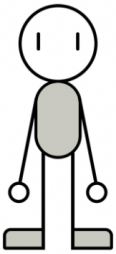
There are three different contestants altogether.

In each contest, two of the three contestants will be given an attempt to lift a box together.

We will change the box at the end of each contest.

Contestant A

Contestant B

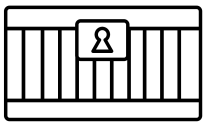


Contestant C

Page Break

The weight of the box is converted onto the same scale as people's strength.

Suppose we rated the weights of objects on a scale of 1 to 10,
where 1 is a weight so light that virtually anyone can lift it, and 10 is a weight so heavy that only the strongest humans can lift it.
A box of weight 1 would be extremely easy to lift, and a box of weight 10 would be extremely hard to lift.
A box of weight 5 would be liftable by an average person (strength 5) if they put in an all-out, 100% effort, and an extremely strong person (strength 10) should be able to lift the box with only a moderate, 50% effort.



Page Break

Contestants can win a certain amount of prize money as an incentive for lifting the box.

We will tell you the strength of each contestant and the minimum incentive they would accept to lift the heaviest box they can lift by themselves.

Your job is to **select two contestants** to lift the box together in each contest, and **decide how much incentive** you would like to provide each contestant you selected.

You will only provide incentives to the two contestants you selected.

The total incentive you can provide to the two contestants of your choice is **$50**.

Contestant A

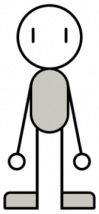
Strength: **9**

Minimum incentive to lift a box of weight 9 alone: **$40**

Contestant B

Strength: **5**

Minimum incentive to lift a box of weight 5 alone: **$20**



Contestant C

Strength: **2**

Minimum incentive to lift a box of weight 2 alone: **$2**

Page Break

For every successful lift, you will get a bonus of $1.

For every dollar of incentive you give, $0.01 will be deducted from your bonus payment, regardless of whether the result is Lift or Fail.

Suppose you provide $30 to one contestant and $10 to the other contestant. If they succeed, then your bonus payment increases by $1-(30+10)x$0.01 = $0.6. If they fail, you lose (30+10)x$0.01 = $0.4.

Your total bonus payment can go up to $3, and will also not be less than $0.

You will see whether the contestants lift the box or not with the incentive you give by the end of each contest.

Page Break

Keep in mind that some people are stronger and some are weaker.

The available contestants are always the same, but **the box changes from contest to contest**.

The contestants know very well their own strength and the other contestant's strength.

They also know how heavy the box is.

Page Break

Also remember that some contestants consider their effort more costly than others.

In other words, some contestants might require more incentive to exert a certain amount of effort.

If the contestant does not feel the prize is worth their effort, they might not lift the box even if they are strong enough to do so.