

LEVERHULME TRUST

Optimisation Uncertainty

Te Bao[1] **Jiaoying Pei**[2]

[1]Nanyang Technological University, Singapore [2]University of Cambridge, U.K.

November 2026

Problem Statement

What we do?

- We introduce optimsiation uncertainty: the difficulty of mapping a representation of the decision problem (e.g., preferences) into an action.

Problem Statement

What we do?

- We introduce optimsiation uncertainty: the difficulty of mapping a representation of the decision problem (e.g., preferences) into an action.
- Optimisation uncertainty is specific to economic experiments: b/c of incentive compatibility, we need to infer preferences from action, when preference is our target of interest.

Problem Statement

What we do?

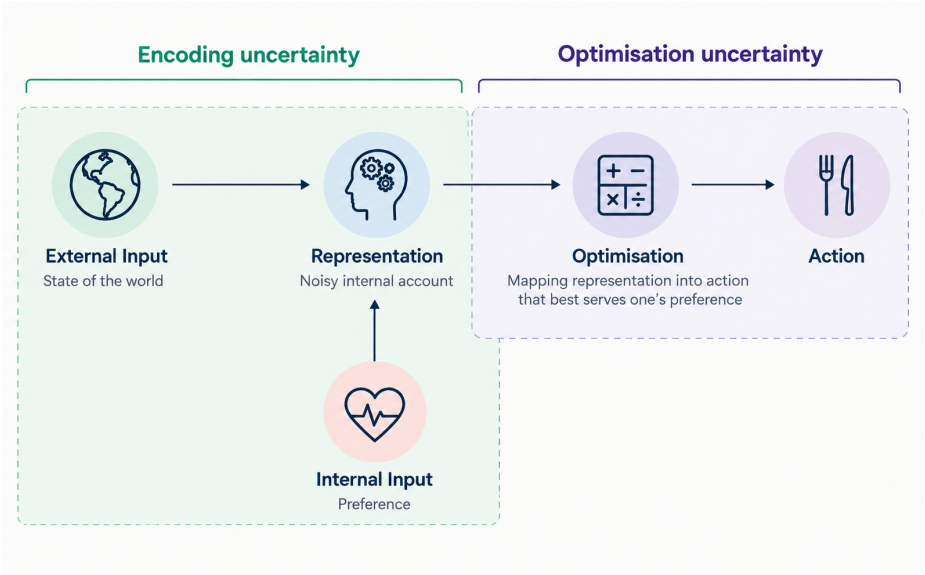
- We introduce optimsiation uncertainty: the difficulty of mapping a representation of the decision problem (e.g., preferences) into an action.
- Optimisation uncertainty is specific to economic experiments: b/c of incentive compatibility, we need to infer preferences from action, when preference is our target of interest.
- We find that confidence mainly captures optimisation uncertainty (rather than other uncertainties that researched and matters in psychological experiments) in economic strategic games.

Problem Statement

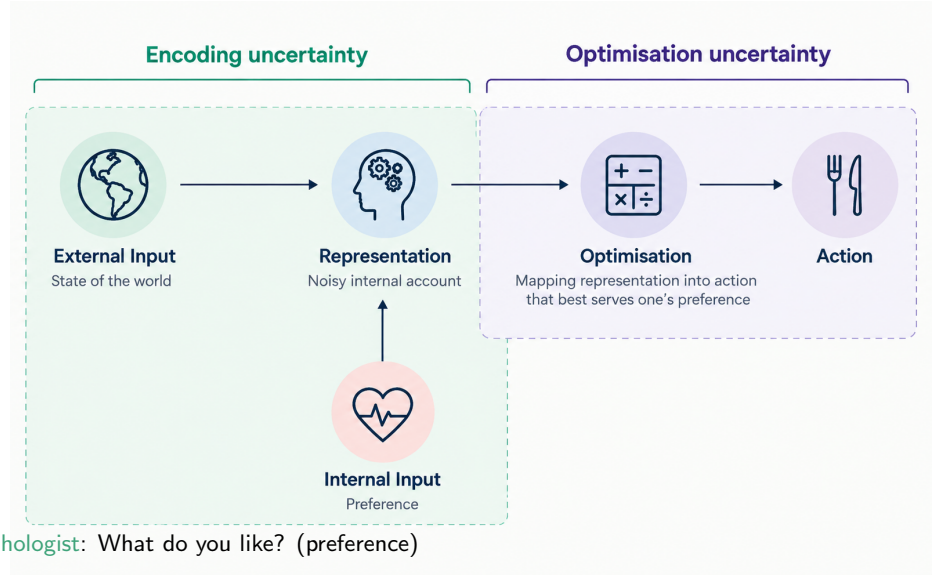
What we do?

- We introduce optimsiation uncertainty: the difficulty of mapping a representation of the decision problem (e.g., preferences) into an action.
- Optimisation uncertainty is specific to economic experiments: b/c of incentive compatibility, we need to infer preferences from action, when preference is our target of interest.
- We find that confidence mainly captures optimisation uncertainty (rather than other uncertainties that researched and matters in psychological experiments) in economic strategic games.
- Because our experiment include this additional translation stage, we as economists have more reason to elicit confidence, when it is already a standard procedure in psychology experiments.

Two stage decision process in economics experiments

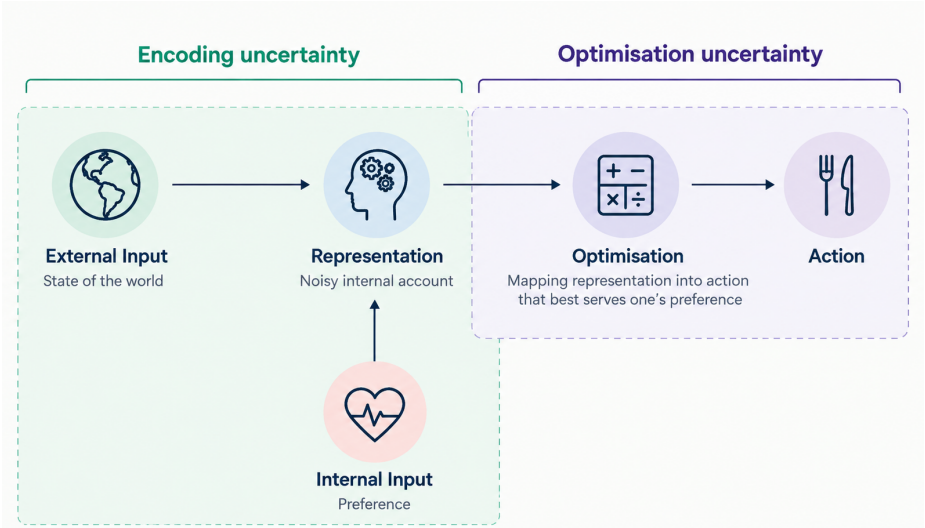


Two stage decision process in economics experiments



- Psychologist: What do you like? (preference)

Two stage decision process in economics experiments



- Psychologist: What do you like? (preference)
- Economists most of the time: What do you choose? (infer preference from action)

Illustration: From what you like to what do you choose - Symmetric matching pennies game

- **Given** the more money the happier, i.e., max YOUR income only.

Illustration: From what you like to what do you choose - Symmetric matching pennies game

- **Given** the more money the happier, i.e., max YOUR income only.
- Two players simultaneously choose a side of a coin. You win by matching, the other wins by mismatching.

Illustration: From what you like to what do you choose - Symmetric matching pennies game

- **Given** the more money the happier, i.e., max YOUR income only.
- Two players simultaneously choose a side of a coin. You win by matching, the other wins by mismatching.
- Preference $\rightarrow$ action, long reasoning alert:

Illustration: From what you like to what do you choose - Symmetric matching pennies game

- **Given** the more money the happier, i.e., max YOUR income only.
- Two players simultaneously choose a side of a coin. You win by matching, the other wins by mismatching.
- Preference $\rightarrow$ action, long reasoning alert:
- You want to do your best, while your opponent tries to guess what you do... now if you choose head for 60%, the opponent will figure it out and choose tail all the time, so that he will win 60% and you will lose 40%. You can improve your situation by decreasing the prob. you choose head, say 50%, in that case it's equal chance of winning for the opponent to choose head and tail, and you win 50%.
- Do I convince you there is difficulty in finding the strategy that optimizes your preference?

Illustration: From what you like to what do you choose - Symmetric matching pennies game

- **Given** the more money the happier, i.e., max YOUR income only.
- Two players simultaneously choose a side of a coin. You win by matching, the other wins by mismatching.
- Preference $\rightarrow$ action, long reasoning alert:
- You want to do your best, while your opponent tries to guess what you do... now if you choose head for 60%, the opponent will figure it out and choose tail all the time, so that he will win 60% and you will lose 40%. You can improve your situation by decreasing the prob. you choose head, say 50%, in that case it's equal chance of winning for the opponent to choose head and tail, and you win 50%.
- Do I convince you there is difficulty in finding the strategy that optimizes your preference?
- Or maybe you memorize it already: optimal action is mixed strategy which is to randomize your action, in this case 50% of head. If now you conclude there is no optimisation uncertainty: would you report 100% confidence?

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.
- You earn \$1 from matching on either side. Your opponent earns from \$2 from mismatching on head and \$1 from mismatching on tail.

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.
- You earn \$1 from matching on either side. Your opponent earns from \$2 from mismatching on head and \$1 from mismatching on tail.
- What's your optimal strategy? $Pr(head) = ?$

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.
- You earn \$1 from matching on either side. Your opponent earns from \$2 from mismatching on head and \$1 from mismatching on tail.
- What's your optimal strategy? $Pr(head) = ?$
- You want to do your best while the opponent tries to guess what you do. Let $p = \Pr(\text{you play H})$ and $q = \Pr(\text{opponent plays H})$. Suppose you play H with probability 50%. The opponent figures it out: playing H he earns \$2 when you play T, so $2(1-p) = 1$, while playing T he earns \$1 when you play H, so $p = 0.5$. He therefore plays H every time, and you would then switch to H, after which he switches to T, and so on: any mix that gives him a favorite can be exploited. You must therefore choose p so that his two options pay the same, $2(1-p) = p$, so $p = \frac{2}{3}$. Notice that you play heads *more* often, even though your own payoff is the same \$1 on either side: his H is tempting because it pays \$2, and a high p is what keeps it from paying too much. Likewise he must keep you indifferent: you earn q from H and $1-q$ from T, so $q = \frac{1}{2}$. His expected payoff is $2(1-\frac{2}{3}) = \frac{2}{3}$ and yours is $\frac{1}{2}$. So your optimal strategy is $\Pr(\text{head}) = \frac{2}{3}$.

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.
- You earn \$1 from matching on either side. Your opponent earns from \$2 from mismatching on head and \$1 from mismatching on tail.
- What's your optimal strategy? $Pr(head) = ?$
- You want to do your best while the opponent tries to guess what you do. Let $p = \Pr(\text{you play H})$ and $q = \Pr(\text{opponent plays H})$. Suppose you play H with probability 50%. The opponent figures it out: playing H he earns \$2 when you play T, so $2(1-p) = 1$, while playing T he earns \$1 when you play H, so $p = 0.5$. He therefore plays H every time, and you would then switch to H, after which he switches to T, and so on: any mix that gives him a favorite can be exploited. You must therefore choose p so that his two options pay the same, $2(1-p) = p$, so $p = \frac{2}{3}$. Notice that you play heads *more* often, even though your own payoff is the same \$1 on either side: his H is tempting because it pays \$2, and a high p is what keeps it from paying too much. Likewise he must keep you indifferent: you earn q from H and $1-q$ from T, so $q = \frac{1}{2}$. His expected payoff is $2(1-\frac{2}{3}) = \frac{2}{3}$ and yours is $\frac{1}{2}$. So your optimal strategy is $\Pr(\text{head}) = \frac{2}{3}$.
- Now do I convince you there is difficulty in finding the strategy that optimizes your preference?

Illustration: Asymmetric matching pennies game

- Now you are playing asymmetric matching pennies, your preference is still max. payoff.
- You earn \$1 from matching on either side. Your opponent earns from \$2 from mismatching on head and \$1 from mismatching on tail.
- What's your optimal strategy? $Pr(head) = ?$
- You want to do your best while the opponent tries to guess what you do. Let $p = \Pr(\text{you play H})$ and $q = \Pr(\text{opponent plays H})$. Suppose you play H with probability 50%. The opponent figures it out: playing H he earns \$2 when you play T, so $2(1-p) = 1$, while playing T he earns \$1 when you play H, so $p = 0.5$. He therefore plays H every time, and you would then switch to H, after which he switches to T, and so on: any mix that gives him a favorite can be exploited. You must therefore choose p so that his two options pay the same, $2(1-p) = p$, so $p = \frac{2}{3}$. Notice that you play heads *more* often, even though your own payoff is the same \$1 on either side: his H is tempting because it pays \$2, and a high p is what keeps it from paying too much. Likewise he must keep you indifferent: you earn q from H and $1-q$ from T, so $q = \frac{1}{2}$. His expected payoff is $2(1-\frac{2}{3}) = \frac{2}{3}$ and yours is $\frac{1}{2}$. So your optimal strategy is $\Pr(\text{head}) = \frac{2}{3}$.
- Now do I convince you there is difficulty in finding the strategy that optimizes your preference?
- Or maybe / Imagine you still get it quickly as a trained economist. Again: would you report 100% confidence?

Deduction A

1. If anyone in the room was unsure about their $Pr(head)$ is optimal, then:
   - $\exists$ optimisation uncertainty.

Deduction A

1. If anyone in the room was unsure about their $Pr(head)$ is optimal, then:
   - $\exists$ optimisation uncertainty.
2. If the below happens:

		Specifically:	Symmetric MP	Asymmetric MP
A	Encoding uncertainty from external input	*Unknown opponent's action*	✓	✓
B	Encoding uncertainty from internal input	*Unknown preference*	✗	✗
C	Optimisation uncertainty		✗	✓
D	Lack of confidence		✗	✓

then, in these games:

- C (optimisation uncertainty) is a sufficient and necessary condition for D (lack of confidence).
- A (unknown opponent's action) is necessary but not sufficient for D (lack of confidence); B (unknown preference) is not necessary for D (lack of confidence).

Deduction B

1. If anyone in the room was unsure about their $Pr(head)$ is optimal, then:
   - $\exists$ optimisation uncertainty.

Deduction B

1. If anyone in the room was unsure about their $Pr(head)$ is optimal, then:
   - $\exists$ optimisation uncertainty.
2. If the below happens:

		Specifically:	Symmetric MP	Asymmetric MP
A	Encoding uncertainty from external input	*Unknown opponent's action*	✓	✓
B	Encoding uncertainty from internal input	*Unknown preference*	✗	✗
C	Optimisation uncertainty		✓	✓
D	Lack of confidence		✓	✓

then, in these games:

- B (encoding uncertainty from internal input) is not a necessary condition for D (lack of confidence).
- We cannot say whether A or C is individually necessary or sufficient for D.

Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.

Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.
- Confidence, as the metacognitive judgement, are defined in psychology literature to track the *representation* stage alone:

Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.
- Confidence, as the metacognitive judgement, are defined in psychology literature to track the *representation* stage alone: I process some uncertain information X about the correct world d, based on that I made the decision a. Then I reflect on whether my decision is correct $P(a = d|\cdot)$:

$$c = P(d = a|X, a, \sigma_X^2), X \sim N(d, \sigma_X^2) \tag{1}$$

Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.
- Confidence, as the metacognitive judgement, are defined in psychology literature to track the *representation* stage alone: I process some uncertain information X about the correct world d, based on that I made the decision a. Then I reflect on whether my decision is correct $P(a = d|\cdot)$:

$$c = P(d = a|X, a, \sigma_X^2), X \sim N(d, \sigma_X^2) \tag{1}$$

- Confidence is a metacognitive judgment ("rater"), different from first-order cognition $1_{\{a=d\}}$ ("actor"); supported by distinct neural systems (Fleming and Daw, 2017).

Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.
- Confidence, as the metacognitive judgement, are defined in psychology literature to track the *representation* stage alone: I process some uncertain information X about the correct world d, based on that I made the decision a. Then I reflect on whether my decision is correct $P(a = d|\cdot)$:

$$c = P(d = a|X, a, \sigma_X^2), X \sim N(d, \sigma_X^2) \tag{1}$$

- Confidence is a metacognitive judgment ("rater"), different from first-order cognition $1_{\{a=d\}}$ ("actor"); supported by distinct neural systems (Fleming and Daw, 2017).
- Recall the matching penny game: according to (1), since you cannot know your opponents' move beforehand ($\sigma_X^2 > 0$), there is no guarantee that $P(d = a) = 1$ of winning you should never report full confidence.

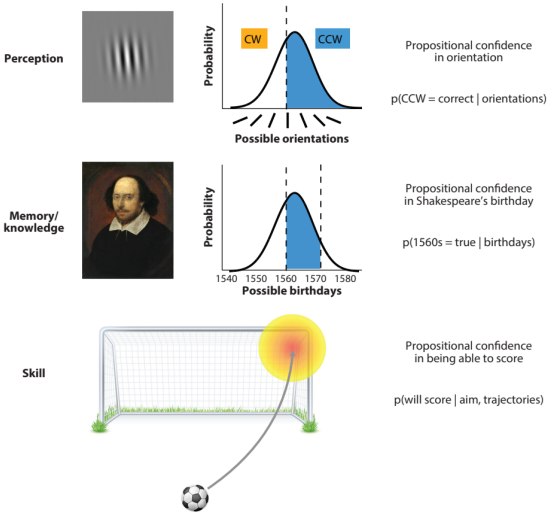
Metacognitive confidence

- In theory, however, confidence elicitation *IS NOT* able to capture optimisation uncertainty.
- Confidence, as the metacognitive judgement, are defined in psychology literature to track the *representation* stage alone: I process some uncertain information X about the correct world d, based on that I made the decision a. Then I reflect on whether my decision is correct $P(a = d|\cdot)$:

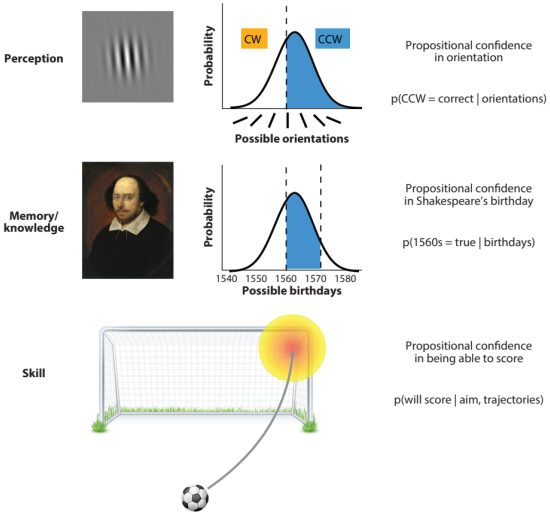
$$c = P(d = a|X, a, \sigma_X^2), X \sim N(d, \sigma_X^2) \tag{1}$$

- Confidence is a metacognitive judgment ("rater"), different from first-order cognition $1_{\{a=d\}}$ ("actor"); supported by distinct neural systems (Fleming and Daw, 2017).
- Recall the matching penny game: according to (1), since you cannot know your opponents' move beforehand ($\sigma_X^2 > 0$), there is no guarantee that $P(d = a) = 1$ of winning you should never report full confidence.
- But (some of you) do! In strategic games, full confidence may represent optimisation over what is within the control (our point), not a forecast of the result in (1)!

Metacognitive confidence: inaccuracy from forecasting the result



Metacognitive confidence: inaccuracy from forecasting the result



Low confidence can only from one place: noisy representation, here **imprecise encoding for the world** reflected as larger σ_x^2 (wider belief distribution, blue and yellow area)

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources $\rightarrow$ apparent irrationality

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources $\rightarrow$ apparent irrationality

Difficulty in encoding external input:

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources $\rightarrow$ apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes $\Rightarrow$ outlier blindness (Payzan-LeNestour and Woodford, 2022)

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources $\rightarrow$ apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes $\Rightarrow$ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information $\Rightarrow$ description-experience gap (Oprea and Vieider, 2024)

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources → apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes ⇒ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information ⇒ description-experience gap (Oprea and Vieider, 2024)
- objective *probabilities* and *payoffs* ⇒ features of prospect theory (Woodford, 2012; Frydman and Jin, 2022)

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources → apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes ⇒ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information ⇒ description-experience gap (Oprea and Vieider, 2024)
- objective *probabilities* and *payoffs* ⇒ features of prospect theory (Woodford, 2012; Frydman and Jin, 2022)

Difficulty in encoding internal input:

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources → apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes ⇒ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information ⇒ description-experience gap (Oprea and Vieider, 2024)
- objective *probabilities* and *payoffs* ⇒ features of prospect theory (Woodford, 2012; Frydman and Jin, 2022)

Difficulty in encoding internal input:

- Subjective representation (*preference*) (Polanía et al., 2019)

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources → apparent irrationality

Difficulty in encoding external input:

- *frequency* of outcomes ⇒ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information ⇒ description-experience gap (Oprea and Vieider, 2024)
- objective *probabilities* and *payoffs* ⇒ features of prospect theory (Woodford, 2012; Frydman and Jin, 2022)

Difficulty in encoding internal input:

- Subjective representation (*preference*) (Polanía et al., 2019)
- the value you feel for a food is not a number stored ; brain reconstructs an estimate from noisy internal information, exploiting the prior distribution of values you usually encounter (e.g., products typically encountered in the supermarket).

Source of imprecise encoding for the world

Efficient coding (Glimcher, 2022): brain cannot encode value with perfect precision and can represent it only noisily, due to limited brain resources → apparent irrationality

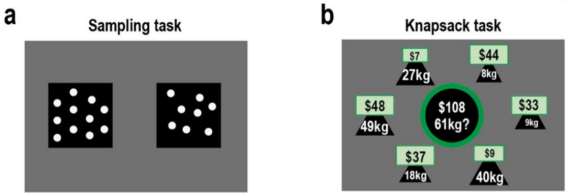
Difficulty in encoding external input:

- *frequency* of outcomes ⇒ outlier blindness (Payzan-LeNestour and Woodford, 2022)
- representing *numerical* information ⇒ description-experience gap (Oprea and Vieider, 2024)
- objective *probabilities* and *payoffs* ⇒ features of prospect theory (Woodford, 2012; Frydman and Jin, 2022)

Difficulty in encoding internal input:

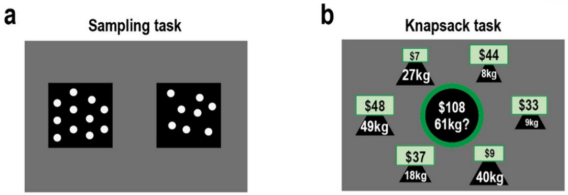
- Subjective representation (*preference*) (Polanía et al., 2019)
- the value you feel for a food is not a number stored ; brain reconstructs an estimate from noisy internal information, exploiting the prior distribution of values you usually encounter (e.g., products typically encountered in the supermarket).
- ↑ degree of noise in the internal representations used for inference when: ↓ exposure time & ↓ # samples can be drawn

Confidence in the literature



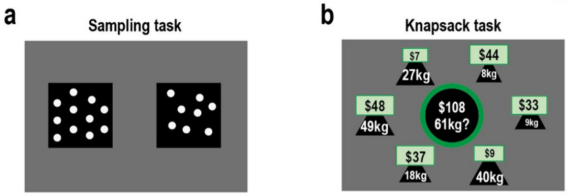
- (External input) Confidence correlates *positively* with performance in perceptual sampling task in (a) (Lu et al., 2025); low confidence predicts higher information searching (Schulz et al., 2023).

Confidence in the literature



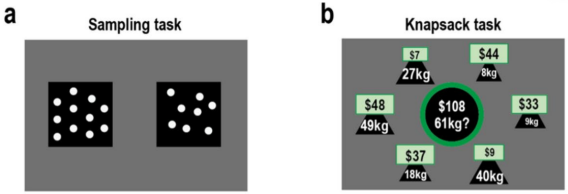
- (External input) Confidence correlates *positively* with performance in perceptual sampling task in (a) (Lu et al., 2025); low confidence predicts higher information searching (Schulz et al., 2023).
  - Meta-cognition accuracy as the moderator: sensitivity: Prob (I think I am right (wrong) when I am actually right (wrong); bias. (Fleming, 2024)

Confidence in the literature



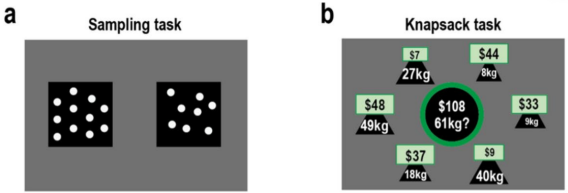
- (External input) Confidence correlates *positively* with performance in perceptual sampling task in (a) (Lu et al., 2025); low confidence predicts higher information searching (Schulz et al., 2023).
  - Meta-cognition accuracy as the moderator: sensitivity: Prob (I think I am right (wrong) when I am actually right (wrong); bias. (Fleming, 2024)
- (Internal input) Confidence correlates *negatively* with choice consistency in preference tasks (Enke and Graeber, 2023; Polanía et al., 2019);

Confidence in the literature



- (External input) Confidence correlates *positively* with performance in perceptual sampling task in (a) (Lu et al., 2025); low confidence predicts higher information searching (Schulz et al., 2023).
  - Meta-cognition accuracy as the moderator: sensitivity: Prob (I think I am right (wrong) when I am actually right (wrong); bias. (Fleming, 2024)
- (Internal input) Confidence correlates *negatively* with choice consistency in preference tasks (Enke and Graeber, 2023; Polanía et al., 2019);
- (Optimisation in NP hard problems) Confidence correlates with performance in Knapsack task in (b). (Lu et al., 2025).

Confidence in the literature



- (External input) Confidence correlates *positively* with performance in perceptual sampling task in (a) (Lu et al., 2025); low confidence predicts higher information searching (Schulz et al., 2023).
  - Meta-cognition accuracy as the moderator: sensitivity: Prob (I think I am right (wrong) when I am actually right (wrong); bias. (Fleming, 2024)
- (Internal input) Confidence correlates *negatively* with choice consistency in preference tasks (Enke and Graeber, 2023; Polanía et al., 2019);
- (Optimisation in NP hard problems) Confidence correlates with performance in Knapsack task in (b). (Lu et al., 2025).
  - These are mathematically intractable vs. optimisation problems here.

This paper

What can confidence capture in economics experiment? Does it also capture optimisation uncertainty?

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief
3. Does confidence capture encoding noise from external input?

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief
3. Does confidence capture encoding noise from external input?
   - Within-subject observe encoding from internal input. $\Rightarrow$ Does confidence increase when repetitively sample from the preference?

Experimental Design

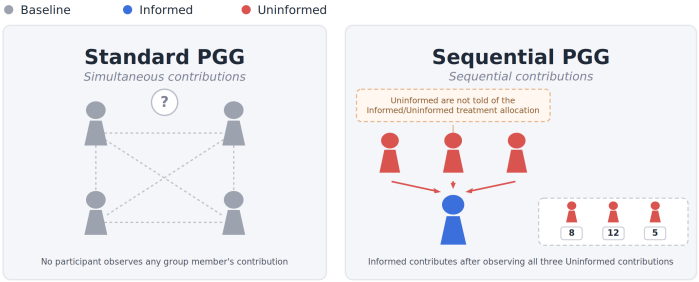
1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief
3. Does confidence capture encoding noise from external input?
   - Within-subject observe encoding from internal input. $\Rightarrow$ Does confidence increase when repetitively sample from the preference?
   - Assumption: preference is not entirely random. Or ...

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief
3. Does confidence capture encoding noise from external input?
   - Within-subject observe encoding from internal input. $\Rightarrow$ Does confidence increase when repetitively sample from the preference?
   - Assumption: preference is not entirely random. Or ...
   - Pooling independent noisy signals $\rightarrow$ precision rise with the number of draws $\rightarrow$ confidence tracks precision should rise

Experimental Design

1. Find an economics experiment involves encoding from 2 sources + the need for optimisation.
   - Public goods game: a social dilemma benefits the group at their own cost.
   - Public goods game (Fischbacher et al., 2001): elicit social preference (*encoding from internal input*), that is affected by others' actions: (*encoding from external input*, strategic uncertainty in e.g., Nagel (1995) & payoff uncertainties); mapping the preference into a contribution (*optimisation*)
2. Does confidence capture encoding noise from external input?
   - Between-subject; remove encoding from external input: when they know exactly how others will do & their payoff under each level of contribution*, before making their own move $\Rightarrow$ Does confidence $\uparrow$?
   - Vs. baseline, re-matching in 44-participant cohort $\rightarrow$ no reliable way to form a precise belief
3. Does confidence capture encoding noise from external input?
   - Within-subject observe encoding from internal input. $\Rightarrow$ Does confidence increase when repetitively sample from the preference?
   - Assumption: preference is not entirely random. Or ...
   - Pooling independent noisy signals $\rightarrow$ precision rise with the number of draws $\rightarrow$ confidence tracks precision should rise
4. Close 2 & 3 leaves optimisation uncertainty as the explanation for lower-than-maximum confidence (observed).



"How certain she is that she would want to contribute the amount she chose", rather than "how certain that this amount maximises her payoff"

Treatment	# Participants	Experiment	Date	# Sessions	Color	Avg Payoff (Points)
Baseline	140	Standard PGG	2023-06-22	3	Gray	22.78
Informed	39	Sequential PGG	2023-10-18, 2025-04-18	4	Blue	24.97
Uninformed	117	Sequential PGG	2023-10-18, 2025-04-18	4	Red	22.99

Results

1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).

* no difference between experiment and treatments.

Results

1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).
2. **Confidence shortfall does not reflect encoding uncertainty arising from external inputs of strategic and payoff uncertainties**: implied by 1.

* no difference between experiment and treatments.

Results

1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).
2. **Confidence shortfall does not reflect encoding uncertainty arising from external inputs of strategic and payoff uncertainties**: implied by 1.
    - Manipulation check where Informed use the additional information: they contributes differently (less); condition their contribution on current group members' contributions rather than previous period's average.

* no difference between experiment and treatments.

Results

1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).
2. **Confidence shortfall does not reflect encoding uncertainty arising from external inputs of strategic and payoff uncertainties**: implied by 1.
   - Manipulation check where Informed use the additional information: they contributes differently (less); condition their contribution on current group members' contributions rather than previous period's average.
3. ... **does not capture encoding uncertainty from internal input (imprecision from preference).**: confidence does not increase through repetition*.

* no difference between experiment and treatments.

Results

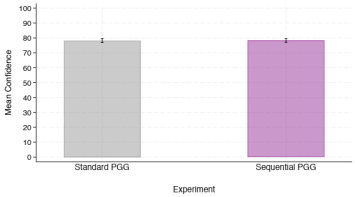
1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).
2. **Confidence shortfall does not reflect encoding uncertainty arising from external inputs of strategic and payoff uncertainties**: implied by 1.
    - Manipulation check where Informed use the additional information: they contributes differently (less); condition their contribution on current group members' contributions rather than previous period's average.
3. ... **does not capture encoding uncertainty from internal input (imprecision from preference).**: confidence does not increase through repetition*.
    - 95% CI of 10 periods' total change in confidence is $[-2.8, 2.6]$, which is negligible against 22% shortfall, robust in a regression with controls and allow per-period change to differ across periods & participants.

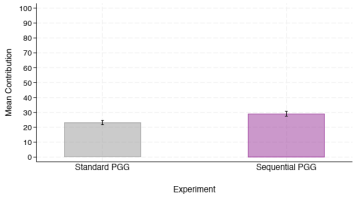
* no difference between experiment and treatments.

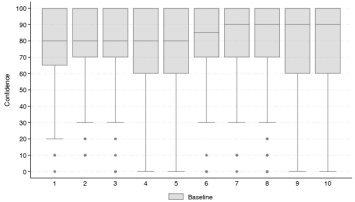
Results

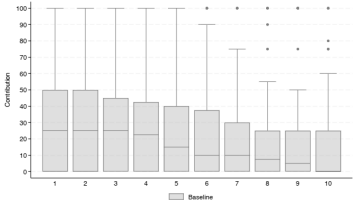
1. **Confidence falls short of its maximum**: Mean 22% below its maximum*; Not due to endpoint avoidance (full confidence is the modal response accounting for 37% frequency, 27% if remove participants who report full confidence throughout, vs. 9% under uniformity).
2. **Confidence shortfall does not reflect encoding uncertainty arising from external inputs of strategic and payoff uncertainties**: implied by 1.
    - Manipulation check where Informed use the additional information: they contributes differently (less); condition their contribution on current group members' contributions rather than previous period's average.
3. ... **does not capture encoding uncertainty from internal input (imprecision from preference).**: confidence does not increase through repetition*.
    - 95% CI of 10 periods' total change in confidence is $[-2.8, 2.6]$, which is negligible against 22% shortfall, robust in a regression with controls and allow per-period change to differ across periods & participants.
    - They do not treat PGG as a bandit problem (or confidence should $\uparrow$, contribution should converge to 0).

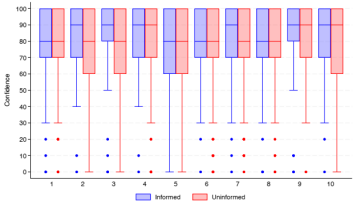
* no difference between experiment and treatments.

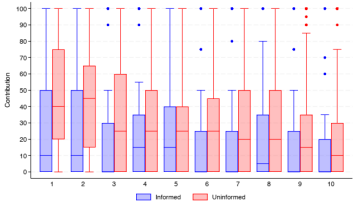

Mean Confidence
100
90
80
70
60
50
40
30
20
10
0
Standard PGG
Sequential PGG
Experiment


Mean Contribution
100
90
80
70
60
50
40
30
20
10
0
Standard PGG
Sequential PGG
Experiment


Confidence
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Baseline


Contribution
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Baseline


Confidence
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Informed
Uninformed


Contribution
100
90
80
70
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
Informed
Uninformed

Results

Confidence in strategic setting point to optimisation uncertainty.

Results

Confidence in strategic setting point to optimisation uncertainty.

- Rule out encoding channel in (2) and (3).

Results

Confidence in strategic setting point to optimisation uncertainty.

- Rule out encoding channel in (2) and (3).

Results

Confidence in strategic setting point to optimisation uncertainty.

- Rule out encoding channel in (2) and (3).
- Estimated contribution continue to vary substantially ($\pm 21.2\%$) at zero reported confidence, when others' contribution does not change.

Results

Confidence in strategic setting point to optimisation uncertainty.

- Rule out encoding channel in (2) and (3).
- Estimated contribution continue to vary substantially ($\pm 21.2\%$) at zero reported confidence, when others' contribution does not change.
- Unincentivized confidence are informative: Low confidence predicts subsequent larger revision in contributions; analogous to what the literature suggests in perception tasks where low confidence $\rightarrow$ costly information searching (Schulz et al., 2023).

Part II: How often do we need to elicit confidence?

We want to screen out the reliable choices, but at some costs of cognitive effort. Is there any shortcut?

Part II: How often do we need to elicit confidence?

We want to screen out the reliable choices, but at some costs of cognitive effort. Is there any shortcut?

1. At the beginning or the end: not in the case of PGG.

Part II: How often do we need to elicit confidence?

We want to screen out the reliable choices, but at some costs of cognitive effort. Is there any shortcut?

1. At the beginning or the end: not in the case of PGG.
2. Minimal elicitation: low confidence compresses decision; cognitive foundation of decision making literature (Enke and Graeber, 2023).

Part II: How often do we need to elicit confidence?

We want to screen out the reliable choices, but at some costs of cognitive effort. Is there any shortcut?

1. At the beginning or the end: not in the case of PGG.
2. Minimal elicitation: low confidence compresses decision; cognitive foundation of decision making literature (Enke and Graeber, 2023).
3. Uniform Distortion

Part II: How often do we need to elicit confidence?

We want to screen out the reliable choices, but at some costs of cognitive effort. Is there any shortcut?

1. At the beginning or the end: not in the case of PGG.
2. Minimal elicitation: low confidence compresses decision; cognitive foundation of decision making literature (Enke and Graeber, 2023).
3. Uniform Distortion
4. Choice-level elicitation

Theory: cognitive foundation of decision-making

- We observe choice (contribution, $p_{i,t}$) from experiment to infer preference (social preference *, $p^*_{i,t}$), given the issue-relevant parameter $(\cdot)$.

Theory: cognitive foundation of decision-making

- We observe choice (contribution, $p_{i,t}$) from experiment to infer preference (social preference *, $p^*_{i,t}$), given the issue-relevant parameter $(\cdot)$.
- The cognitive foundations programme (Enke and Graeber, 2023) proposes a "dual process" model that $p^*_{i,t}$ is a weighted function between $p^*_{i,t}$ and a default (seemly "safer" strategy) $p^o_{i,t}$

$$p_{i,t}(\cdot) = \frac{c_{i,t}}{100} p^*_{i,t}(\cdot) + (1 - \frac{c_{i,t}}{100}) p^o_{i,t}(\cdot)$$

so that

$$p_{i,t}(\cdot, c_{i,t} = 0) = p^o_{i,t}(\cdot); p_{i,t}(\cdot, c_{i,t} = 100) = p^*_{i,t}(\cdot)$$

Theory: cognitive foundation of decision-making

- We observe choice (contribution, $p_{i,t}$) from experiment to infer preference (social preference *, $p^*_{i,t}$), given the issue-relevant parameter $(\cdot)$.
- The cognitive foundations programme (Enke and Graeber, 2023) proposes a "dual process" model that $p^*_{i,t}$ is a weighted function between $p^*_{i,t}$ and a default (seemly "safer" strategy) $p^o_{i,t}$

$$p_{i,t}(\cdot) = \frac{c_{i,t}}{100} p^*_{i,t}(\cdot) + (1 - \frac{c_{i,t}}{100}) p^o_{i,t}(\cdot)$$

so that

$$p_{i,t}(\cdot, c_{i,t} = 0) = p^o_{i,t}(\cdot); p_{i,t}(\cdot, c_{i,t} = 100) = p^*_{i,t}(\cdot)$$

- The weighting function resembles the elaboration likelihood model (ELM) in the persuasion literature (Petty and Cacioppo, 1986): the extent of elaboration on the information (high = systematic thinking over issue-relevant material, low = employs some simple decision rule/heuristics) lead to different outcomes.

Theory: cognitive foundation of decision-making

- We observe choice (contribution, $p_{i,t}$) from experiment to infer preference (social preference *, $p^*_{i,t}$), given the issue-relevant parameter $(\cdot)$.
- The cognitive foundations programme (Enke and Graeber, 2023) proposes a "dual process" model that $p^*_{i,t}$ is a weighted function between $p^*_{i,t}$ and a default (seemly "safer" strategy) $p^o_{i,t}$

$$p_{i,t}(\cdot) = \frac{c_{i,t}}{100} p^*_{i,t}(\cdot) + (1 - \frac{c_{i,t}}{100}) p^o_{i,t}(\cdot)$$

so that

$$p_{i,t}(\cdot, c_{i,t} = 0) = p^o_{i,t}(\cdot); p_{i,t}(\cdot, c_{i,t} = 100) = p^*_{i,t}(\cdot)$$

- The weighting function resembles the elaboration likelihood model (ELM) in the persuasion literature (Petty and Cacioppo, 1986): the extent of elaboration on the information (high = systematic thinking over issue-relevant material, low = employs some simple decision rule/heuristics) lead to different outcomes.
- Confidence is in place of elaboration here. The weighting function has been tested to be valid in investment decision in (Charles et al., 2024).

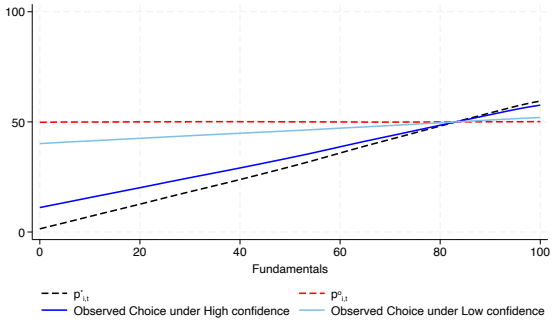
- The Central argument in the cognitive foundation literature is that low confidence decision are compressed towards fundamentals $(\cdot)$, which is essentially assuming $\frac{\partial p^{o}_{i,t}}{\partial(\cdot)} < \frac{\partial p^{*}_{i,t}}{\partial(\cdot)}$.

- The Central argument in the cognitive foundation literature is that low confidence decision are compressed towards fundamentals $(\cdot)$, which is essentially assuming $\frac{\partial p^o_{i,t}}{\partial(\cdot)} < \frac{\partial p^*_{i,t}}{\partial(\cdot)}$.
- The cognitive foundation literature estimates it as a single population-level default, of central tendency:

$$p^o_{i,t}(\cdot) = 50, \frac{\partial p^o_{i,t}}{\partial(\cdot)} = 0 < \frac{\partial p^*_{i,t}}{\partial(\cdot)}$$

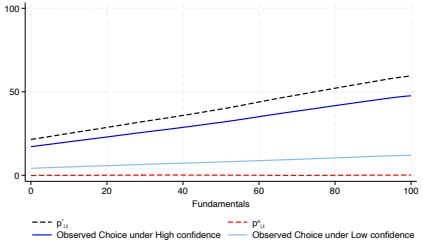
- The Central argument in the cognitive foundation literature is that low confidence decision are compressed towards fundamentals $(\cdot)$, which is essentially assuming $\frac{\partial p^o_{i,t}}{\partial(\cdot)} < \frac{\partial p^*_{i,t}}{\partial(\cdot)}$.
- The cognitive foundation literature estimates it as a single population-level default, of central tendency:

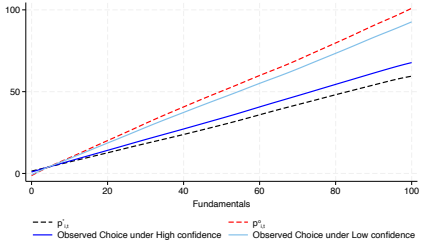
$$p^o_{i,t}(\cdot) = 50, \frac{\partial p^o_{i,t}}{\partial(\cdot)} = 0 < \frac{\partial p^*_{i,t}}{\partial(\cdot)}$$



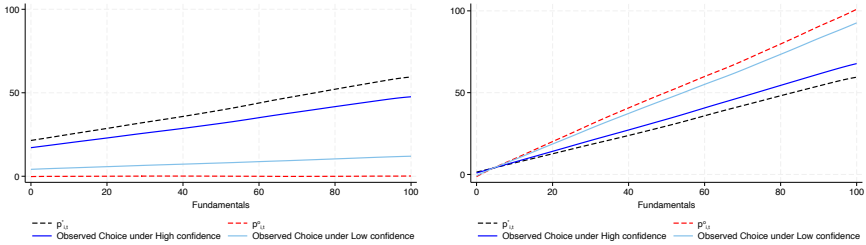
- Attempt to measure default at individual level directly finds that variation is mostly explained by subject level FE (Charles et al., 2024).

- Attempt to measure default at individual level directly finds that variation is mostly explained by subject level FE (Charles et al., 2024).





- Attempt to measure default at individual level directly finds that variation is mostly explained by subject level FE (Charles et al., 2024).



- Psychologists like Krosnick (1991) has proposed that non-differentiation is *only one* of the strategies without substantial cognitive effort; it can be anything, e.g., what if

$$\frac{\partial p^o_{i,t}}{\partial(\cdot)} > \frac{\partial p^*_{i,t}}{\partial(\cdot)}$$

- Then low confidence decisions NOT compressed towards to fundamentals!

Estimation

- We take the weighting function as given, and estimate $(\frac{\partial p^*_{i,t}}{\partial(\cdot)} - \frac{\partial p^o_{i,t}}{\partial(\cdot)})$.

Estimation

- We take the weighting function as given, and estimate $(\frac{\partial p^*_{i,t}}{\partial(\cdot)} - \frac{\partial p^o_{i,t}}{\partial(\cdot)})$.
- In PGG, $(\cdot)$ is the tendency for conditional cooperation, i.e., others' contribution $\bar{p}^j$; we thus estimate, for each participant:

$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}.$$

where

$$\begin{cases} p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0 \\ p^o_{i,t} = \phi^i_1 \bar{p}^j + \phi^i_0, \end{cases} \tag{2}$$

Estimation

- We take the weighting function as given, and estimate $(\frac{\partial p^*_{i,t}}{\partial(\cdot)} - \frac{\partial p^o_{i,t}}{\partial(\cdot)})$.
- In PGG, $(\cdot)$ is the tendency for conditional cooperation, i.e., others' contribution $\bar{p}^j$; we thus estimate, for each participant:

$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}.$$

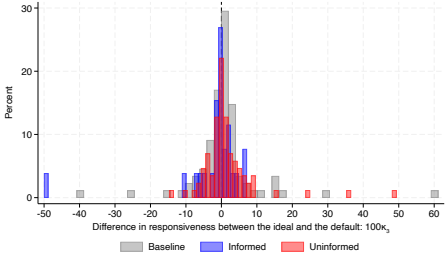
where

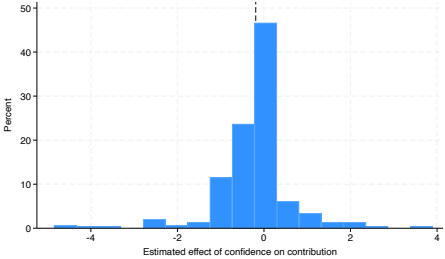
$$\begin{cases} p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0 \\ p^o_{i,t} = \phi^i_1 \bar{p}^j + \phi^i_0, \end{cases} \tag{2}$$

- By substitution, $(\frac{\partial p^*_{i,t}}{\partial \bar{p}^j} - \frac{\partial p^o_{i,t}}{\partial \bar{p}^j}) > 0$ if $\hat{\kappa}_3 > 0$ from the following:

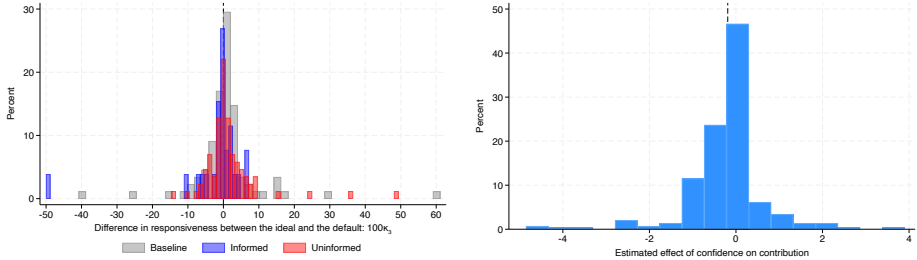
$$p^i_t = \kappa^i_0 + \kappa^i_1 c^i_t + \kappa^i_2 \bar{p}^j + \kappa^i_3 \left(c^i_t \times \bar{p}^j\right) + \varepsilon^i_t \tag{3}$$

Results


Percent
30
20
10
0
-50
-40
-30
-20
-10
0
10
20
30
40
50
60
Difference in responsiveness between the ideal and the default: 100κ_3
Baseline
Informed
Uninformed

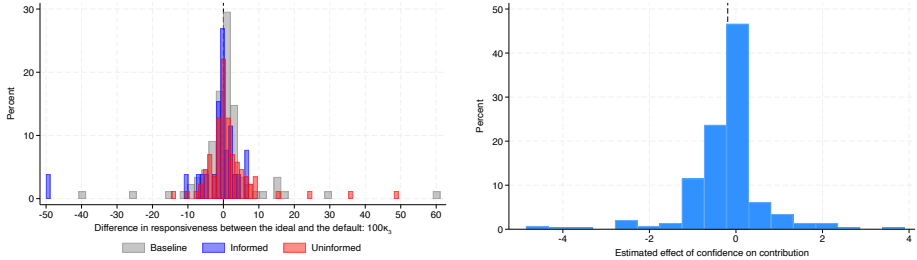

Percent
50
40
30
20
10
0
-4
-2
0
2
4
Estimated effect of confidence on contribution

Results



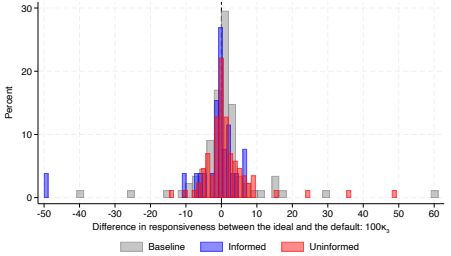
- Left: The responsiveness is the same ($\frac{\partial p_{i,t}^{*}}{\bar{p}^{j}} = \frac{\partial p_{i,t}^{o}}{\bar{p}^{j}}$), implying that default strategy does not have to be a constant. Confirmed by the results from the same analyses in (Enke and Graeber, 2023) assuming single-population-level default from pooled regression. coef. $= 0.00[0.530]$

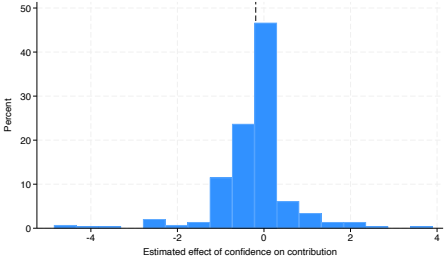
Results



- Left: The responsiveness is the same ($\frac{\partial p_{i,t}^{*}}{\bar{p}^{j}} = \frac{\partial p_{i,t}^{o}}{\bar{p}^{j}}$), implying that default strategy does not have to be a constant. Confirmed by the results from the same analyses in (Enke and Graeber, 2023) assuming single-population-level default from pooled regression. coef. $= 0.00[0.530]$
- Right: Similarly, contributions are not uniformly distorted (upwards / downwards) by low confidence ($\beta_1^i = -0.19(p = 0.5)$ in $p_t^i = \mu^i + \beta_1^i c_t^i + \beta_4 \bar{p}^j + \varepsilon_t^i$).

Results





- Left: The responsiveness is the same ($\frac{\partial p_{i,t}^{*}}{\bar{p}^{j}} = \frac{\partial p_{i,t}^{o}}{\bar{p}^{j}}$), implying that default strategy does not have to be a constant. Confirmed by the results from the same analyses in (Enke and Graeber, 2023) assuming single-population-level default from pooled regression. coef. $= 0.00[0.530]$
- Right: Similarly, contributions are not uniformly distorted (upwards / downwards) by low confidence ($\beta_1^i = -0.19(p = 0.5)$ in $p_t^i = \mu^i + \beta_1^i c_t^i + \beta_4 \bar{p}^j + \varepsilon_t^i$).
- No shortcut exists. We have to elicit confidence on choice-level.

Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.

Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.
- Recall:

$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}, p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0.$$

Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.
- Recall:

$$p_{i,t} = \frac{c_{i,t}}{100} p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}, p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0.$$

- When $c = 100$ (for confidence captures optimisation uncertainty): closest to the removing "confusion" in the literature, defined as "a failure to find the payoff/utility-maximizing strategy" (Andreoni, 1995), e.g., mistakenly believe that matching others lead to higher payoff (Burton-Chellew et al., 2016);

Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.
- Recall:

$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}, p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0.$$

- When $c = 100$ (for confidence captures optimisation uncertainty): closest to the removing "confusion" in the literature, defined as "a failure to find the payoff/utility-maximizing strategy" (Andreoni, 1995), e.g., mistakenly believe that matching others lead to higher payoff (Burton-Chellew et al., 2016);
- and when $\bar{p}^j = 0$, remove confusion-independing reading: imprecise belief about others', norm following disposition, reciprocal preference $\rightarrow$ zero contribution,

Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.
- Recall:

$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}, p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0.$$

- When $c = 100$ (for confidence captures optimisation uncertainty): closest to the removing "confusion" in the literature, defined as "a failure to find the payoff/utility-maximizing strategy" (Andreoni, 1995), e.g., mistakenly believe that matching others lead to higher payoff (Burton-Chellew et al., 2016);
- and when $\bar{p}^j = 0$, remove confusion-independing reading: imprecise belief about others', norm following disposition, reciprocal preference $\rightarrow$ zero contribution,
- the remaining contribution (ϕ^i_0) is our best attempt to resemble kindness-versus-confusion debate has been sought

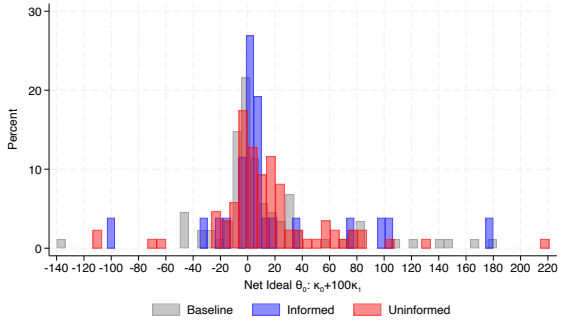
Part III: Correcting Decisions

- Here, we demonstrate the use of elicit choice-level confidence.
- Recall:

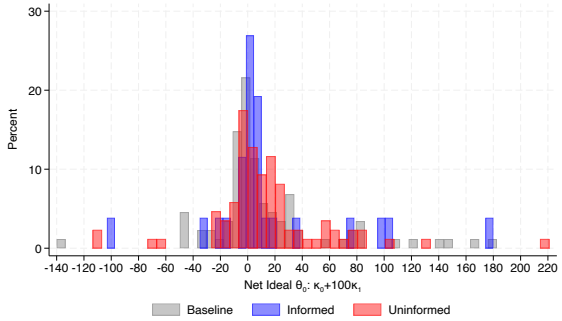
$$p_{i,t} = \frac{c_{i,t}}{100}\, p^*_{i,t} + \left(1 - \frac{c_{i,t}}{100}\right) p^o_{i,t}, p^*_{i,t} = \theta^i_1 \bar{p}^j + \theta^i_0.$$

- When $c = 100$ (for confidence captures optimisation uncertainty): closest to the removing "confusion" in the literature, defined as "a failure to find the payoff/utility-maximizing strategy" (Andreoni, 1995), e.g., mistakenly believe that matching others lead to higher payoff (Burton-Chellew et al., 2016);
- and when $\bar{p}^j = 0$, remove confusion-independing reading: imprecise belief about others', norm following disposition, reciprocal preference $\rightarrow$ zero contribution,
- the remaining contribution (ϕ^i_0) is our best attempt to resemble kindness-versus-confusion debate has been sought
- Ours contrasts with the either-or approach to settle "confusion or kindness" in the literature (Andreoni, 1995; Burton-Chellew and West, 2013; Houser and Kurzban, 2002; Wang et al., 2024)

Results

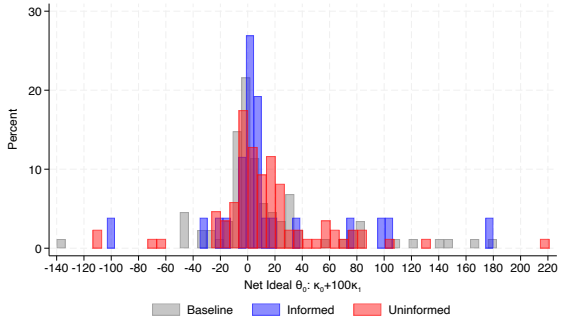

Percent
30
20
10
0
-140 -120 -100 -80 -60 -40 -20 0 20 40 60 80 100 120 140 160 180 200 220
Net Ideal θ_0: κ_0+100κ_1
Baseline
Informed
Uninformed

Results



- The net ideal (ϕ_0^i) is positive: positive for 61% of the participants at mean of 14.4 %(signed rank $p < 0.001$, t test $p < 0.001$) $\rightarrow$ Positive contributions in our data are **not** made out of confusion *alone*. Align with contribution does not converge to zero and confidence does not rise over the game.

Results



- The net ideal (ϕ_0^i) is positive: positive for 61% of the participants at mean of 14.4 %(signed rank $p < 0.001$, t test $p < 0.001$) $\rightarrow$ Positive contributions in our data are **not** made out of confusion *alone*. Align with contribution does not converge to zero and confidence does not rise over the game.
- Robustness check: net ideal does not vary across treatments. $\rightarrow$ What is expected to see when (θ_0^i) measures preference, as random assignment should leave the three groups with same preference to begin with; else, this estimation would be capturing something about information environment.

Conclusion

- We show that the 22-point shortfall from maximal confidence is not explained by encoding uncertainties as psychologists suggest, but optimisation uncertainty.
- This gives us more reasons to elicit choice-level confidence than psychologists, especially when we do not find any shortcut.
- We elaborate the use of the choice-level of confidence: the recovered contributions that is free from "confusion" and reciprocity concern is still positive for 61% participants and stable across treatments that vary information, implies that positive contributions in PGG are not made out of confusion alone.



Thank you !

Jiaoying Pei: jp2127@cam.ac.uk

References

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *The American Economic Review*, 891–904.

Burton-Chellew, M. N., C. El Mouden, and S. A. West (2016). Conditional cooperation and confusion in public-goods experiments. *Proceedings of the National Academy of Sciences 113*(5), 1291–1296.

Burton-Chellew, M. N. and S. A. West (2013). Prosocial preferences do not explain human cooperation in public-goods games. *Proceedings of the National Academy of Sciences 110*(1), 216–221.

Charles, C., C. Frydman, and M. Kilic (2024). Insensitive investors. *The Journal of Finance 79*(4), 2473–2503.

Enke, B. and T. Graeber (2023). Cognitive uncertainty. *The Quarterly Journal of Economics 138*(4), 2021–2067.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? evidence from a public goods experiment. *Economics letters 71*(3), 397–404.

Fleming, S. M. (2024). Metacognition and confidence: A review and synthesis. *Annual Review of Psychology 75*(1), 241–268.

Fleming, S. M. and N. D. Daw (2017). Self-evaluation of decision-making: A general bayesian framework for metacognitive computation. *Psychological review 124*(1), 91.

Frydman, C. and L. J. Jin (2022). Efficient coding and risky choice. *The Quarterly Journal of Economics 137*(1), 161–213.

Glimcher, P. W. (2022). Efficiently irrational: deciphering the riddle of human choice. *Trends in cognitive sciences 26*(8), 669–687.

Houser, D. and R. Kurzban (2002). Revisiting kindness and confusion in public goods experiments. *American Economic Review 92*(4), 1062–1069.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied cognitive psychology 5*(3), 213–236.

Lu, X., C. Murawski, P. Bossaerts, and S. Suzuki (2025). Estimating self-performance when making complex decisions. *Scientific Reports 15*(1), 3203.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American economic review 85*(5), 1313–1326.

Oprea, R. and F. M. Vieider (2024). Minding the gap: On the origins of probability weighting and the description-experience gap. Technical report, working paper, UC Santa Barbara.

Payzan-LeNestour, E. and M. Woodford (2022). Outlier blindness: A neurobiological foundation for neglect of financial risk. *Journal of Financial Economics 143*(3), 1316–1343.

Petty, R. E. and J. T. Cacioppo (1986). The elaboration likelihood model of persuasion. In *Advances in experimental social psychology*, Volume 19, pp. 123–205. Elsevier.

Polanía, R., M. Woodford, and C. C. Ruff (2019). Efficient coding of subjective value. *Nature neuroscience 22*(1), 134–142.

Schulz, L., S. M. Fleming, and P. Dayan (2023). Metacognitive computations for information search: Confidence in control. *Psychological Review 130*(3), 604.

Wang, G., J. Li, W. Wang, X. Niu, and Y. Wang (2024). Confusion cannot explain cooperative behavior in public goods games. *Proceedings of the National Academy of Sciences 121*(10), e2310109121.

Woodford, M. (2012). Inattentive valuation and reference-dependent choice.