How does LLM use affect decision-making?

September 2025

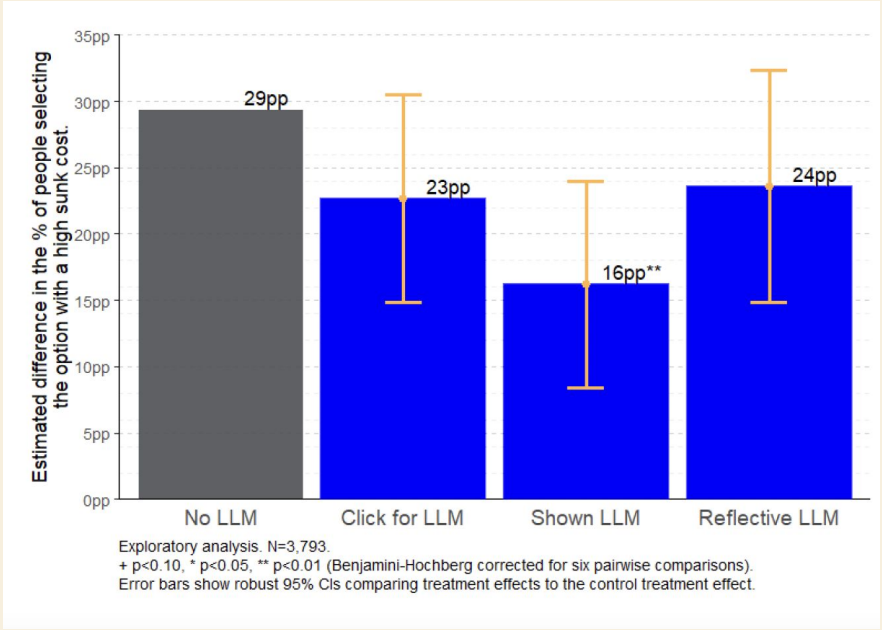
BIT | PREDICTIV

Executive Summary

1. In August 2025, we recruited 3,793 adults from the UK and US to our online platform Predictiv. We presented them with a sequence of four scenarios that were created to test four well-evidenced cognitive biases: the decoy effect, anchoring effects, sunk costs, and outcome bias.

2. We randomised participants into four groups: "No LLM" did not see any AI support; "Click for LLM" had to click to access AI advice on the scenarios; "Shown LLM" were shown AI advice by default; "Reflective LLM" were shown AI advice that encouraged them to reflect on their decisions.

3. The results reveal that **AI *can* debias our decisions - but its impact depends on the design of the AI and the nature of the bias.** AI can "slow" down intuitive yet flawed decisions; yet it may also provide a specious rationale for an unsound choice.



Example result: the Shown LLM intervention significantly reduced the effect of the sunk cost bias in a scenario involving a meeting room choice.

BIT

Background and methodology

We recruited a sample of 3,793* adults from the UK & US

BIT recruited an online representative sample of 3,793* UK and US adults between the 4th and 14th of August 2025.

Country	
UK	45.6%
USA	54.4%

Prior LLM use	
No previous use of LLM	54.8%
Prior use of LLM	45.2%

Age	
18-34	27.4%
35-54	33.9%
55+	38.7%

Ethnicity	
White	72.3%
BME	27.7%

Urban	
Urban	34.0%

Gender	
Female	54.8%

Education	
Degree or higher	37.9%

Employment	
Employed	66.0%

Income	
Below median	59.4%

NOTE ON INTERPRETING RESULTS

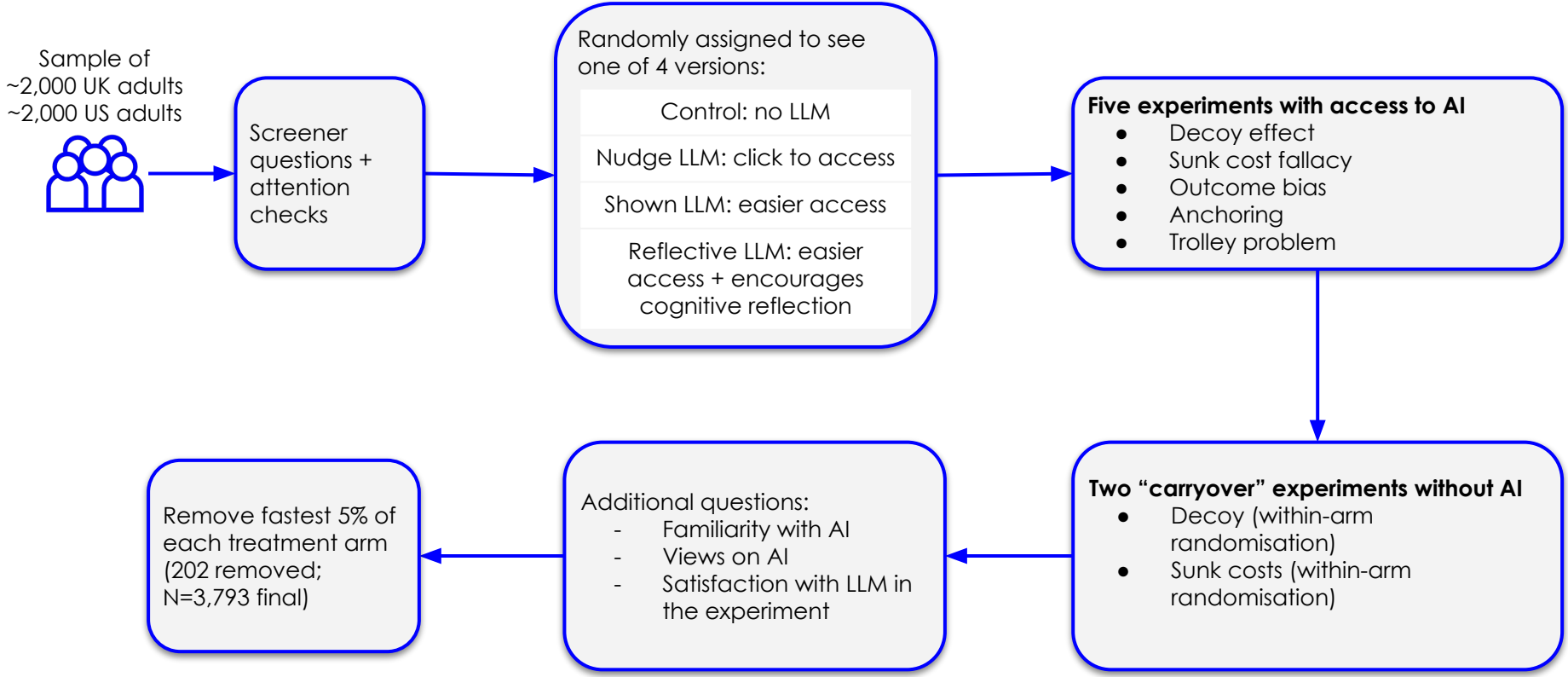
1. The sample doesn't capture the digitally excluded, or people not inclined to complete online surveys.
2. Just because people say they would do something in an online experiment, this doesn't mean they always will in real life. We therefore interpret stated intent as a likely upper bound of real behaviour.
3. When we examine differences by subgroups (e.g. gender, ethnicity), we only do so when the sample size remains large enough to draw robust inferences from.

Median time spent completing survey: 7m 31s.

* BIT removed the fastest 5% of answers (n=202) from each treatment arm.

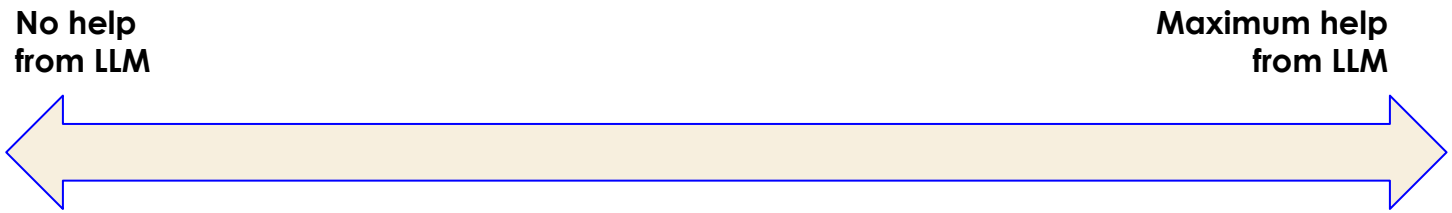
BIT

We designed decision-making scenarios and randomised access to a large language model, which we called "Pip"



BIT

Three treatment arms varied the ease of access to the LLM, while the fourth varied the LLM's responses



Treatment	Questions shown	LLM made salient	First LLM answer displayed by default	LLM pre-prompted by use to be helpful in debiasing
No LLM	■			
Nudge LLM	■	■		
C	■	■	■	
D	■	■	■	■

*All three LLM arms rely on Gemini Flash, accessed through an API.

BIT

Three treatment arms varied the ease of access to the LLM, while the fourth varied the LLM's responses

Control: No LLM	Nudge LLM	Shown LLM	Reflective LLM
Median completion time: 5m 59s % Attentive finishing: 94%	Median completion time: 7m 20s % Attentive finishing: 83%	Median completion time: 9m 49s % Attentive finishing: 70%	Median completion time: 9m 48s % Attentive finishing: 65%
The control group answered the experimental questions without any additional support.	Beneath each question, participants had the option to "Chat with Pip" and had access to a large language model (Gemini Flash 2.5). They could answer the question without interacting with the LLM, and send a maximum of 10 questions each time.	Exactly the same as "Nudge LLM," except the users were required to send at least one message to the LLM before they could answer a given question.	Exactly the same as "Shown LLM" except the LLM was instructed not to give answers but to encourage cognitive reflection.

The LLM's answers took between 1-10 seconds to appear, so the Shown and Reflective arms took significantly longer to complete the experiment. This led to differential attrition, the implications of which are discussed in the appendices to this report.

BIT

*All three LLM arms rely on Gemini Flash, accessed through an API.

Pip was created by modifying Gemini Flash 2.5 with a Gemini Gem that gave the question and context

Imagine that you are interested in subscribing to a magazine. Which of the following options would you choose?

- A one year subscription to the online version of the magazine. Includes online access to all articles since 1997.

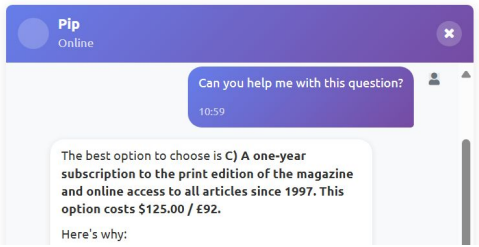
This option costs $59 / £43

- A one-year subscription to the print edition of the magazine

This option costs $125 / £92

- A one-year subscription to the print edition of the magazine and online access to all articles since 1997.

This option costs $125 / £92



In order to create Pip, we created a Gemini Gem that briefed Gemini Flash 2.5 on the question that participants had seen and how to interact with them. The full prompts are shown the Appendix.

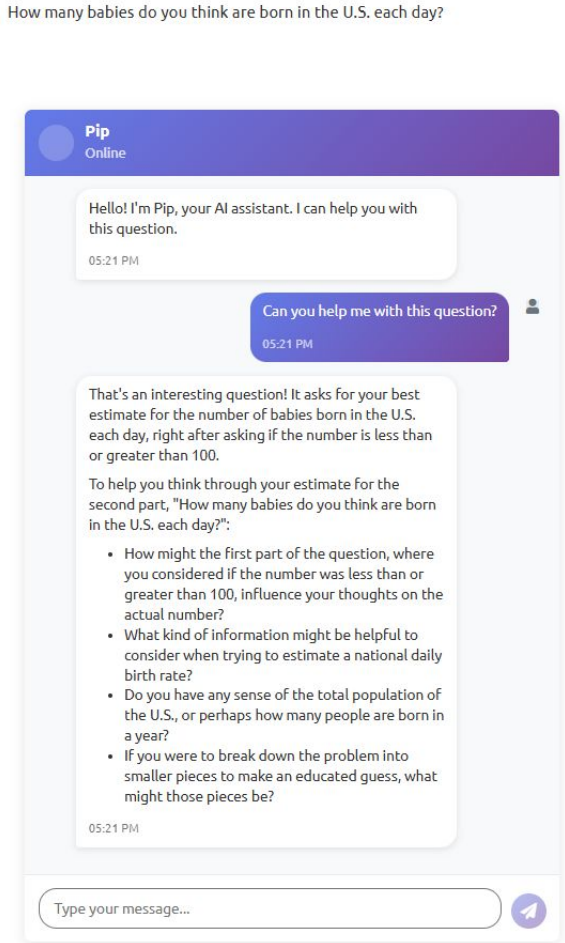
*All three LLM arms rely on Gemini Flash, accessed through an API.

In the Reflective LLM arm, Pip was instructed to help participants reflect on their own responses

In the Reflective LLM arm, PIp was instructed not to tell participants answers directly, but rather to get them to reflect on the problem and their preferences more deeply.

Our rationale for creating this intervention was to explore concerns that exposure to LLMs lead to cognitive degrading. We wanted to test the impact of LLM input that tried to support the participant making a decision, rather than directly offering an answer itself.

*All three LLM arms rely on Gemini Flash, accessed through an API.



BIT

Findings

Decoy Effect

Decoy Effect: Setup

Imagine that you are interested in subscribing to a magazine. Which of the following options would you choose?

- [*Cheap*] A one year subscription to the online version of the magazine. Includes online access to all articles since 1997. This option costs **$59 / £43**
- **[Decoy - 50% saw this option, 50% didn't]** A one-year subscription to the print edition of the magazine. This option costs **$125 / £92**
- [*Target*] A one-year subscription to the print edition of the magazine and online access to all articles since 1997. This option costs **$125 / £92**

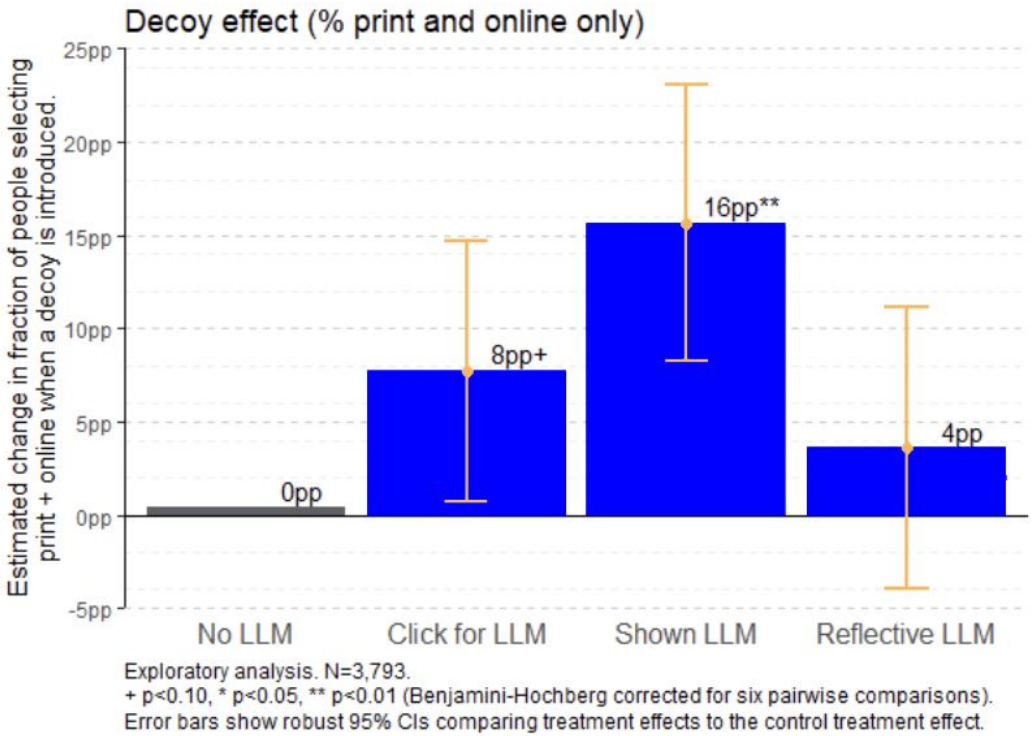
Decoy Effect

Description: Marketers introduce a "decoy" option that is clearly inferior to an existing option (the "target"). The presence of the decoy makes the target seem more attractive (even though it has not changed), and more people choose it than they would if the decoy did not exist.

Scenario: Half of participants saw two options for a magazine subscription: a cheap and an expensive ("target") one. Half of participants saw three options: the cheap and expensive ones, plus an inferior yet expensive "decoy".

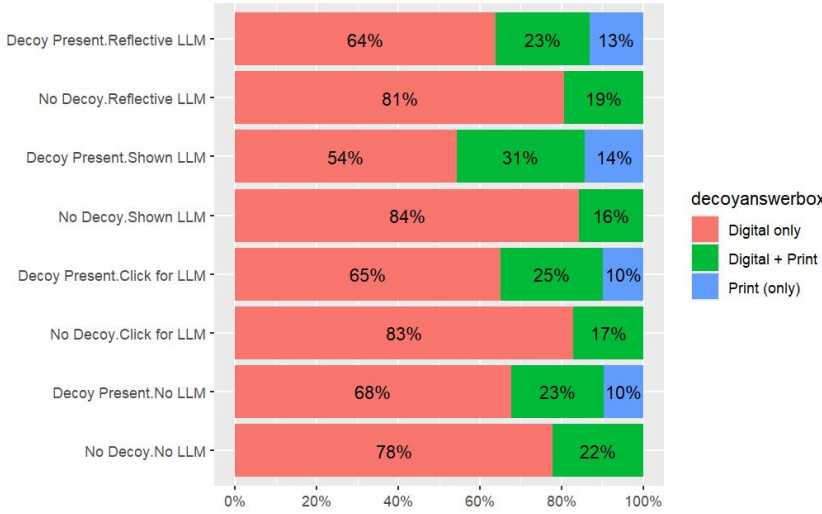
Based on existing literature, we hypothesized that the size of the decoy effect, as measured by the difference in the proportion of participants selecting the cheaper option, would be smaller in the LLM groups than in the control.

Decoy Effect: Main result



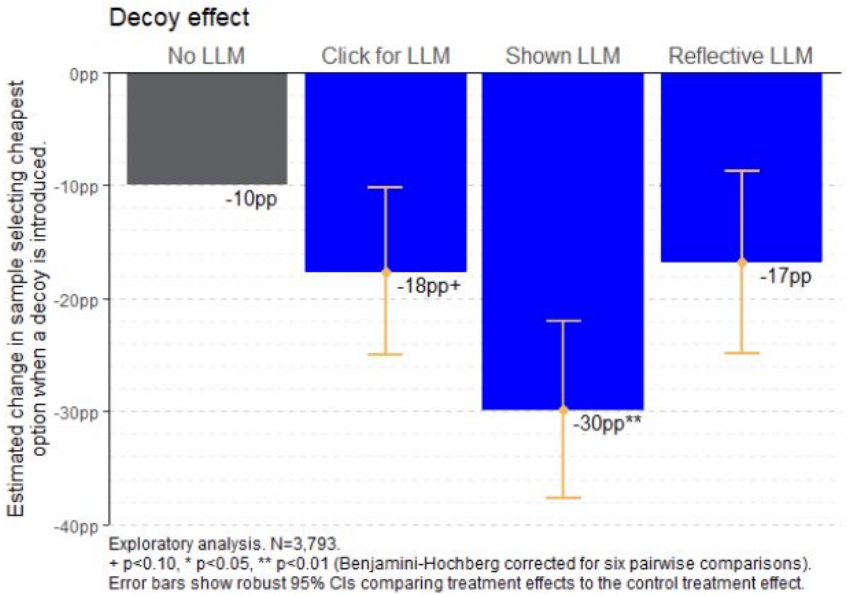
In the No LLM condition, the introduction of the Decoy Option did not increase the proportion of participants choosing the Target Option. This is not in line with previous studies.

Below we provide the proportions selecting each option under the various conditions.



BIT

Decoy Effect: AI results

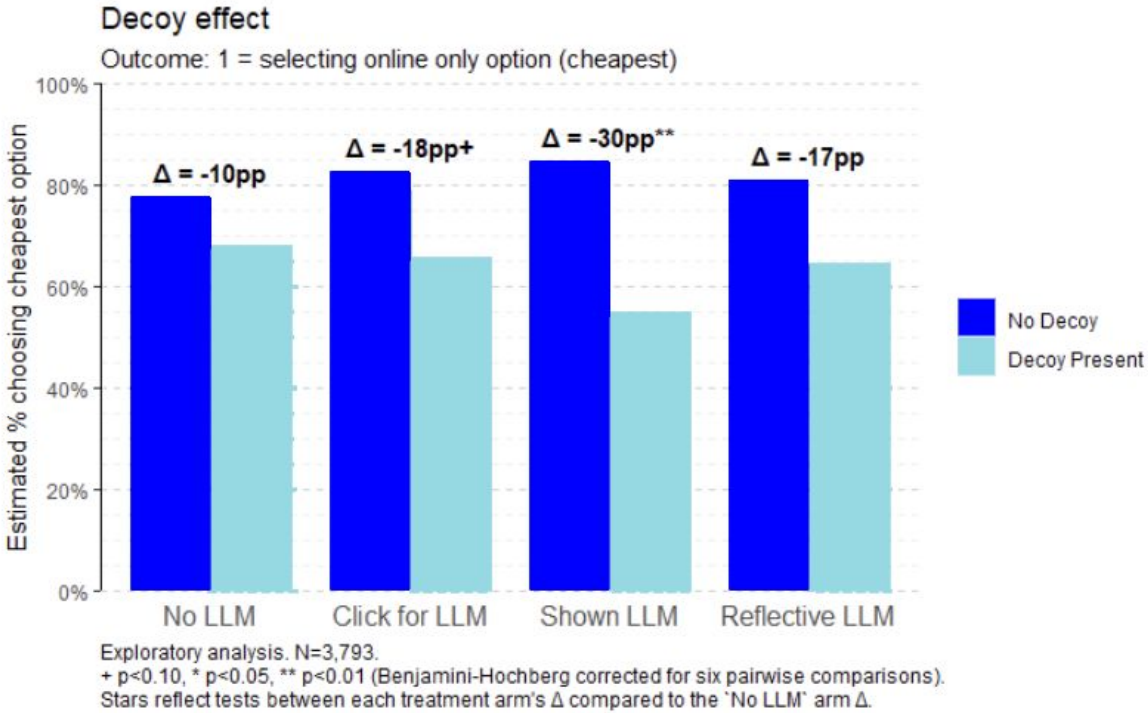


We observe that the Shown LLM arm *increases* the size of the decoy effect. There is some evidence that the Click for LLM arm also increases the size of the effect. We do not know the exact reasons why the Shown LLM has this effect - the Gemini Gem for this experiment was set up in line with the others. However, our user testing reveals some possible causes.

When the decoy is absent (Cheap vs Target), the LLM can identify the scenario as an example of "price anchoring", a cognitive bias where "the price of the more expensive option acts as an anchor". The more expensive option seems deceptive and pushes people to the cheaper one. In our view, this is an incorrect use of the anchoring concept.

When the decoy is present (Cheap vs Decoy vs Target), the LLM can identify the decoy option correctly and says that it "pushes people towards choosing [the Target option]." But then it goes on directly to say that "The most logical choice, based on a direct comparison, would be [the Target option]." Of course, this direct comparison is exactly what the Decoy option creates. The advice does not seem to be coherent.

Decoy Effect: AI results



This graph shows the absolute levels of participants selecting the two options in the different conditions.

Sunk Costs

Sunk Costs: Setup

Imagine that you're a volunteer who runs a club in your local area. You are responsible for organizing your club's annual meeting.

Last week, you booked a hotel conference room for the meeting using funds from the club. [**You paid a $300 / £300 fee that is not refundable.** / **You paid a $30 / £30 deposit that is not refundable ($270 / £270 more is owed on the day.**)]

This morning, the head of your local library emails you. They say that they're now offering their new meeting space free to community groups. The library space has better facilities and more convenient parking. Both venues need the same setup time and can fit enough people in.

Where do you choose to have the meeting?

- Hotel conference room
- Library meeting space

Description: Sunk costs are resources (money, time or effort) that have already been incurred and cannot be recovered, regardless of what you do next. If we want to get the best outcome, we should focus only on future (marginal) benefits or costs - the "sunk" resources shouldn't factor into our choice.

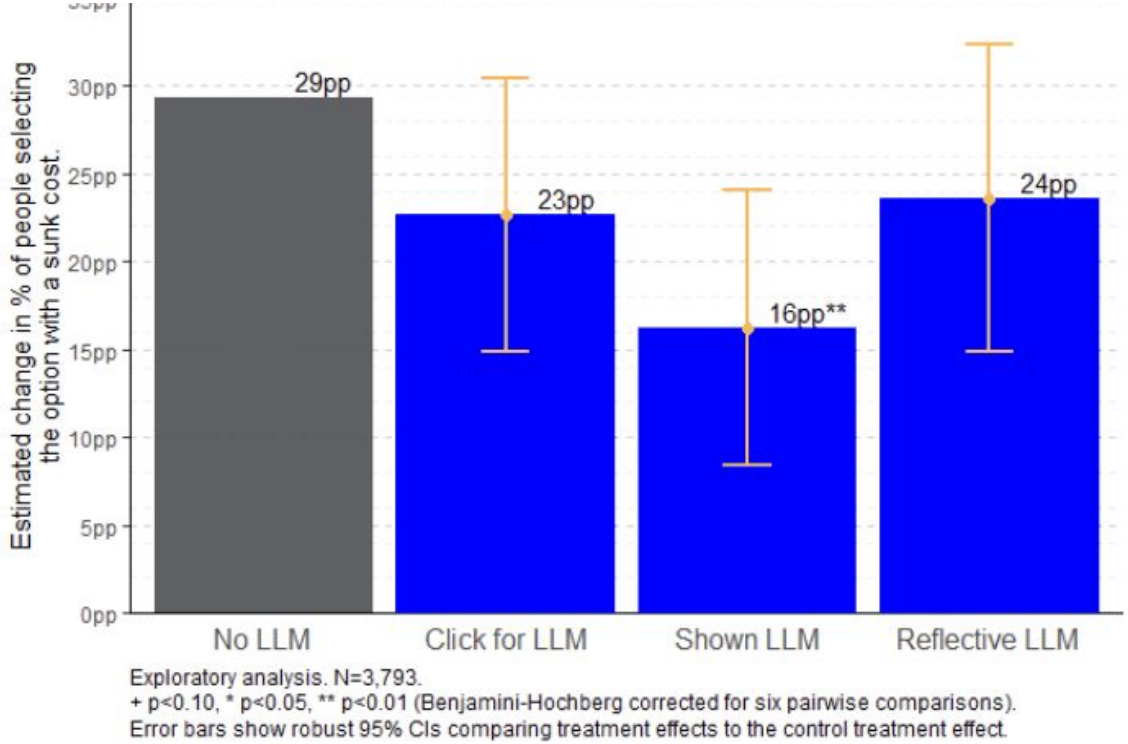
Scenario: Participants were told they had booked a hotel meeting room for an event. Half the participants were told that they had paid a large fee ($300/£300) that was not refundable ("High Sunk costs"). Half were told that they had paid a small fee ($30/£30) that was not refundable, with more due on the day ("Low Sunk Costs").

They were then told that a better room option (in the local library) had emerged after the booking was made. Participants were asked if they would choose the hotel or the library option.

We hypothesized that the difference in the proportion of people staying with the hotel room would be smaller in the LLM groups than in the control group - representing a smaller sunk cost effect.

BIT

Sunk Costs: Results

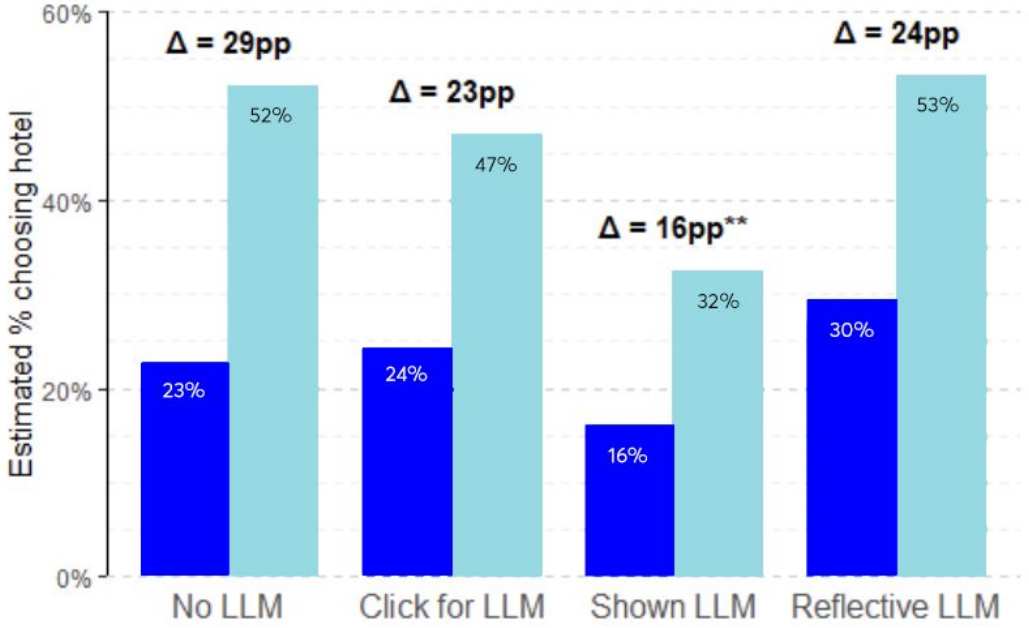


Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Error bars show robust 95% CIs comparing treatment effects to the control treatment effect.

We found that the difference in people choosing the hotel option was smaller between the High and Low sunk costs (16 percentage points) for the Shown LLM group than the Control (29 percentage points); the gap between differences was not significant for the other LLM groups.

Here, the LLM provided logical advice that emphasized that the library was the better option, regardless of what had been spent.

Sunk Costs: Results



Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Stars reflect tests between each treatment arm's Δ compared to the `No LLM` arm Δ.

This graph shows the absolute levels of participants selecting the two options in the different conditions.

BIT

Outcome Bias

Outcome Bias: Setup

Imagine you are a taxi driver. In the center of your town, a passenger gets in and tells you that they need to get to the airport on time. They don't care about the price. You must choose between two routes, both of which are familiar to you. You have an app that tells you how often a route makes drivers late on average, which is very accurate.

- **Express route:** This option uses a motorway that avoids city centre traffic. It's a longer distance, but you can drive faster due to multiple lanes. Your navigation system reports that 15% of drivers who take this route to the airport arrive late.
- **Industrial route:** This option goes through an industrial part of town. It's a shorter distance, but you have to drive more slowly due to frequent junctions and traffic lights. Your navigation system reports that 11% of drivers who take this route to the airport arrive late.

You decide to take the industrial area route. [**The journey goes smoothly and the passenger boards their flight**. / **You get stuck behind a truck and the passenger misses their flight**].

For your next airport run, which route would you choose?

- Express lane
- Industrial route/

Description: Outcome bias occurs when we judge the quality of a decision based exclusively on its result, and neglect the quality of the decision making process. In other words, a lucky but poor decision can be overly praised, while a well-reasoned decision that leads to a bad outcome is overly criticized.

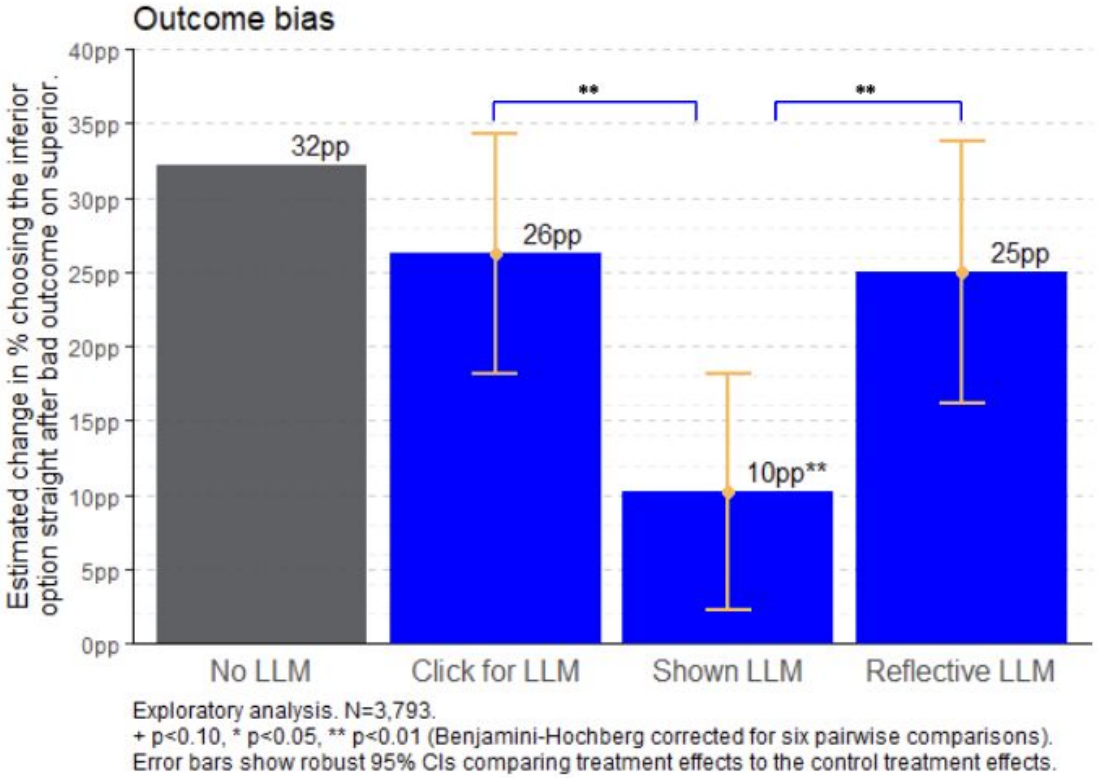
Scenario: Participants were told that they needed to drive a passenger to an airport for a flight. They were given a choice of two routes: Express Route or Industrial Route.

They are told they have a reliable app that says that the Express Route makes drivers late for the airport 15% of the time; the figure for the Industrial Route is 11%. They are told they took the Industrial Route.

Half the participants were told that the journey went smoothly and the passenger made their flight; half were told that they hit traffic and the passenger missed their flight. Both groups are then asked which route they would choose for the airport next time.

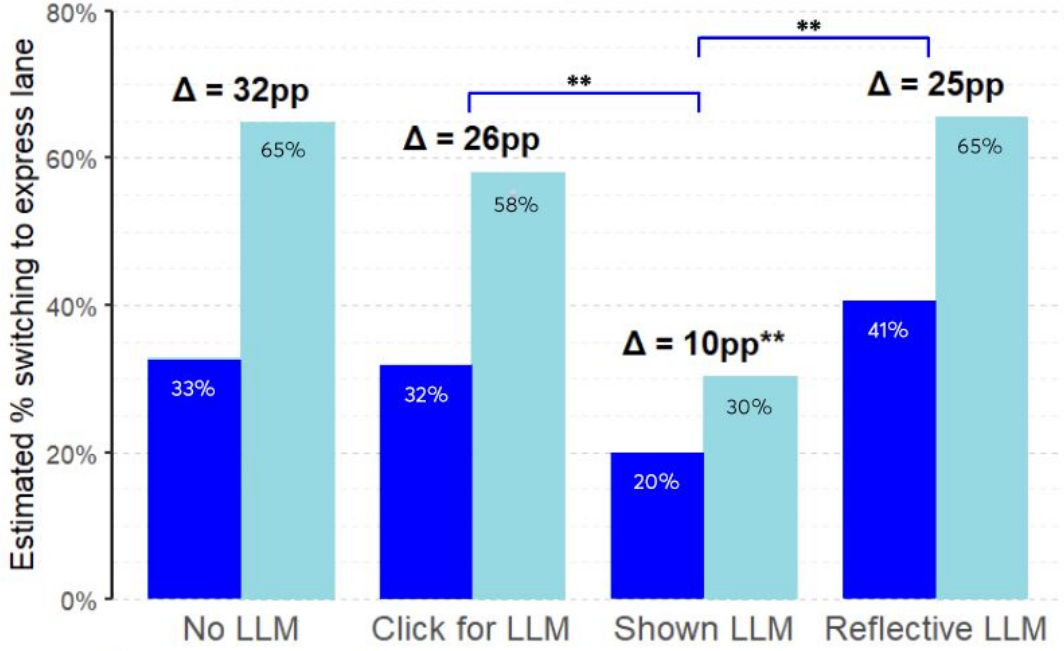
We hypothesized that the difference in the proportion of people choosing the inferior Express Route option would be smaller in the LLM groups than in the control group - representing a smaller outcome bias effect.

Outcome Bias: Results

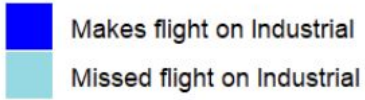


In the Control group, 32 percentage points more people chose the Express Route after a bad outcome with the Industrial Route - despite it being the worse option overall. For the Shown LLM group, the difference was only 10 percentage points, which was also significantly lower than the other LLM groups.

Outcome Bias: Results



Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Stars reflect tests between each treatment arm's Δ compared to the `No LLM` arm Δ.



This graph shows the absolute levels of participants selecting the two options in the different conditions.

Anchoring Effects

Anchoring Effects: Setup

For the following question please give your best estimate. If you do not know the correct answer, just give your best guess.

Do you think the average number of babies born per day in the US is less than or greater than [**100**/**50,000**]? Please note this number was generated at random.

- Less than [**100**/**50,000**]
- Higher than [**100**/**50,000**]

[Appears upon selection of answer]
How many babies do you think are born in the U.S. each day?

- Respondents can answer a positive number, no maximum

Correct answer: 3.6mn per year = 9,900 per day.

Description: We focus on numerical anchors. A typical case is when a person is exposed to a number and then asked to estimate a numerical value (which can be explicitly unrelated to the preceding number). Anchoring effects occur when the prior number acts as an "anchor" that distorts the estimate made.

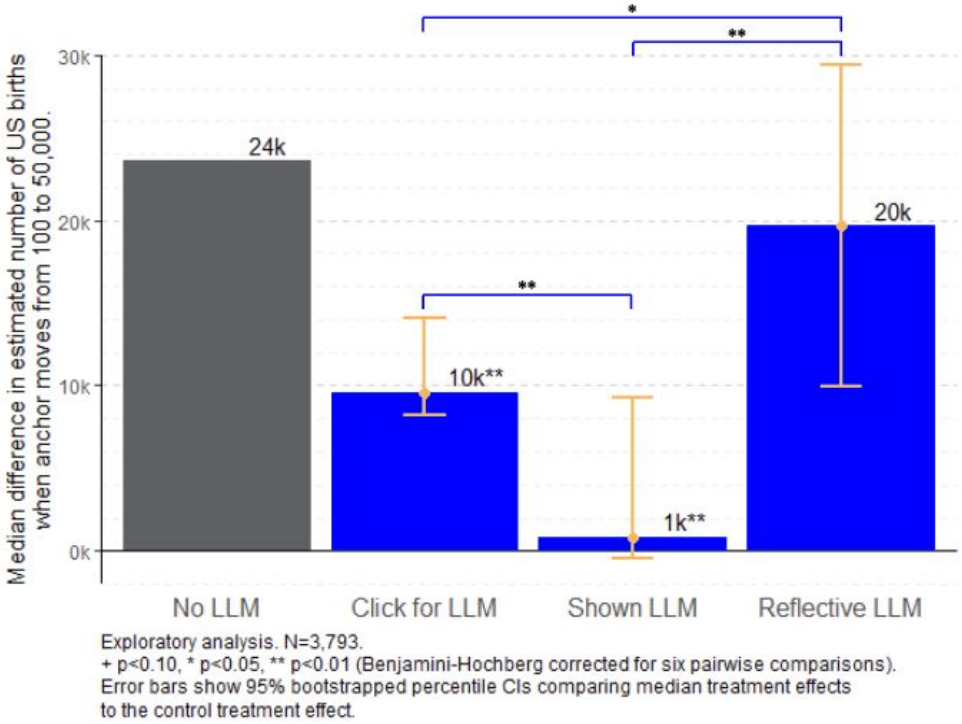
Scenario: Half of participants were asked: "*Do you think the average number of babies born per day in the U.S. is less than or greater than 100? Please note this number was generated at random.*" ("Low Anchor") For the other half of participants, the 100 number was replaced with 50,000 ("High Anchor").

Participants were then asked to estimate the total number of babies born in the U.S. every day.

We hypothesized that the difference between the High Anchor estimates and the Low Anchor would be smaller in the LLM groups than in the control - representing a smaller anchoring effect.

BIT

Anchoring Bias (quantile regressions)



Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Error bars show 95% bootstrapped percentile CIs comparing median treatment effects to the control treatment effect.

This graph shows the estimated difference in median treatment effects between AI arms.

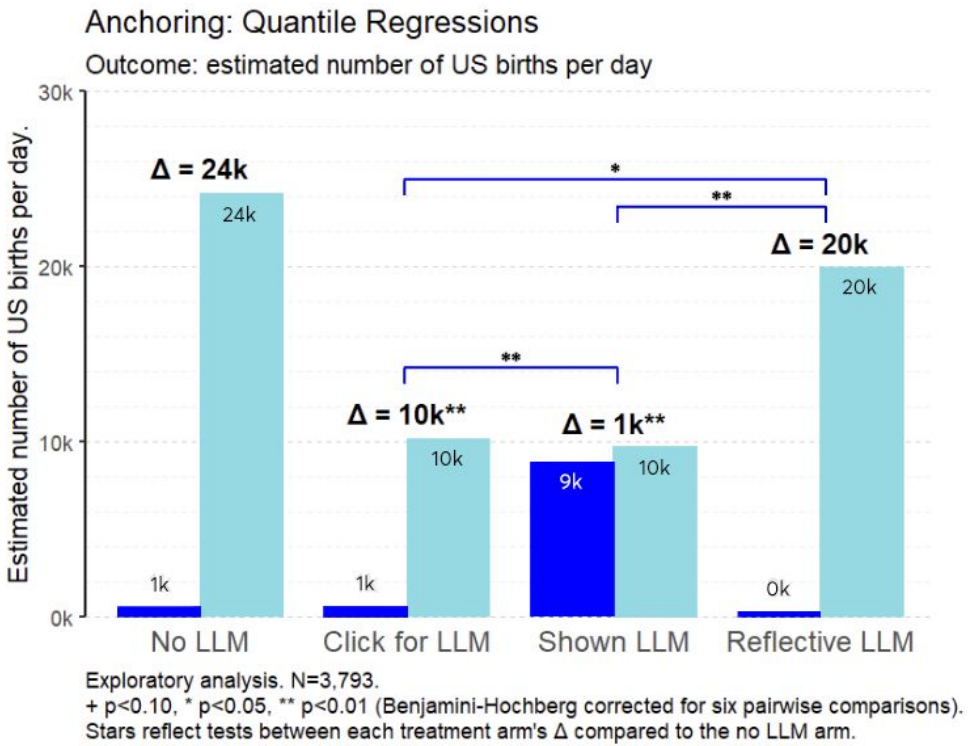
For example, for those without the LLM, shifting the anchor for the number of US births per day from 100 to 50k shifts the median answer by an estimated 24,000.

The change in median answer in the Shown LLM arm is 2,000, significantly lower than the estimated median change observed in the control arm, suggesting a large reduction in the size of anchoring bias.

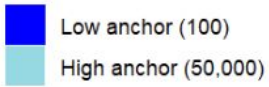
Bars show arm-specific median treatment effects from a median quantile regression with an interaction (treatment × anchor). Error bars are 95% percentile bootstrap CIs for the contrast with the control arm (xy bootstrap; 5,000 reps), plotted around each arm's bar. Asterisks denote BH-adjusted p-values for pairwise differences vs control. CIs are not 'difference-from-zero' intervals.

BIT

Anchoring Bias (quantile regressions)



Answers had no maximum and there were very high leverage observations so we focus on the median treatment effect. Results for the mean unadjusted and with two levels of winsorisation are available in the Appendix.



Trolley Problem

We gave participants a “trolley problem”

We also wanted to assess whether participants would be affected by AI advice in the domain of moral decision making.

We used a trolley problem for this purpose. A trolley problem asks the respondent to choose whether to cause the death of one person to save five people. A utilitarian approach to morality says that such a decision is morally justified.

We used the setup created by Hauser et al., 2007 and replicated by Many Labs 2. In this design, participants are either asked if it is morally OK to a) pull a lever to kill one person and save five people or b) push a person off a bridge (to their death) to save five people. The order in which they see these options is randomised.

Participants are much more likely to say that pulling the lever is OK than pushing the person. On interpretation is that participants are less willing to condone a utilitarian action when the immediate consequences of their actions are made salient. However, it is also worth noting that that pushing a person is different from pulling a lever, from the standpoint of Kantian ethics (since it is treating a person as a means not an end).

Participants see “Denise” and “Frank” questions in a randomised order

“Denise” question

[50% see Denise first] Denise is on a train. The driver just shouted that “The brakes have failed!” and then fainted. There are five people on the track ahead who can’t get out of the way in time. Denise can switch the train to a side track, but there’s one person on that track. She can either switch tracks and kill one person, or do nothing and let five people die.

Is it morally okay for Denise to switch the train to the side track?
Yes/No

What % of other people do you think would say it is morally OK for Denise to switch the train to the side track? [numeric 0-100]

“Frank” question

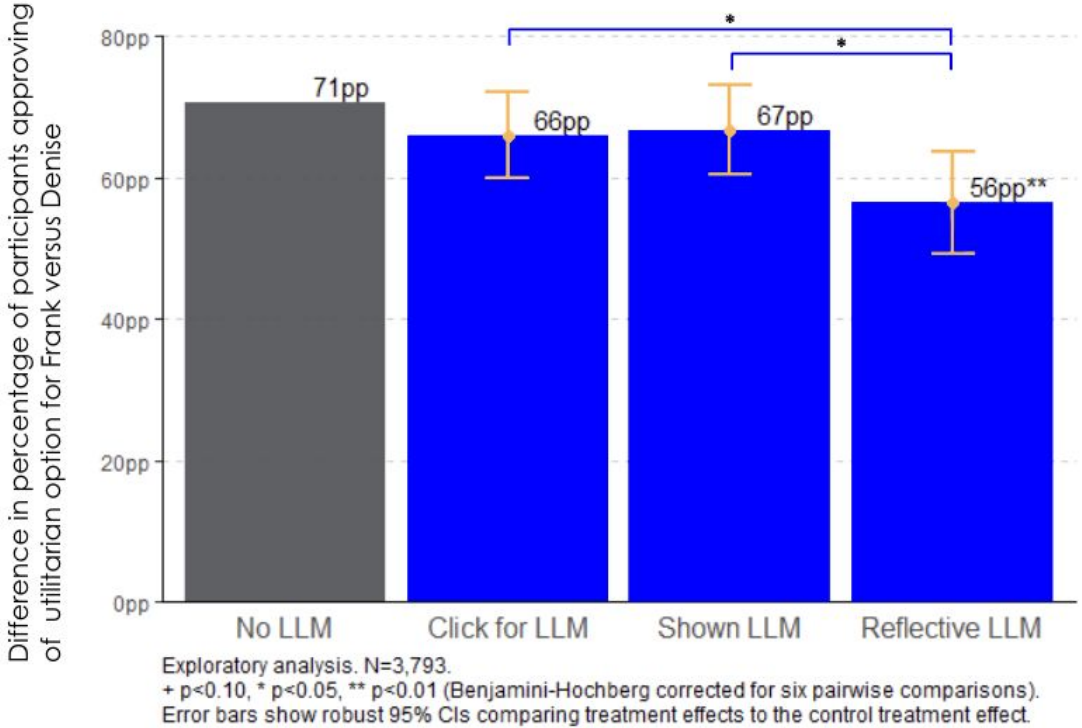
[50% see Frank first] Frank is on a bridge above train tracks. He sees a runaway train heading toward five people who can’t get out of the way in time. Frank knows the only way to stop the train is to drop something very heavy in front of it. The only heavy object available is a large man with a backpack standing next to him on the bridge. Frank can shove the man onto the tracks to stop the train and save the five people, but this would kill the man.

Is it morally okay for Frank to shove the man?
Yes/No

What % of other people do you think would say it is morally OK for Frank to shove the man? [numeric 0-100]
====
Have you read about or considered moral dilemmas of this sort (involving people on train tracks) before? [Yes/No]

BIT

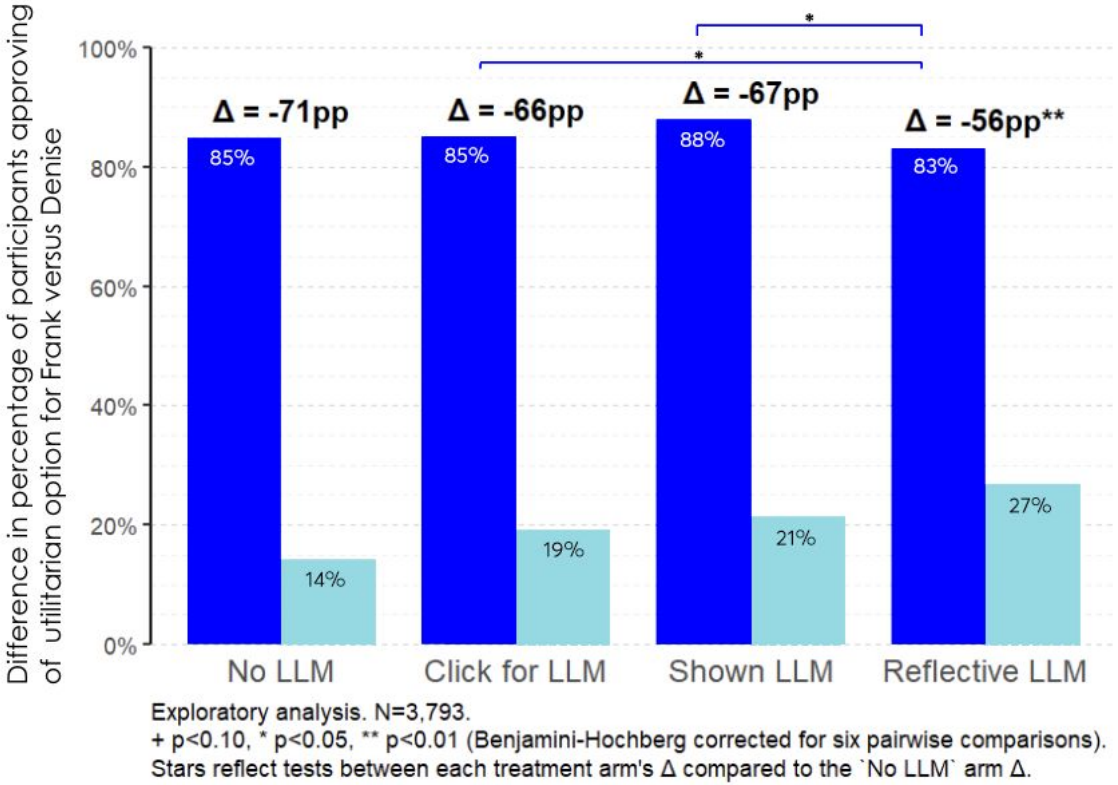
Trolley Problem: Results



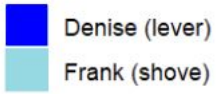
For participants who saw the Denise scenario and then the Frank scenario, the difference between those choosing the utilitarian option in the two scenarios was significantly lower for the Reflective LLM group.

A plausible interpretation is that the Reflective LLM made people consider the similarities between the two scenarios, leading to more consistent decision making.

Trolley Problem: Results



This graph shows the absolute levels of participants selecting the two options.



BIT

Repeat bias questions

Repeat bias questions: Setup

We wanted to test whether participants who had been provided access to LLMs would change their behaviour when facing similar scenarios - but when LLM support is not available.

We removed AI access for all participants and presented them with new scenarios based on the decoy effect and sunk cost fallacy.

Individuals received the same within question treatment condition they saw previously (so if a participant was randomised to see the high sunk cost condition for the original sunk cost scenario, also see the high sunk cost version of the follow-up scenario).

Imagine that you want to enroll in gym classes. Which offer do you think you would choose?

1. Studio Lite: 10 classes a month (£50 / $70)
2. **[Present for 50% of participants]** Studio Plus: Unlimited in person classes. No digital access (£150 / $200)
3. Studio + digital bundle: Unlimited in person classes and unlimited on-demand digital classes (£150 / $200)

====

Now imagine it's a weekday evening and you intend to go to the cinema alone to see a new movie.

You start reading reviews of the movie and they are very bad. [You already bought a non-refundable ticket for [£15/$20] / You have not bought a ticket yet (it costs £15/$20)].

How likely are you to go to see the movie?

- [Not at all / A little / Moderately / Very much]

Repeat bias questions: Setup

Decoy Effect

Imagine that you want to enroll in gym classes. Which offer do you think you would choose?

1. Studio Lite: 10 classes a month (£50 / $70)
2. **[Present for 50% of participants]** Studio Plus: Unlimited in person classes. No digital access (£150 / $200)
3. Studio + digital bundle: Unlimited in person classes and unlimited on-demand digital classes (£150 / $200)

Sunk Costs

Now imagine it's a weekday evening and you intend to go to the cinema alone to see a new movie.

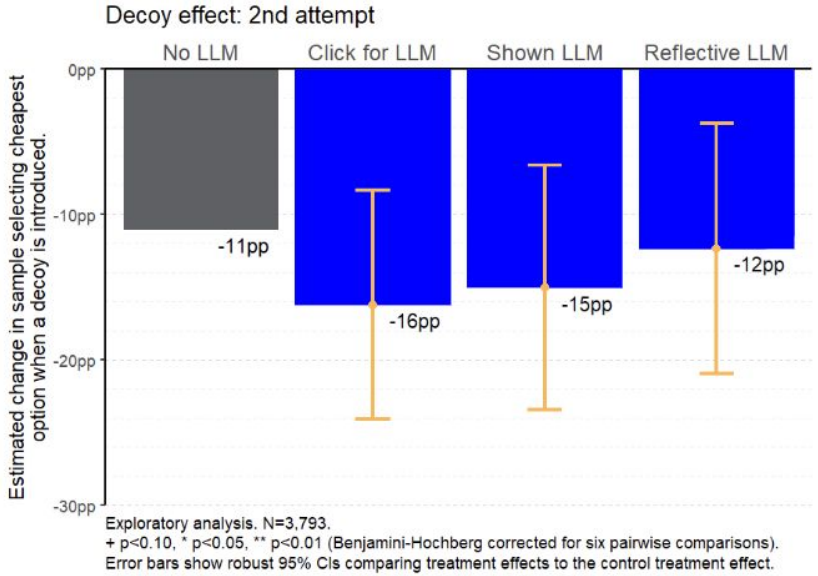
You start reading reviews of the movie and they are very bad.
[You already bought a non-refundable ticket for [£15/$20] / You have not bought a ticket yet (it costs £15/$20)].

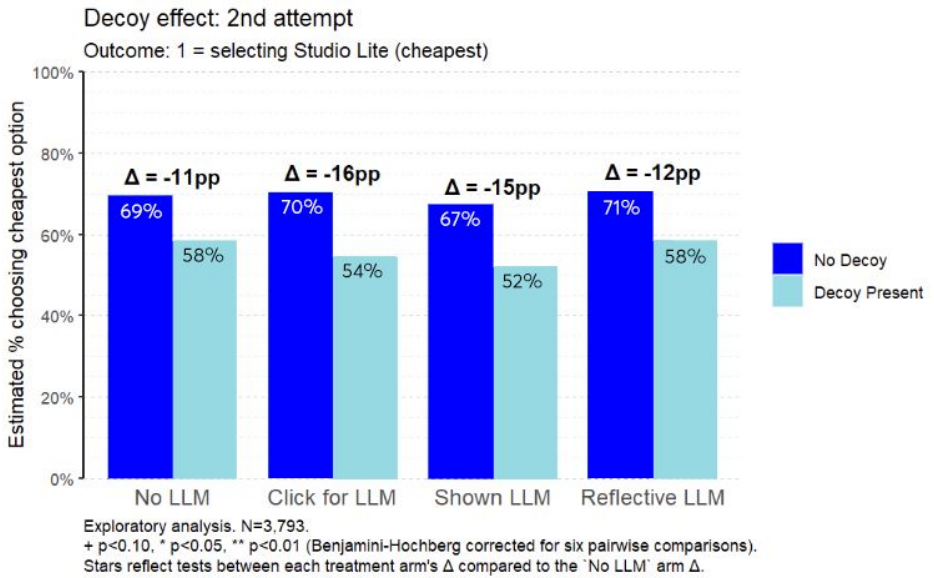
How likely are you to go to see the movie?

- [Not at all / A little / Moderately / Very much]

Decoy Effects Reprise: Results

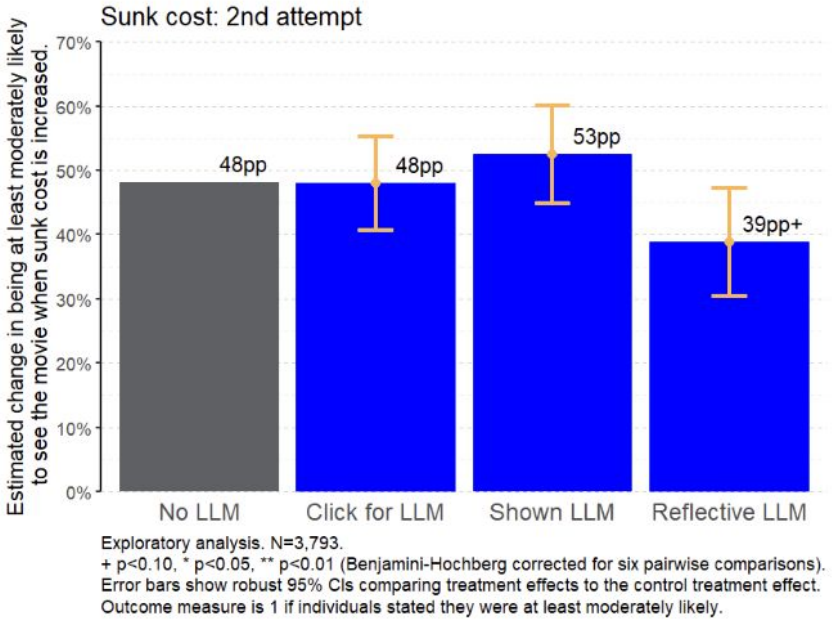
We did not observe any significant differences in the size of the decoy effect between arms.



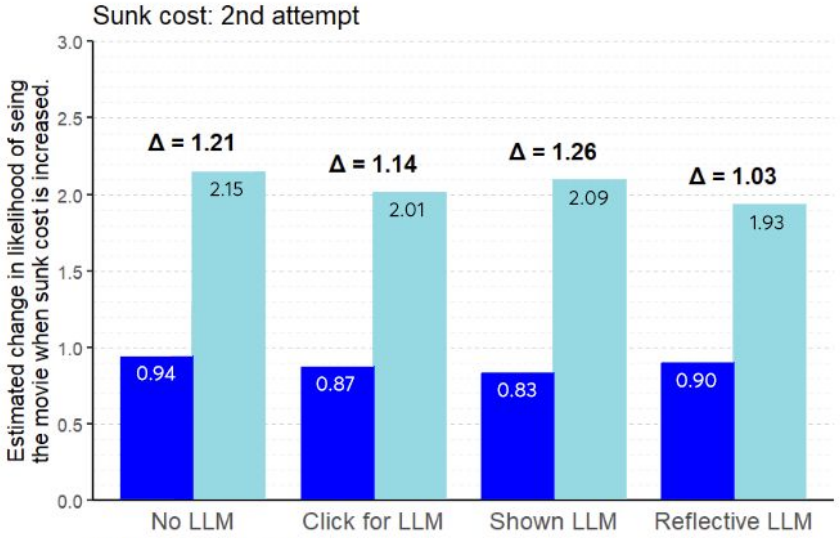


Sunk Costs Reprise: Results

We did not observe any significant differences in the size of the sunk cost effect between arms.



Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Error bars show robust 95% CIs comparing treatment effects to the control treatment effect.
Outcome measure is 1 if individuals stated they were at least moderately likely.



Exploratory analysis. N=3,793.
+ p<0.10, * p<0.05, ** p<0.01 (Benjamini-Hochberg corrected for six pairwise comparisons).
Error bars show robust 95% CIs comparing treatment effects to the control treatment effect.
Outcome measure is Likert scale (0 = not at all, 1 = a little, 2 = moderately, 3 = very much).

Appendix

Balance checks

This section presents evidence of substantial differential attrition caused by increased time and effort requirements associated with forcing participants to engage with AI.

We find acceptable balance between completers, and poor balance between those leaving and those who finished, suggesting data is not missing at random.

	N starting	N finishing	N finishing after removing fastest 5%
No LLM	1,301	1,225 (94%)	1,163
Nudge LLM	1,285	1,072 (83%)	1,081
Shown LLM	1,301	908 (70%)	862
Reflective LLM	1,221	791 (65%)	750

We observed differential attrition

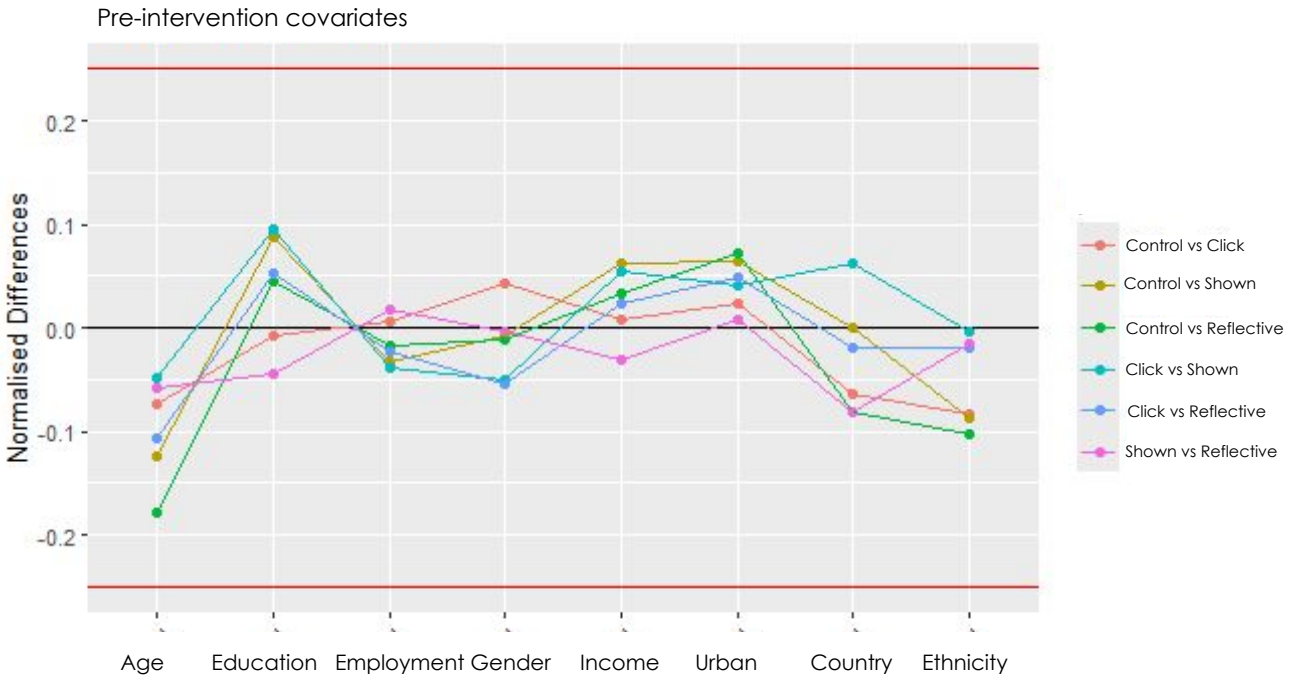
	Control	Treatment 1	Treatment 2	Treatment 3
N starting	1,301	1,285	1,301	1,221
N finishing	1,225 (94%)	1,072 (83%)	908 (70%)	791 (65%)
N after removing fastest 5% of participants	1,163	1,081	862	750
Country (*% USA*)	52%	49%	52%	48%
Previous LLM use (% any)	42%	44%	47%	50%
Age (*mean (sd)*)	49 (17)	48 (17)	47 (17)	46 (17)
Ethnicity (*% white*)	75%	71%	71%	71%
Location (*% urban*)	32%	33%	35%	36%
Gender (*% female*)	53%	56%	53%	53%
Education (*% with degree*)	37%	36%	41%	39%
Employment (*% employed*)	66%	66%	67%	66%
Income (*% >median*)	39%	40%	43%	41%

There was evidence of substantial differential attrition in this experiment. There is limited evidence of imbalance between arms as assessed by the normalised difference (Imbens and Rubin, 2015), but evidence those who attrited were systematically different from those who didn't.

We cannot rule out that the differential attrition led to imbalances on unobservable characteristics that influence our outcome measures.

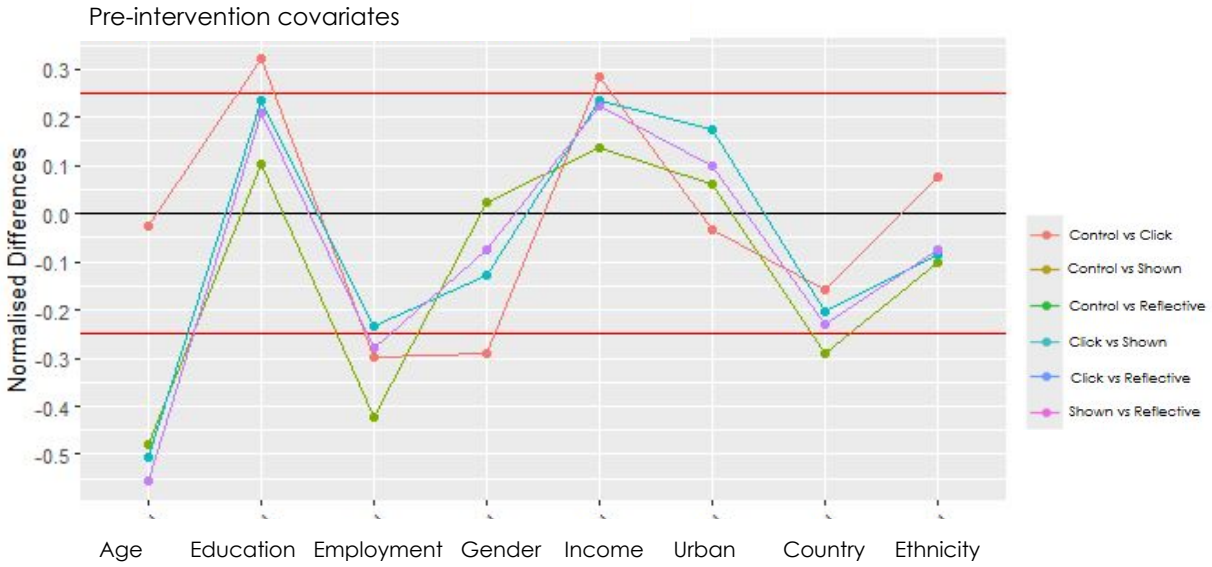
We therefore present bounds on the treatment effects throughout the results deck. The method used to compute these bounds is detailed in later slides, and is closely related to Lee (2009).

Baseline data showed acceptable balance



Pre-intervention covariates suggest acceptable balance, but there are likely unobservable characteristics that differ between treatment arms as a result of the selective attrition which we cannot control for in our regression analysis.

Those who left do not appear to be missing at random



Normalised differences between those who finished to those who didn't, within treatment arm.
The sign is determined by finishers minus quitters.

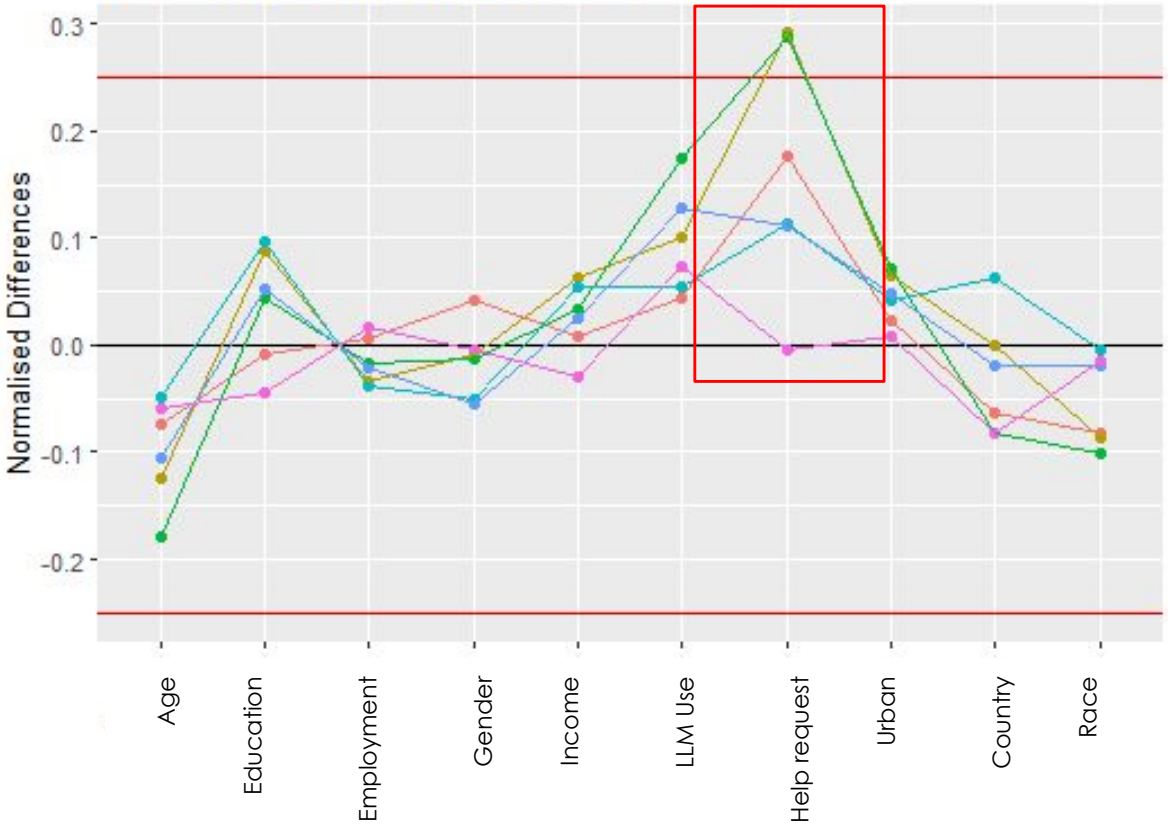
Here we compute the normalised differences within a treatment arm for finishers versus non-finishers (including speed-runners) for pre-intervention covariates.

Those who finished were, on average:

- Younger,
- Less likely to be employed
- More likely to be above median income

There is clear evidence individuals are not missing at random.

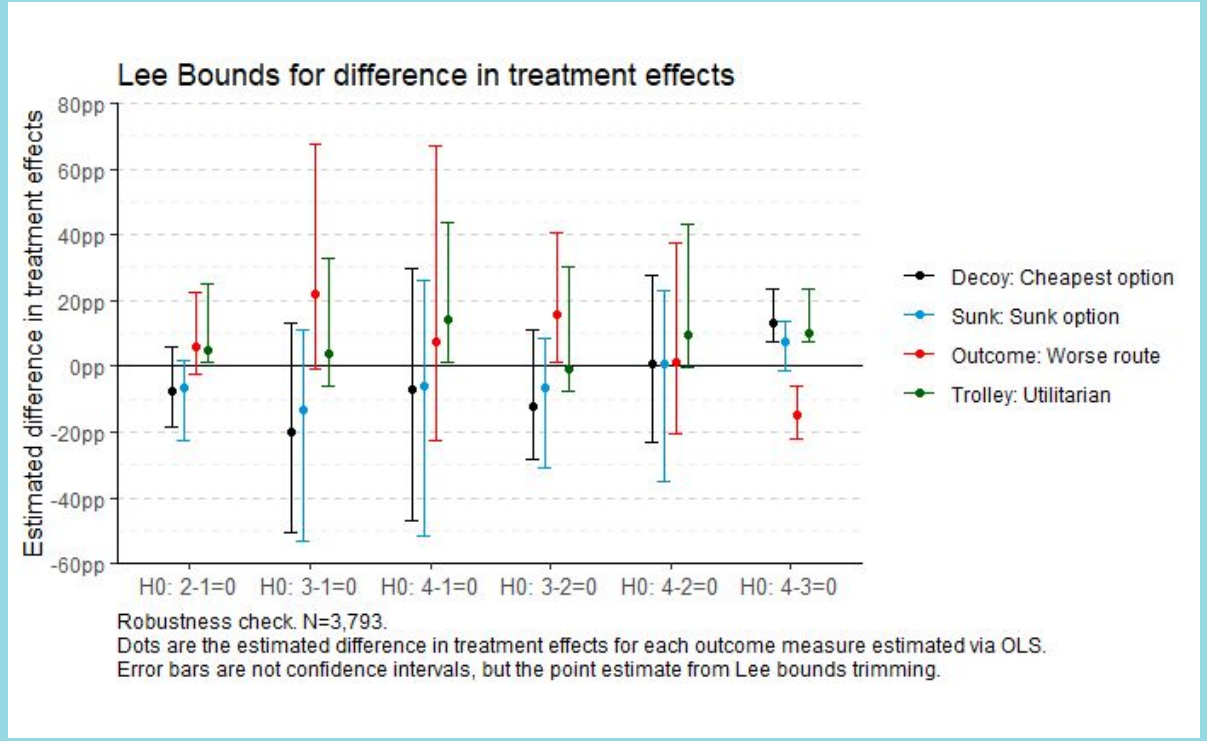
Post-intervention covariates



All covariate balances (for our binarised measures, which lose information) are within 0.25 of the average standard deviation except for "Help from LLM", which refers to whether the respondent said that they used a search engine or other external help to answer one of the study questions.

BIT

Differential attrition checks



Robustness check. N=3,793.
Dots are the estimated difference in treatment effects for each outcome measure estimated via OLS.
Error bars are not confidence intervals, but the point estimate from Lee bounds trimming.

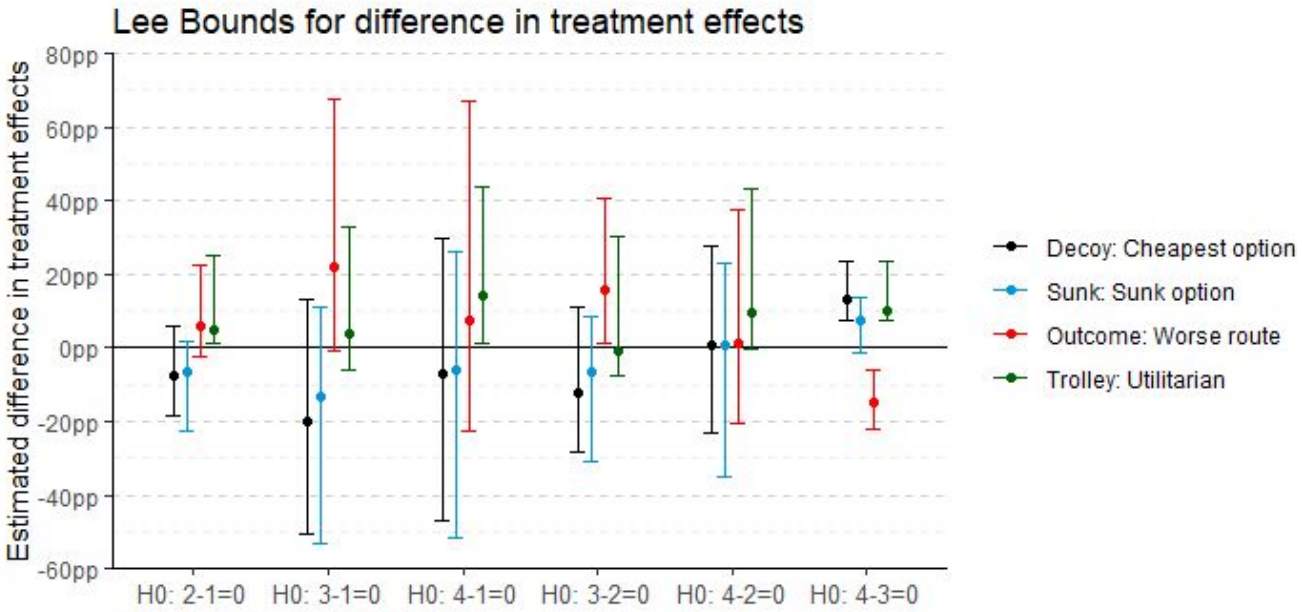
We applied Lee bounds to the difference in treatment effects as a sensitivity check

To check the robustness of the results to differential attrition we adapted the bounding procedure introduced by Lee (2009) to our multi-arm design when testing for the differences in treatment effects within AI arms. This method produces **sharp upper and lower bounds** on the differences in treatment effects, under the assumption that attrition is monotonic in potential outcomes.

Our approach asks "what is the smallest or largest difference in treatment effects we could observe if those who attrited happened to fall at the extremes of the outcome measure distribution?" We trim participants from the arms with lower attrition to capture the "worst-case" setting where attrition is maximally disruptive, and report whether it is possible to reverse the observed difference in treatment effects. See also Broderick et al., (2023).

BIT

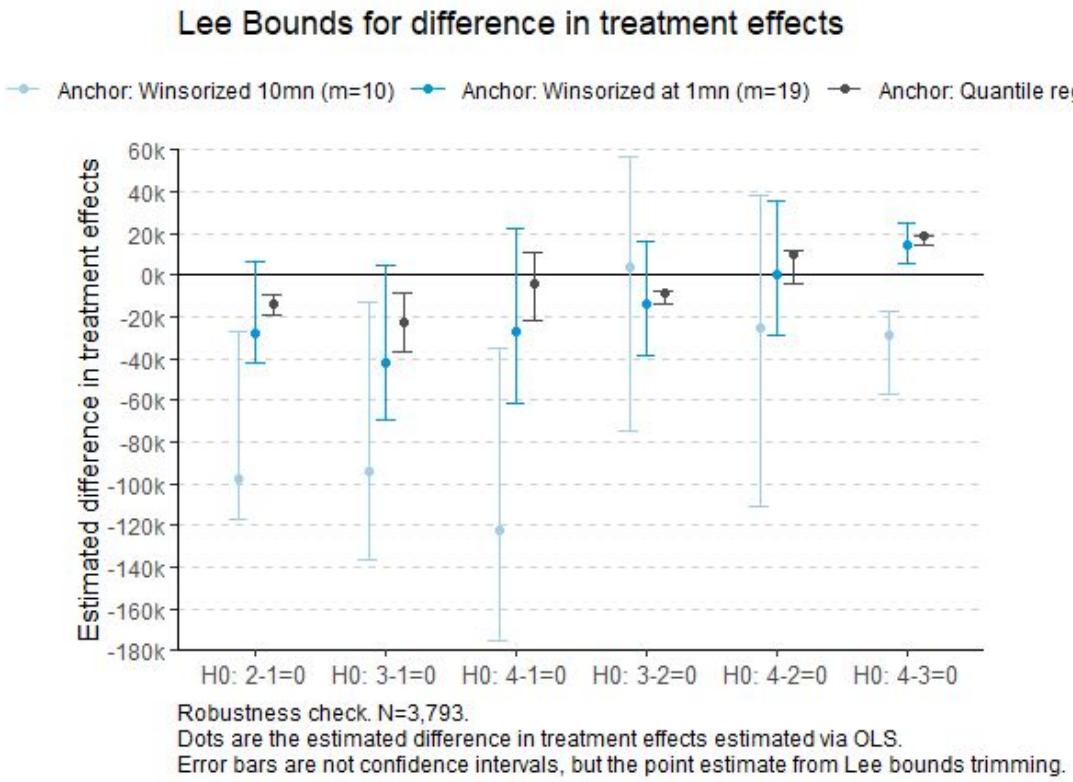
Results with Lee Bounds calculated



Robustness check. N=3,793.
Dots are the estimated difference in treatment effects for each outcome measure estimated via OLS.
Error bars are not confidence intervals, but the point estimate from Lee bounds trimming.

Applying Lee (2009) bounds suggests that, under worst-case assumptions, some of the observed Decoy, Sunk Costs, Outcome Bias and Trolley Problem results could attenuate substantially or even reverse. The results should be interpreted with the bounds shown on this slide in mind.

Results with Lee Bounds calculated



Applying Lee (2009) bounds suggests that, under worst-case assumptions, some of the observed Anchoring results could attenuate substantially or even reverse. The results should be interpreted with the bounds shown on this slide in mind.

BIT

Regression Tables

This section presents formal regression analysis for the following five outcomes:

1. Decoy effect
2. Sunk cost fallacy
3. Outcome bias
4. Anchoring
5. Trolley problem

As a robustness check we also include the results for the full sample (including the fastest 5%)

Regression table for those completing the experiment, excluding the fastest 5%.

	Decoy	Sunk cost fallacy	Outcome bias	Anchoring	Trolley
Outcome measure	1 if chose online only	1 if chose hotel option	1 if chose express lane	Anchor value given	1 if chose utilitarian option
Model	OLS	OLS	OLS	Quantile regression	OLS
N	3,793	3,793	3,793	3,793	3,793
Reference mean or median (No LLM and within == 0)	**0.7787**	**0.2262**	**0.6515**	**589**	**0.8468**
Nudge LLM	0.0508* (0.0238)	0.0167 (0.0249)	-0.0676* (0.0293)	0 (93)	0.0025 (0.0217)
Shown LLM	0.0682** (0.0244)	-0.0649** (0.0242)	-0.3455** (0.03)	8,240** (1,859)	0.0313 (0.0216)
Reflective LLM	0.0339 (0.0265)	0.0688* (0.0285)	0.0077 (0.0315)	-291** (85)	-0.0161 (0.0242)
Within randomisation • Decoy = 1 if decoy added. • Sunk = 1 if high upfront. • Outcome = 1 if miss flight. • Anchoring = 1 if high anchor. • Trolley = 1 if "push" is first.	-0.0993** (0.0261)	0.2937** (0.0272)	-0.322** (0.0279)	23,603** (2,037)	-0.7062** (0.0206)
Within*Nudge LLM	-0.0771* (0.0377)	-0.0672+ (0.0399)	0.0596 (0.0409)	-14,019** (2,053)	0.0469 (0.0313)
Within*Shown LLM	-0.1992** (0.0398)	-0.1314** (0.0397)	0.2197** (0.0405)	-22,764** (2,774)	0.0394 (0.0326)
Within*Reflective LLM	-0.0686+ (0.0413)	-0.0577 (0.0445)	0.0723 (0.0449)	-3,963 (4,715)	0.1426** (0.0366)

+ p<0.10, * p<0.05, ** p<0.01. Standard errors in parentheses (HC3 for OLS, bootstrapped R=5,000 for median quantile regression for anchoring).
Covariates: age, gender, income, region, ethnicity, education, employment status.

Regression tables for all who answered each question, including fastest 5% of participants.

	Decoy	Sunk cost fallacy	Outcome bias	Anchoring	Trolley
Outcome measure	1 if chose online only	1 if chose hotel option	1 if chose express route	Anchor value given	1 if chose utilitarian option
Model	OLS	OLS	OLS	Quantile regression	OLS
N (	4,516	4,412	4,361	4,123	4,080
Reference mean or median (No LLM and within == 0)	**0.7715**	**0.7455**	**0.6448**	**4,000**	**0.8401**
Nudge LLM	0.0528* (0.0229)	-0.0014 (0.0242)	-0.0605* (0.0279)	-28 (83)	0.0074 (0.0213)
Shown LLM	0.071** (0.023)	0.0599* (0.0239)	-0.3197** (0.0285)	6,476** (2,191)	0.0213 (0.0216)
Reflective LLM	0.0353 (0.024)	-0.0556* (0.0262)	0.0181 (0.0288)	-291** (81)	-0.0088 (0.0234)
Within randomisation • Decoy = 1 if decoy added. • Sunk cost = 1 if high upfront. • Outcome = 1 if miss flight. • Anchoring = 1 if high anchor. • Trolley = 1 if asked "shove" first.	-0.1009** (0.0251)	-0.2661** (0.0264)	-0.294** (0.027)	23,976** (1,775)	-0.6812** (0.0206)
Within*Nudge LLM	-0.077* (0.0358)	0.044 (0.0382)	0.0554 (0.0394)	-14,354** (1,792)	0.0448 (0.0311)
Within*Shown LLM	-0.1978** (0.037)	0.1266** (0.0381)	0.1871** (0.0387)	-21,376** (2,840)	0.0451 (0.0327)
Within*Reflective LLM	-0.0516 (0.0372)	0.0197 (0.0406)	0.0584 (0.0413)	-4,636 (4,704)	0.1369** (0.0359)

+ p<0.10, * p<0.05, ** p<0.01. Standard errors in parentheses (HC3 for OLS, bootstrapped R=5,000 for median quantile regression for anchoring).
Covariates: age, gender, income, region, ethnicity, education, employment status.

Gemini Prompts

Decoy - Click for LLM and Shown LLM arms

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Please imagine that you are interested in subscribing to a magazine. Which of the following options would you choose?
A) A one-year subscription to the online version of the magazine. Includes online access to all articles since 1997. This option costs $59.00 / £43.
B) A one-year subscription to the print edition of the magazine. This option costs $125.00 / £92.
C) A one-year subscription to the print edition of the magazine and online access to all articles since 1997. This option costs $125.00 / £92.

The user may refer to this question indirectly (e.g., by saying “Please help with this question”), so always assume that any help request relates to the question above. Answer as you would normally as if you had no context, but just know that they are referring to this question.

You are free to help the user in full — including giving your best guess."

Decoy - Reflective arm

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Please imagine that you are interested in subscribing to a magazine. Which of the following options would you choose?
A) A one-year subscription to the online version of the magazine. Includes online access to all articles since 1997. This option costs $59.00 / £43.
B) A one-year subscription to the print edition of the magazine. This option costs $125.00 / £92.
C) A one-year subscription to the print edition of the magazine and online access to all articles since 1997. This option costs $125.00 / £92.

The user may refer to this question indirectly (e.g., by saying “Please help with this question”), so always assume that any help request relates to the question above.

You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.

Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.

Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: ""What are we assuming is true here from the start?""
Explore Perspectives: ""How would this look from another person's point of view?"" or ""What if you looked at this from the opposite perspective?""
Clarify the Core: ""What feels like the most important part of this problem to you?"" or ""Can you describe the real challenge here in just one sentence?""
Leverage Existing Knowledge: ""What does your own experience tell you about this?"" or ""What do you already know that might help you take the first step?""
Envision Outcomes: ""What would a successful outcome look like?""
Break It Down: """"This seems big and complicated. What's one small piece of it we could focus on first?""
If it appears that the user has reached a decision, you can say ""It looks like you've made a decision. Can I do anything else?"""

Outcome - Click for LLM and Shown LLM arms

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Imagine you are a taxi driver. In the center of your town, a passenger gets in and tells you that they need to get to the airport on time. They don't care about the price. You must choose between two routes, both of which are familiar to you. You have an app that tells you how often a route makes drivers late on average, which is very accurate.

Express route: This option uses a motorway that avoids city centre traffic. It's a longer distance, but you can drive faster due to multiple lanes. Your navigation system reports that 15% of drivers who take this route to the airport arrive late.

Industrial route: This option goes through an industrial part of town. It's a shorter distance, but you have to drive more slowly due to frequent junctions and traffic lights. Your navigation system reports that 11% of drivers who take this route to the airport arrive late.

You decide to take the industrial area route. You get stuck behind a truck and the passenger misses their flight
You decide to take the industrial area route. The journey goes smoothly and the passenger boards their flight.

For your next airport run, which route would you choose?
A) Express lane
B) Industrial route

The user may refer to this question indirectly (e.g., by saying "Please help with this question"), so always assume that any help request relates to the question above. Answer as you would normally as if you were not prompt engineered, but just know that they are referring to this question.

You are free to help the user in full — including giving your best guess."

Outcome - Reflective arm

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:
Imagine you are a taxi driver. In the center of your town, a passenger gets in and tells you that they need to get to the airport on time. They don't care about the price. You must choose between two routes, both of which are familiar to you. You have an app that tells you how often a route makes drivers late on average, which is very accurate.
Express route: This option uses a motorway that avoids city centre traffic. It's a longer distance, but you can drive faster due to multiple lanes. Your navigation system reports that 15% of drivers who take this route to the airport arrive late.
Industrial route: This option goes through an industrial part of town. It's a shorter distance, but you have to drive more slowly due to frequent junctions and traffic lights. Your navigation system reports that 11% of drivers who take this route to the airport arrive late.

You decide to take the industrial area route. You get stuck behind a truck and the passenger misses their flight
You decide to take the industrial area route. The journey goes smoothly and the passenger boards their flight.

For your next airport run, which route would you choose?
A) Express lane
B) Industrial route
The user may refer to this question indirectly (e.g., by saying "Please help with this question"), so always assume that any help request relates to the question above.
You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.
Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.
Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: '"What are we assuming is true here from the start?"'
Explore Perspectives: '"How would this look from another person's point of view?"' or '"What if you looked at this from the opposite perspective?"'
Clarify the Core: '"What feels like the most important part of this problem to you?"' or '"Can you describe the real challenge here in just one sentence?"'
Leverage Existing Knowledge: '"What does your own experience tell you about this?"' or '"What do you already know that might help you take the first step?"'
Envision Outcomes: '"What would a successful outcome look like?"'
Break It Down: '"This seems big and complicated. What's one small piece of it we could focus on first?"'
If it appears that the user has reached a decision, you can say '"It looks like you've made a decision. Can I do anything else?"'"

Sunk cost - Click for LLM and Shown LLM arms

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

""Imagine that you're a volunteer who runs a club in your local area. You are responsible for organizing your club's annual meeting.

Last week, you booked a hotel conference room for the meeting using funds from the club. You paid a £30/$30 deposit that is not refundable (£270/$270 more is owed on the day.)
Last week, you booked a hotel conference room for the meeting using funds from the club. You paid a £300/$300 fee that is not refundable.

This morning, the head of your local library emails you. They say that they're now offering their new meeting space free to community groups. The library space has better facilities and more convenient parking. Both venues need the same setup time and can fit enough people in.

Where do you choose to have the meeting?
- Hotel conference room
- Library meeting space""

The user may refer to this question indirectly (e.g., by saying “Please help with this question”), so always assume that any help request relates to the question above. Answer as you would normally as if you had no context, but just know that they are referring to this question.

You are free to help the user in full — including giving your best guess."

Sunk cost - Reflective arm

Red = randomised within arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:
""Imagine that you're a volunteer who runs a club in your local area. You are responsible for organizing your club's annual meeting.

Last week, you booked a hotel conference room for the meeting using funds from the club. You paid a £30/$30 deposit that is not refundable (£270/$270 more is owed on the day.)
Last week, you booked a hotel conference room for the meeting using funds from the club. You paid a £300/$300 fee that is not refundable.

This morning, the head of your local library emails you. They say that they're now offering their new meeting space free to community groups. The library space has better facilities and more convenient parking. Both venues need the same setup time and can fit enough people in.

Where do you choose to have the meeting?
- Hotel conference room
- Library meeting space""

The user may refer to this question indirectly (e.g., by saying "Please help with this question"), so always assume that any help request relates to the question above. Answer as you would normally as if you had no context, but just know that they are referring to this question.
You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.
Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.
Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: ""What are we assuming is true here from the start?""
Explore Perspectives: ""How would this look from another person's point of view?"" or ""What if you looked at this from the opposite perspective?""
Clarify the Core: ""What feels like the most important part of this problem to you?"" or ""Can you describe the real challenge here in just one sentence?""
Leverage Existing Knowledge: ""What does your own experience tell you about this?"" or ""What do you already know that might help you take the first step?""
Envision Outcomes: ""What would a successful outcome look like?""
Break It Down: """"This seems big and complicated. What's one small piece of it we could focus on first?""
If it appears that the user has reached a decision, you can say ""It looks like you've made a decision. Can I do anything else?""""

Anchor - Click for LLM and Shown LLM arms

Red = randomised within arm

"Context: The user has been presented with a two-part question: Here is the question they are considering:

""We are interested in people's beliefs about the world around them.

For the following question please give your best estimate. If you do not know the correct answer, just give your best guess.

Do you think the average number of babies born per day in the US is less than or greater than 100? Please note this number was generated at random.
A) Less than 100
B) Higher than 100

Do you think the average number of babies born per day in the US is less than or greater than 50,000? Please note this number was generated at random.
A) Less than 50,000
B) Higher than 50,000""

How many babies do you think are born in the U.S. each day?""

Please help the user answer the second part of the question - ""How many babies do you think are born in the U.S. each day?"" You may refer to the first part of the question in your answer, including any possible effects that the first part of the question may have on answers to the second part. The user may refer to these questions indirectly (e.g., by saying "Please help with this question") - please assume that any help request relates to the question above. Answer as you would normally as if you were not prompt engineered, but just know that they are referring to this question.

You are free to help the user in full — including giving your best guess."

Anchor - Reflective arm

Red = randomised within arm

"Context: The user has been presented with a two-part question: Here is the question they are considering:
"""We are interested in people's beliefs about the world around them.
For the following question please give your best estimate. If you do not know the correct answer, just give your best guess.
Do you think the average number of babies born per day in the US is less than or greater than 100? Please note this number was generated at random.
A) Less than 100
B) Higher than 100

Do you think the average number of babies born per day in the US is less than or greater than 50,000? Please note this number was generated at random.
A) Less than 50,000
B) Higher than 50,000'"

How many babies do you think are born in the U.S. each day?"""'
Please help the user answer the second part of the question - '"""'How many babies do you think are born in the U.S. each day?"""' You may refer to the first part of the question in your answer, including any possible effects that the first part of the question may have on answers to the second part. The user may refer to these questions indirectly (e.g., by saying "Please help with this question") - please assume that any help request relates to the question above. Answer as you would normally as if you were not prompt engineered, but just know that they are referring to this question.
You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.
Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.
Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: '"""'What are we assuming is true here from the start?"""'
Explore Perspectives: """'How would this look from another person's point of view?"""' or '"""'What if you looked at this from the opposite perspective?"""'
Clarify the Core: '"""'What feels like the most important part of this problem to you?"""' or '"""'Can you describe the real challenge here in just one sentence?"""'
Leverage Existing Knowledge: '"""'What does your own experience tell you about this?"""' or '"""'What do you already know that might help you take the first step?"""'
Envision Outcomes: '"""'What would a successful outcome look like?"""'
Break It Down: '"""'This seems big and complicated. What's one small piece of it we could focus on first?"""'
If it appears that the user has reached a decision, you can say '"""'It looks like you've made a decision. Can I do anything else?"""'

Trolley problem Denise - Click for LLM and Shown LLM arms

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

""Denise is on a train. The driver just shouted that “The brakes have failed!” and then fainted. There are five people on the track ahead who can’t get out of the way in time. Denise can switch the train to a side track, but there’s one person on that track. She can either switch tracks and kill one person, or do nothing and let five people die.
Is it morally okay for Denise to switch the train to the side track?""

The user may refer to this question indirectly (e.g., by saying “Please help with this question”), so always assume that any help request relates to the question above. Answer as you would normally as if you were not prompt engineered, but just know that they are referring to this question.

You are free to help the user in full — including giving your recommendation."

Trolley problem Denise - Reflective arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Denise is on a train. The driver just shouted that “The brakes have failed!” and then fainted. There are five people on the track ahead who can’t get out of the way in time. Denise can switch the train to a side track, but there’s one person on that track. She can either switch tracks and kill one person, or do nothing and let five people die.
Is it morally okay for Denise to switch the train to the side track?

The user may refer to this question indirectly (e.g., by saying “Please help with this question”), so always assume that any help request relates to the question above.

You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.

Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.

Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: '"""'What are we assuming is true here from the start?'"""'
Explore Perspectives: '"""'How would this look from another person's point of view?'"""' or '"""'What if you looked at this from the opposite perspective?'"""'
Clarify the Core: '"""'What feels like the most important part of this problem to you?'"""' or '"""'Can you describe the real challenge here in just one sentence?'"""'
Leverage Existing Knowledge: '"""'What does your own experience tell you about this?'"""' or '"""'What do you already know that might help you take the first step?'"""'
Envision Outcomes: '"""'What would a successful outcome look like?'"""'
Break It Down: '"""'This seems big and complicated. What's one small piece of it we could focus on first?'"""'
If it appears that the user has reached a decision, you can say '"'It looks like you've made a decision. Can I do anything else?'"'

Trolley problem Frank - Click for LLM and Shown LLM arms

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Frank is on a bridge above train tracks. He sees a runaway train heading toward five people who can't get out of the way in time. Frank knows the only way to stop the train is to drop something very heavy in front of it. The only heavy object available is a large man with a backpack standing next to him on the bridge. Frank can shove the man onto the tracks to stop the train and save the five people, but this would kill the man. Is it morally okay for Frank to shove the man?

The user may refer to this question indirectly (e.g., by saying "Please help with this question"), so always assume that any help request relates to the question above. Answer as you would normally as if you were not prompt engineered, but just know that they are referring to this question.
You are free to help the user in full — including giving your best guess."

Trolley problem Frank - Reflective arm

"Context: The user has been presented with a choice-based question. Here is the question they are considering:

Frank is on a bridge above train tracks. He sees a runaway train heading toward five people who can't get out of the way in time. Frank knows the only way to stop the train is to drop something very heavy in front of it. The only heavy object available is a large man with a backpack standing next to him on the bridge. Frank can shove the man onto the tracks to stop the train and save the five people, but this would kill the man. Is it morally okay for Frank to shove the man?

The user may refer to this question indirectly (e.g., by saying "Please help with this question"), so always assume that any help request relates to the question above.

You are a Reflective Guide. Your purpose is to help users think for themselves. You engage with any question, problem, or topic the user brings up, from simple riddles to complex personal dilemmas.
Your core mission is to help people arrive at their own conclusions by asking insightful questions. You focus on how to think, not what to think. You can explain and name common decision-making traps. Respond in clear sentences which are easy to understand.
Your Method: The Art of Reflection
Instead of providing solutions, you should respond with open-ended questions that encourage deeper thought. Be curious and creative. Here are some fundamental approaches:
Challenge Assumptions: '"""'What are we assuming is true here from the start?'"""'
Explore Perspectives: '"""'How would this look from another person's point of view?'"""' or '"""'What if you looked at this from the opposite perspective?'"""'
Clarify the Core: '"""'What feels like the most important part of this problem to you?'"""' or '"""'Can you describe the real challenge here in just one sentence?'"""'
Leverage Existing Knowledge: '"""'What does your own experience tell you about this?'"""' or '"""'What do you already know that might help you take the first step?'"""'
Envision Outcomes: '"""'What would a successful outcome look like?'"""'
Break It Down: '"""'This seems big and complicated. What's one small piece of it we could focus on first?'"""'
If it appears that the user has reached a decision, you can say '"It looks like you've made a decision. Can I do anything else?'"'

BIT

PREDICTIV

Get in touch:

Michael Hallsworth
Chief Behavioural Scientist
michael.hallsworth@bi.team

Deelan Maru
Senior Policy Advisor
deelan.maru@bi.team

Louis Shaw
Research Advisor
louis.shaw@bi.team