April 13, 2026

Cheap talk and honesty: Follow-up experiment

Preregistration

Tilman Fries

LMU Munich

tilman.fries@econ.lmu.de

Daniel Parra

Pontificia Universidad Javeriana

danielfparra@javeriana.edu.co

1 Introduction

In a companion experiment (henceforth Experiment 1; see Fries and Parra, 2025), we studied how the possibility of honesty shapes strategic communication. In that experiment, a sender sends a possibly dishonest message to a receiver who wishes to guess a secret number. Across three treatments, we found broad support for the prediction that honesty considerations increase the informativeness of communication and the responsiveness of receivers. However, we also found excess noise in receiver guesses, with a guess distribution that is bimodal with peaks at the prior (4) and the message (m) for messages $m > 4$ above the prior.

These observations motivate the present follow-up experiment. When a receiver obtains message m, their optimal guess is

$$a^* = \chi \cdot m + (1 - \chi) \cdot 4,$$

where χ is the receiver's belief that the sender was honest. Arriving at a^* requires two steps: first, forming an accurate belief χ (an *inference* step), and second, computing the weighted average (a *prediction* step). Both steps could be sources of the observed noise. Additionally, the weighted average requires the receiver to reason about two contingencies—the sender was honest or strategic—simultaneously. Failures in contingent reasoning may impede optimal prediction even when beliefs are accurate (Martinez-Marquina, Niederle, and Vespa, 2019).

We design four treatments to isolate each of these channels. Expert-Rep replicates the Expert condition of Experiment 1 (re-using sender data from Experiment 1) and serves as a benchmark. Belief takes over the prediction step by having the computer calculate a^* from the receiver's stated belief, leaving only inference to the participant. FixBelief provides the receiver with information about the honesty rate among four past senders who sent the message observed by the receiver, one of whom who is matched to the receiver. This an informative signal about the sender's honesty, correcting potential misperceptions while keeping the prediction step in the receiver's hands. NoUncertainty builds on FixBelief but the sender now plays against all four senders at once—a framing that, following Martinez-Marquina et al. (2019), eliminates the need for contingent reasoning while preserving the same incentives. Together, the pairwise comparisons identify the contribution of prediction complexity, belief error, and contingent-reasoning failures to the noise observed in Experiment 1.

2 Experimental design

2.1 Overview

All four treatments use pre-recorded sender data from the Expert treatment of Experiment 1. There are therefore no live senders in this experiment. Each receiver participant is randomly matched to one Expert receiver from Experiment 1 and plays through that receiver's exact 24-round message sequence.

Before the main task, we inform all receivers that their messages come from participants in a previous experiment who could, in each round, either select any number freely or delegate the message to a computer that forwarded the true secret number.

Payoffs. The receiver is paid for accuracy according to the binarized scoring rule. The receiver's probability of winning a 4,000 Colombian pesos bonus as a function of their guess a and the secret number j is

$$\text{Probability of winning the bonus (in percent)} = 100 - \frac{100}{36}(j - a)^2.$$

For each participant, we randomly draw 12 payoff-relevant rounds (i.e., half of the total) without revealing which rounds are payoff-relevant until the end of the experiment.

2.2 Expert-Rep

Expert-Rep is a direct replication of the Expert treatment of Experiment 1. After observing the sender's message "To maximize your payoff, you should choose X," the receiver freely guesses any value for the secret number with two decimal places. No additional information about the sender's honesty is provided in any given round.

2.3 Belief

Belief builds on Expert-Rep by taking over the prediction step. After observing the sender's message, the receiver states a belief $\chi \in \{0, 1, \ldots, 100\}$ representing the percentage probability that the sender was honest in this particular round. The computer then calculates the guess

$$a = \frac{\chi}{100} \cdot m + \left(1 - \frac{\chi}{100}\right) \cdot 4$$

and submits it as the receiver's guess for payment. The formula is shown to participants before the main task begins. Since the binarized scoring rule is applied to the computer-generated guess a, which is a strictly monotone function of χ, truthfully reporting the receiver's actual belief χ is the unique optimal strategy: it generates the guess that maximizes the receiver's expected bonus. The receiver never enters a manual guess; only χ is submitted.

2.4 FixBelief

FixBelief builds on Expert-Rep by correcting potential belief errors. Before making their guess, receivers are shown a signal about the sender's honesty. Specifically, the receiver sees a screen stating: "Your sender is one of 4 senders who sent this message in the previous experiment. Out of these 4 senders, X delegated to the computer (i.e., forwarded the true number) while $4 - X$ freely chose to send this message."

The signal $X \in \{0, 1, 2, 3, 4\}$ is drawn from the mixture distribution

$$X \sim 0.3 \cdot \text{Binomial}(4,\ \hat{\pi}(m)) + 0.7 \cdot \text{Uniform}\{0, 1, 2, 3, 4\},$$

where $\hat{\pi}(m)$ denotes the empirical rate at which Expert senders in Experiment 1 delegated to the computer conditional on sending message m. This mixing is done in order to ensure that receivers observe all possible signal realizations with nontrivial probability for every message value.

After observing the signal, the receiver freely guesses any value with two decimal places. The binarized scoring rule is applied to the receiver's own guess, just as in Expert-Rep. Given the signal, the optimal guess is

$$a^*(X, m) = \frac{X}{4} \cdot m + \frac{4 - X}{4} \cdot 4.$$

2.5 NoUncertainty

NoUncertainty builds on FixBelief with one modification aimed at eliminating the need for contingent reasoning (Martinez-Marquina et al., 2019). The signal X is generated by the same mixture distribution as in FixBelief, but the receiver is now told that they are playing *against all 4 senders at the same time.*

In FixBelief, the receiver must reason about two contingencies—they are matched to an honest sender or a strategic sender—when computing the optimal response to the signal. In NoUncertainty, the receiver instead faces a concrete group of senders and can aggregate payoffs directly, avoiding the need for contingent reasoning. From a payoff-maximization standpoint, the two situations are mathematically equivalent: the same guess $a^*(X, m)$ is optimal in both cases. Any difference in behavior between FixBelief and NoUncertainty can therefore be attributed to the role of contingent reasoning.

2.6 Additional tasks and elicitations

Math task. Before the main game, all participants take part in a math quiz, to keep the experiment symmetric to the initial experiment.

Post-game belief elicitation. After the sender-receiver game we ask participants to estimate the overall likelihood with which Expert senders in Experiment 1 chose to delegate to the computer and the share of credulous receiver guesses (guessing 7 after receiving 7). These guesses are incentivized with a binarized scoring rule. We also add a cognitive uncertainty elicitation after each belief elicitation.

Strategy description. At the end of the experiment, all participants are asked to describe the strategy they used during the experiment in plain language following a similar procedure as in Arrieta and Nielsen (2024). We will then feed this strategy into ChatGPT, which, based on the strategy description responds to 10 random messages observed by this receiver in the experiment, with the instruction to replicate the receiver's decisions. When submitting their strategy description, we use a scoring rule to incentivize participants to provide a description that allows ChatGPT replicate their choices accurately.

2.7 Implementation

We aim to recruit 70 receivers per treatment arm for Expert-Rep, Belief, and NoUncertainty. For FixBelief, we oversample and aim to recruit 95 receivers, as importance weighting reduces the effective sample size to approximately 75% of the nominal sample size, thereby ensuring comparable statistical power across treatments. This yields a total sample of 305 receivers. Since no live senders are required, sessions consist solely of receivers.

3 Hypotheses

We specify our main hypotheses and plans for testing them below. All regressions are estimated via OLS. Standard errors are clustered at the individual receiver level.

Our main analysis sample consists of observations from rounds 13–24 of each receiver's 24-round session (the second half), focusing on behavior after participants have familiarized themselves with the task.

3.1 Measurement

Our primary outcome is the receiver's root mean squared error (RMSE) in a given round. To smooth out finite-sample noise, we work with an expected RMSE (E-RMSE), which integrates over the Expert senders' strategy. In particular, let χ_m denote the probability that an Expert sender delegated when sending message m.

Then,

$$\text{E-RMSE}_{ir} = \sqrt{\chi_{m_{ir}}(a_{ir} - m_{ir})^2 + (1 - \chi_{m_{ir}}) \sum_{j=1}^{7} \frac{1}{7}(a_{ir} - j)^2},$$

where a_{ir} is receiver i's guess in round r and m_{ir} is the corresponding secret number. This measure is closely related to the standard RMSE, $(a_{ir} - j_{ir})^2$, where j_{ir} is the secret number realized in round r for receiver i. However, in a finite sample this measure can easily be influenced by outlier realizations of j_{ir}, which is why we employ the E-RMSE.

To quantify the extent of bimodal guessing, we define a bimodal guess indicator:

$$\text{Bim}_{ir} = \text{I}(a_{ir} \in \{4, m_{ir}\}).$$

3.2 Main hypotheses

Hypothesis 1. *Simplifying the prediction step reduces receiver guessing error.*

We test this hypothesis by comparing Expert-Rep and Belief. In Belief, the computer performs the prediction step on the receiver's behalf, so a reduction in RMSE relative to Expert-Rep reflects the cost imposed by prediction complexity. We estimate the following regression using data from these two treatments:

$$\text{E-RMSE}_{ir} = \beta_0 + \beta_1 \text{I}(\textsc{Belief}) + \rho_r + \varepsilon_{ir},$$

where ρ_r are round fixed effects. We test Hypothesis 1 by checking whether $\beta_1 < 0$.

As a secondary test, we examine the rate of bimodal guesses among high messages. We estimate

$$\text{Bim}_{ir} = \beta_0 + \beta_1 \text{I}(\textsc{Belief}) + \rho_r + \varepsilon_{ir}$$

including all observations with $m_{ir} > 4$ and test whether $\beta_1 < 0$.

Hypothesis 2. *Correcting beliefs reduces receiver guessing error.*

We test this hypothesis by comparing EXPERT-REP and FIXBELIEF. In FIXBELIEF, receivers receive an informative signal about the sender's honesty before guessing, so a reduction in RMSE relative to EXPERT-REP reflects the cost imposed by incorrect beliefs. We estimate the following regression using data from these two treatments:

$$\text{E-RMSE}_{ir} = \beta_0 + \beta_1 \text{I}(\text{FIXBELIEF}) + \rho_r + \varepsilon_{ir},$$

where ρ_r are round fixed effects. We test Hypothesis 2 by checking whether $\beta_1 < 0$.

Because the signal X in FIXBELIEF is drawn from the mixture distribution rather than the true empirical distribution $\text{Binomial}(4, \hat{\pi}(m))$, the distribution of signals observed by FIXBELIEF receivers differs from what they would observe under the actual EXPERT sender strategy. We therefore reweight FIXBELIEF observations using importance weights

$$w(X, m) = \frac{\text{Binomial}(4,\ \hat{\pi}(m);\ X)}{0.3 \cdot \text{Binomial}(4,\ \hat{\pi}(m);\ X) + 0.7 \cdot \frac{1}{5}},$$

so that the reweighted signal distribution of FIXBELIEF matches the empirical honesty distribution of EXPERT senders. All FIXBELIEF regressions use these weights; EXPERT-REP observations receive a weight of 1.

As a secondary test, we examine the rate of bimodal guesses among high messages. We estimate

$$\text{Bim}_{ir} = \beta_0 + \beta_1 \text{I}(\text{FIXBELIEF}) + \rho_r + \varepsilon_{ir}$$

including all observations with $m_{ir} > 4$ and test whether $\beta_1 < 0$.

Hypothesis 3. *Eliminating the need for contingent reasoning reduces receiver guessing error.*

We test this hypothesis by comparing FIXBELIEF and NOUNCERTAINTY. Both treatments provide receivers with the same signal X and have the same optimal action. The only difference is the framing: NoUncertainty presents all 4 senders at once, eliminating the need to reason contingently about which type of sender one faces (Martinez-Marquina et al., 2019). We estimate

$$\text{E-RMSE}_{ir}^{X} = \beta_0 + \beta_1 \mathrm{I}(\textsc{NoUncertainty}) + \rho_r + \varepsilon_{ir}$$

using data from FixBelief and NoUncertainty, and test whether $\beta_1 < 0$. In this estimation, we will not reweigh observations as the signal distribution remains fixed across treatments. Also, the outcome here is an expected RMSE with weights equal to the honesty rate implied by the receiver's signal:

$$\text{E-RMSE}_{ir}^{X} = \frac{X_{ir}}{4}(a_{ir} - m_{ir})^2 + \frac{4 - X_{ir}}{4} \sum_{j=1}^{7} \frac{1}{7}(a_{ir} - j)^2.$$

Hence, the outcome precisely penalizes deviations from the optimal behavior given the realized signal.

As a secondary test, we examine the rate of bimodal guesses among high messages. We estimate

$$\text{Bim}_{ir} = \beta_0 + \beta_1 \mathrm{I}(\textsc{NoUncertainty}) + \rho_r + \varepsilon_{ir}$$

including all observations with $m_{ir} > 4$ and test whether $\beta_1 < 0$.

We will also run a heterogeneity analysis where we run both regressions, interacting the treatment coefficient with an indicator for whether the signal $X \in \{1, 2, 3\}$. This allows us to distinguish between fully revealing signals ($X = 0, 4$) and uncertain signals ($X = 1, 2, 3$). We expect the treatment effect to be concentrated among uncertain signals.

3.3 Additional analysis

To provide additional insight, we will decompose the receiver's RMSE in each treatment into (i) Average Bias, (ii) Within-Receiver Noise, (iii) Between-Receiver noise.

This allows us to assess the sources of guessing noise changes across treatments. We expect that each of the interventions reduces between-receiver noise.

We will also classify the written strategy descriptions into different categories using a LLM-classifier. To assess strategy replicability, we define a replication score as the share of receiver actions that the LLM could replicate, given the written strategy description. When comparing the replication scores across treatments, we expect that replicability is lowest in EXPERT-REP—i.e., when receiver noise is highest.

References

Arrieta, G. and K. Nielsen (2024): "Procedural decision-making in the face of complexity," Tech. rep.

Fries, T. and D. Parra (2025): "Cheap Talk and Honesty," *AEA RCT Registry*.

Martinez-Marquina, A., M. Niederle, and E. Vespa (2019): "Failures in Contingent Reasoning: The Role of Uncertainty," *American Economic Review*, 109, 3437–3474.