Additional Information: High salience experiment

Author(s)

Submission
1) Have any data been collected for this study already?
No

2) What's the main question being asked or hypothesis being tested in this study?

We examine whether the process of transmitting information through speech distorts the information and how such distortions differ between different kinds of information content. In this experiment, we increase the salience of the distinction between information about the level and reliability of the prediction as well as the incentives to transmit both kinds of information. We hypothesize that in this high salience set-up, differential information loss will decrease, meaning the degree of loss of level and reliability information will be relatively similar to each other.

3) Describe the key dependent variable(s) specifying how they will be measured.

Our experiments involve transmission of information about two unknown states: (i) the change in home price growth in a US city and (ii) the change in revenue growth of a US retailer, both for the upcoming year.

Each state has three key dependent variables associated with it: the respondent's belief movement about the unknown state (difference between posterior, elicited after the respondent listens to a message about the state, and prior, elicited beforehand); the respondent's belief about the message originator's prediction about the state of the world; and the respondent's belief about the originator's reliability, as measured on a scale ranging from 0 (extremely unreliable) to 100 (extremely reliable).

For the state beliefs, we will drop outliers (as specified below), and the question entry box forces answers to be between -100% and +100%. For all three outcomes, we will z-score within topic*manipulation type quadrants and then pool together the observations from both topics.

All three beliefs are incentivized through random incentives. Respondents will be told at the start that 1 in 10 respondents will be randomly chosen to be eligible for a bonus payment and have one of the incentivized tasks be paid out according to the formulas below.
  (a) Beliefs about the true state of the world are incentivized with the following formula:
      (i)   Probability of winning $20 [in %] = 100 − 10*(Estimate [in %] - True state of the world in 12 months [in %])^2
  (b) Beliefs about the message originator's beliefs and reliability are incentivized as follows:
      (i)   For 50% of respondents, there will be no incentive and the question will be phrased as a direct question about the message originator's beliefs.

     (ii)   For 50% of respondents, the question will be phrased as a second-order question ("your job is to predict what people would on average respond to the direct question") and responses will be incentivized with the following formula:
- (1) For beliefs about the originator's prediction: Probability of winning $20 [in %] = 100 − 10*(Response [in %] - Average response to direct question [in %])^2
- (2) For beliefs about the originator's reliability: Probability of winning $20 [in %] = 100 − 2*(Response - average response to the direct question)^2

If a respondent is selected to be eligible for a bonus, one of the incentivized beliefs of that respondent will be randomly selected to be the one that counts for payment.

4) How many and which conditions will participants be assigned to?

The full experiment comprises two separate data collections that build on each other, a transmitter experiment and a listener experiment. The two experimental collections rely on different respondent samples.

Transmitter experiment:
Participants listen to two short recordings played consecutively and without a break, each one of an opinion piece providing a qualitative narrative about the future path of a different economic variable. Then, they record their own summary of these recordings, separately for the first and second variable.

A randomly chosen 50% of transmitters will be asked their prior belief about each variable before hearing the recordings, and all transmitters will be asked for the three beliefs described above after recording their transmitted message for each topic.

Recording treatment arms:
Within each topic, we randomize three key features of the original recordings:

Level of variable: We randomize whether the piece argues for an increase or a decrease in the level of the variable.

Reliability of message: Second, we randomize the reliability of the original message. We randomly assign respondents to one of two different types of reliability manipulations:
- Naturalistic (combination of explicit statements about confidence, source quality and speaker competence, as well as implicit markers of reliability): Respondents in the naturalistic condition are assigned to one of the following 2 conditions: (i) Strong reliability; (ii) Weak reliability.
- Modular (Insertion of explicit markers indicating high or low reliability (e.g., definitely vs. possibly, will vs. might, etc.): Respondents in the modular condition are assigned to one of the following 3 conditions: (i) Strong reliability; (ii) No reliability markers; (iii) Weak reliability.

Sex of transmitter voice: We randomize whether the recording is a male voice or a female voice. This is not a focus of analysis and we randomize simply for symmetry.

Randomization is stratified: each transmitter hears two recordings, one with an "increase" and one with a "decrease," one with "strong reliability" and one with "weak reliability," and one with a male voice and one with a female voice. Then, if exactly one of the two topics is in the modular condition, that topic has a 33% chance of getting switched to "no reliability markers." If both topics are in the modular condition, there is a 66% chance that one of the two topics is randomly switched to "no reliability markers."

Respondents receive incentives for transmitting all information contained in the original messages. Respondents are informed that one in 10 people will be selected for bonus eligibility and that, if selected, a different group of participants will score their recordings on a scale of 0 to 10, where 0 corresponds to "Nothing conveyed in meaning" and 10 corresponds to "Everything conveyed in meaning". If the average score their recordings receive is at least an 8, they will receive a $20 bonus payment. Respondents are explicitly told that the other participants will answer two questions, one about the point forecast implied by the message and one about the reliability of the message. In particular, they receive the following instruction: "*The other participants will answer two questions about your voice message. How accurately was the level of the speaker's prediction conveyed in the voice message? How accurately was the reliability of the prediction conveyed in the voice message?*"

To increase the salience of reliability information we add three features compared to our baseline design:
- We add a comprehension question about the incentives facing respondents.
- *Just before the recording we ask respondents: "*What do you have to pass on well to maximize your chances of receiving a bonus? Tick all that apply" with the following response options: (i) level of the speaker's prediction; (ii) reliability of the speaker's prediction.
- On the screen on which people make the recording we add the following reminder: "Remember: Your bonus payment is based equally on how well you pass on **both** of the following:
  - The **level** of the speaker's prediction.
  - The **reliability** of the speaker's prediction"

Listener experiment:

This involves a separate set of respondents. For each of the two topics, respondents first state their prior belief about the outcome variable of interest and then listen to a recording about the variable. As before, the order of the topics is randomized. For each topic, respondents are randomly matched to a transmitter and listen either to the same original recording as the one the transmitter heard, or that transmitter's transmitted recording. There is a 30% chance of hearing the original and 70% chance of hearing the transmitted recording.

After listening to a recording, respondents are incentivized to forecast the future development of the variable as well as to guess the prediction of the message originator and the reliability of the original message. (The same 3 outcomes described above, incentivized in the same way).

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

Our three key outcome variables are belief movements about the state of the world (posterior minus prior), beliefs about the message originator's prediction, and beliefs about the message originator's reliability. Our main analyses compare the sensitivity of these outcomes to the *increase prediction/decrease prediction* and *strong reliability/weak reliability* randomizations, between the original and transmitted recordings.

Specifically: to examine distortions of information about the level of a variable, we compare the following two quantities: (1) the difference between average beliefs among listeners who heard **original** recordings in the "increase" condition and average beliefs among listeners who heard **original** recordings in the "decrease" condition, and (2) the difference between average beliefs among listeners who heard **transmitted** recordings in the "increase" condition and average beliefs among listeners who heard **transmitted** recordings in the "decrease" condition. By "belief" here, we mean *the belief about the message originator's prediction*, not *the listener's own belief about the state of the world* (because the latter is a function of transmission of both level and reliability information).

Similarly, to examine distortions of information about the reliability of a message, we will conduct an identical comparison except using listeners' beliefs about the reliability of the message originator and comparing between the strong reliability/weak reliability conditions.

Finally, to examine distortions in downstream beliefs about the world, we will compare listeners' average belief movements (posterior minus prior) between original and transmitted messages across our four key quadrants (decrease + strong reliability, decrease + weak reliability, increase + weak reliability, increase + strong reliability). In supplementary analyses focusing on the Modular manipulation, we will add the intermediate sextants "decrease + no reliability information" and "increase + no reliability information".

We will also examine the transcripts of transmitted messages in the following ways.
- Using handcoding and prompting of a large language model (such as GPT-4) we will classify each transcript according to whether it contains explicit statements about the original message's level prediction and the original message's reliability. We will repeat the first analysis above separately by whether the script contains an explicit level statement or not, and the second analysis separately by whether the script contains an explicit reliability statement or not.
- For the modular scripts only, we will count the number of uncertainty/weak-reliability denoting terms and certainty/strong-reliability denoting terms in the original versus corresponding transmitted scripts. We will check if listener reliability beliefs are correlated with the number of uncertainty or certainty denoting terms in the transmitted transcripts, and check if controlling for the number of these terms eliminates any difference between the transmitted and original recordings in the second analysis step. Again, we will use handcoding and prompting of large language models.