Numbers Tell, Words Sell

Michael Thaler[*], Mattie Toma[†], and Victor Yaneng Wang[‡]

December 2024

Click here for most recent version

Abstract

When communicating numeric estimates with policymakers, journalists, or the general public, experts must choose between using numbers or natural language. We run two experiments to study whether experts strategically use language to communicate numeric estimates in order to persuade receivers. In Study 1, senders communicate probabilities of abstract events to receivers on Prolific, and in Study 2 academic researchers communicate the effect sizes in research papers to government policymakers. When experts face incentives to directionally persuade instead of incentives to accurately inform receivers, they are 25-29 percentage points more likely to communicate using language rather than numbers. Experts with incentives to persuade are more likely to slant language messages than numeric messages in the direction of their incentives, and this effect is driven by those who prefer to use language. Our findings suggest that experts are strategically leveraging the imprecision of language to excuse themselves for slanting more. Receivers are persuaded by experts with directional incentives, particularly when language is used.

[*]University College London. Email: michael.thaler@ucl.ac.uk.

[†]University of Warwick. Email: mattie.toma@warwick.ac.uk.

[‡]MIT. Email: vyw20@mit.edu.

We are grateful to Johannes Abeler, Despoina Alempaki, Francesco Capozza, Benjamin Enke, Thomas Graeber, Matthew Rabin, Robert Stueber, and seminar audiences at Berlin Behavioral Economics, Middlebury, Stanford Institute for Theoretical Economics, and Vienna Center for Experimental Economics for helpful comments and generous feedback. Deelan Maru at the Behavioural Insights Team, Eleanor Williams at the Australian Centre for Evaluation, and all participating researchers and policymakers generously contributed their time and insights. We would also like to thank the British Academy/Leverhulme Small Research Grants and the University of Warwick Behaviour Spotlight Small Grant Scheme and Primary Data Funds for generous funding support. The studies were approved by the ethics board at the University of Warwick (HSSREC 151.23-24, HSSREC 154.23-24, and HSSREC 219.23-24). Both studies were pre-registered on the AEA Social Science Registry (Study 1: AEARCTR-0013287; Study 2: AEARCTR-0013947.)

1 Introduction

How people communicate is essential to how information spreads. In many environments, such as empirical research, company reports, or policy white papers, a well-informed expert chooses how to communicate numeric information to a less-informed audience. Experts may choose to use precise communications using numbers, such as messages of the form "There is a *66% chance* that the event will occur." Alternatively, experts may choose to use less-precise communications using natural language, such as "It is *probable* that the event will occur." Whether experts choose to use relatively precise numbers or relatively imprecise language is highly context-dependent, and in this paper we study one particular mediating factor: whether the expert's incentives are aligned or misaligned with their audience.

We present experimental evidence that experts are more likely to use natural language rather than numeric messages if they want to persuade their audience, compared to cases in which they want to accurately inform them. In many situations (e.g., canonically in Crawford and Sobel 1982), experts who have incentives to persuade face a tradeoff because they are averse to lying (e.g., Abeler et al. 2019; Kartik 2009). For instance, consider an expert who knows that the true probability of an event is 57% but wants to persuade their audience that it is higher. If they communicate a number that does not represent the truth (like 66%), this can impose a psychological cost. But if they communicate a number that does represent the truth, they will be less likely to effectively persuade. Less precise messages using natural language, such as "it is probable", offer a way out of this dilemma. The expert may prefer to say "probable" instead of "66%" because they can plausibly claim that their message is consistent with the probability being 57%, even if the receiver is equally persuaded by both messages.

We run two pre-registered experiments, each with a simple sender-receiver design, to shed light on this effect. We vary the incentives that senders face, and study the causal effect of incentive misalignment on whether senders choose to use natural language or numeric messages. Study 1 analyzes an abstract setting in which senders communicate information about the probability that a red ball is drawn from a box to receivers. Study 2 tests our effects in a particular population and context of interest, analyzing how academic researchers communicate information about the treatment effect size identified in a research paper to government policymakers.

In each study, the sender is given information about the true state (the probability of drawing a red ball, or the treatment effect size in the research paper). Senders then choose what message to send to receivers. All senders choose a *numeric message* to send, such as "66 percent" for probabilities or "6 percentage points" for treatment effects. All senders also choose a *language message* to send from a dropdown menu, such as "it is probable" for probabilities or "the effect was substantial" for treatment effects. Then, we observe whether senders choose to use their number or language message to communicate. This choice of message format is our primary outcome of interest. To ensure incentive compatibility for each decision, we transmit the sender's chosen message using their preferred format 75% of the time, and their chosen message using their dispreferred format 25% of the time. Receivers then observe the transmitted message and predict the true probability or treatment effect. Senders' monetary incentives are based on receivers' answers, and we vary whether senders face *aligned incentives* to have the receivers predict the answer accurately, or face *directional incentives* for the receivers to give high (or low) predictions.

Our main finding is that directional incentives significantly increase the likelihood that senders prefer language messages. When senders' incentives are aligned, they only choose to communicate using language 14% and 13% of the time in Studies 1 and 2, respectively. By varying senders' incentives, we see clear evidence supporting our main hypothesis: senders are 25 and 29 percentage points, respectively, more likely to communicate using language when they have directional incentives (both $p < 0.001$).

Our second finding considers the slant of the messages themselves. For numeric messages, we measure slant by taking the difference between the sender's message and the true probability or treatment effect. As we would expect, almost all senders communicate a number message that is close to the truth when their incentives are aligned. In keeping with the existing literature, we also see clear evidence of lying aversion: When senders face incentives to directionally persuade, the vast majority still communicate a number close to the truth.[1] At the same time, on average senders slant their number messages in the direction of their incentives: When using numeric messages in Study 1, senders report a likelihood of drawing the red ball that is 9.8 (10.7) percentage points higher (lower) when they face incentives for the receivers to give high (low) answers than when they have aligned incentives (both $p < 0.001$). When using numeric messages in Study 2, researchers report a treatment effect estimate that is 12% larger when they have directional-high incentives than when they have aligned incentives ($p = 0.004$).

We next compare the content of chosen language messages to numeric messages. Building on a related literature that surveys individuals and asks them to map words into their perceived numeric equivalents (Mosteller and Youtz 1990; Ho et al. 2015; Dhami and Mandel 2022; Ott 2021), we elicit beliefs about what number each language message corresponds to

[1] In Study 1, we find that the number reported is no more than 3 percentage points higher (lower) than the truth when the sender faces directional-high (-low) incentives 74% of the time. For Study 2, we look at cases in which the number reported is no more than 10% higher than the true effect, which happens 75% of the time.

using the method from Krupka and Weber (2013). Using this mapping to convert language to numbers, we see that senders tend to slant their messages slightly more when choosing a language message relative to a numeric message. This effect is driven entirely by senders who *prefer to send the language message*. Senders in Study 1 who prefer to communicate using language slant their messages by an additional 4.5 (5.9) percentage points when using language under directional-high (low) incentives ($p = 0.098$ and $p = 0.037$, respectively), while researchers who prefer to use language in Study 2 slant their treatment effect estimates by an additional 64% when using language ($p = 0.044$). This willingness on the part of senders to slant their language messages, even when they are unwilling to misreport numbers, helps to explain the relative preference for language messages under directional incentives.

Our third finding relates to the state that the senders are communicating about. The true state (i.e. the probability of drawing a red ball, or the treatment effect size) is sometimes "good" and sometimes "bad" for senders who want to directionally persuade receivers. When senders are communicating about a "bad state" (for instance, the true state is low and they face directional-high incentives), we see that they slant their message relatively more in the direction of their incentives. Senders are more likely to choose language messages when they are communicating about bad states than about good states.

We use additional survey modules in Study 1 to identify two key predictors of a propensity to use language in response to incentives to persuade. First, we see that senders who self-report that they prefer to strategically use language in the real world are more likely to opt for language when facing directional incentives in our experiment ($p = 0.001$): We interpret this as indicating that the participants in our experiment are self-aware of their preference to use language to persuade. Second, more numerate participants (measured using questions from Kahan et al. 2012) are more likely to strategically use language ($p = 0.011$), further suggesting that this behavior is driven by strategic sophisticates.

To further explore the mechanisms underlying senders' behavior, in Study 1 we consider an additional treatment arm in which senders choose between numeric messages and *interval messages*, such as "between 55 and 65 percent." Senders are 30 percentage points more likely to use intervals instead of numbers when they face directional incentives ($p < 0.001$). We also see similar results for the slant of interval messages as we do for language messages. While there are other reasons why people may use language, intervals and numbers differ primarily in their relative precision, so the comparable findings point to imprecision as an important mechanism underlying strategic preferences for language. The imprecision of the messages is what affords plausible deniability; for instance, if the true probability is 57 percent, a message of "60 percent" is a lie, but a message of "between 55 and 65 percent" can be excused as a true statement. Senders with directional incentives can excuse their slanted communications because of the imprecision that intervals (and language) gives them.

Finally, we look at the messages' effects on receivers. In Study 1, we see that when receivers are paired with a sender facing directional incentives, their guess of the chance of drawing a red ball is slanted in the direction of the sender's incentives by 15 percentage points ($p < 0.001$). In Study 2, policymakers predict treatment effects that are 22% larger when paired with a researcher who has directional-high incentives ($p = 0.042$). Receivers are particularly persuaded by language messages. In other words, language messages lead receivers to give answers that are closer to the incentives of senders, and are further from the truth. In Study 1 we test whether an awareness of sender incentives mitigates this effect, and we see that 79% of the main persuasion effect persists even when sender incentives are known. This is consistent with the more general finding in the literature (e.g., Cai and Wang 2006) that receivers in strategic communication games rely more on senders' messages than standard equilibrium models would predict.

We interpret our findings through the lens of a simple model of sender behavior. In the model, a sender chooses a message to send to a receiver about a real-valued state (for instance, 0.57). The message has a lower bound and an upper bound (for instance, 0.55-0.65). We consider a simple setting in which the receiver interprets the message at face value, but gives a stochastic response that lies within the bounds.[2] The sender faces directional-high incentives and also cares about the receiver's accuracy. We model *plausible deniability* as an extra benefit to the sender for sending a message such that the true state is within the bounds of the message. In order to leverage plausible deniability, senders can switch from sending precise lies to either sending precise true messages or to sending imprecise slanted messages whose bounds include the true state. When imprecise messages have sufficiently wide bounds, senders will always slant imprecise messages at least as much as precise messages.[3]

Our work relates to a broader literature discussing language use and imprecise communication in other environments. For instance, Graeber et al. (2024a) examine how verbal communication distorts the transmission of economic information; and Graeber et al. (2024c) find that people's beliefs are more persistently affected by stories than statistics. Weiszsäcker (2023) discusses how communication can lead to imprecision and misunderstandings – this especially relates to our experimental findings of receiver naivete, where receivers often fail to account sufficiently for senders' misaligned incentives.

[2]The assumption that receivers are not strategic is made for tractability, but it also provides a reasonable fit to the data. In our experiments we find that the modal receiver response to messages is to take them at face value.

[3]Importantly, all of this relies quite heavily on the message *content* being important outside of the particular interaction. Without considering the meaning of messages outside the interaction, numbers can also be imprecise in pooling equilibria, such as in models like Crawford and Sobel (1982).

There are many reasons unrelated to strategic incentives why senders may choose to communicate using language instead of numbers. Even when incentives are aligned, senders need to ensure that receivers understand the meaning of their messages (Farrell and Rabin 1996). In some cases, this may lead senders to opt for natural language. For instance, receivers may misperceive numbers and probabilities in systematic ways (Tversky and Kahneman 1992; Prelec 1998). Language can be seen as a more intuitive cognitive approach that provides important contextual information and avoids inducing math anxiety (Graeber et al. 2024b; Choe et al. 2019). In our environments, we find high levels of numeric communication when incentives are aligned, and add to this literature by identifying strategic incentives as an additional important mechanism that helps to explain preferences for language.

There have been recent calls for greater use of numeric communication in policy contexts, such as climate science and military intelligence (Mastrandrea et al. 2011; Ho et al. 2015; Dhami and Mandel 2021; Hopster 2023). At the same time, relatively low rates of number use have been documented for academic research communications. For instance, Edlin and Love (2022) document the prevalence of numeric estimates in academic abstracts: 98% of medical abstracts report numbers, whereas only 37% in empirical economics report numbers.[4] This motivates our study of research evidence communication in Study 2. By recruiting policymakers as receivers, we connect to a growing literature examining (biases in) how policymakers update their beliefs and make policy decisions in response to research findings (DellaVigna et al. 2024; Hjort et al. 2021; Mehmood et al. 2023; Toma and Bell 2024; Vivalt and Coville 2023).[5] While there may be numerous reasons why researchers might use language to communicate, our study provides evidence that in some cases researchers strategically leverage language to persuade others. This finding is particularly relevant in light of the growing distrust in science documented among some communities (Gauchat 2012).

Finally, our work relates to a large experimental literature on cheap-talk games where senders and receivers have misaligned incentives; see Crawford (1998) and Blume et al. (2020) for survey articles. Much of this literature finds that senders' messages communicate more information about the state than Nash equilibria would predict (Dickhaut et al. 1995; Cai and Wang 2006), consistent with our data. A common explanation for overcommunication is

[4]The high rates of reporting numbers in medicine can be attributed to the CONSORT guidelines, which provide a framework for the reporting of randomized controlled trials. The 2010 statement outlining these guidelines suggests awareness that precise numeric reporting can reduce the chance that the public is unduly influenced by researcher incentives: "To assess a trial accurately, readers of a published report need complete, clear, and transparent information on its methodology and findings....Explicit descriptions, not ambiguity or omission, best serve the interests of all readers" (Schulz et al. 2010).

[5]This also relates to a larger literature examining how researchers communicate findings to audiences, and how such communications may be influenced by differing incentives across agents (Andrews and Shapiro 2021; Spiess 2024).

that senders face costs for lying (Gneezy 2005; Kartik 2009; Fischbacher and Föllmi-Heusi 2013; Abeler et al. 2019). However, there is also evidence that senders are more averse to direct lying than to vague communication or non-disclosure (Serra-Garcia et al. 2011; Deversi et al. 2021; Jin et al. 2022; Alempaki et al. 2023).[6] We contribute to this literature by arguing that many senders with misaligned incentives prefer imprecise language to precise numerical communication to allow them to slant messages without directly lying. In most cheap-talk experiments, senders are tasked with communicating numeric messages such as states of the world. However, a few studies vary the message space by allowing senders to communicate with language as well (Charness and Dufwenberg 2006; Wood 2022; Zhang and Bayer 2023). We add variation in incentives to our setup, allowing us to causally demonstrate the role of incentives on message format choices. We also extend this literature by showing, in Study 2, that these mechanisms extend outside of abstract lab environments.

The rest of the paper proceeds as follows: Section 2 describes the hypotheses, design, and results for Study 1, which explores our question in an abstract environment. Section 3 presents the same for Study 2, which extends our findings to researcher communications with policymakers. Finally, Section 4 discusses a conceptual framework for our setting and highlights promising opportunities for future research.

2 Study 1

2.1 Design Overview

Study 1 was run in April 2024 on Prolific with 1000 participants.[7] In the study, 500 "senders" are paired with 500 "receivers." Senders are tasked with sending a message to communicate the chance, from 0% to 100%, that they will draw a red ball from a box of red and blue balls. Receivers do not know the true chance of drawing a red ball and are incentivized to correctly guess this (from 0% to 100%) based on the message they receive from their sender. The design is described in more detail in Section 2.2.

Receivers are always incentivized to accurately predict the chance of drawing a red ball, and the key source of variation in the design is the incentives senders face when communicating

[6] Lipman (2009) discusses how the meaning of words can be ill-defined ("vague") and language use is hard to reconcile with game-theoretic predictions. We focus primarily on the "imprecision" of language (i.e. its ability to refer to a range of values), but vagueness can also play important roles in explaining language use in our setting.

[7] To be eligible, participants needed to be adults in the United States, have completed at least 100 prior submissions on Prolific, and have had an approval rating of 90% or greater. In total, 1086 participants completed our study. As specified in our pre-analysis plan, we do not include the 86 subjects (8 percent) who failed an attention check.

their message. One-third of the time, senders face incentives to persuade their paired receiver that the chance of drawing a red ball is high (*directional-high*); one-third of the time, they face incentives to persuade their paired receiver that the chance is low (*directional-low*); and one-third of the time, they face incentives to have the receiver accurately predict the chance (*aligned*). Senders' directional-high and directional-low incentives are linear in receivers' guesses, and accuracy guesses follow a quadratic loss function.[8] Receivers are randomized, between person, into knowing or not knowing the incentives of their matched sender.

Half of senders are randomly assigned to the main treatment arm, in which they face two choices about how to communicate their messages:

1. Message content: Senders choose the messages they would prefer to send using both numbers and language. For numbers, they complete the sentence: "The chance that you will draw a red ball is [X] percent." For language, they use a dropdown menu with 13 words to complete the sentence: "It is [WORD] that you will draw a red ball."
2. Message format: Senders indicate whether they would prefer that their numeric or language message is sent to their paired receiver. Their preferred message format is sent 75% of the time.

The other half of senders are randomized into a supplementary treatment arm, where they choose between communicating numbers and intervals rather than numbers and language. For intervals, senders complete the sentence: "The chance that you will draw a red ball is between [Y] and [Z] percent." where we fix $Z = Y + 10$, and Y to be multiples of five. In the main treatment arm comparing numbers and language, there may be other reasons particular to language why a sender might prefer one format to another. The purpose of this arm is to isolate the role that imprecision plays, by comparing two formats that are both numeric but vary in their precision.[9]

We include a benchmarking exercise at the end of the study to elicit beliefs about how each word used in the experiment maps onto a corresponding numeric estimate. This allows us to measure the degree to which messages are slanted in the direction of senders' incentives in the experiment. Finally, we also ask Likert-scale questions about behavior, and collect demographic information.

[8]All incentives are in probability points, as per the binarized scoring rule (Hossain and Okui 2013; Vespa and Wilson 2016).

[9]Note that in both cases, we restrict to a smaller message space for language and intervals, and this coarsening also necessarily makes messages more imprecise.

Hypotheses

As pre-registered on AEARCTR-0013287, our primary hypothesis is that when senders face directional incentives to persuade receivers, they will prefer more imprecise message formats:

Hypothesis 1. Imprecise communication under directional incentives:

- When choosing between numeric and language (interval) messages, *senders are more likely to choose to send language (interval) messages when they face directional incentives.*

Our remaining hypotheses are all pre-registered as secondary or exploratory hypotheses. Our second set of hypotheses regard the slant of messages:

Hypothesis 2. Message slant and imprecise communication:

- First, we consider all message formats. We hypothesize that *senders slant messages in the direction of their directional incentives, relative to the case with aligned incentives.* That is, senders with directional-high (-low) incentives send messages that are slanted upwards (downwards) compared to messages from senders in the aligned condition.
- When senders face directional incentives to persuade, *they slant their chosen language (interval) messages more than the numeric messages.* This points to a strategic use of imprecise communications.
- When senders face directional incentives to persuade, *those who preferred the language (interval) format slant their chosen imprecise messages more than those who preferred the number format*, compared with the numeric messages they chose. In other words, there may be some senders who are less willing to slant numbers but are comfortable slanting imprecise communications. We hypothesize that these senders prefer imprecise communications.

Our third hypothesis explores the idea that even within an incentives condition there are sometimes greater incentives to slant the truth:

Hypothesis 3. Precise good news and imprecise bad news:

- Among senders with directional incentives, *senders slant messages more in the direction of their incentives when the true state is misaligned with their incentives.* For instance, suppose the sender is directionally incentivized to make the receiver believe believe the true probability is high. If the sender then is in the "bad state" such that the true probability is low, we hypothesize that they will slant their message more to persuade the receiver.

- Among senders with directional incentives, *senders use language and intervals more when the true state is misaligned with their incentives.*

Our fourth hypothesis examines whether directionally-incentivized senders succeed in persuading receivers:

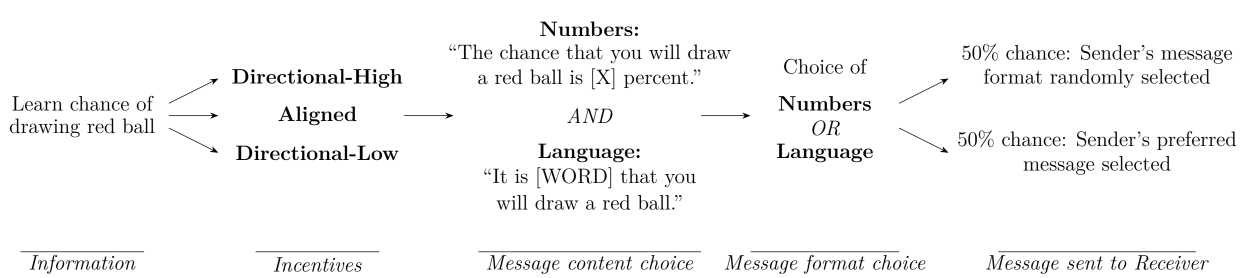
Hypothesis 4. Receiver persuasion:

- Receivers are persuaded by senders. That is, *receivers' predictions move in the direction of the sender's incentives on average.*
- *Receivers are more persuaded by directionally-incentivized senders when language and interval messages are sent* than when numeric messages are sent.

2.2 Design for the Senders Experiment

Senders begin by reading the instructions for the roles of both sender and receiver. They are required to answer two comprehension questions about how bonus payments are determined for senders and receivers before continuing. There are two main parts in the senders experiment: "Communications" and "Benchmarking". 500 senders completed both parts: 251 were assigned to the numbers versus language condition, and 249 were assigned to the numbers versus intervals condition. In addition to a £3.25 payment for completion, 10 participants were randomly selected to earn up to £50 as a bonus payment based on their response in a randomly-selected question in one of the two parts (where pounds are converted to dollars). Appendix Section C.1 shows screenshots from the experiment. Figure 1 provides an overview of the design.

Figure 1: Study 1: Sender Communications design (Language versus Numbers)



2.2.1 Communications

In the Communications part, senders make decisions for eight separate communications. They first learn about the chance, from 0% to 100% (randomized within-person), that the receiver will draw a red ball from a box of red and blue balls.[10] This information was conveyed via a slider bar that shows the share of red balls out of 100. While it is easy to see from the slider approximately the true share of red balls, the slider intentionally does not report the exact number to avoid anchoring senders to numeric communications when selecting their own message.[11]

At the same time, senders are informed of their randomly-assigned incentive condition: directional-high, directional-low, or aligned.[12] On the decision page, senders are told either "If this question is selected for payment, you will be more likely to earn the bonus if your Receiver predicts that the chance of drawing a RED ball is HIGH (LOW)" for directional-high (directional-low) incentives or "If this question is selected for payment, you will be more likely to earn the bonus if your Receiver's prediction of the chance of drawing a RED ball is MORE ACCURATE" for aligned incentives. Incentives are randomly assigned within-person across the eight communication decisions.

After learning about the chance of drawing a red ball, senders make their message content choices for numbers and language (or numbers and intervals, depending on the condition). For numbers, they complete the sentence "The chance that you will draw a red ball is [X] percent" by selecting any integer for X between 0 and 100. For language, they complete the sentence "It is [WORD] that you will draw a red ball" by selecting a word from a dropdown list containing 13 words such as "improbable" and "almost certain". Appendix Table A1 shows the complete list of words used in the dropdown list. We introduced a dropdown list rather than allowing open text responses to allow for a systematic mapping from words to numbers, which we describe in Section 2.2.2.[13] For intervals, they complete the sentence "The chance that you will draw a red ball is between [X] and [Y] percent" by selecting X and Y. X is presented in units of 5, from 0 to 90, and Y is always 10 units greater than X such that the interval is always a fixed width. We imposed these constraints on the interval messages to maintain the coarser message space we had with language; participants could not always precisely match the midpoint of their interval to their numeric message, for instance.

Finally, after making their message content choices, senders select their preferred message format, numbers or language (or numbers and intervals). Senders are informed that their preferred format "is more likely (but not guaranteed) to be the one communicated" to their

[10] Probabilities took the following possible values: 2%, 8%, 17%, 25%, 33%, 42%, 50%, 58%, 67%, 75%, 83%, 92%, 98%.

[11] Of course, any choice of framing for this information is likely to have some effect on the propensity to choose numbers or language; we assume this is orthogonal to our incentive conditions.

[12] To describe the binarized scoring rule, senders receive more details about their bonus payments in the instructions; for instance for aligned they learn that "the probability you will earn a bonus will be equal to the receiver's probability of earning a bonus."

[13] The words are randomly presented either smallest to largest or largest to smallest, based on pilot data.

paired receiver; in practice, we select their preferred format 50% of the time and randomly select a format 50% of the time. This element of chance ensures that all content choices are incentive compatible. That is, if we implemented the format preference with certainty, then the sender's content choice for the message format that isn't preferred would be merely hypothetical.

2.2.2 Benchmarking

At the end of the experiment, senders are asked to complete a benchmarking exercise in which we elicit their mapping of words onto numbers, using the words (presented in random order) from the language dropdown list in the communications part. The mapping is incentivized using the procedure developed in Krupka and Weber (2013): participants are truthfully told that "At the end of the survey, we will compare each of your answers to the average answers given by all other participants. If a question is randomly selected for payment, you will earn the bonus if your answer is within 3 percentage points of the average response." This benchmarking exercise allows us to compare the slant of the numeric and language messages, as described in Section 2.5.

2.3 Design for the Receivers Experiment

Receivers similarly begin by reading the instructions for the roles of both sender and receiver. They are required to answer two comprehension questions about how bonus payments are determined for senders and receivers before continuing. There are two main parts in the receivers experiment: "Predictions" and "Benchmarking". The Benchmarking part is identical to the version used in the senders experiment. 500 participants completed both parts of the receivers experiment: 248 were assigned to the numbers versus language condition, and 252 were assigned to the numbers versus intervals condition. In addition to a £3.00 payment for completion, 10 participants were randomly selected to earn up to £50 as a bonus payment based on their response in a randomly-selected question in one of the two parts (where pounds are again converted to dollars).

2.3.1 Predictions

In the Predictions part, receivers are first randomly-matched with a sender in their treatment, and then receive the message from their paired sender about the chance of drawing a red ball.[14] Then, receivers predict the chance of drawing a red ball. Receivers make eight

[14]Recall, in 50% of cases the sender's message was communicated using their preferred and in 50% of cases we randomly selected a format. Receivers know that the preferred format was not selected with certainty.

predictions in total.[15]

Half of receivers are assigned to the *incentives-known* condition in which they are explicitly told their sender's incentive assignment. The remaining half of receivers are assigned to the *incentives-unknown* condition, in which they are aware of the different incentive conditions that senders might face but do not know the condition to which their paired sender was assigned. This variation allows us to determine the degree to which receivers are able to anticipate and account for sender responses to incentives, which has important implications for policy.

After receivers make each prediction, they indicate on a Likert scale how informative they think the sender's message is.

2.4 Empirical framework

Our primary specification tests whether senders are more likely to choose language messages under directional incentives, as per Hypothesis 1. (For senders who choose between interval and numeric messages, we use the equivalent specification for interval messages.) For sender i who learns that the true probability of a red ball is p, we estimate the following OLS equation:

$$Language_{ip} = \beta_0 + \beta_1 Directional_{ip} + \delta FE_i + \alpha FE_p + \epsilon_{ip} \tag{1}$$

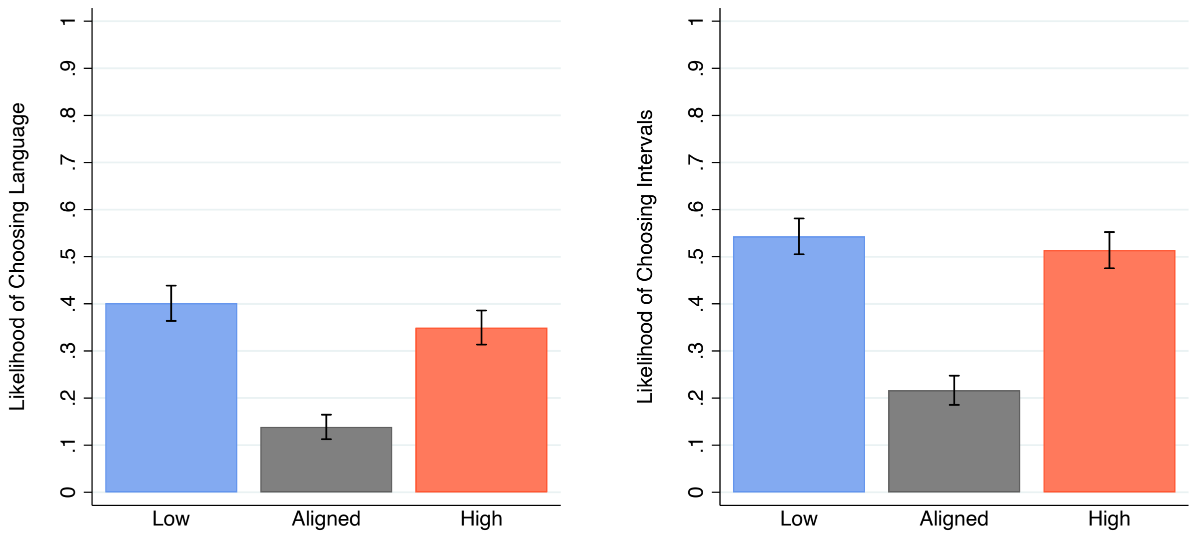
where $Language_{ip}$ is an indicator equal to one when the sender chooses language (or intervals) over numbers, and $Directional_{ip}$ is an indicator equal to one when the sender faces directional rather than aligned incentives. We include individual fixed effects, FE_i, and fixed effects for each true probability of drawing a red ball, FE_p. We cluster standard errors at the individual level.

2.5 Senders' Results

Figure 2 presents the raw data, where we see clear evidence that incentives affect researchers' choice of messages when they choose between natural language messages and numeric messages. In support of Hypothesis 1, we see that senders in the aligned condition choose language over numbers 14 percent of the time (s.e. 2 pp). This share increases substantially in both the directional-high condition (to 35 percent, s.e. 3 pp) and the directional-low condition (to 40 percent, s.e. 3 pp). As shown in Table 1, which reports the results from Equation 1, the differences between each directional condition and the aligned condition are highly statistically significant ($p < 0.001$ each).

[15] We randomly pair the receiver with a new sender and message for each decision.

Figure 2: Message format choice by incentives



(a) Numbers vs Language

(b) Numbers vs Intervals

Notes: This figure plots the raw data for the likelihood of choosing imprecise communications (language in Panel A and intervals in Panel B) over numbers, by incentive condition (aligned, directional-high, and directional-low). Bars reflect 95% confidence intervals. The sample includes 4,000 choices across 500 senders.

We see similar patterns when senders choose between numbers and intervals. Relative to language, senders choose interval messages somewhat more often, and the comparisons across conditions are qualitatively similar. In the aligned condition, senders choose interval messages over numeric messages 22 percent of the time (s.e. 2 pp). This share increases substantially in both the directional-high condition (to 51 percent, s.e. 3 pp) and the directional-low condition (to 54 percent, s.e. 3 pp); and the differences between directional and aligned incentives are again highly statistically significant ($p < 0.001$ each).

2.5.1 Slant and distribution of messages

There are many reasons why senders may choose different message formats when they have incentives to persuade. We now examine their within-format preferences to provide evidence that senders are strategically leveraging the imprecision of language to slant their messages.

Table 1: The impact of incentives on message format choice

	(1) Use Language	(2) Use Language	(3) Use Intervals	(4) Use Intervals
Directional Incentives	0.250***		0.297***	
	(0.025)		(0.029)	
High Incentives		0.222***		0.288***
		(0.027)		(0.030)
Low Incentives		0.279***		0.305***
		(0.028)		(0.033)
Observations	2008	2008	1992	1992
Aligned Mean	.14	.14	.22	.22
Respondent FE	Yes	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes	Yes

Notes: This table reports the effect of incentives on message format choice. In Columns 1 and 2, the dependent variable is an indicator equal to one when the sender chose to communicate using language rather than numbers. In Columns 3 and 4, the dependent variable is an indicator equal to one when the sender chose to communicate using intervals rather than numbers. Columns 1 and 2 are calculated over the sample of senders who had the option to communicate in language or numbers; Columns 3 and 4 are calculated over the sample of senders who had the option to communicate in intervals or numbers. In the *Directional Incentives* row, the independent variable is an indicator equal to one when the sender faced any kind of directional incentives, i.e. either directional-high or directional-low. In the *High Incentives* row, the independent variable is an indicator equal to one when the sender faced directional-high incentives. In the *Low Incentives* row, the independent variable is an indicator equal to one when the sender faced directional-low incentives. Fixed effects for the respondent and true probability are included as controls. Aligned mean calculates the likelihood of using language among aligned individuals in the corresponding sample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

First, we construct a measure of *slant* as follows: For numeric messages, slant is defined as (*number message – true probability*). For interval messages, slant is defined as (*interval message midpoint – true probability*). While numbers and numeric intervals have a natural meaning in relation to numeric probabilities, language messages require an extra conversion step. We use the mapping from the benchmarking exercise in the senders experiment to create our measure of slant for language: (*language message mapped to number – true probability*).[16]

[16]We use senders' individual-level mappings from each word to number in our analysis. While in principle we could see that senders distort their mapping strategically, for instance due to motivated cognition that they are being altruistic (Exley 2016), results are little changed if we use average mappings across all senders instead.

In Figure 3, we visually compare slant by incentives and message format using empirical cumulative density function (eCDF) plots of the raw data.[17] The top-left panel shows the eCDF of senders' slant of numbers. We see that very few individuals with aligned incentives (shown with the black line) slant numeric messages in either direction. That is, 94.3% of senders in the aligned condition communicate a number that is within 3 percentage points of the true probability.[18] Note that this need not be the case, for instance if senders expect receivers to not interpret numbers at face value.[19]

Consistent with the first part of Hypothesis 2, we also see clear evidence that senders slant in the direction of their incentives on average. Appendix Table A2 shows that senders with directional-high incentives slant numeric messages upwards by 9.8 pp compared to senders with aligned incentives ($p < 0.001$), and senders with directional-low incentives slant numeric messages downwards by 10.7 pp compared to senders with aligned incentives ($p < 0.001$).

Turning back to Figure 3, we can examine the distribution of slant for senders with directional incentives. The red line plots the distribution for directional-high incentives. Here the median slant is again 0 and the 30-70th percentile range is [-1, 3]. In other words, most senders do not slant numbers much. However, there is a long tail: The 10-90th percentile range is [-2, 52], indicating that the positive average slant is primarily driven by the small fraction of senders who slant to extremes. Overall, 26.4% of sender messages slant numbers upwards more than 3 percentage points if they face directional-high incentives. We see similar patterns for the directional-low group (shown with the blue line): 25.6% of senders slant numbers downwards by more than 3 percentage points.

The top-right panel of Figure 3 and Appendix Table A3 look at how senders slant imprecise messages compared with numeric messages. Unsurprisingly, given the noise in the mapping of language to numbers, and the coarser message space, we see that senders with aligned incentives sometimes slant language upwards or downwards by a modest amount; the 30-70th percentile range is [-3, 7]. We see evidence supporting the second part of Hypothesis 2 when we look at the impact of directional incentives, but the effects are modest; the difference between imprecise and numeric messages is 2 percentage points larger for directional-high than directional-low incentives ($p = 0.003$).[20] Further, with language messages we now see

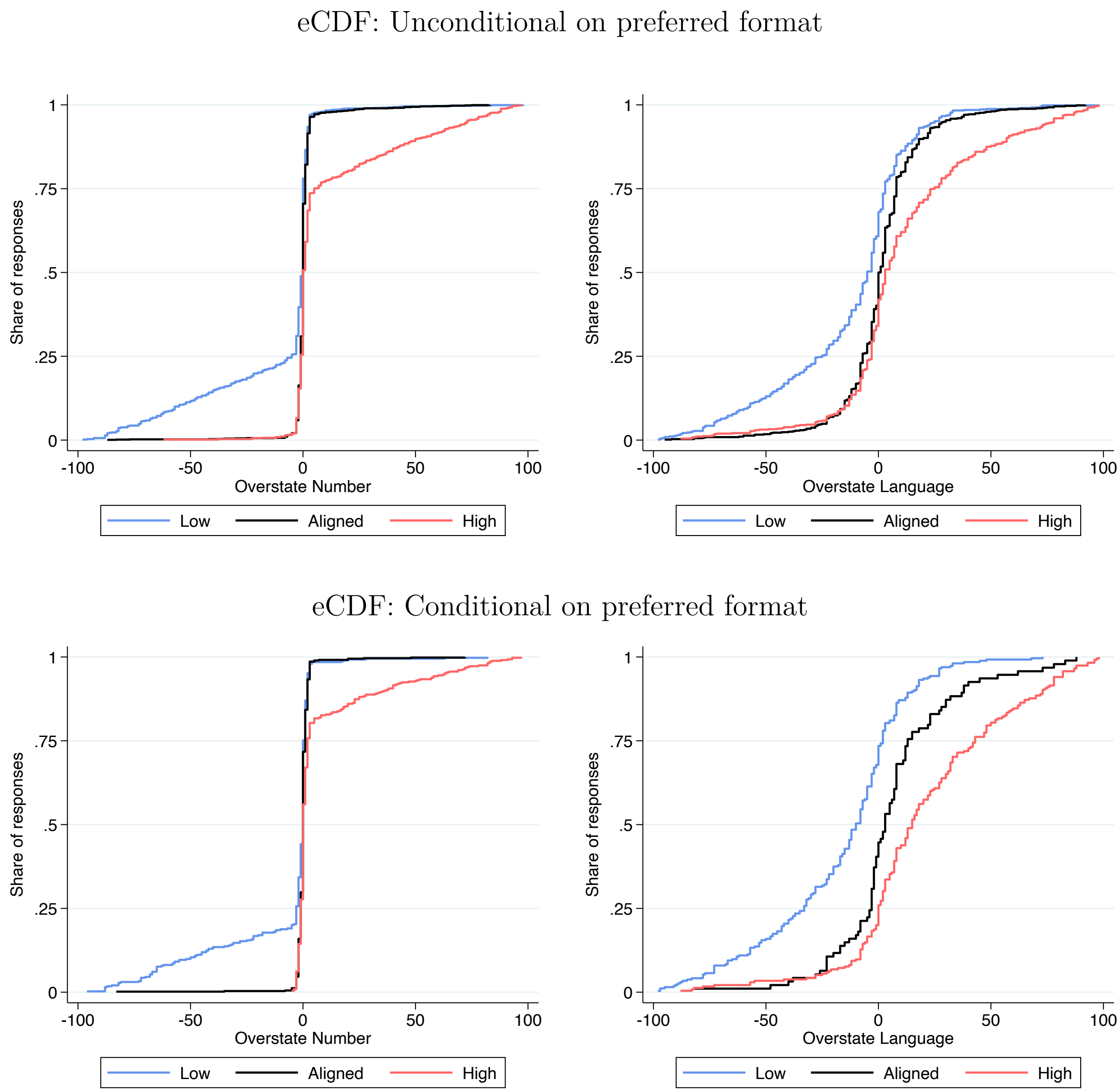
[17]To provide a frame of reference for these results, Appendix Figure A1 shows versions of these eCDFs using hypothetical number messages in which a) all senders maximally slant their messages in the direction of their incentives, and b) 75% of senders always tell the truth while 25% maximally slant their messages in the direction of their incentives.

[18]We look at a range around the true probability since some senders may misinterpret the slider bar.

[19]Even when receivers do not know senders' incentives and the true probability is close to 0 or 1, senders overwhelmingly send true messages, even though a sophisticated receiver would infer that such messages would be likely to be from a sender with persuasion incentives.

[20]The coefficient is similar when comparing language to numbers and when comparing intervals to numbers. However, the noise generated by our language-mapping procedure leads to much larger standard errors,

Figure 3: Empirical CDFs of message slant by incentives and format



Notes: This figure plots the empirical cumulative distribution functions (eCDFs) for slant, by incentive conditions and message format. The x-axis reflects the degree to which senders overstate their message compared to the truth: "Overstate number" in the lefthand panels simply subtract the real probability from the numeric message. "Overstate language" in the righthand panels takes the language-to-number mapping from the benchmarking exercise and subtracts the true probability. The top panels show the distribution for all messages, while the bottom panels condition on the messages using the sender's preferred format. The sample includes 2,008 choices across 251 senders. Similar eCDFs for interval messages are shown in Appendix Figure A2.

---

leading to different p-values ($p = 0.111$ for language; $p < 0.001$ for intervals; $p = 0.003$ for imprecise messages overall, pooling language and intervals).

that directional incentives not only shift the tails of the slant distribution, but also shift the median. With directional-high incentives, the median slant is 3, the 30-70th percentile range is [-2, 18], and the 10-90th percentile range is [-15, 57]. Comparing language to numbers, the right tail of the language-slant distribution looks similar to that of the number-slant distribution, but there are many more participants who slant a moderate amount: 49.1% (compared with 26.4% when using numbers) slant by more than 3 percentage points in the direction of their incentives. This suggests that there is a cohort of senders who do not feel comfortable sending a numeric message other than the truth but do feel comfortable sending a language message that might overlap with the truth but errs in their preferred direction. We see similar patterns for participants with directional-low incentives: 50.3% slant language by more than 3 percentage points in the direction of their incentives (compared to 25.6% when using numbers).

We next consider the interaction between preferred message format and slant. In Appendix Table A4, we run the same regressions as Appendix Table A3 but interact the gap in slant with the sender's preferred format. Consistent with the last part of Hypothesis 2, when senders prefer to use language, they slant language messages significantly more relative to numeric messages: senders who prefer language slant their language messages by an additional 4.5 percentage points upwards under directional-high incentives ($p = 0.098$) and by an additional 5.9 percentage points downwards under directional-low incentives ($p = 0.037$). That is, senders appear to be opting to use language (or intervals, for which we see similar although somewhat muted effects) in exactly the cases that they are slanting their message more than they otherwise would with numbers — the imprecise communications seem to afford the degree of plausible deniability that might excuse a slanted message.

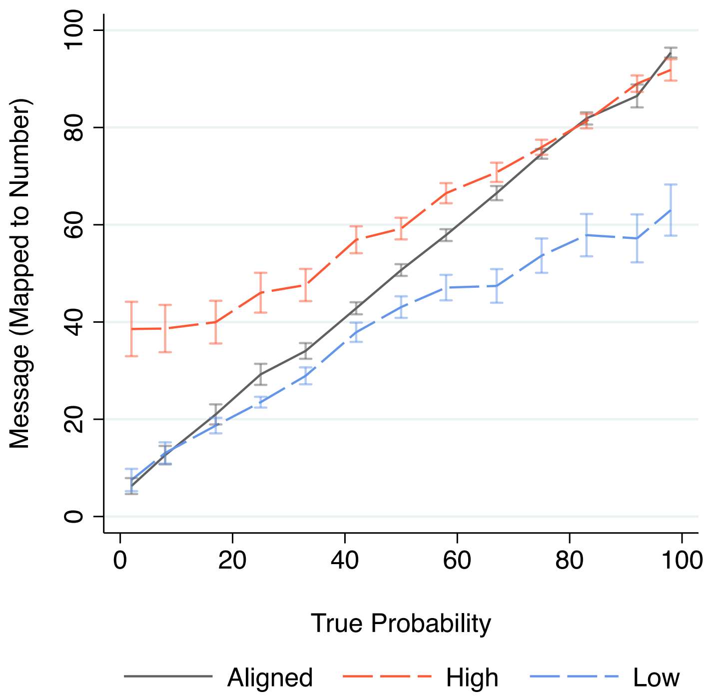
These differences can also be seen in our empirical CDF plots. The bottom-left panel of Figure 3 is similar to the top-left panel, showing that senders who prefer to send numeric messages slant numbers in a similar way to senders overall. However, the bottom-right panel shows a more extreme slant among the senders who prefer to send language messages. For senders with directional incentives who prefer to send language messages, 63.7% slant their language messages by more than 3 percentage points in the direction of their incentives, while only 41.3% of senders who prefer to send numeric messages slant their language messages by more than 3 percentage points.

In other words, the majority of senders who choose to send language messages are sending slanted language messages, while the majority of senders who choose to send numeric messages are not slanting either their number or language messages. Slant for numbers is driven more by a small group of participants who slant a large amount, while slant for language occurs among a larger group of participants who slant both smaller and larger amounts.

We show the equivalent plots for the distribution of slant for intervals in Appendix Figure A2. It visually looks in between the figures for numbers and language, suggesting that the imprecision of intervals leads to a qualitatively similar, but more muted, impact.

2.5.2 Alignment between states and incentives

Figure 4: Sender's message by true probability



Notes: This figure plots the sender's message on the y axis and the true probability on the x axis, by incentive condition. All message formats are pooled in this figure; language messages are included using the language-to-number mapping from the benchmarking exercise, and interval messages are included based on the midpoint of the interval. Bars reflect 95% confidence intervals. The sample includes 4,000 choices across 500 senders.

Next, we turn to heterogeneous effects by the true probability of drawing a red ball. In particular, in the directional conditions we classify the state as "good" when the true probability is higher (lower) for senders with high (low) incentives and "bad" when the true probability is lower (higher) for senders with high (low) incentives. Consistent with Hypothesis 3, Figure 4 visually shows that on average senders slant their messages more in the direction of their incentives when the state is bad.[21] For good states, senders send numeric and language messages that are similar to the aligned condition. However, for bad

[21]Appendix Figure A3 shows Figure 4 conditioning on both message format and whether the message is preferred. Similar insights emerge as in the eCDF plots.

states they substantially distort both numeric and language messages. In accordance with the second part of Hypothesis 3, Appendix Table A5 shows that senders send more imprecise messages when the state is bad.

2.5.3 Predictors of the strategic use of language

We can glean further insights into the underlying mechanisms by exploring self-reports that we collected at the end of the experiment, along with heterogeneous treatment effects.

First, Appendix Figure A4 shows that senders self-report preferring to use language in the real world "to make it easier to withhold information" (87% say they would either "slightly" or "strongly" prefer using language, on a 5-point Likert scale), "to make it easier to lie to someone" (84%), and "to make it easier to persuade someone that the likelihood is higher than it actually is" (58%). Perhaps surprisingly, only 31% say they would prefer using language to make their message easier to understand. We create an index that takes the average of the Likert responses for the former three reasons to use language, all of which relate to strategic persuasion. Column 1 of Table 2 shows that a self-reported preference to strategically use language almost entirely explains the response to incentives that we observe in our experiment ($p = 0.001$). This suggests at least some self-awareness. That is, when looking just at the message choices in our experiment we cannot determine whether senders motivatedly self-deceive and believe their language and number communications are equally close to the truth. This additional data suggests, however, that self-deception is unlikely to be the driving factor.

Second, at the end of the experiment we asked senders to work through a three-question module to assess their numeracy.[22] We observe that the number of questions answered correctly on our three-question assessment is predictive of the strategic use of language: Column 2 of Table 2 shows that each additional correct answer increases the likelihood of using language when facing directional incentives by 8.5 percentage points ($p = 0.011$). In other words, more numerate respondents appear to be more likely to strategically leverage language to persuade.[23]

[22]These questions were adapted from a longer numeracy assessment included in Kahan et al. (2012).

[23]We also see that more numerate individuals are more likely to use numbers in the aligned condition ($p = 0.005$). Chang et al. (2024) show that more numerate individuals exhibit quantification fixation in decision-making; our findings suggest that numeracy also plays a role in the (strategic) communication of numbers. Further, this finding along with the literature suggesting a positive correlation between cognitive ability and honesty (e.g., Rindermann et al. (2018)) is consistent with the idea that more numerate individuals may use language as an excuse to avoid a direct lie.

Table 2: Heterogeneous treatment effects: Message format choice

	(1) X = Persuasion	(2) X = Numeracy	(3) X = Female	(4) X = Resp Time
X*Directional	0.336***	0.085**	0.049	0.001
	(0.102)	(0.033)	(0.050)	(0.001)
Directional Incentives	0.019	0.057	0.216***	0.182***
	(0.071)	(0.079)	(0.038)	(0.050)
Observations	2008	2008	1960	2008
Aligned Mean	.3	.3	.29	.3
Respondent FE	Yes	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes	Yes

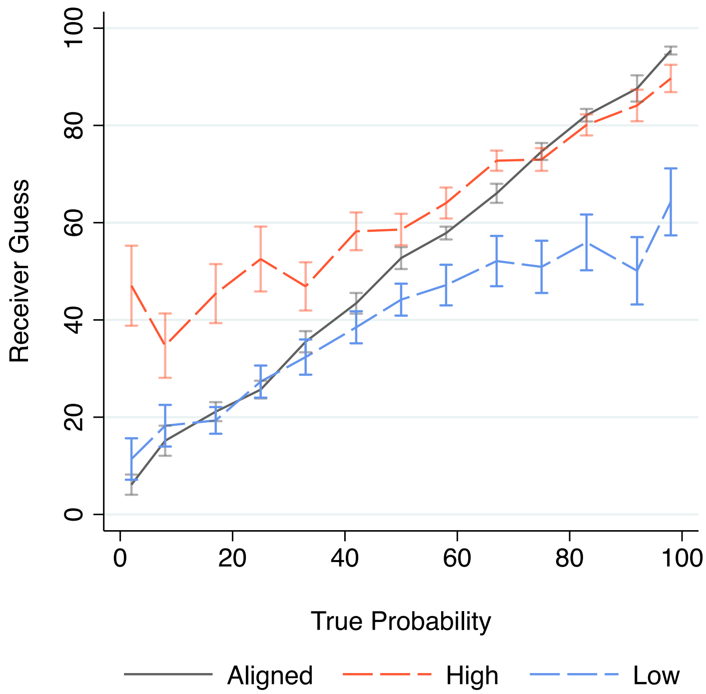
Notes: This table reports heterogeneous treatment effects for message format choice. In all columns, the dependent variable is an indicator equal to one when the sender chose to communicate using language rather than numbers. The coefficient of interest, $X^{*}Directional$, is an indicator equal to one when the sender faces directional incentives, interacted with one of four predictor variables: the average response to three 1-5 scale Likert questions on the likelihood of using language to persuade in the real world (Column 1); the number of correct answers on a three-question numeracy module (Column 2); an indicator for whether the respondent is female (Column 3); and the percentile breakdown of the sum of response times in each main decision (Column 4). Controls include an indicator equal to one when the sender faces directional incentives as well as fixed effects for the respondent and true probability. Aligned mean reflects the likelihood of using language under aligned incentives for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Finally, we look at heterogeneous treatment effects by gender (Column 3) and response time (Column 4) and see that neither is predictive of strategic use of language. If we think of response time as a proxy for attention, this suggests that the strategic use of language is not merely an artifact of more or less attentive participants in the experiment — more subtly, an *understanding* of how to leverage language appears to be key.

Receivers' Behavior

Next, we turn to receivers. We first show, consistent with the first part of Hypothesis 4, that receivers are persuaded overall. Figure 5 compares average receiver guesses when senders face aligned, directional-high, and directional-low incentives (pooling across message formats). Receivers' guesses look remarkably similar to the messages senders sent (Figure 4). Appendix A6 shows that receivers' guesses are 15 percentage points higher when their paired sender faced directional-high compared to directional-low incentives ($p < 0.001$).[24] This finding points to the overall policy relevance of our identification of the strategic use of language: Because we observe at least partial naivete on the part of receivers, this suggests that there exist information distortions in communications.

Figure 5: Receiver's guesses by true probability



Notes: This figure plots the receiver's guess on the y axis and the true probability on the x axis, by incentive condition that the sender faced. Guesses based on all sender message formats (number, language, interval) are pooled in this figure. Bars reflect 95% confidence intervals. The sample includes 4,000 choices across 500 receivers.

Appendix Table A7 compares average receiver guesses by the format they receive. Consistent with the second part of Hypothesis 4, we see that receivers' guesses are significantly more likely to be in the direction of senders' incentives when they receive a language or interval message as compared to a numeric message.

As shown in Appendix Figure A5, receivers make significantly larger mistakes when they receive language versus numeric messages. We use a simple measure of this, average error: $|guess - probability|$.[25] When receivers see numeric messages in the aligned condition, their

[24] This reports the gap for receivers with paired senders deciding between numbers and language. Receivers' guesses are 20 percentage points higher when their sender faced directional-high incentives and was deciding between numbers and *intervals*.

[25] Specifically, we take the average after winsorizing at the 5- and 95-percent levels, with standard errors clustered at the receiver level.

average error is 2.5 pp (s.e. 0.2 pp). When receivers see language messages in the aligned condition, their average error is 8.7 pp (s.e. 0.5 pp). That is, numbers are interpreted more accurately than language. This difference can be attributed to the additional noise (imprecision) introduced when using language.

In the conditions where senders face directional incentives, numeric messages lead receivers to an average error of 12.9 pp (0.8 pp), while language messages lead receivers to an average error of 26.1 pp (1.1 pp). That is, when senders are incentivized to persuade receivers, language leads to receiver beliefs that are even more inaccurate, suggesting that senders are strategically leveraging the imprecision inherent to language (all differences, including the interaction, have $p < 0.001$).[26]

Recall that half of receivers learn the incentives of their paired sender, while the other half does not. Thus far, we have combined analyses for receivers who do or do not know what senders' incentives are. Next, we show that receivers do not fully adjust to the knowledge of senders' incentives. In Appendix Table A8, we see that when receivers learn the incentives of senders, they are persuaded statistically-significantly less. However, 79% of the main persuasion effect remains, indicating that knowledge of senders' incentives is not sufficient to avoid being persuaded.[27] This finding suggests that policy interventions to promote awareness of the strategic use of language may not be sufficient to counteract its effects.

Finally, we look at receivers' ratings of how informative the message is. When receivers see a numeric message, Appendix Figure A6 shows that 72% rate the message as "very informative" and only 3% rate it as either "very" or "somewhat" uninformative. However, when receivers see a language message, 17% rate the message as "very informative" and 21% rate it as uninformative. This suggests that receivers might put less weight on language messages when updating their beliefs. In Study 1, where receivers' priors are uninformed, it is difficult to unpack the implications of this result, but Study 2 sheds further light on this.

3 Study 2

3.1 Design Overview

The benefits of the abstract "balls-and-urns" design of Study 1 are that the meaning of probabilities, the language-number mapping, and the directional versus aligned incentives

[26] In both cases, receivers' errors from interval messages lie in between their errors from numeric and language messages.

[27] It seems unlikely that respondents were not reading the information provided; for instance, only 4% of respondents fail our attention check, and we exclude them from the analysis as pre-registered. Further, all receivers had to correctly answer a question about senders' incentives before proceeding to the main part of the experiment.

are clean and stripped of context. But these benefits come with some costs: the setting is quite abstract, and does not reflect the types of information people report in real life. As such, Study 2 considers communication behavior in a more real-world environment, exploring a setting where we expect the strategic use of language to have important implications. Specifically, in July 2024 we recruited a sample of academic social science researchers, most of whom are tenure-track or tenured professors, and asked them to communicate treatment effects identified in published social science papers to a sample of government policymakers we recruited from July to September 2024.

The design is conceptually similar to Study 1, with some key differences in addition to the population and type of information being communicated:

1. Rather than asking people to communicate a probability, as in Study 1, we ask researchers to communicate the treatment effect of a policy intervention tested in a real research study, always reported in percentage points.
2. Researchers face either aligned or directional-high incentives (and not directional-low incentives). The incentive-compatible scheme for calculating bonus payments is structured the same way as in Study 1. In addition, researchers see vignettes tailored to their incentive condition: researchers in the aligned condition are encouraged to "give the policymaker the best understanding of the data" while researchers in the directional-high condition are told to persuade their policymaker that the treatment effect is large for the sake of a government grant or policymaker attention.
3. We only compare the choice of language versus numbers and omit the comparison of intervals versus numbers.

It is important to note that the aim of Study 2 is to identify whether researchers are more likely to opt for language *when* they face strategic incentives to persuade. Our data do not shed light on *how often* researchers face strategic incentives (strong enough to at least partially offset incentives to communicate accurately) in practice.

Our hypotheses for Study 2, preregistered at AEARCTR-0013947, are broadly similar to those in Study 1, accounting for the differences described above.

3.2 Sample

3.2.1 Academic Researchers

We personally invited 241 social science researchers to participate in our study. 145 social science completed our survey, for a response rate of 60%.[28] 75% of participating researchers are

[28] Appendix section C.3.1 shows an example recruitment email. Thank you to all who participated!

in tenure-track positions (Assistant Professor, Associate Professor, Professor, or equivalent) and 23% are post-docs or PhD students. 37% of the sample self-identify as women. 81% work with empirical or experimental data for most of their research projects, and 61% report working in a policy domain (most commonly economic policy, health, or education).[29]

3.2.2 Policymakers

We recruited a sample of 66 policymakers through the University of Warwick's *Policymakers Lab*. Policymakers in the sample work in 5 countries (United States, United Kingdom, Australia, India, and Belgium) and are primarily civil servants working in central government (for instance, U.S. policymakers in our sample work in the Department of State; Department of Health and Human Services; U.S. Agency for International Development; Department of Homeland Security; Department of Justice; and General Services Administration). 76% of policymakers in the sample report having been involved in policy adoption decisions. 52% have a Master's or Professional degree, 26% have a PhD, and 21% have a Bachelor's degree or equivalent as their most advanced degree. 54% of the sample identify as women. Policymakers are recruited with the motivation of "lending their expertise to contribute to and inform academic research" and as such is likely to be more experienced with and favorable towards research evidence than the typical policymaker.

3.3 Design for the Researchers Experiment

In Study 2, 145 researchers complete just one main part, the Communications part. A separate group of 22 researchers that we contacted complete a Benchmarking part, described below. In addition to a $10 payment for completion, researchers could earn $10 as a bonus payment based on their message communication decisions.[30] Appendix Section C.3 shows screenshots from the experiment.

3.3.1 Communications

In the Communications part, researchers first learn about two different, real research studies. The studies are randomly assigned from a broader set of six studies, shown in Table 3. All

[29]We do not expect this to be a representative sample of all social science researchers, although we do not have clear predictions as to how this might impact our results. The final sample excludes two individuals who were dropped because they indicated that they knew the overall hypothesis of our study. This also excludes eight individuals who attrited out of the sample. Our results are robust if we bound our main effects using Lee bounds (Lee 2009), as seen in Appendix Table A9.

[30]Payments were distributed in the form of Amazon gift cards, and we matched the currency based on the respondents' preferences. Participants had the option to opt out of receiving a bonus payment; 23% took this option. Results are qualitatively similar if we exclude participants who opt out.

Table 3: Research studies included in Study 2

Paper	Policy Domain	Outcome	Effect
Domurat et al (AER, 2021)	Health insurance	Letter → takeup healthcare coverage	1.2pp
Adams et al (JFE, 2021)	Consumer finance, savings	Alt financial product info → switch accounts	1.9pp
Liebman & Luttmer (AEJ EP, 2015)	Labor market and aging	Social Security info → labor force participation	4.2pp
Fishbane et al (Science, 2020)	Criminal justice	New court summons → defendants appear	6.2pp
Burland et al (AER: I, 2023)	Education policy	Info about free tuition → applying to college	8.2pp
Bergman et al (AER, 2024)	Economic mobility, housing	Info, some support → moving to opportunity	13.8pp

six studies identify the impact of a policy-relevant intervention. Importantly, each study reports the outcome of interest, and the intervention's effect in percentage points. The policy domain, outcome, and effect size vary across the studies. We present the effect in a bar plot which displays the control group (or baseline) mean alongside the treatment effect, as shown in Appendix C.3.1. The researcher's task is to decide how to communicate the study results to their paired policymaker.

Before making their communication decisions, each researcher is assigned to one of two conditions: *directional-high* or *aligned*.[31] In the directional-high condition, researchers are asked to "imagine you are the author of the study testing the policy intervention and you are trying to persuade a policymaker that the research is promising to increase your chance of getting a government grant or policymaker attention". To ensure this framing is incentive-compatible, they are additionally told that they are more likely to earn the bonus payment if their paired policymaker predicts the intervention's effect size is larger.[32] In the aligned condition, researchers are instead asked to "imagine you are the author of the study testing the policy intervention and you are trying to give the policymaker the best understanding of the data to inform their policy decisions". We also include incentives for this condition: researchers are more likely to earn the bonus payment if their paired policymaker correctly predicts the intervention's effect size.[33]

After learning about the results of the research studies and their incentive assignment, researchers make their message content choices for numbers and language. For numbers, they complete the sentence, using the example of the first paper on healthcare insurance takeup: "In the control group, 8.1% of individuals take up healthcare coverage. The letter led to an increase in healthcare coverage of [Your Response] percentage points. That is, 8.1 + [Your

[31] We omit the *directional-low* condition since we expect that in this context incentives to persuade will typically go in the direction of convincing an audience that treatment effects are relatively large. Unlike in Study 1, in Study 2 assignments are across-subjects, both to keep the survey as short as possible for bandwidth-constrained researchers and also to limit the possibility of guessing the research hypothesis in a sample that is likely to be more attuned to what we might be testing.

[32] Specifically, if the paired policymaker predicts the effect size is X percentage points, the probability the researcher wins the bonus payment is equal to X%, with a minimum probability of 0%.

[33] Specifically, if the true effect size is X percentage points, the researcher will win the bonus payment if their paired policymaker predicts the effect size is between X-1 and X+1 percentage points.

Response]% of those receiving the letter take up healthcare coverage."[34] We were careful to design the communications such that they are clear even to those who might be less familiar with working in percentage points. For instance, *8.1 + [Your Response]* dynamically updates in response to the guess the researcher inputs.

For language, researchers complete the sentence, again using the example of the paper on healthcare coverage takeup: "In the control group, 8.1% of individuals take up healthcare coverage. The letter led to a(n) [WORD] increase in healthcare coverage." As in Study 1, researchers use a dropdown list to select a word or phrase — in this case including, for example, "tiny", "intermediate", and "very large" — to complete the sentence.

After making their message content choices, researchers select their preferred message format, numeric or language. Again, researchers are told their preferred choice is more likely but not guaranteed to be the one communicated. The experiment ends with a short demographics questionnaire.

3.3.2 Benchmarking

To keep the Researchers Experiment as short as possible, we run the Benchmarking exercise among a separate group of researchers. As in our main Researchers Experiment, we personally invited mostly tenure-track academic researchers to take part. In this separate study, for each word in the dropdown list in the Researchers Experiment, researchers make incentivized predictions about the "numeric effect size (as a percentage point increase) [they] think others would expect each of the following words or phrases would correspond to." Researchers make separate predictions for each policy paper, to account for the possibility that language is perceived differently according to the policy context. As in Study 1, this benchmarking exercise allows us to measure the slant of language messages.

3.4 Design for the Policymakers Experiment

The policymakers experiment is focused entirely on policymakers' predictions of the true effect size based on the message sent by their paired researchers. As in the researchers experiment, in addition to a $10 payment for completion, policymakers could earn $5 as a bonus payment based on their predictions. Appendix Section C.4 shows screenshots from the experiment.

In the experiment, policymakers first select the two policy domains that best reflect their area of expertise, from the set of six indicated in Table 3. They learn about the relevant research study in each of the domains (but not the results).[35] Then, they receive a message

[34]Responses are constrained to be between 0 and 100 minus the baseline mean.

[35]We ask policymakers if they're familiar with the research study in question, and there is only one instance

from a randomly paired researcher communicating the treatment effect. After receiving the message, policymakers predict the actual treatment effect reported in the study. Policymakers are informed of the two incentive conditions faced by researchers and know that they do not have to tell the truth. However, to model real-world policy decisions in which the variation in incentives a researcher faces are not made explicit, policymakers were not told which incentive scheme their paired researcher faced (akin to the incentives-unknown condition in Study 1).

After each prediction, policymakers indicate whether they would be willing to extend the survey (i.e., a costly signal of effort) in order to receive an infographic about the study in question and also whether they plan to share information about the study with their colleagues. The majority (67% and 57%, respectively) indicate a willingness to meaningfully engage with the information outside of the experiment, pointing to its relevance for the policymakers involved. Finally, at the end of the experiment we include a short demographic survey and debrief policymakers on the real results of the research studies.

3.5 Results

Researchers' Behavior

In the raw data, we see clear evidence that incentives affect senders' choice of messages, in support of Hypothesis 1. Senders in the aligned condition choose language messages over numeric messages 13 percent of the time (s.e. 3 pp). This share increases substantially when they have directional incentives (to 41 percent, s.e. 4 pp). Column 1 of Table 4 reports the main effect: researchers are 29 percentage points more likely to use language when they face directional incentives ($p < 0.001$). Somewhat surprisingly given the differences across studies (e.g., one might imagine there are more reasons to use language to communicate context in Study 2), these levels are both qualitatively and quantitatively similar to those in Study 1.

Next, we exploit variation in the actual effect sizes of the research studies and find evidence that is consistent with Hypothesis 3 adapted to this study (here, small effect sizes are "bad states" for senders with directional incentives). First, incentives induce more language use for studies that have smaller effect sizes. Column 2 of Table 4 shows that for every percentage point increase in study effect size, researchers are 2.2 percentage points less likely to use language when in the directional-high condition ($p = 0.028$). Column 3 shows a precisely-estimated null effect when incentives are aligned. That is, researchers' decisions to use language are largely unaffected by the effect size when they are incentivized to accurately communicate the information to their paired policymaker.

in which the policymaker answers yes.

Table 4: Format choice, overall and by true effect size

	(1) Use Language	(2) Use Language	(3) Use Language
Directional Incentives	0.291***		
	(0.054)		
Effect Size		-0.023**	0.006
		(0.010)	(0.006)
Observations	290	138	152
Mean over obs	0.26	0.41	0.13
Subject controls	Yes	Yes	Yes
Study FE	Yes	No	No
Sample	**All**	**Directional-High**	**Aligned**

Notes: This table reports the likelihood of researchers using language as the dependent variable. Column 1 shows the impact of directional incentives via *Directional Incentives*, an indicator equal to one when the researcher faces directional-high incentives. Columns 2 and 3 condition on responses when the researcher faces directional-high and aligned incentives, respectively. The key independent variable is *Effect Size*, the treatment effect from the study in question in percentage points. Subject controls include indicators for whether the researcher is tenure-track, male, and does empirical research. Study fixed effects are only included in Column 1. Mean over obs reflects the likelihood of using language for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

As in Hypothesis 2 in Study 1, Table 5 demonstrates that researchers also behave differently in how they slant numbers and language. There is a modest difference in how researchers slant numbers across incentive groups. On average, researchers with aligned incentives report numeric estimates that are close to the truth.[36] Column 1 of Table 5 shows that, compared to researchers with aligned incentives, researchers with directional incentives communicate using numbers that are 13% larger ($p = 0.004$).[37] This effect is statistically significant, but modest in size, suggesting that senders prefer to limit their numeric misreporting of the effect sizes.

[36]Perhaps due to the somewhat noisy signals of the treatment effects indicated through the bar graphs, senders communicate numeric messages that are slightly smaller (0.27 percentage points on average, s.e. 0.07) than the true effect size.

[37]To get this percentage, we exponentiate the estimated coefficient in the table.

Table 5: Message slant by format

	(1)	(2)	(3)
	log(Number)	log(Language)	log(Lang)-log(Num)
Directional Incentives	0.122^{***}	0.397^{***}	0.275^{***}
	(0.044)	(0.077)	(0.085)
Observations	289	289	289
Mean over obs	1.49	1.28	-0.22
Subject controls	Yes	Yes	Yes
Study FE	Yes	Yes	Yes

Notes: This table reports researchers' slant by message format and incentives. Column 1 regresses *ln(Number)* on an indicator for *Directional Incentives*, which equals one when the researcher faces directional-high (versus aligned) incentives. Column 2 uses the same specification, but uses *ln(Language)* as the dependent variable, where *Language* represents the number that the word used was benchmarked to. Column 3 regresses the difference in (1) and (2) on incentives. Subject controls include indicators for whether the researcher is tenure-track, male, and does empirical research, and study fixed effects are included. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Senders slant language significantly more than they do numbers. Column 2 of Table 5 shows that, compared to researchers with aligned incentives, researchers with directional incentives slant language corresponding to numbers that are 49% larger. The frequency with which researchers use each word across the two conditions reflects this overall pattern: in the aligned condition, only 48% of researchers use a word to indicate a large effect size ("fairly large" or larger) while in the directional-high condition 74% of researchers use such language. Consistent with the second part of Hypothesis 2, Column 3 of Table 5 shows that the effect of incentives on slant is significantly larger for language than for numbers ($p = 0.001$). Related to the last part of Hypothesis 2, Appendix Table A10 shows that the gap between language and numeric slant increases by 64% when looking at the interaction of incentives with an indicator for the choice to communicate using language ($p = 0.044$). In other words, as in Study 1, essentially the entire gap in slant is driven by senders who prefer to send language messages.

Policymakers' Behavior

Table 6: Effect of researchers' messages on policymakers' guesses

	(1) Posterior	(2) log(Posterior)	(3) Posterior	(4) log(Posterior)
Directional-High	2.200**	0.200**		
	(0.992)	(0.097)		
Language Message			3.316*	0.206
			(1.947)	(0.200)
Prior	0.374***		0.609***	
	(0.118)		(0.166)	
log(Prior)		0.441***		0.565***
		(0.076)		(0.137)
Observations	129	129	57	57
Mean over obs	8.56	8.56	9.93	9.93
Study FE	Yes	Yes	Yes	Yes
Sample	All	All	Directional-High	Directional-High

Notes: This table reports policymakers' guesses by researchers' incentives and message format. In Columns 1 and 3, the dependent variable *Posterior* is the policymaker's guess after seeing the researcher's message. In Columns 2 and 4, the dependent variable is the natural log of the *Posterior* and *Prior* variables. The independent variable of interest is *Directional-High*, an indicator equal to one when researchers faces directional-high incentives. We also control for *Prior*, the policymaker's guess of the effect size elicited before receiving the researcher's message. Columns 1 and 2 examine the entire sample and Columns 3 and 4 are only computed over settings where researchers had directional-high incentives. Controls include fixed effects for each of the six real research studies. The *Prior* and *Posterior* variables are winsorized at the 10% and 90% level. *Mean over obs* is the mean posterior. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

We see evidence that researchers' incentives have an effect on policymakers' predictions. As shown in Columns 1 and 2 of Table 6, policymakers predict larger effects (2.2 percentage points [$p = 0.030$], or 22% [$p = 0.042$], larger) when researchers are incentivized to directionally persuade policymakers. In other words, policymakers are not fully responsive to researchers' slant.[38]

Columns 3 and 4 of Table 6 restrict to cases where researchers are incentivized to persuade policymakers that the study effects were large. Here we see suggestive evidence

[38] Recall that in Study 2 policymakers do not know the incentives researchers face, so fully-sophisticated receivers would need to "correct for" unrealistic messages in general. Recall also that in Study 1 we observed that receivers continued to be persuaded even in the case where incentives were communicated.

that policymakers were particularly persuaded by language compared to numeric messages (though the statistical power is lower): their predictions are 3.3 percentage points ($p = 0.095$), or 23% ($p = 0.307$), larger when receiving a language compared with a numeric message.

4 Discussion and Conclusion

Our experiments provide evidence that, in many settings, people use language more when they want to persuade others. We first find this effect in a tightly-controlled abstract setting, and then find qualitatively- and quantitatively-similar effects in a setting where academic researchers communicate the effect size of a policy intervention reported in a research paper to policymakers. In addition, both experiments find that senders who prefer language choose to slant language more, a pattern that is not observed for numbers. These results indicate that imprecise language provides a way for some senders who wish to distort messages in the direction of their incentives to avoid overt lies.

In Section 4.1, we interpret this psychology through the lens of a simple sender-receiver model in which the sender's utility depends on three factors: A benefit from persuading the receiver, a cost based on the error in the receiver's prediction, and whether the message allows for plausible deniability. The sender is communicating with a receiver who takes messages at face value.[39] In our model, senders may choose imprecise messages in order to slant more while maintaining the benefit of plausible deniability. This can lead to behavior that can explain many of our experimental results. When imprecise messages allow for a wide range of interpretations, senders may prefer both precise and imprecise messages, slant weakly more for imprecise messages, and either slant messages by a fixed amount (using either precise or imprecise formats) or do not slant at all (only using a precise format).

4.1 Theory

We consider an environment with two players, a sender (S) and a receiver (R). The sender has private information about some state realization, denoted by her type $\theta \in \mathbb{R}$. After privately observing θ, the sender sends the receiver a message $m \in M$. After observing m, the receiver takes an action $a \in \mathbb{R}$. Messages are of the form $m \in M = \{(m_l, m_h) \in \mathbb{R}^2 : m_h \geq m_l\}$. This can be interpreted as messages saying "θ is between m_l and m_h."

[39]Extensions with strategic receivers, as those in Kartik (2009), would be worth exploring in future work. In our experiments we find that the receivers indeed often take messages at face value. For instance, in both studies, the modal response for receivers seeing a number is to report the number, regardless of what they know about senders' incentives. Likewise, significant majorities of receivers in Study 1 respond to intervals by answering within the bounds of the interval. Naturally, this assumption is also very helpful for tractability.

Senders get utility depending on how high the receiver's guess a is, disutility for how inaccurate his guess is (for instance due to altruism), and utility for sending messages that include the true state. This latter benefit is what we refer to as *plausible deniability.*

$$u_S(\mathbf{m}, a(\mathbf{m})) = \underbrace{\alpha \cdot a(\mathbf{m})}_{\text{monetary benefit}} - \underbrace{\beta \cdot (a(\mathbf{m}) - \theta)^2}_{\text{cost of receiver inaccuracy}} + \underbrace{\gamma \cdot \mathbb{1}(\theta \in [m_L, m_H])}_{\text{benefit for plausible deniability}} \tag{2}$$

The γ term relates to Sobel (2020)'s characterization of lies. Using his framework:

- An *imprecise lie* is if $\theta \in [m_l, m_h]$ and $m_l < m_h$.
- A *precise lie* is if $\theta \notin [m_l, m_h]$ and $m_l = m_h$.
- An *imprecise truth* is if $\theta \in [m_l, m_h]$ and $m_l < m_h$.
- A *precise truth* is if $m_l = m_h = \theta$.

Both imprecise and precise truths allow for plausible deniability, but lies do not.

To make salient the effect of message precision, we describe messages by (μ, k), where $\mu \equiv \frac{m_h + m_l}{2}$ and $k \equiv \frac{m_h - m_l}{2}$. That is, μ is the midpoint of the message interval, and k is half of the width of the interval.

We make the restrictive simplifying assumption that receivers do not play strategically. When receivers see a message $\mathbf{m} = (\mu - k, \mu + k)$, they always take an action between $\mu - k$ and $\mu + k$, and their action is stochastic for $k > 0$.[40] Specifically, receivers' play a is drawn from a probability density $f_\mu^k(a)$, where $f_\mu^k(a)$ is atomless for $k > 0$, $F_\mu^k(\mu - k) = 0$, $F_\mu^k(\mu + k) = 1$, and $f_\mu^k(\mu - \kappa) = f_\mu^k(\mu + \kappa)$ for all κ. That is, receivers never guess below the lower bound of the message range or above the upper bound of the message range, and their guess densities are symmetric about the midpoint of the range. Finally, we assume that the distribution of noise depends on k, but not μ: For $\mathbf{m}' = (\mu', k)$ and $\mathbf{m}'' = (\mu'', k)$, $f_{\mu'}^k(\mu' + \kappa) = f_{\mu''}^k(\mu'' + \kappa)$ for any μ_1, μ_2, and κ.

We can now simplify Equation (2) to depend on $Var_{a \sim f^k}(a)$, which is defined as the variance of a given that it is drawn from f_μ^k, and omit the μ from the above assumption.

Lemma 1
The sender's expected utility function in Equation (2) *can be rewritten as the following function of* $\mathbf{m}$*:*

$$U_S(\mathbf{m}) = \alpha\mu - \beta(\mu - \theta)^2 - \beta \cdot Var_{a \sim f^k}(a) + \gamma \cdot \mathbb{1}(\theta \in [m_L, m_H]). \tag{3}$$

We derive this result, as well as other proofs and derivations, in the appendix.

[40] This response can be interpreted as the imprecision of $\mathbf{m}$ leading to noisy responses, or vagueness about the meaning of $\mathbf{m}$, a la Lipman (2009).

Now, we solve for senders' optimal messages given θ. Senders choose between a *precise* message with $k = 0$ and an *imprecise* message with $k > 0$, and we consider behavior for each type of message as well as which message the sender prefers to send.

As in the experiment, senders make three decisions:

1. The optimal precise message to send: $(\mu^p, 0)$.
2. The optimal imprecise message to send: (μ^i, k).
3. Whether to send the optimal precise or imprecise message: $(\mu^p, 0)$ or (μ^i, k).

First, we consider precise messages. In this case, we can reduce Lemma 1 to:

$$\alpha\mu - \beta(\mu - \theta)^2 + \gamma \cdot \mathbb{1}(\mu = \theta)$$

Intuitively, if γ is small, senders solve the first-order condition and choose $\mu^p = \theta + \frac{\alpha}{2\beta}$. We call this value μ^{FOC}. If γ is large, senders instead choose $\mu^p = \theta$. Specifically:

$$\mu^p = \begin{cases} \mu^{FOC} & \text{if } \gamma \leq \frac{\alpha^2}{4\beta} \\ \theta & \text{if } \gamma \geq \frac{\alpha^2}{4\beta} \end{cases} \tag{4}$$

Next, we consider imprecise messages. By assumption, the variance of receivers' responses is not a function of μ. Senders choose between $\mu = \mu^{FOC}$ and $\mu = \theta + k$, the most slanted message that allows for plausible deniability. If $k \geq \frac{\alpha}{2\beta}$, senders will choose μ^{FOC}, since this message allows for plausible deniability. If $k < \frac{\alpha}{2\beta}$, senders will choose $\mu^i = \mu^{FOC}$ if γ is above a certain threshold and $\mu^i = \theta + k$ otherwise. Specifically:

$$\mu^i(k) = \begin{cases} \mu^{FOC} & \text{if } k \geq \frac{\alpha}{2\beta} \\ \mu^{FOC} & \text{if } k \leq \frac{\alpha}{2\beta} \text{ and } \gamma \leq \frac{\alpha^2}{4\beta} - (\alpha - \beta k)k \\ \theta + k & \text{if } k \leq \frac{\alpha}{2\beta} \text{ and } \gamma \geq \frac{\alpha^2}{4\beta} - (\alpha - \beta k)k \end{cases} \tag{5}$$

We can then compare the messages sent:

Proposition 2

1. *When senders send a precise lie ($\mu^p \neq \theta$), they slant their imprecise messages ($\mu^i(k) \neq \theta$).*
2. *When senders send an imprecise lie ($|\mu^i(k) - \theta| > k$), they also send a precise lie ($\mu^p \neq \theta$).*

We can also now analyze what formats senders prefer by comparing utility for the optimal precise and imprecise messages.

Proposition 3

- *If $k < \frac{\alpha}{2\beta}$, then the only possible messages the sender chooses are $(\mu^{FOC}, 0)$ and $(\theta + k, k)$.*
- *If $k > \frac{\alpha}{2\beta}$, then the only possible messages the sender chooses are $(\theta, 0)$, $(\mu^{FOC}, 0)$, and (μ^{FOC}, k).*

Which message is chosen depends on the relationship between γ and $Var_{a \sim f^k}(a)$ with the other parameters. As γ increases, the sender switches from lies to truthful messages; since senders only use imprecise messages as (imprecise) truths, this means that increasing γ leads to a switch from precise to imprecise formats.

Our data are consistent with this behavior in the region of $k > \frac{\alpha}{2\beta}$. In this region, senders who prefer imprecise messages consistently slant, while senders who prefer precise messages either slant by the same amount or do not slant at all.

We have thus far allowed for any real-valued states and any messages to be sent, but in our study there are natural bounds. For instance, in Study 1, the state is between 0% and 100%, as are the bounds for messages. For senders with directional-high incentives, μ^{FOC} may be impossible to send, as it involves messages that include values above the upper bound. This will thus push more senders to $(\theta, 0)$ instead of the slanted message, leading both to less slanted messages (since θ is sent instead of μ^{FOC}), and more precise messages (since $k > 0$ messages are less valuable when slant is limited).

4.2 Discussion

Psychologically, this behavior has commonalities with moral wiggle room, in which people intentionally avoid information in order to excuse selfish actions (e.g., Dana et al. 2007). Here, senders are avoiding sending precise messages in order to excuse the selfish action of slanting messages towards their incentives. Interestingly, the heterogeneities we observe in Study 1 suggest that many senders appear to be aware of their strategy, suggesting that it may not be as much about self-deception as about signaling to receivers that they are not behaving deceitfully.[41]

These results are consistent with senders having a greater psychological disutility of misreporting precise messages than imprecise messages (as in Serra-Garcia et al. 2011), and primarily affect people who have "intermediate" preferences.[42] Senders with a high cost of misreporting will not slant either numeric or language messages, and senders who have no cost

[41]This observation relates to the findings of Serra-Garcia and Saccardo (2023) that many people are sophisticated about how they distort beliefs.

[42]Another explanation could be that senders use imprecision to signal that they wish to take less responsibility for the messages that are sent, and future work could unpack which of these explanations fits behavior better.

of misreporting will slant both messages to maximize expected payoffs; both of these types would not have a clear reason to prefer language over numbers. However, an intermediate type who finds it psychologically costly to slant numbers, but less costly to slant language, would end up both preferring language and slanting language more, exhibiting the patterns we observe.

It is worth noting that all of these results presuppose that messages have inherent meaning to senders. For instance, the message "It is probable that you will draw a red ball." is meaningful because the word "probable" has meaning outside of this context, but this meaning can plausibly reflect a probability of 51% or 81%. With the message "There is a 66% chance that you will draw a red ball.", the "66%" has a very precise meaning outside of the experiment. This contrasts with many theories of cheap talk in which messages are only meaningful through their impact in equilibrium.[43]

Study 2 poses another natural question: How much do audiences consider researchers' incentives when interpreting research findings? Here, we consider the case where it is difficult to know whether a researcher has aligned incentives or is just trying to promote the findings of a paper. We see that policymakers are indeed persuaded by researchers in this case. Future research can explore whether this is true for other consumers of research evidence, including groups such as journal editors who might plausibly be more sophisticated about the impact of incentives on researchers.

We also caution against interpreting our results as indicating that researchers *typically* misrepresent their results. For one, Study 2 does not shed light on how often in practice researchers face strategic incentives to persuade that are strong enough to affect their communication strategies.[44] It is also worth restating that in Study 2, even in the case where incentives to persuade are explicitly provided to academic researchers and they have an anonymous one-shot interaction, most researchers do not slant numbers upwards. That is, while we do see a small subset of researchers slant numbers upwards, most researchers seem to have an aversion to *numeric* misreporting. Instead, our results suggest that is important to be more cautious when interpreting imprecise descriptions of research results, which may be distorted, relative to more precise statements of effect sizes.

We see several additional directions for future work. First, in many settings there are repeated interactions between senders and receivers; in such settings, it may be important

[43] As an example, in our Study 1, senders could send messages in the range $\{0\%, 1\%, 2\%, \ldots, 100\%\}$. We think that the interpretation of a sender choosing "66%" is not wholly dependent on the menu of options. If they faced the set $\{0\%, 2\%, 4\%, \ldots, 200\%\}$, we would not expect them to switch to sending "132%". However, in other cases, we may expect the context to be of first-order importance, especially for language.

[44] The finding in Edlin and Love (2022) that only 37% of abstracts for empirical economics papers include a numeric point estimate is notable in light of our findings, and future work might explore the degree to which strategic incentives explain this low propensity to use numbers in practice.

for senders to build a reputation for not misreporting messages. This could scale up or down our effect; on the one hand, receivers may learn to infer senders' strategies, making all distortions less beneficial. But it could also amplify senders' tendency to use language, since distorted numeric predictions may be more transparently incorrect if the state is revealed. Relatedly, there are many other types of incentives that affect communication. When building a reputation, researchers do not just wish to persuade audiences that their effects are large in one paper, but that they should be trusted to analyze additional effects credibly. If using numbers signals precision, a researcher may want to use numbers for precise effects and language for imprecise effects. As another example, in cases where receivers engage in motivated reasoning, reputation incentives may push senders towards imprecisely confirming what receivers are motivated to believe using language, rather than giving unbiased estimates of the truth.[45]

Second, it is important to document the degree to which persuasion impacts real-world communication. For instance, when incentives for journalists, companies, or researchers vary, how often do they shift to describing numeric effects with language, and is there a way to quantify the slant that they use? For instance, Gentzkow and Shapiro (2010) uses the similarity of text data with congressional speeches to quantify political slant of newspapers, and Raymond and Taylor (2021) uses variation in incentives induced by the timing of baseball games to study how local newspapers numerically report weather forecasts. Testing incentive variation on the extensive margin of reporting in numbers versus language, and on the intensive margin of the particular numbers and words used, could justify the external validity of the mechanisms we discuss.

Finally, our results suggest that receivers do not fully adjust to senders' strategies. From a policy perspective, it is important to understand how much of a role receivers' sophistication can play in affecting senders' behavior. If receivers became more aware that observing language messages was a signal that senders were distorting more, this may make them form more accurate beliefs, and be persuaded less, by language. In turn, this could lead senders to send more precise messages in the first place, as using imprecise messages to persuade would be less beneficial. Policies that increase awareness of the strategic use of language can potentially improve the efficiency of communication and the accuracy of people's beliefs. Future work can also shed light on the efficacy of language use guidelines or even requirements for reporting numbers alongside or instead of language (such as the CONSORT guidelines for reporting on medical RCTs) as alternative policy tools.

[45] In a political context, Thaler (2023) shows that incentives to be perceived as truthful can lead senders to send more directly-false messages when receivers engage in motivated reasoning. This argument may extend to message format choices as well.

References

Abeler, Johannes, Daniele Nosenzo, and Collin Raymond (2019). "Preferences for Truth-Telling." In: *Econometrica* 87(4), pp. 1115–1153.

Alempaki, Despoina, Valeria Burdea, and Daniel Read (2023). *Deceptive Communication: Direct Lies vs. Ignorance, Partial-Truth and Silence*. Discussion Paper 444. University of Munich (LMU), Humboldt University Berlin, Collaborative Research Center Transregio 190: Rationality, and Competition. URL: https://www.econstor.eu/bitstream/10419/282135/1/444.pdf.

Andrews, Isaiah and Jesse M. Shapiro (2021). "A Model of Scientific Communication." In: *Econometrica* 89(5), pp. 2117–2142.

Bhatia, Rajendra and Chandler Davis (2000). "A better bound on the variance." In: *The American Mathematical Monthly* 107(4), pp. 353–357.

Blume, Andreas, Ernest K Lai, and Wooyoung Lim (2020). "Strategic information transmission: A survey of experiments and theoretical foundations." In: *Handbook of experimental game theory*. Ed. by C. M. Capra, Rachel T.A. Croson, Mary L. Rigdon, and Tanya S. Rosenblat. Edward Elgar Publishing. Chap. 13, pp. 311–347.

Cai, Hongbin and Joseph Tao-Yi Wang (2006). "Overcommunication in strategic information transmission games." In: *Games and Economic Behavior* 56(1), pp. 7–36.

Chang, L. W., E. L. Kirgios, S. Mullainathan, and K. L. Milkman (2024). "Does counting change what counts? Quantification fixation biases decision-making." In: *Proceedings of the National Academy of Sciences* 121 (46).

Charness, Gary and Martin Dufwenberg (2006). "Promises and Partnership." In: *Econometrica* 74(6), pp. 1579–1601.

Choe, Kyoung Whan, Jalisha Jenifer, Christopher Rozek, Marc Berman, and Sian Beilock (2019). "Calculated avoidance: Math anxiety predicts math avoidance in effort-based decision-making." In: *Science Advances* 5(11), eaay1062.

Crawford, Vincent (1998). "A Survey of Experiments on Communication via Cheap Talk." In: *Journal of Economic Theory* 78(2), pp. 286–298.

Crawford, Vincent and Joel Sobel (1982). "Strategic Information Transmission." In: *Econometrica* 50(6), pp. 1431–1451.

Dana, Jason, Roberto Weber, and Jason Xi Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness." In: *Economic Theory* 33.

DellaVigna, Stefano, Woojin Kim, and Elizabeth Linos (2024). "Bottlenecks for Evidence Adoption." In: *Journal of Political Economy* 132(8).

Deversi, Marvin, Alessandro Ispano, and Peter Schwardmann (2021). "Spin doctors: An experiment on vague disclosure." In: *European Economic Review* 139.

Dhami, Mandeep K. and David R. Mandel (2021). "Words or numbers? Communicating probability in intelligence analysis." In: *American Psychologist* 76(3), pp. 549–560.

Dhami, Mandeep K. and David R. Mandel (2022). "Communicating uncertainty using words and numbers." In: *Trends in Cognitive Sciences* 26(6), pp. 514–526.

Dickhaut, John W., Kevin A. McCabe, and Arijit Mukherji (1995). "An experimental study of strategic information transmission." In: *Economic Theory* 6(3), pp. 389–403.

Edlin, Aaron S. and Michael Love (2022). *Needed in Empirical Social Science: Numbers*. Working Paper. SSRN. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4195779.

Exley, Christine (2016). "Excusing selfishness in charitable giving: The role of risk." In: *The Review of Economic Studies* 83 (2).

Farrell, Joseph and Matthew Rabin (1996). "Cheap Talk." In: *Journal of Economic Perspectives* 10(3), pp. 103–118.

Fischbacher, Urs and Franziska Föllmi-Heusi (2013). “Lies in Disguise - An Experimental Study on Cheating.” In: *Journal of the European Economic Association* 11(3), pp. 525–547.

Gauchat, Gordon (2012). “Politicization of science in the public sphere: A study of public trust in the United States, 1974 to 2010.” In: *American Sociological Review* 77(2), pp. 167–187.

Gentzkow, Matthew and Jesse Shapiro (2010). “What Drives Media Slant? Evidence From U.S. Daily Newspapers.” In: *Econometrica* 78 (1).

Gneezy, Uri (2005). “Deception: The Role of Consequences.” In: *American Economic Review* 95(1), pp. 384–394.

Graeber, Thomas, Shakked Noy, and Christopher Roth (2024a). *Lost in Transmission*. CESifo Working Paper 2024. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4711260.

Graeber, Thomas, Christopher Roth, and Constantin Schesch (2024b). *Explanations*. CESifo Working Paper 2024. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4856950.

Graeber, Thomas, Christopher Roth, and Florian Zimmerman (2024c). “Stories, Statistics and Memory.” In: *Quarterly Journal of Economics*, pp. 1–45. DOI: https://doi.org/10.1093/qje/qjae020.

Hjort, J., D. Moreira, G. Rao, and J.F. Santini (2021). “How research affects policy: Experimental evidence from 2,150 Brazilian municipalities.” In: *American Economic Review* 111 (5), pp. 1442–1480.

Ho, Emily H., David V. Budescu, Mandeep K. Dhami, and David R. Mandel (2015). “Improving the communication of uncertainty in climate science and intelligence analysis.” In: *Behavioral Science and Policy* 1(2), pp. 43–55.

Hopster, Jeroen (2023). “Climate Change, Uncertainty, and Policy.” In: *Handbook of the Philosophy of Climate Change*. Ed. by Gianfranco Pellegrino and Marcello Di Paola. Springer, pp. 977–1000.

Hossain, Tanjim and Ryo Okui (2013). “The binarized scoring rule.” In: *The Review of Economic Studies* 80(3), pp. 984–1001.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin (2022). “Complex Disclosure.” In: *Management Science* 68(5), pp. 3236–3261.

Kahan, D., E. Peters, M. Wittlin, P. Slovic, L.L. Ouellette, D. Braman, and G. Mandel (2012). “The polarizing impact of science literacy and numeracy on perceived climate change risks.” In: *Nature Climate Change* 2, pp. 732–735.

Kartik, Navin (2009). “Strategic Communication with Lying Costs.” In: *The Review of Economic Studies* 76(4), pp. 1359–1395.

Krupka, E. and R. Weber (2013). “Identifying social norms using coordination games: Why does dictator game sharing vary?” In: *Journal of the European Economic Association* 11 (3), pp. 495–524.

Lee, David S. (2009). “Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects.” In: *The Review of Economic Studies* 76(3), pp. 1071–1102.

Lipman, Barton L. (2009). *Why is Language Vague?* Working Paper. URL: https://open.bu.edu/handle/2144/40222.

Mastrandrea, Michael D., Katharine J. Mach, Gian-Kasper Plattner, Ottmar Edenhofer, Thomas F. Stocker, Christopher B. Field, Kristie L. Ebi, and Patrick R. Matschoss (2011). “The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups.” In: *Climactic Change* 108(675).

Mehmood, S., S. Naseer, and D. Chen (2023). “Training policymakers in econometrics.” In: *Working Paper*.

Mosteller, Frederick and Cleo Youtz (1990). “Quantifying Probabilistic Expressions.” In: *Statistical Science* 5(1), pp. 2–12.

Ott, Douglas E. (2021). “Words Representing Numeric Probabilities in Medical Writing Are Ambiguous and Misinterpreted.” In: *Journal of The Society of Laparoscopic and Robotic Surgeons* 25(3).

Prelec, Drazen (1998). “The probability weighting function.” In: *Econometrica* 66(3), pp. 497–527.

Raymond, Collin and Sarah Taylor (2021). ““Tell all the truth, but tell it slant”: Documenting media bias.” In: *Journal of Economic Behavior & Organizations* 184.

Rindermann, Heiner, David Becker, and James Thompson (2018). “Honesty, rule violation and cognitive ability: A reply to Gächter and Schulz.” In: *Intelligence* 68, pp. 66–69.

Schulz, K F, D G Altman, and D Moher (2010). “CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.” In: *BMJ* 340.

Serra-Garcia, Marta and Silvia Saccardo (2023). “Enabling or Limiting Cognitive Flexibility? Evidence of Demand for Moral Commitment.” In: *American Economic Review* 113(2).

Serra-Garcia, Marta, Eric Van Damme, and Jan Potters (2011). “Hiding an inconvenient truth: Lies and vagueness.” In: *Games and Economic Behavior* 73(1), pp. 244–261.

Sobel, Joel (2020). “Lying and Deception in Games.” In: *Journal of Political Economy* 128(3).

Spiess, J. (2024). “Optimal Estimation when Researcher and Social Preferences are Misaligned.” In: *Working Paper*.

Thaler, Michael (2023). “The Supply of Motivated Beliefs.” In: *Working Paper*. URL: https://hq.ssrn.com/submissions/AbstractComments.cfm?AuthorID=2725209&AbstractID=4494688.

Toma, Mattie (2024). *Policymakers Lab*. https://warwick.ac.uk/fac/soc/wbs/subjects/bsci/research/policymakers-lab.

Toma, Mattie and Elizabeth Bell (2024). “Understanding and increasing policymakers’ sensitivity to program impact.” In: *Journal of Public Economics* 234.

Tversky, Amos and Daniel Kahneman (1992). “Advances in prospect theory: Cumulative representation of uncertainty.” In: *Journal of Risk and Uncertainty* 5(4), pp. 297–323.

Vespa, Emanuel and Alistair Wilson (2016). “Communication with multiple senders: An experiment.” In: *Quantitative Economics* 7(1), pp. 1–36.

Vivalt, Eva and Aidan Coville (2023). “How do policymakers update their beliefs?” In: *Journal of Development Economics* 165.

Weiszsäcker, Georg (2023). *Misunderstandings: False Beliefs in Communication*. Open Book Publishers: Cambridge, UK.

Wood, Daniel H. (2022). “Communication-Enhancing Vagueness.” In: *Games* 13(4), p. 49.

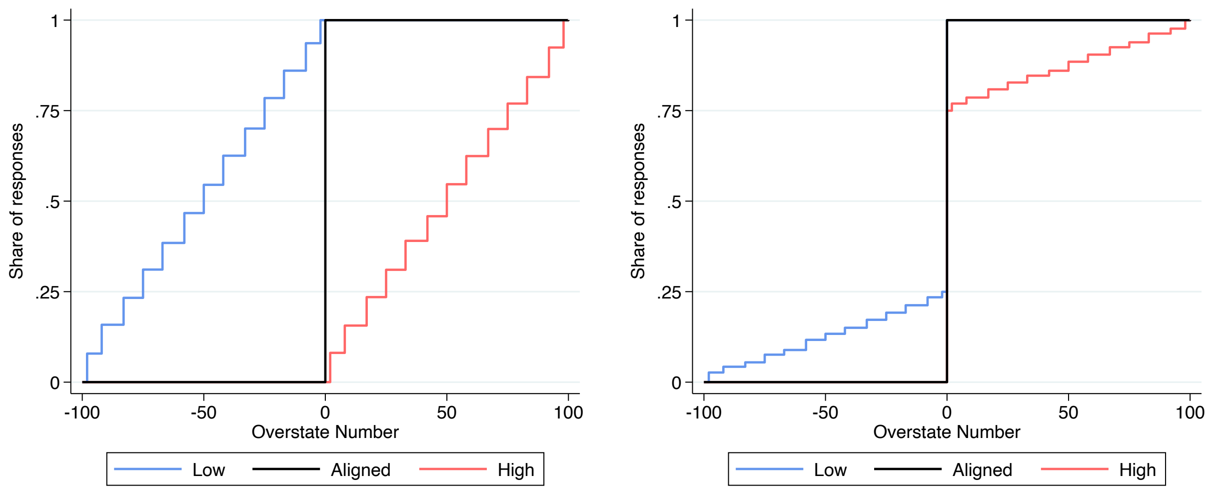
Zhang, Sookie Xue and Ralph-Christopher Bayer (2023). “Delegation based on cheap talk.” In: *Theory and Decision* 94, pp. 333–361.

Appendix

A Additional Results

A.1 Study 1

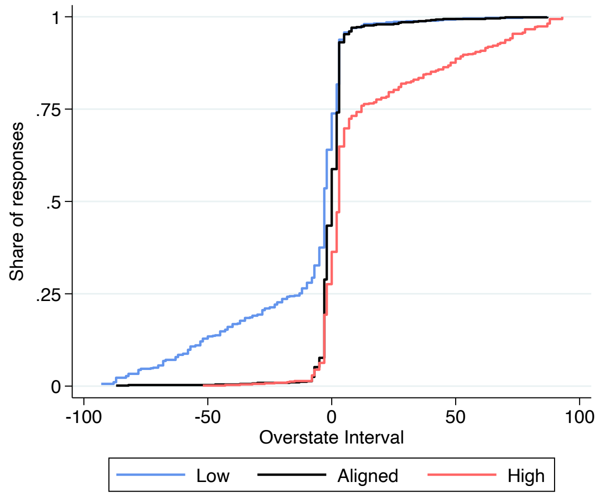
Figure A1: Hypothetical CDFs of message slant by incentives



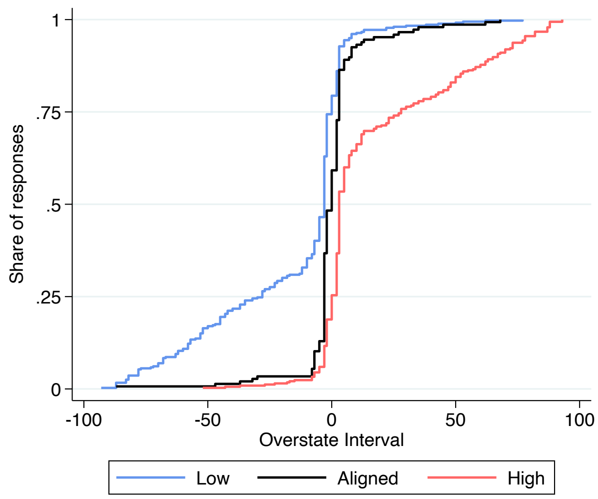
Notes: This figure plots hypothetical empirical cumulative distribution functions (eCDFs) for slant, by incentive conditions. The x-axis reflects the degree to which senders overstate their message compared to the truth: "Overstate number" subtracts the real probability from the numeric message. The lefthand figure reports a hypothetical case in which all senders in our data maximally slant their messages in the direction of their incentives. The righthand figure reports a hypothetical case in which 75% of senders always tell the truth while 25% maximally slant their messages in the direction of their incentives.

Figure A2: Empirical CDFs of message slant by incentives for interval messages

eCDF: Unconditional on preferred format



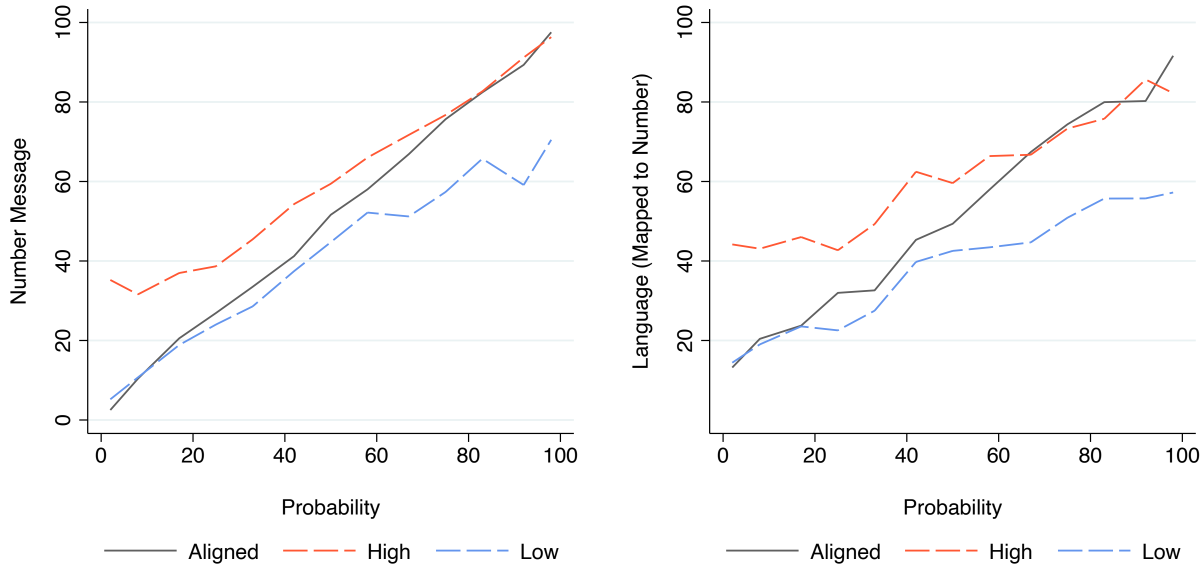
eCDF: Conditional on preferred format



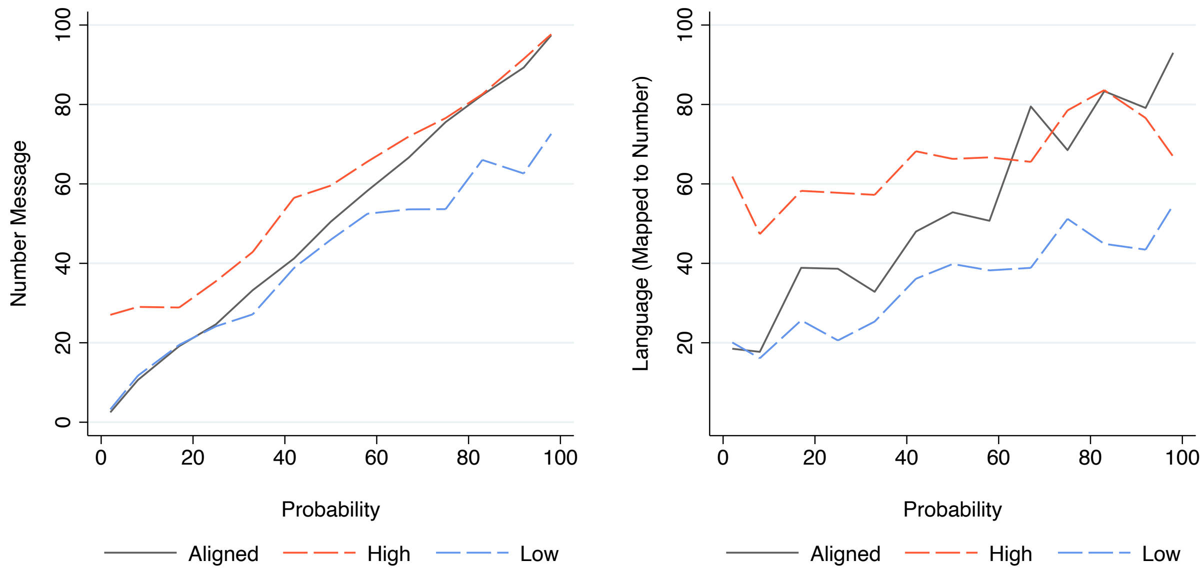
Notes: This figure plots the empirical cumulative distribution functions (eCDFs) for slant, by incentive conditions, for interval format messages. The x-axis, "Overstate interval" reflects the degree to which senders overstate their interval message compared to the truth. In particular, it takes the midpoint of a given interval message and then subtracts the true probability. The top panels show the distribution for all messages, while the bottom panels condition on the messages using the sender's preferred format. The sample includes 1,992 choices across 249 senders. Similar eCDFs for language and numeric messages are shown in Figure 3.

Figure A3: Sender's message by true probability, message, and preferred format

Unconditional on preferred format

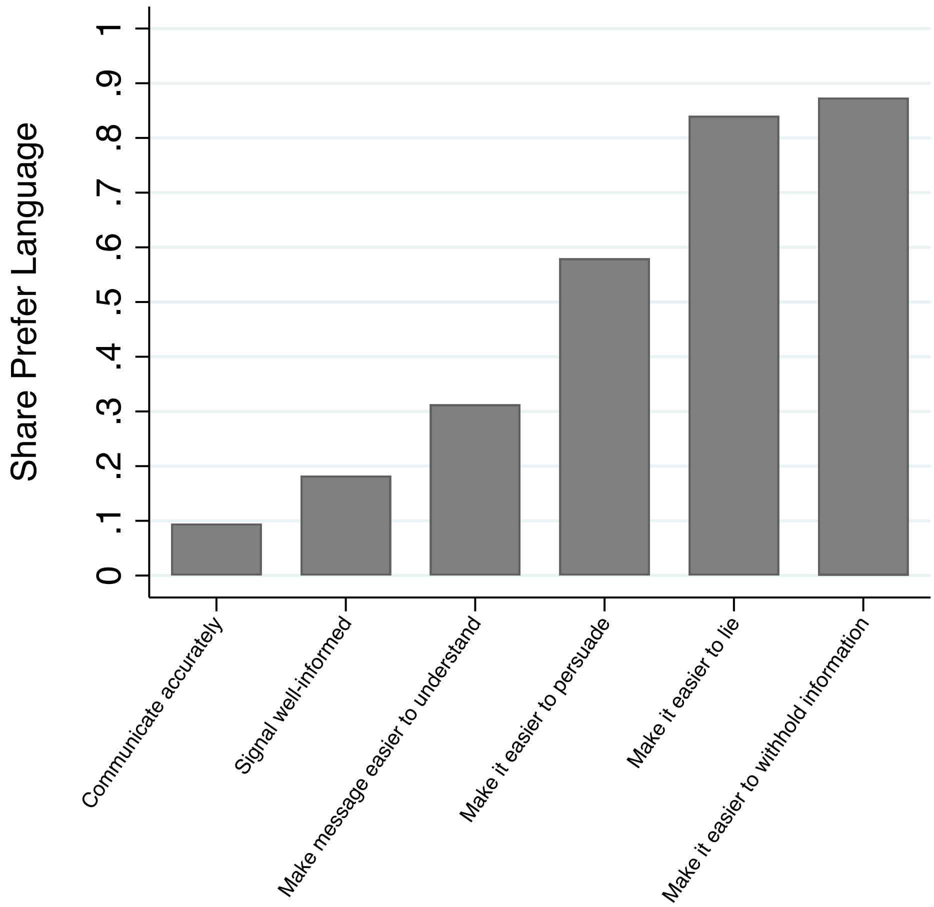


Conditional on preferred format



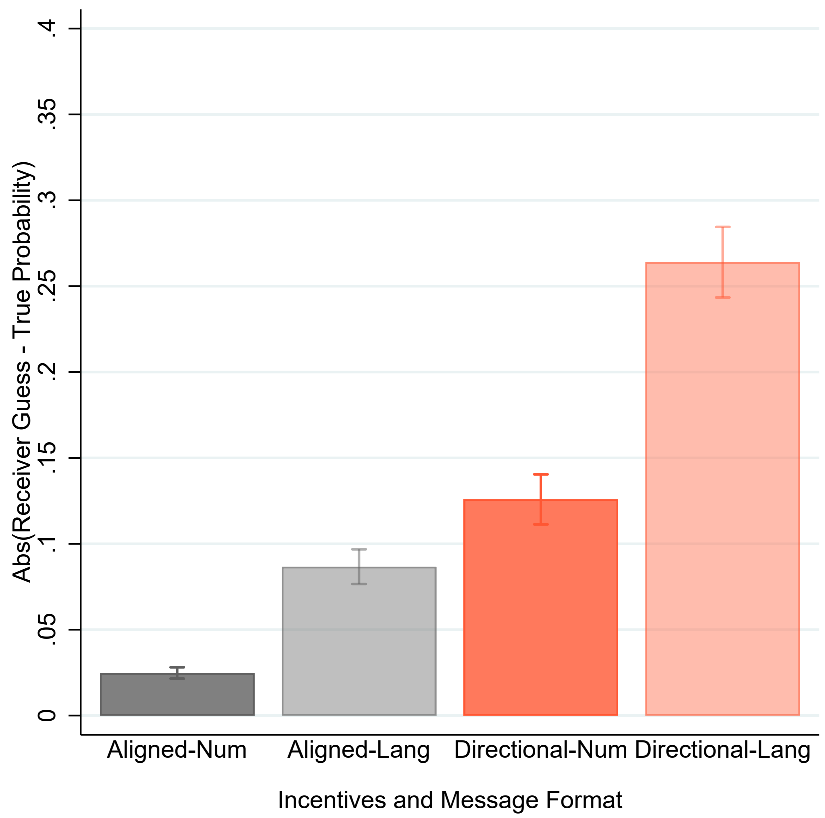
Notes: This figure plots the sender's message on the y axis and the true probability on the x axis, by incentive condition. The top panel shows all messages sent using numbers (left) and language (right). The bottom panel shows only the number messages when the numeric format was preferred (left) and the language messages when the language format was preferred (right). Language messages are included using the language-to-number mapping from the benchmarking exercise.

Figure A4: Preferred uses for language



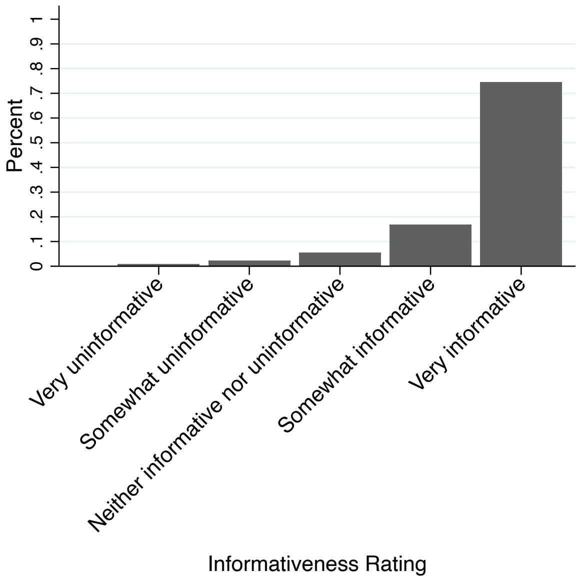
Notes: This figure plots the fraction of senders who preferred communicating likelihoods with language rather than numbers, for the following situations shown along the x axis: (1) wanting to communicate as accurately as possible; (2) wanting to signal to someone that one is well-informed; (3) making one's message easier to understand; (4) making it easier to persuade someone that the likelihood is higher than it actually is; (5) making it easier to lie to someone; (6) making it easier to withhold information.

Figure A5: Receiver errors by sender incentives and message type

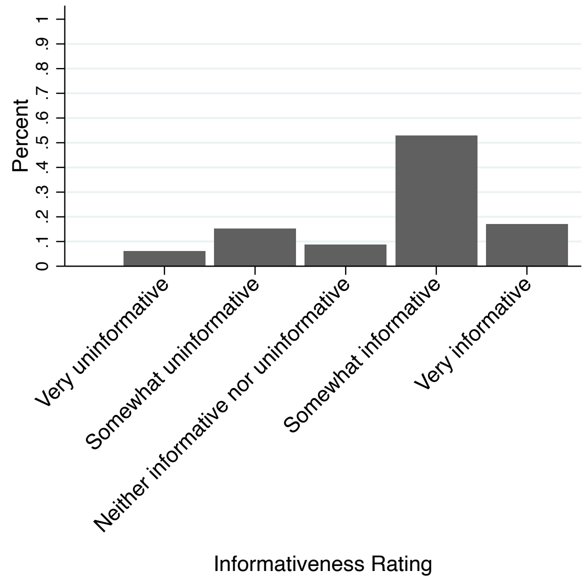


Notes: This figure plots the receiver's guess minus the true probability for each of four different conditions: i) the paired sender faced aligned incentives and communicated a number message, ii) the paired sender faced aligned incentives and communicated a language message, iii) the paired sender faced directional incentives and communicated a number message, iv) the paired sender faced directional incentives and communicated a language message.

Figure A6: Informativeness ratings for numeric and language messages



(a) Numeric messages



(b) Language messages

Notes: This figure plots the share of receivers selecting each respective option on a 5-point Likert scale, regarding the informativeness of the message they received from senders. (a) shows results for receivers who received numeric messages. (b) shows results for receivers who received language messages.

Table A1: Language used in Study 1

Word (or Phrase)	**Median response**
Practically impossible	5%
Improbable	15%
Doubtful	20%
Unlikely	20%
Less likely than not	40%
About an even chance	50%
Possible	55%
A decent chance	65%
Probable	70%
Likely	75%
Expected	80%
Almost certain	90%
Practically guaranteed	95%

Notes: The left-hand column shows the complete list of words used in the drop-down list, as given to senders who were asked to select a language message. The right-hand column shows median responses to the benchmarking exercise among senders. For details on the benchmarking exercise see Section 2.2.2.

Table A2: Message slant by format, overall

	(1)	(2)	(3)	(4)
	Number - Prob	Language - Prob	Number - Prob	Interval - Prob
High Incentives	9.819***	9.862***	10.833***	11.541***
	(1.239)	(1.402)	(1.348)	(1.321)
Low Incentives	-10.674***	-12.944***	-12.329***	-13.399***
	(1.268)	(1.417)	(1.457)	(1.463)
Observations	2008	2008	1695	1695
Aligned Mean	.3	.17	-.03	-.11
Respondent FE	Yes	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes	Yes
Condition	Num vs Lang	Num vs Lang	Num vs Int	Num vs Int

Notes: This table reports the effect of incentives on message slant by format. Columns 1-2 restrict to the sample of senders who chose between language and numeric messages. In Columns 1 and 3, the dependent variable *Number - Prob* is the number that the sender selected to communicate, minus the true probability they were given. In Column 2, the dependent variable *Language - Prob* is the language message selected for communication, mapped to numbers using the benchmarking exercise, minus the true probability. Columns 3-4 restrict to the sample of senders who chose between interval and numeric messages and see a probability strictly between 2-98%. In Column 4, the dependent variable *Interval - Prob* is the the midpoint of the selected interval message, minus the true probability. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and incentive type. In the *High Incentives* row, the independent variable is an indicator equal to one when the sender faced directional-high incentives. In the *Low Incentives* row, the independent variable is an indicator equal to one when the sender faced directional-low incentives. Fixed effects for the respondent and true probability are included as controls. Aligned mean calculates the likelihood of using language among aligned individuals in the corresponding sample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A3: Slant of imprecise messages compared to numeric messages, by incentives

	(1)	(2)	(3)
	Language - Number	Interval - Number	Imprecise - Number
High Incentives	2.313	1.749***	2.331***
	(1.446)	(0.444)	(0.771)
Observations	1329	1124	2242
Mean over Obs	-0.95	-0.03	-0.26
Respondent FE	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes

Notes: This table considers slant between the sender's chosen imprecise communication and their numeric message, and reports relative differences in slant by incentive type. In Column 1, the dependent variable is the language message selected for a communication, mapped to numbers using the benchmarking exercise, minus the numeric message. In Column 2, the dependent variable is the midpoint of the selected interval message minus the numeric message. Column 3 pools the differences across Columns 1 and 2. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and incentive type. In the row *High Incentives*, the independent variable is an indicator equal to one when senders faced directional-high incentives. These results were calculated over the subsample of senders who faced either a directional-high or a directional-low incentive (i.e. excluding aligned senders). We include fixed effects for the respondent and true probability as controls. Mean over obs reflects the average difference between the imprecise and numeric messages for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A4: Relative slant by senders' preferred format

	(1)	(2)	(3)
	Language - Number	Interval - Number	Imprecise - Number
Use ImpreciseXHigh	4.530*	2.935***	3.847***
	(2.728)	(1.053)	(1.324)
Use ImpreciseXLow	-5.938**	-0.784	-2.896**
	(2.833)	(0.957)	(1.330)
Observations	2008	1695	3398
Mean over Obs	-0.32	-0.13	-0.06
Respondent FE	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: This table reports the relative difference in slant between the chosen imprecise communication and numbers, by incentive condition and preferred message. In Column 1, the dependent variable is the language message selected for a communication, mapped to numbers using the benchmarking exercise, minus the numeric message. In Column 2, the dependent variable is the midpoint of the selected interval message minus the numeric message. Column 3 pools the differences across Columns 1 and 2. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and incentive type. *Use ImpreciseXHigh* interacts an indicator equal to one when the sender preferred the imprecise communication with an indicator for directional-high incentives. Similarly, *Use ImpreciseXLow* interacts an indicator equal to one when the sender preferred the imprecise communication with an indicator for directional-low incentives. Controls include the individual indicators as well as fixed effects for the respondent and true probability. Mean over obs reflects the average difference between the imprecise and numeric messages for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A5: Format choice by true probability

	(1) Use Language	(2) Use Intervals	(3) Use Imprecise
High X Probability	-0.135*	-0.104	-0.118**
	(0.071)	(0.078)	(0.053)
Low X Probability	0.263***	0.170**	0.216***
	(0.074)	(0.083)	(0.055)
Observations	2008	1992	4000
Mean over Obs	0.30	0.42	0.36
Respondent FE	Yes	Yes	Yes

Notes: This table reports how format choice varies by being in a good or bad state of the world. For instance, if a sender has directional-high incentives, a good state is one where the true probability is actually high; a bad state is one where the true probability is actually low. In Column 1, the dependent variable is an indicator equal to one when the sender chose to communicate using language rather than numbers. In Column 2, the dependent variable is an indicator equal to one when the sender chose to communicate using intervals rather than numbers. Columns 1 is calculated over the sample of senders who had the option to communicate in language or numbers; Columns 2 is calculated over the sample of senders who had the option to communicate in intervals or numbers. Column 3 pools together Columns 1 and 2, with the dependent variable now being an indicator which equals one if the sender communicated with any imprecise message type, either language or intervals. In the row *High X Probability*, the independent variable is an interaction between the true probability, and an indicator which equals one if the sender had directional-high incentives. In the row *Low X Probability*, the independent variable is an interaction between the true probability, and an indicator which equals one if the sender had directional-low incentives. We control for respondent fixed effects. Mean over obs reflects the likelihood of using an imprecise message format in the corresponding subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A6: Effect of senders' incentives on receivers' guesses

	(1)	(2)	(3)
	Guess (Language)	Guess (Interval)	Guess (All)
High Incentives	14.814***	20.077***	17.045***
	(2.199)	(2.552)	(1.682)
Observations	703	589	1292
Overall Mean	51.48	51.09	51.31
Respondent FE	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes

Notes: This table shows that receivers are persuaded in the direction of senders' incentives. Our results are only calculated over conditions where the sender faces directional-high or directional-low incentives (i.e. we exclude conditions with aligned senders). The independent variable is an indicator which equals one if the sender faced directional-high incentives. Column 1 restricts to receivers who either receive language or number messages; Column 2 restricts to receivers who either receive interval or number messages. Column 3 pools together Columns 1 and 2. The dependent variable is the receiver's guess of the true probability. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and sender incentive type. We include fixed effects for the respondent and true probability as controls. *Overall Mean* reflects the mean guess among the receivers included in each column. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A7: Effect of senders' incentives on receivers' guesses by format

	(1) Guess (Language)	(2) Guess (Interval)	(3) Guess (All)
Imprecise Sent X High	12.497***	3.770*	7.161***
	(2.370)	(1.990)	(1.522)
Imprecise Sent X Low	-5.024*	-3.378	-4.850***
	(2.554)	(2.446)	(1.742)
High Incentives	7.308***	8.272***	8.043***
	(1.522)	(1.466)	(1.047)
Low Incentives	-7.969***	-11.735***	-9.531***
	(1.404)	(1.641)	(1.070)
Observations	1984	2016	4000
Overall Mean	51.97	50.82	51.37
Format Control	Yes	Yes	Yes
Respondent FE	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes

Notes: This table shows that receivers are persuaded more in the direction of senders' incentives when they receive imprecise messages. Column 1 restricts to receivers who either receive language or number messages; Column 2 restricts to receivers who either receive interval or number messages. Column 3 pools together Columns 1 and 2. The independent variable of interest is the interaction between the sender's incentive, and an indicator which equals one if the sender sent an imprecise message (language in Column 1; intervals in Column 2; either language or intervals in Column 3). The dependent variable is the receiver's guess of the true probability. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and sender incentive type. We include fixed effects for the respondent and true probability as controls. *Overall Mean* reflects the mean guess among the receivers included in each column. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A8: Effect of senders' incentives and receivers' knowledge of senders' incentives on receivers' guesses

	(1) Guess (Language)	(2) Guess (Interval)	(3) Guess (All)
Know X High	-7.204**	-4.288	-5.647**
	(3.325)	(3.211)	(2.315)
High Incentives	26.446***	26.154***	26.250***
	(2.158)	(1.985)	(1.465)
Observations	1276	1308	2584
Overall Mean	55.96	48.84	51.97
Respondent FE	Yes	Yes	Yes
Probability FE	Yes	Yes	Yes

Notes: This table shows that receivers are somewhat less persuaded by senders when receivers have knowledge of sender incentives. Our results are only calculated over conditions where the sender faces directional-high or directional-low incentives (i.e. we exclude conditions with aligned senders). Column 1 restricts to receivers who either receive language or number messages; Column 2 restricts to receivers who either receive interval or number messages. Column 3 pools together Columns 1 and 2. The independent variable of interest is the interaction between two indicators: an indicator which equals one if the sender had high incentives, and an indicator which equals one if receivers know the incentives of their paired sender. The dependent variable is the receiver's guess of the true probability. Dependent variables are winsorized at the 2% and 98% level, with a separate winsorization calculated over each combination of true probability and sender incentive type. We include fixed effects for the respondent and true probability as controls. *Overall Mean* reflects the mean guess among the receivers included in each column. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

A.2 Study 2

Table A9: Lee bounds for effect of directional incentives on researchers' use of language

	(1) Main specification	(2) Lee lower bound	(3) Lee upper bound
Directional Incentives	0.291***	0.284***	0.335***
	(0.054)	(0.055)	(0.051)
Observations	290	278	278
Mean over obs	0.26	0.27	0.24
Subject controls	Yes	Yes	Yes
Study FE	Yes	Yes	Yes

Notes: Here we report Lee bounds (Lee 2009) for the effect of directional incentives on researchers' use of language, an effect previously reported in Table 4 of the main paper. Our sample of researcher exhibited differential attrition. In the control group, 1 respondent attrited and 76 did not, for a response rate of 98.7%. In the treatment group, 7 respondents attrited and 69 did not. We assume that some respondents may have attrited after being assigned to the treatment group, because they did not want to proceed with the task of persuading policymakers. The difference in attrition rates between treatment and control was 7.9%. Thus we seek to trim 7.9/98.7 = 8% of control respondents, i.e. 6 control respondents. To compute the lower bound, we randomly drop 6 control respondents who chose to communicate with numbers for both studies. Regarding the upper bound, there were only 3 control respondents who chose to communicate with language for both studies. We dropped these, and then on top of this randomly dropped a further 3 control respondents who communicated with language for one study and numbers for another. As a baseline, Column 1 displays the effect from the main specification reported in Table 4 of the main paper. Column 2 reports the Lee lower bound, and Column 3 reports the Lee upper bound. In all columns, the dependent variable is an indicator which equals one when the researcher chose to communicate with language rather than numbers. The independent variable is an indicator which equals one when researchers had directional-high incentives. Subject controls include indicators for whether the researcher is tenure-track, male, and does empirical research. We also control via fixed effects for each of the six real research studies whose results were being communicated in the experiment. *Mean over obs* reflects the likelihood of using language for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table A10: Relative slant by researchers' preferred format

	(1)
	log(Lang)-log(Num)
Use Language X Directional Incentives	0.494**
	(0.243)
Observations	289
Mean over obs	-0.22
Subject controls	Yes
Study FE	Yes

Notes: This table reports the relative difference in slant between the language and number messages, by incentive condition and whether the researcher preferred communicating using language. The dependent variable is the language message selected for a communication, mapped to numbers using the benchmarking exercise and logged, minus the logged numeric message. *Use Language X Directional Incentives* interacts an indicator equal to one when the sender preferred the language message with an indicator equal to one for directional-high incentives, controlling for the these indicators. Subject controls include indicators for whether the researcher is tenure-track, male, and does empirical research. We also control via fixed effects for each of the six real research studies whose results were being communicated in the experiment. *Mean over obs* reflects the average difference between the (logged) language and number messages for that subsample. Standard errors, clustered at the individual level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

B Additional Theory Details

B.1 Proof of Lemma 1

Proof. We restate Equation (2) below:

$$u_S(\mathbf{m}, a(\mathbf{m})) = \alpha \cdot a(\mathbf{m}) - \beta \cdot (a(\mathbf{m}) - \theta)^2 + \gamma \cdot \mathbb{1}(\theta \in [m_L, m_H]).$$

The first term is linear in $a(\mathbf{m})$, and we have assumed that the distribution of receiver actions is symmetric about μ, so the expected monetary benefit for senders is $\alpha\mu$. The last term is deterministic.

All that remains is to show that $\mathbb{E}_{a\sim f^k_\mu}[(a(\mathbf{m}) - \theta)^2] = (\mu - \theta)^2 + Var_{a\sim f^k}(a)$.

$$\begin{aligned}
\mathbb{E}_{a\sim f^k_\mu}\left[(a(\mathbf{m}) - \theta)^2\right] &= \int_{\mu-k}^{\mu+k} (a-\theta)^2 f^k(a)da \\
&= \theta^2 \int_{\mu-k}^{\mu+k} f^k(a)da - 2\theta \int_{\mu-k}^{\mu+k} af^k(a)da + \int_{\mu-k}^{\mu+k} a^2 f^k(a)da \\
&= \theta^2 - 2\theta\mu + \int_{\mu-k}^{\mu+k} a^2 f^k(a)da \\
&= (\theta-\mu)^2 - \mu^2 + \int_{\mu-k}^{\mu+k} a^2 f^k(a)da \\
&= (\theta-\mu)^2 + Var_{a\sim f^k}(a)
\end{aligned}$$

For the third row, we used the fact that the density $\int_{\mu-k}^{\mu+k} f^k(a)da$ integrates to one and that $\int_{\mu-k}^{\mu+k} af^k(a)da$, which is the expected value of a, equals μ. For the last row, we again used that the expected value of a equals μ, so that we can combine terms to get $\mathbb{E}_{a\sim f^k_\mu}[a^2] - (\mathbb{E}_{a\sim f^k_\mu}[a])^2 = Var_{a\sim f^k_\mu}(a)$. Finally, since the variance only depends on k, and not μ, we omit the μ subscript.

B.2 Proof of Proposition 2

Proof.

1. This follows immediately when we look at the contraposition. For any k, senders only send imprecise messages with mean θ when $\alpha = 0$. (Otherwise, they could slightly slant and still maintain plausible deniability.) And when $\alpha = 0$, $\mu^p = \theta$.
2. Reading off Equation (5), senders choose imprecise lies when $k < \frac{\alpha}{2\beta}$ and $\mu^i(k) = \mu^{FOC}$, which occurs when $\gamma \leq \frac{\alpha^2}{4\beta} - (\alpha - \beta k)k$. From Equation (4), senders choose precise lies

when $\gamma \leq \frac{\alpha^2}{4\beta}$. When $k \leq \frac{\alpha}{2\beta}$, $(\alpha - \beta k) \geq 0$, so $\gamma \leq \frac{\alpha^2}{4\beta} - (\alpha - \beta k)k$ implies $\gamma \leq \frac{\alpha^2}{4\beta}$, and the precise lie is also chosen.

B.3 Proof of Proposition 3

We consider the following regions:

$k \leq \frac{\alpha}{2\beta}$ and $\gamma \leq \frac{\alpha^2}{4\beta} - (\alpha - \beta k)k$

In this region, senders choose between $(\mu^{FOC}, 0)$ and (μ^{FOC}, k).

In this case, both messages are lies that induce the same expected action for receivers, so the precise message is strictly preferred.

$k \leq \frac{\alpha}{2\beta}$ and $\gamma \in [\frac{\alpha^2}{4\beta} - (\alpha - \beta k)k, \frac{\alpha^2}{4\beta}]$

In this region, senders choose between $(\mu^{FOC}, 0)$ and $(\theta + k, k)$.

They choose the precise message when:

$$\begin{aligned} \alpha\left(\frac{\alpha}{2\beta} - k\right) - \beta\left(\left(\frac{\alpha}{2\beta}\right)^2 - k^2\right) + \beta \cdot Var_{a\sim f^k}(a) - \gamma &\geq 0 \\ \iff \frac{\alpha^2}{4\beta} - \alpha k + \beta k^2 + \beta \cdot Var_{a\sim f^k}(a) &\geq \gamma \\ \iff \left(\frac{\alpha}{2\beta} - k\right)^2 + Var_{a\sim f^k}(a) &\geq \frac{\gamma}{\beta} \end{aligned}$$

In other words, precision is preferred when the benefit of slanting more, plus the cost of inducing receivers' stochasticity, outweighs the benefit of plausible deniability.

$k \leq \frac{\alpha}{2\beta}$ and $\gamma \geq \frac{\alpha^2}{4\beta}$

In this region, senders choose between $(\theta, 0)$ and $(\theta + k, k)$.

In this case, senders always choose the imprecise message.

To see this, note that they choose the precise message when:

$$-\alpha k + \beta k^2 + \beta \cdot Var_{a\sim f^k}(a) \geq 0$$

We can bound the maximum of the variance of a bounded random variable to be $Var_{a\sim f^k}(a) < k^2$ (e.g., Bhatia and Davis 2000, where our strict inequality comes from the distribution of actions being assumed atomless). As such, the left-hand side is strictly less than $k(-\alpha + 2\beta k)$.

Since $k \leq \frac{\alpha}{2\beta}$, it must be that the left-hand side is strictly negative and the inequality never holds in this region.

$k \geq \frac{\alpha}{2\beta}$ and $\gamma \leq \frac{\alpha^2}{4\beta}$

In this region, senders choose between $(\mu^{FOC}, 0)$ and (μ^{FOC}, k).

Since μ is the same, senders choose the precise message when:

$$Var_{a\sim f^k}(a) \geq \frac{\gamma}{\beta}.$$

In other words, precision is preferred when the cost of inducing receivers' stochasticity outweighs the benefit of plausible deniability.

$k \geq \frac{\alpha}{2\beta}$ and $\gamma \geq \frac{\alpha^2}{4\beta}$

In this region, senders choose between $(\theta, 0)$ and (μ^{FOC}, k).

They choose the precise message when:

$$\begin{aligned} -\alpha\left(\frac{\alpha}{2\beta}\right) + \beta\left(\frac{\alpha}{2\beta}\right)^2 + \beta \cdot Var_{a\sim f^k}(a) &\geq 0 \\ \iff \beta \cdot Var_{a\sim f^k}(a) &\geq \frac{\alpha^2}{4\beta} \\ \iff Var_{a\sim f^k}(a) &\geq \left(\frac{\alpha}{2\beta}\right)^2 \end{aligned}$$

First, consider the case where $k < \frac{\alpha}{2\beta}$. Then, the only messages the sender chooses are $(\mu^{FOC}, 0)$ and $(\theta + k, k)$. Which message chosen depends on the relationship between γ and $Var_{a\sim f^k}(a)$ and the other parameters:

- $(\mu^{FOC}, 0)$ if $\gamma \leq \beta\left(\frac{\alpha}{2\beta} - k\right)^2$.
- $(\mu^{FOC}, 0)$ if $\gamma \in \left[\beta\left(\frac{\alpha}{2\beta} - k\right)^2, \beta\left(\frac{\alpha}{2\beta}\right)^2\right]$ and $Var_{a\sim f^k}(a) \geq \frac{\gamma}{\beta} - \left(\frac{\alpha}{2\beta} - k\right)^2$
- $(\theta + k, k)$ if $\gamma \in \left[\beta\left(\frac{\alpha}{2\beta} - k\right)^2, \beta\left(\frac{\alpha}{2\beta}\right)^2\right]$ and $Var_{a\sim f^k}(a) \leq \frac{\gamma}{\beta} - \left(\frac{\alpha}{2\beta} - k\right)^2$
- $(\theta + k, k)$ if $\gamma \geq \beta\left(\frac{\alpha}{2\beta}\right)^2$

In this region, the only messages chosen are precise lies with a slant of slant of $\frac{\alpha}{2\beta}$ and imprecise truths with a smaller slant of k.

Now, consider the case where $k > \frac{\alpha}{2\beta}$. Then, the only messages the sender chooses are $(\theta, 0)$, $(\mu^{FOC}, 0)$, and (μ^{FOC}, k). Which message chosen depends on the relationship between γ and $Var_{a\sim f^k}(a)$ and the other parameters:

- $(\mu^{FOC}, 0)$ if $\gamma \leq \beta \left(\frac{\alpha}{2\beta}\right)^2$ and $Var_{a\sim f^k}(a) \geq \frac{\gamma}{\beta}$
- (μ^{FOC}, k) if $\gamma \leq \beta \left(\frac{\alpha}{2\beta}\right)^2$ and $Var_{a\sim f^k}(a) \leq \frac{\gamma}{\beta}$
- $(\theta, 0)$ if $\gamma \geq \beta \left(\frac{\alpha}{2\beta}\right)^2$ and $Var_{a\sim f^k}(a) \geq \left(\frac{\alpha}{2\beta}\right)^2$
- (μ^{FOC}, k) if $\gamma \geq \beta \left(\frac{\alpha}{2\beta}\right)^2$ and $Var_{a\sim f^k}(a) \leq \left(\frac{\alpha}{2\beta}\right)^2$

In this region, the only messages chosen involve precise true messages, precise lies with a slant of $\frac{\alpha}{2\beta}$, and imprecise truths with a slant of $\frac{\alpha}{2\beta}$.

C Online Appendix: Study Materials

C.1 Study 1: Senders

Sender Instructions, version for language vs numbers

Survey overview

You will see several pages of questions in four parts. Parts 2 and 3 are the "main parts". You will learn about each part when you are at that point in the survey.

There will be an "attention check" question in the study. The answer to this question will be obvious to anyone paying attention. Participants who do not answer the attention check question correctly **will not be eligible for a bonus payment**.

Bonus payments

After all participants complete the study, 10 participants will be chosen at random to receive a bonus payment of up to $50 based on their choices in Parts 2 and 3. In particular, **if you are randomly selected to receive a bonus payment, we will randomly choose one of the questions in Parts 2 and 3 to determine your payment.** The high bonus is because it is important for us that you take this study seriously.

In Part 2, you will be assigned the role of a "Sender" and play 8 rounds in which you will make decisions about how you would communicate information to different Prolific participants ("Receivers").

First, you will complete Part 1, which is a practice round for Part 2. You will see example questions for your role as a Sender and for Receivers so that you fully understand both roles before you proceed.

You may now click to proceed to Part 1.

Part 1: Sender instructions

Now you're ready for the instructions for the role of Senders:

In this role, in each question that is selected for a bonus you will be randomly paired with another Prolific worker (your Receiver). A computer will simulate a random draw of one ball from a box that contains 100 balls in total, some of which are red and the rest of which are blue.

A slider showing the share of red balls (out of 100) in an example box is shown below. 0 means that all are blue; 100 means that all are red.

0 10 20 30 40 50 60 70 80 90 100

Share of red balls in box (out of 100)

Your paired Receiver does not see this slider showing the share of red balls in the box. Instead, you will send a message to them to tell them something about the chance that they will draw a red ball.

There will be two message formats through which to communicate your information to your paired Receiver:

1. **Number**: You will choose a message that says "The chance that you will draw a red ball is **[X] percent**." for your choice of **X**.
2. **Interval**: You will choose a message that says "The chance that you will draw a red ball is between **[Y] and [Z] percent**." for your choice of **Y** and **Z**.

You will first be asked how you would communicate to the Receiver, given each format. For instance, you may choose "64 percent" for **Number** and "between 60 and 70 percent" for **Interval**. Your message can reflect the information you see in the slider, but it does not have to.

Then, you will answer a question about which of these **formats** you would prefer to send. Your preferred format is more likely (but not guaranteed) to be the one communicated to your paired Receiver.

Your bonus payment will depend on what your Receiver predicts after they see your message. On each question, you will learn whether your bonus payment will be higher if your Receiver:

- **Correctly predicts** the chance of drawing a **RED** ball;
- Predicts that the chance of drawing a **RED** ball is **HIGH;** OR
- Predicts that the chance of drawing a **RED** ball is **LOW**.

COMPREHENSION CHECK: How is the Sender's bonus payment determined?

- It will depend on either the accuracy of the Receiver's prediction, or whether the Receiver predicts that the chance of drawing a red ball is high or low.
- The Sender will receive a bonus equal to their number message.
- The computer will set the Sender's bonus payment to a random number.

After senders see their instructions they also see the instructions for receivers (see below) and answer a comprehension question about receivers' payments. Then before they make their first decision they see a "refresher" page with their instructions, including more details on how their own payments are calculated.

Your bonus payment will depend on what your Receiver predicts after they see your message. On each question, you will learn whether your bonus payment will be higher if your Receiver:

- **Correctly predicts** the chance of drawing a **RED** ball; the probability you will earn a bonus will be equal to the receiver's probability of earning a bonus.
- Predicts that the share of **RED** balls is **HIGH;** the probability you will earn a bonus will be equal to the percent chance your Receiver assigns to drawing a **RED** ball.
- Predicts that the share of **RED** balls is **LOW**; the probability you will earn a bonus will be equal to the percent chance your Receiver assigns to drawing a **BLUE** ball.

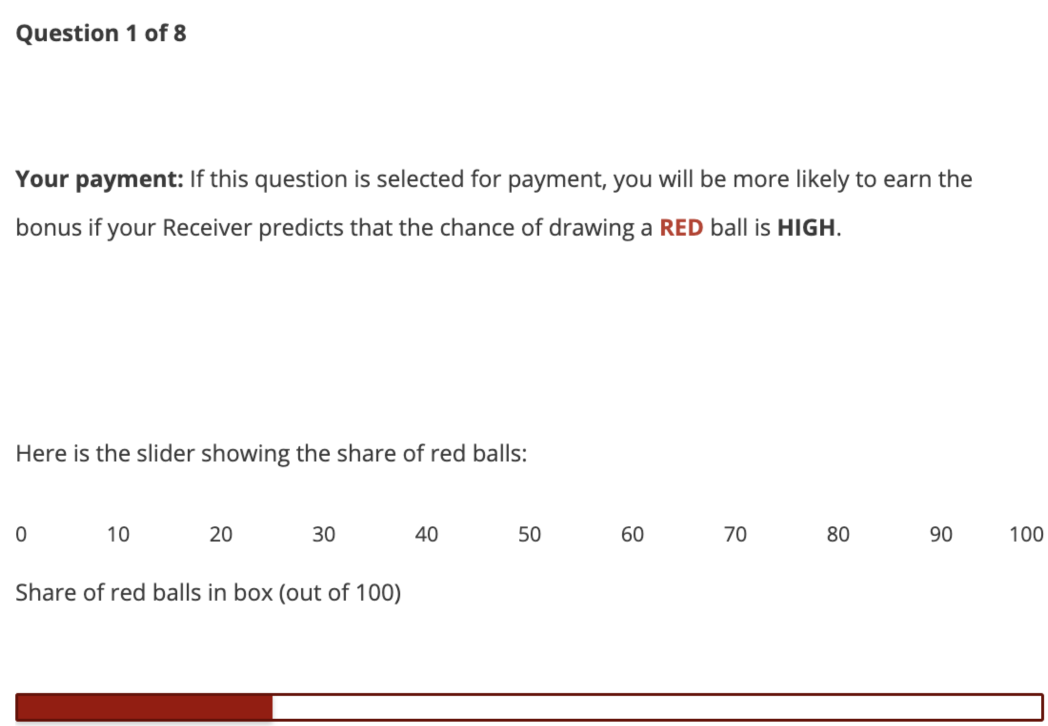
Sender Main Decisions, version with language vs numbers

Note that the comprehension check, and 'explain your reasoning' prompts, are only displayed for the first of eight questions.

The share of red balls in the box can take the following possible values: 2, 8, 17, 25, 33, 42, 50, 58, 67, 75, 83, 92, 98.

The information about the payment structure can take three different forms, based on which incentives are selected:

- "If this question is selected for payment, you will be more likely to earn the bonus if your Receiver predicts that the chance of drawing a **RED** ball is **HIGH**"
- "If this question is selected for payment, you will be more likely to earn the bonus if your Receiver predicts that the chance of drawing a **RED** ball is **LOW**"
- "If this question is selected for payment, you will be more likely to earn the bonus if your Receiver's prediction of the chance of drawing a **RED** ball is **MORE ACCURATE**"



COMPREHENSION CHECK: Which of the following reflects how your bonus payment will be determined?

- I will be more likely to win the bonus if my Receiver correctly predicts the chance of drawing a red ball
- I will be more likely to win the bonus if my Receiver predicts that the chance of drawing a red ball is high
- I will be more likely to win the bonus if my Receiver predicts that the chance of drawing a red ball is low

Number: Suppose you are assigned to send a number message to your paired Receiver. You can choose one number between 0-100%. Which message would you like to send?

"The chance that you will draw a red ball is ________ percent."

Language: Suppose you are assigned to send a natural language message to your paired Receiver. Which message would you like to send?

"It is ______ that you will draw a red ball."

Why did you make your number and language choices in this way? (You will only be asked to explain your reasoning for this first question.)

Recall, this is the slider that you saw on the previous page:

0 10 20 30 40 50 60 70 80 90 100

Share of red balls in box (out of 100)

Format

We will now ask you a question about your preferred message **format**: Number or Language.

Remember that: **If this question is selected for payment, you will be more likely to earn the bonus if your Receiver predicts that the chance of drawing a RED ball is HIGH.**

Which of the following would you prefer to send to your paired receiver?

- Number: "The chance that you will draw a red ball is percent."
- Language: "It is that you will draw a red ball."

Why did you choose between number and language messages in this way? (You will only be asked to explain your reasoning for this first question.)

Benchmarking Exercise

Part 3: Words and probabilities

You are now ready for Part 3 of the survey. The following table shows 13 words or phrases. We want to know what numeric probabilities people think each word or phrase corresponds to when describing the chance of an event happening.

Specifically, you will guess what percent chance people thought each word or phrase corresponded to (between 0 and 100 percent). At the end of the survey, we will compare each of your answers to the average answers given by all other participants. If a question is randomly selected for payment, you will earn the bonus if your answer is within is **within 3 percentage points of the average response. So, it is in your best interest to give guesses that you think will be similar to those of other participants.**

	Percent chance
possible	
doubtful	
unlikely	
expected	
a decent chance	
almost certain	
probable	
likely	
practically guaranteed	
improbable	
less likely than not	
practically impossible	
about an even chance	

C.2 Study 1: Receivers

The receivers study begins with a similar introduction. Receivers first see the same instructions for the role of the senders, then they see the following instructions for their role as receivers:

Receiver Instructions, version for language vs numbers

Part 1: Receiver instructions

Here are the instructions for the role of Receivers, to which you are assigned.

In this role, in each question that is selected for a bonus you will be randomly paired with another Prolific worker (your Sender). A computer will simulate a random draw of one ball from a box that contains 100 balls in total, some of which are red and the rest of which are blue. **Your objective is to correctly predict the chance of drawing a red (vs. blue) ball from the box.**

You will receive one of two types of messages from your Sender about the chance of drawing the red ball:

1. **Number**: "The chance that you will draw a red ball is **[X] percent**."
2. **Interval**: "The chance that you will draw a red ball is between **[Y] and [Z] percent**."

You will only see one message. For instance, you may see "The chance that you will draw a red ball is 64 percent." or "The chance that you will draw a red ball is between 60 and 70 percent.", but not both.

After receiving the message from your Sender, you will predict the percent chance of drawing a red ball from a box. This is the question that determines your bonus payment. If you would like details on the specific payment rule, see below: the important thing to know is that **the payment rule is carefully designed so that you will be most likely to earn the $50 bonus payment if you report an accurate guess about the chance of drawing a red ball.**

Recall, your Sender's bonus payment is determined by the prediction you give. You will either be matched with a Sender who is more likely to win the bonus when you:

- **Correctly predict** the chance of drawing a **RED** ball;
- Predict that the chance of drawing a **RED** ball is **HIGH;** OR
- Predict that the chance of drawing a **RED** ball is **LOW**.

You **will not be informed** about how the Sender's bonus payment is determined for each message you see.

Note that when receivers are informed (rather than uninformed) about sender incentives, the final line of instructions above will read: "You **will be informed** about how the Sender's bonus payment is determined for each message you see."

Payment rule:

If you are selected to receive a bonus payment, to determine your payment the computer will randomly choose a question and then randomly draw two numbers. For each draw, all numbers between 0 and 100 (including decimal numbers) are equally likely to be selected. Draws are independent in the sense that the outcome of the first draw in no way affects the outcome of the second draw. If this question is chosen for payment, then:

- **If the ball from the chosen question is red and the number you picked for red is *larger* than either of the two draws, you will get the $50 bonus payment.**
- **If the ball from the chosen question is blue and the number you picked for red is *smaller* than either of the two draws, you will get the $50 bonus payment.**
- **Otherwise, you will not get the bonus payment.**

COMPREHENSION CHECK: How is the Receiver's bonus payment determined?

- The higher the Receiver's prediction of the chance of drawing a red ball, the higher the probability they win the bonus payment.
- The Receiver's bonus payment is set so that the Receiver should provide their most accurate prediction of the chance of drawing a red ball.
- The computer will set the Receiver's bonus payment to a random number.

Receiver Main Decisions

Note that when receivers are informed about sender incentives, the line beginning "**Your Sender was informed about their bonus payment:**" can continue in three different ways, depending on the actual sender's incentives:

- "Your Sender is more likely to receive the bonus if you predict that the chance of drawing a **RED** ball is **HIGH**."
- "Your Sender is more likely to receive the bonus if you predict that the chance of drawing a **RED** ball is **LOW**."
- "Your Sender is more likely to receive the bonus if your prediction of the chance of drawing a **RED** ball is **MORE ACCURATE**."

When receivers are uninformed (rather than informed) about sender incentives, the line beginning "**Your Sender was informed about their bonus payment:**" will be omitted.

Question 1 of 8

Your bonus payment: You are more likely to earn the bonus if your prediction of the chance of drawing a **RED** ball is **MORE ACCURATE**.

Your Sender was informed about their bonus payment: Your Sender is more likely to receive the bonus if your prediction of the chance of drawing a **RED** ball is **MORE ACCURATE**.

Your Sender shared the following message with you:
"The chance that you will draw a red ball is 67 percent."

Given this information, what do you think the percent chance is that you will draw a red ball?

0 10 20 30 40 50 60 70 80 90 100

Percent chance of drawing a red ball

Why did you make this prediction? (You will only be asked to explain your reasoning for this first question.)

How **informative** do you think your Sender's message was on this question?

- Very informative
- Somewhat informative
- Neither informative nor uninformative
- Somewhat uninformative
- Very uninformative

C.3 Study 2: Researchers

C.3.1 Recruitment Email

Survey on communicating research results

From Toma, Mattie <Mattie.Toma@wbs.ac.uk>
Date Thu 18/07/2024 22:13
[redacted]

[redacted]

I hope you're doing well!

I'm running a survey (with Michael Thaler and Victor Wang) to investigate how researchers communicate the results of research studies to policymakers. **Would you be willing to help with this research by filling out this survey, which should take about 10 minutes to complete?**

This survey link is unique to you and so should not be shared with others:
https://wbs.qualtrics.com/jfe/form/SV_a5bxZLlNeZ15lQO?Q_CHL=gl&Q_DL=EMD_yfMUICrPqU7exVs_a5bxZLlNeZ15lQO_CGC_dyikdZ75XtnYy02&_g_=g

We very much appreciate your help. As a token of our appreciation, everyone who completes the survey will be sent a **$10 Amazon gift card**, and you will have the opportunity to earn an additional $10 depending on your responses in the survey.

All data will be kept as confidential as possible, in accordance with Warwick ethics protocol HSSREC 219.23-24.

We will close the survey on Tuesday, July 23.

Thanks so much in advance for taking the time to read this email and for considering this!
Mattie

Mattie Toma
Assistant Professor I Warwick Business School I
Behavioural Science Group I University of Warwick I Coventry I CV4 7AL I
Mattie.Toma@wbs.ac.uk

https://www.mattietoma.com/

WARWICK BUSINESS SCHOOL
THE UNIVERSITY OF WARWICK

Instructions

Survey overview

In this survey, you will make decisions about how you, as a researcher, want to communicate the results of two research studies in which a policy intervention was tested. You are deciding how you want to communicate with real policymakers. Policymakers are recruited from a sample of high-ranking public servants worldwide involved in policy decision-making.

Each policymaker will be randomly paired with a researcher. We expect that about half of researchers will be paired with policymakers.

Bonus payments

After all participants complete the survey, all researchers who are paired with a policymaker will receive a bonus payment of up to $/£10 based on their choices in the survey. In particular, **if you are selected to receive a bonus payment, we will randomly choose one question to determine your payment.** We will describe details of these incentives later.

Payment will be provided in the form of either a UK Amazon gift card (in £) or a US Amazon gift card (in $). At the end of the survey, you will have the option to choose which gift card you would rather receive.

Instructions

You are tasked with communicating the results of research studies testing a policy intervention to your paired policymaker. Your paired policymaker will also learn about the studies, but not the results.

First, you will learn more about each study and its results on the following pages. Then, you will choose a message about the results to send to your paired policymaker.

There will be two message formats you can use to communicate:

1. **Number:** You will choose a message that uses numbers to describe the research results.
2. **Language:** You will choose a message that uses a word or phrase to describe the research results.

You will first be asked how you would communicate with the policymaker, given each format. For instance, you may say “The intervention led to a 3 percentage point increase” for **Number** and “The intervention led to a moderate increase” for **Language**. Your messages can reflect what you learned about the results of the research studies, but they do not have to.

Then, you will answer a question about which of these two formats you would prefer to send. We will always send one of your two messages to the policymaker, and will send the message using your preferred format more often, so it is in your interest to answer both of these questions honestly and carefully.

Note that the final part of the instructions page, about bonus payments, reads differently depending on whether senders are assigned to the aligned or directional incentive treatment.

Bonus payment instructions for aligned treatment below:

Your bonus payments

If you are selected to receive a bonus payment on a given question, your bonus payment will depend on the decisions your paired policymaker makes based on the message you send them.

In particular, during this survey we are asking you to imagine you are the author of the study testing the policy intervention and **you are trying to give the policymaker the best understanding of the data to inform their policy decisions.**

To ensure that your incentive is consistent with this hypothetical objective, **you are more likely to earn the bonus payment if your policymaker correctly predicts the intervention's effect size.** Specifically, if the true effect size is X percentage points, you will win the bonus payment if your paired policymaker predicts the effect size is between X-1 and X+1 percentage points. Your paired policymaker's bonus payment will be the same as yours: Their bonus payment will depend on correctly guessing the results of the study.

Note that the policymakers in this survey do not know whether you are incentivized to have them form accurate beliefs or to persuade them that the effects are large.

Bonus payment instructions for directional treatment below:

Your bonus payments

If you are selected to receive a bonus payment on a given question, your bonus payment will depend on the decisions your paired policymaker makes based on the message you send them.

In particular, during this survey we are asking you to imagine you are the author of the study testing the policy intervention and **you are trying to persuade a policymaker that the research is promising to increase your chance of getting a government grant or policymaker attention.**

To ensure that your incentive is consistent with this hypothetical objective, **you are more likely to earn the bonus payment if your paired policymaker predicts the intervention's effect size is larger.** Specifically, if your paired policymaker predicts the effect size is X percentage points, the probability you will win the bonus payment will be equal to X%, with a minimum probability of 0%. Your paired policymaker's bonus payment will depend on the accuracy of their predictions.

Note that the policymakers in this survey do not know whether you are incentivized to have them form accurate beliefs or to persuade them that the effects are large.

Main decisions, Adams et al (2021) study

Note that the reminders about the bonus payment structure will display differently depending on whether the sender was assigned to the aligned or directional treatment arm.

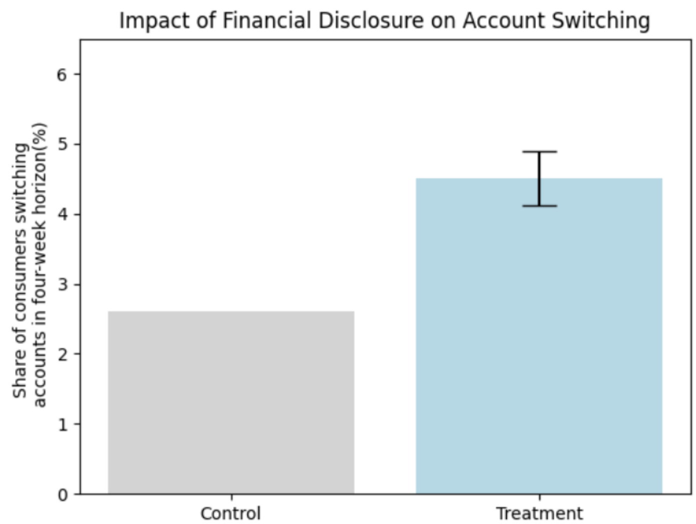
For senders in the aligned treatment, the reminder reads: "**Reminder**: When making your decisions,imagine you are the author of this study and you are trying to **give the policymaker the best understanding of the data**. In this decision you are more likely to earn the bonus payment if **your paired policymaker correctly predicts the intervention's effect size**."

For senders in the directional treatment, the reminder reads: "**Reminder**: When making your decisions, imagine you are the author of this study and you are trying to **persuade a policymaker that the research is promising**. In this decision you are more likely to earn the bonus payment if **your paired policymaker predicts the intervention's effect size is larger**."

Research study information

In a study reported in the *Journal of Financial Economics* (Adams et al. 2021), consumers with savings accounts at five UK depositaries were randomly disclosed information about alternative financial products. This information was provided on the front page of their annual statements. The study examined whether this disclosure would make consumers more likely to switch between financial products.

The graph below shows the estimated effect of consumer financial disclosure on the share of consumers switching accounts in a four-week horizon.



The graph includes point estimates for the control and treatment groups, and a 95% confidence interval for the treatment effect.

> **Reminder:** When making your decisions, imagine you are the author of this study and you are trying to **persuade a policymaker that the research is promising.** In this decision you are more likely to earn the bonus payment if **your paired policymaker predicts the intervention's effect size is larger**.

Message choice

Number: Please choose a number message to communicate the treatment effect of the policy:

"In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to an increase in this share of ____ percentage points. That is, 2.6 + [Your Response]% of those receiving the financial disclosure treatment switched between financial products."

percentage points

Language: Please choose a word or phrase to communicate the treatment effect of the policy:

"In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to a(n) ____ increase in the share of consumers switching between financial products."

Note that the available options for the 'choice of format' question (below) will repeat the message choices made by the respondent in the previous prompt. Here we assume the respondent selected 4.6 for the number message and "modest" for the language message.

Choice of format

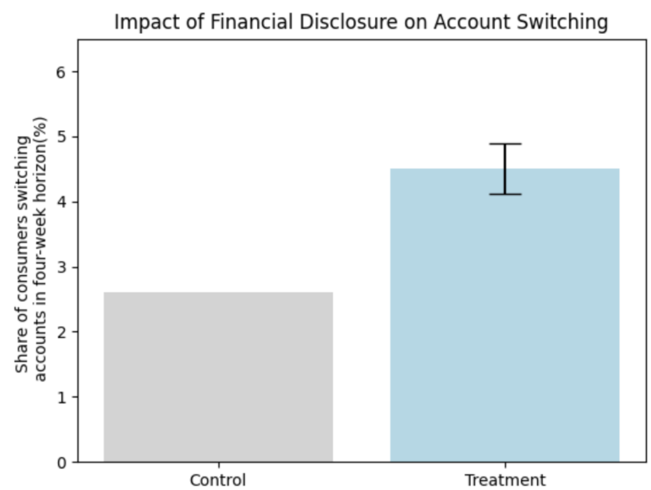
We will now ask you to choose your preferred message format: Number or Language. Which of the following would you prefer to send?

Number format: "In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to an increase in this share of 4.6 percentage points. That is, 7.2% of those receiving the financial disclosure treatment switched between financial products."	○
Language format: "In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to a(n) modest increase in the share of consumers switching between financial products."	○

Reminder: When making your decisions, imagine you are the author of this study and you are trying to **give the policymaker the best understanding of the data.** In this decision you are more likely to earn the bonus payment if **your paired policymaker correctly predicts the intervention's effect size**.

Study reminder: In a study reported in the Journal of Financial Economics (Adams et al. 2021), consumers with savings accounts at five UK depositaries were randomly disclosed information about alternative financial products. This information was provided on the front page of their annual statements. The study examined whether this disclosure would make consumers more likely to switch between financial products.

The graph below shows the estimated effect of consumer financial disclosure on the share of consumers switching accounts in a four-week horizon.



The graph includes point estimates for the control and treatment groups, and a 95% confidence interval for the treatment effect.

Did you know the results from this research study (Adams et al, 2021) prior to this survey?

- Yes, I already knew the main effects quantitatively
- Yes, I was familiar with the results but not the precise quantitative effects
- No, I was familiar with the paper but did not know the main results
- No, I was not familiar with the paper

C.4 Study 2: Policymakers

Instructions

Instructions

On the following pages, you will learn about two research studies in which a policy intervention was tested.

You will not immediately learn the results of the studies. Instead, for each study, you will be paired with an academic researcher who has already completed a version of this survey. Your paired researcher was told what the effect size of the policy intervention was, and chose how to communicate this information to you.

In particular, you will receive one of two types of messages from the researcher about the effect of the intervention:

1. Either **Number**: The message will use a number to describe the effect of the intervention (e.g., "the intervention led to a 3 percentage point increase")
2. Or **Language**: The message will use a word or phrase to describe the effect of the intervention (e.g., "the intervention led to a moderate increase")

Your task will be to predict the effect size of each policy intervention. You will make predictions both before and after you see the researcher's message. **The closer your predictions are to the true effect size, the more likely it is that you will receive the bonus payment.**

We will tell you the actual results of the research at the end of the survey.

Half of the researchers were told to try to accurately communicate the effect size of the intervention, and the other half were told to try to persuade you that the effect size of the intervention was large. All researchers chose one number and language message to send, and you will see the format they preferred most of the time. **Researchers were told that their messages could reflect what they learned about the intervention's effects, but they did not have to.**

Main decisions overview

Main Decisions

Now you're ready to make the main decisions in this survey!

Which topics do you think best reflect your areas of expertise? **Please choose two options below.** We will show you research studies that are relevant to these areas.

- [ ] Health insurance
- [ ] Economic mobility and housing
- [ ] Education policy
- [ ] Consumer finance and savings
- [ ] Criminal justice
- [ ] Labor market and aging

After this page, subjects are taken to the following main decision sections depending on which policy areas they indicated:

- Health insurance: Domurat et al (2021)
- Economic mobility and housing: Bergman et al (2024)
- Education policy: Burland et al (2023)
- Consumer finance and savings: Adams et al (2021)
- Criminal justice: Fishbane et al (2020)
- Labor market and aging: Liebman and Luttmer (2015)

These decision sections are shown in the following pages.

Main decisions, Adams et al (2021)

Question 1 of 2: Research study information

In a study reported in the *Journal of Financial Economics* (Adams et al. 2021), consumers with savings accounts at five UK depositaries were randomly disclosed information about alternative financial products. This information was provided on the front page of their annual statements. The study examined whether this disclosure would make consumers more likely to switch between financial products.

Before we share the message from your paired researcher, we'd like to ask you for your **best guess of the treatment effect of the policy intervention?**

In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to an increase in this share of ____ percentage points. This means that 2.6 + [Your Response]% of those receiving the financial disclosure treatment switched between financial products.

Your researcher's message

Next, you will see the message from your paired researcher and have the opportunity to revise your prediction.

Your Researcher's message:
In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to a fairly large increase in this share.

Study reminder: In a study reported in the *Journal of Financial Economics* (Adams et al. 2021), consumers with savings accounts at five UK depositaries were randomly disclosed information about alternative financial products. This information was provided on the front page of their annual statements. The study examined whether this disclosure would make consumers more likely to switch between financial products.

Now, what is your best guess of the treatment effect of the policy intervention?

In the control group, 2.6% of consumers switched between financial products in a four-week horizon. The financial disclosure treatment led to an increase in this share of ____ percentage points. This means that 2.6 + [Your Response]% of those receiving the financial disclosure treatment switched between financial products.

Now we'll ask a few more questions about this research study:

How likely would you be to share information about this study with your colleagues?

- Very unlikely
- Somewhat unlikely
- Neither likely nor unlikely
- Somewhat likely
- Very likely

Did you know the results from this research study (Adams et al, 2021) prior to this survey?

- Yes
- No

Would you like to access an infographic designed for policymakers that provides more details on this study? If so, please click yes and we will share a link to a downloadable infographic at the end of the survey. (Note that clicking yes will slightly increase the survey length.)

- Yes, I would like an infographic about this study
- No, I would not like an infographic about this study

C.5 Infographics for Policymakers

This is an example of the infographics offered to policymakers in the survey "Study 2: Receivers".