# Lying Aversion and Vague Communication: An Experimental Study

Keh-Kuan Sun*
Washington University in St. Louis

Guangying Chen†
Washington University in St. Louis

November 23, 2020

The latest version of the draft is available here.
(Preliminary draft; please do not circulate.)

**Abstract**

We study the interplay between the strategic and the behavioral aspects of potentially vague communication through an online experiment governed by a novel theoretical analysis. A sender may benefit from misleading the receiver's belief about the state of the world, but at the cost of losing a reputation for honesty, in addition to internal costs of dishonesty. Thus, a rational agent must balance the degree of truthfulness and vagueness of the message. We find that the use of vague messages endogenously arises from one's aversion toward lying. The data shows that subjects strategically exploit vagueness so as to be consistent with the truth, yet at the same time leveraging the imprecision to their own benefit. The results support an important conjecture that the vagueness of a message affects the agent's utility through her reputation, not her internal utility for being honest.

Keywords: Lying; Vagueness; Communication; Experiments; Behavioral Economics
JEL codes: C91, D90

*PhD Candidate, Department of Economics. Email: sun.k@wustl.edu
†Marketing PhD Student, Olin Business School. Email: guangyingchen@wustl.edu

# 1 Introduction

Interpretation of a message is an essential part of communication. The sender, who chooses the content of a message, often plays a more active role in the mutual process of forming an interpretation. In many cases, the sender has the liberty to choose a vague message and benefit from misleading the receiver's belief about the state of the world. However, the sender is equally equipped to blatantly lie to mislead the receiver in the same or even more effective manner. A natural question that follows is why and when the sender should use a vague message over a blatant lie. This issue is crucial for a unified understanding of misleading behavior in many applications, including public good provision problem (Serra-Garcia, Damme, and Potters 2011), sender-receiver disclosure game (Hagenbach and Perez-Richet 2018), and persuasion game (Deversi, Ispano, and Schwardmann 2020).

The answer to the question requires a careful examination of both the strategic and behavioral aspects of vague communication. First, a message not only affects the receiver's belief about the state of the world, but also the belief about how honest the sender is. When the sender cares about her reputation from being interpreted as honest, this external concern should impact her choice of strategies. Second, as witnessed by the recent development in the literature of lying in economics(Serra-Garcia 2018), people exhibit a non-trivial degree of aversion toward lying. If the sender intrinsically prefers to be honest, then this internal cost of lying should inhibit the use of a blatant lie and encourage the use of a vague yet relatively truthful message. The nontrivial cost of lying suggests that the communication is no longer a cheap-talk game and the misleading messages may bear strategic significance as a credible signal to the receiver about the behavioral types of the sender. On the other hand, the strategic incentive may influence people's behavior at the same time. Thus, a rational agent must balance the degree of truthfulness and vagueness of the message.

We explore the extent to which people's attitude toward lying and misleading behavior affects their strategic use of vague messages and vice versa in a simple experimental setting. Our main results provide empirical evidence for the two-way interaction. We find that the use of vague messages endogenously arises from one's aversion toward lying. The majority of the subjects used vague messages when allowed. The data shows that subjects strategically exploit vagueness so as to be consistent with the truth, yet at the same time leveraging the imprecision to their own benefit. We observe greater monetary payoff and less lying behavior when vague messages are used, suggesting that subjects send a vague message to remain truthful and increase their monetary payoff. The results support an important conjecture that the vagueness of a message affects the agent's utility through her reputation, not her internal utility for being honest. The result also partially supports the claim that the anonymity of agents would entail more lying behavior and higher monetary payoff.

To study vague communication with lying aversion, we compare treatments from an online experiment in which subjects face a variant of Fischbacher-Föllmi-Heusi type of reporting task. In this experiment, subjects privately observe an integer randomly drawn from a uniform distribution from 1 to 10. The subjects are asked to report the number to the experimenter, and their monetary payment depends solely on the report. The basic idea of the experiment is that the discrepancy between the true observation and the report should capture one's lying aversion for otherwise everyone should report 10.

We analyze the experimental data through the predictions of a model of costly lying. In this model, an agent may report a vague, set-valued message, or a precise, single-valued message to an audience after privately observing the state of the world. The agent's utility depends on the monetary payoff, which is determined solely by the reported message, and her preferences for truth-telling. The model assumes two separable motivations for honesty: the internal motivation for being honest and the reputation/external concern for being seen as honest.

The experiment varies the availability of vague messages and the anonymity of subjects to test the predictions and, in particular, to separate the internal and external costs of dishonesty. There are two types of experiment sessions that represent the variation in the anonymity of agents. In the anonymous session, the responses are recorded under screen names so that the experimenter cannot map a subject's identity to her response. In the non-anonymous session, on the other hand, the experimenter knows each subject's response. Within each session, a subject confronts two reporting tasks. The observation process is identical and independent in the two tasks. When reporting, however, the available set of messages differs between the two tasks. In the task with restricted communication, the set is restricted to the single-valued messages only. In the task with unrestricted communication, the subjects are allowed to use both single-valued and set-valued messages. The combination of these between-subject and within-subject variations yields four treatments in our experiment.

If the vagueness of a message does not interfere with one's preferences for truth-telling, the variation in the anonymity should not impact the reporting behavior. We find that the pattern of the reported messages, however, differs significantly between the two types of sessions: our subjects used more vague messages in non-anonymous sessions compared to the anonymous counterpart. Furthermore, we can isolate the effect on the lying behavior by comparing the propensity for lying across the two tasks within the anonymous session. We observe that subjects lie less when vague messages are allowed. As the anonymous environment should suppress the subjects' concern for their reputation, the decrease in lying behavior suggests that the difference arises from the relation between the use of a vague message and one's internal concern for honesty. These results suggest that the observed aversion for monetary-payoff-maximization in similar experiments conducted in the literature may be decomposed into different components, including but not limited to the signaling function of the message.

We find a result analogous to the "warm glow" giving behavior in the vague communication. That is, as long as the message can remain even remotely truthful, most agents exhibit no hesitation in increasing their monetary payoff. On the other hand, there exists a group of truth-tellers whose reports are not only truthful but also precise, thus relinquishing possible monetary gain opportunities. The non-trivial size of this group suggests a possibility of another motivation for truth-telling in addition to the internal and external concern for honesty. Potential candidates include a concern for the consequence of their choices or concern for good intentions.

Conceptually, we interpret the lying behavior by Becker (1968) in the broadest context. We view the lying behavior as an optimization process in which a rational agent chooses the optimal amount of dishonest act that balances its marginal monetary benefit and the non-monetary costs that depend on the probability of being found out as a liar. We employ the notion of the psychological game by Geanakoplos, Pearce, and Stacchetti (1989) to embed the audience's belief into the agent's utility.

Structurally, we study the lying behavior as a subset of signaling games. Since Crawford and Sobel (1982), a stream of research has tested the assumption of costless lying by which individuals would lie whenever there was a material incentive to do so. The empirical evidence continues to make a case against the notion (e.g. Dickhaut, McCabe, and Mukherji (1995), Blume et al. (1998), Gneezy (2005), Sánchez-Pagés and Vorsatz (2007) and many more). They argue that individuals certainly express aversion toward lying and take significantly lower payoffs than the theories predicted.

In terms of the experiment design, our paper belongs to the rolling-a-die paradigm first introduced by Fischbacher and Föllmi-Heusi (2013) (FFH). The idea, with the assumptions of a non-atomic game, simplifies the signaling game into an individual decision-making problem. In these experiments, a subject draws a random number (e.g. rolling a die), reports the observed number, and receives a monetary payoff based on the report. As the monetary payoff is independent of the drawn number, the subject should report a non-maximal number only if aversion to lying is present, for otherwise this becomes a simple case of a cheap-talk game. A consensus of the experiments is that people do not lie as much as they could, potentially because lying is costly. Abeler, Nosenzo, and Raymond (2019) summarizes 90 such experiments into common findings that describe the average reporting behavior. They find the behavior is indeed bounded away from the maximal report but also departs from a complete truth-telling scenario. Their paper concludes that a preference for being seen as honest and a preference for being honest are the main motivations for truth-telling.

However, little is known about how the message space in communication plays a role in one's reporting decision in relation to lying cost. For instance, as Abeler, Nosenzo, and Raymond note in their conclusion, the FFH paradigm has focused on subjects reporting a single number and excludes lies by omission or vagueness. We generalize the FFH model by allowing subjects to transmit a set-valued message to understand the effect of vagueness in communication.

The closest papers to our work are Gneezy, Kajackaite, and Sobel (2018) and Khalmetski and Sliwka (2019) in which an agent cares not only about her monetary payoff but also whether she lies and how others interpret her report. We inherit their assumptions of these two motivations for truth-telling, the internal guilt and the external concern for reputation, and extend by allowing vague communication.

Our paper provides a bridge between the literature of lying behavior to a broader set of studies that involves vague communication. For instance, Serra-Garcia, Damme, and Potters uses the assumption that a set-valued vague message that contains the true state of the world would incur much less lying cost compared to a single-valued precise outright lie in analyzing a public good provision game between two agents with asymmetric information. Our experiment tests this assumption and provides further insight into the lying cost with respect to the reputational concern as well.

We find a similarity between our design and that of Deversi, Ispano, and Schwardmann (2020) in which subjects are allowed to send an interval as a message. However, their research focuses on the strategic use of vague messages without considering the behavioral aspect of lying aversion. We also provide subjects with more flexibility in their choice of messages by allowing any subset of the state space.

The remainder of the paper is organized as follows. Section 2 defines terminologies and presents the model setting. Section 3 provides theoretical analysis, and Section 4 lists the experiment hypotheses based on the theoretical predictions. Section 5 describes our experiment design, Section 6 summarizes the experiment outcome, and Section 7 concludes with remarks. We list all the proofs and additional experiment details in the Appendix.

# 2 Model

## 2.1 A model of lying aversion with vague communication

We study lying aversion with vague communication by considering a variant of Fischbacher and Föllmi-Heusi cheating game with a population of agents and one audience. An agent privately observes the state of the world $i \in \Omega$ where $\Omega = \{1, 2, \ldots, N\}$ is finite. We assume $i$ is drawn i.i.d. from a uniform distribution over $\Omega$ across agents. Each agent has a private type $t$ that represents her intrinsic aversion toward lying. We also assume $t$ is i.i.d. across agents. Its CDF $F(t)$ is strictly increasing, continuous, and has the support $[0, T]$. Denote an agent who observes the state $i$ and has the intrinsic aversion type $t$ as a type $(i, t)$ agent.

An agent reports a message $J$ to the audience after observing the true state of the world $i$. This reporting takes place only once and there is no repeated interactions. Similarly to Gneezy, Kajackaite, and Sobel (2018)(GKS) and Khalmetski and Sliwka (2019)(KS), we assume that an agent's utility consists of three components: monetary payoff, the internal concern for being honest, and the external concern for being seen as honest. Formally, we write a type $(i, t)$ agent's utility for reporting a message $J$ as:

$$U(i, J, t) = \underbrace{\bar{\pi}(J)}_{\text{monetary payoff}} - \underbrace{\mathbb{1}(i \notin J)[t + c(\pi(\{i\}), \bar{\pi}(J))]}_{\text{internal guilt}} + \underbrace{\gamma\rho(J)}_{\text{external reputation}}. \tag{1}$$

A message $J$ is a nonempty[1] subset of the state space: $J \in M^\Omega \equiv 2^\Omega \setminus \emptyset$. An important

[1] An empty set is often interpreted as silence and plays an interesting role in the literature of vague communication. We abstract away from the silent message to focus on the direct comparison of lying aversion between

distinction of our model from GKS and KS is that the message space admits set-valued messages instead of being isomorphic to the state space. This generalization allows messages to be categorized in multiple ways. We define the relevant terminologies below.

**Definition 1.** A message $J$ is truthful if $i \in J$. A message is a lie if it is not truthful.

**Definition 2.** A message is called precise if it is a singleton set, and vague otherwise.

Let $M_P^\Omega \equiv \{\{1\},\{2\},\ldots,\{N\}\}$ and $M_V^\Omega \equiv M^\Omega \backslash M_P^\Omega$ denote the set of precise messages and that of vague messages, respectively. For example, if $\Omega = \{1,2,3\}$, the set of possible precise messages is $\{\{1\},\{2\},\{3\}\}$, and the set of possible vague messages is $\{\{1,2\},\{2,3\},\{1,3\},\{1,2,3\}\}$.

The monetary payoff of an agent maps her message to utility: $\pi : M^\Omega \to \mathbb{R}$. For simplicity, we assume $\pi(J)$ is a uniform draw over $J$ and that the agent is a risk-neutral, expected-utility maximizer. Denote $\bar{\pi}(J) = E[\pi(J)]$. Note that when a message is precise, i.e. $J = \{x\}$ with $x \in \Omega$, her monetary payoff $\pi(J)$ is simply $x$.

In addition to her monetary payoff, the agent also has two different motivations for honesty. First, she has the internal motivation for being honest. When her report $J$ is not truthful, that is, $i \notin J$, the dishonesty incurs the internal cost of $t + c(\pi(\{i\}), \bar{\pi}(J))$. The agent's private type $t$ captures her sensitivity toward the intrinsic (fixed) cost of lying. The function $c(\pi(\{i\}), \bar{\pi}(J)) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ represents the variable cost of lying depending on the size of the lie. The size of the lie is measured as the ex-ante difference in the monetary payoff of a report $J$ and that of the true and precise report $\{i\}$. We assume that i) $c(\cdot) \geq 0$; ii) $c(\pi(\{i\}), \pi(\{i\})) = 0$; iii) $c(\pi(\{i\}), \bar{\pi}(J))$ is weakly increasing in $|\pi(\{i\}) - \bar{\pi}(J)|$; iv) $c(\pi(\{i\}), \pi(\{i\}) + 1) < 1$[2]; and v) $c(\pi(\{i\}), \bar{\pi}(J)) + c(\bar{\pi}(J), \bar{\pi}(K)) \geq c(\pi(\{i\}), \bar{\pi}(K))$.

She also has the external motivation for being seen as honest; that is, her utility depends on the audience's belief on how honest the agent is. We assume the audience is a rational Bayesian who forms his posterior belief based on her report $J$. This belief, in turn, depends on the agent's mixed strategy in equilibrium.

**Definition 3.** A type $(i, t)$ agent's mixed strategy is a mapping $\sigma : \Omega \times [0, T] \to [0,1]^{2^N - 1}$ such that

$$\sigma(i,t) = \left(\sigma_{it}^{\{1\}}, \sigma_{it}^{\{2\}}, \ldots, \sigma_{it}^{\{1,2,\ldots,N\}}\right) \tag{2}$$

the cases of vague and precise language.

[2]This assumption excludes the trivial case where the variable cost of lying is too high that no agent choose to lie

where $\sigma_{it}^{J}$ is the probability that the intrinsic aversion type $t$ agent with the true observation $i$ assigns to the report $J \in M_{\Omega}$.

In equilibrium, the audience's posterior belief about whether an agent's report $J$ is truthful is computed using Bayes rule:

$$\rho(J) = \frac{P(\text{agent is honest} \wedge \text{agent reports } J)}{P(\text{agent reports } J)} = \frac{\sum_{k \in J}(\int_0^T \sigma_{kt}^{J} df(t))}{\sum_{k=0}^{N}(\int_0^T \sigma_{kt}^{J} df(t))} \tag{3}$$

We normalize the posterior belief $\rho$ in terms of the agent's utility by a parameter $\gamma$. The parameter measures the agent's sensitivity toward her reputation by reporting a message $J$. We further assume that i) $\gamma$ is homogeneous across agents and is a common knowledge; and ii) $N + \gamma < T$. The latter condition ensures that $F(N + \gamma) < 1$, or there always exists a positive mass of agents who will be truth-telling for any observation $i$.

Because the agent's payoff depends on the belief of the audience, we adopt the notion of sequential equilibria in the induced psychological game in the sense of Battigalli and Dufwenberg (2009) and Geanakoplos, Pearce, and Stacchetti (1989).

An equilibrium is a set of mixed strategies and beliefs satisfying the following conditions:

$$\forall (i, J, t) : \sigma_{it}^{J} > 0 \text{ only if } J \in \operatorname*{argmax}_{J'} U(i, J', t), \tag{4}$$

$$\forall (i, t) : \sum_{J \in M^{\Omega}} \sigma_{it}^{J} = 1, \tag{5}$$

$$\forall J : \rho(J) = \frac{\sum_{k \in J}(\int_0^T \sigma_{tk}^{J} df(t))}{\sum_{k=0}^{N}(\int_0^T \sigma_{tk}^{J} df(t))}. \tag{6}$$

As noted by GKS, the existence of an equilibrium follows Schmeidler (1973).

## 2.2 Communication environment and anonymity of agents

The baseline model assumes no restriction on the message space $M^{\Omega}$. By restricting the message space to $M_{P}^{\Omega}$, on the other hand, we obtain the model with restricted communication. We refer to the baseline model as the model with unrestricted communication. With an abuse of notation, let us denote a message as $j$ in the models with restricted communication.

Furthermore, for a more comprehensive understanding of the relation between the use of vague messages and the lying costs, it is helpful to isolate its effect on the internal guilt from that on the external reputation. We can achieve this by employing an anonymous environment where the audience cannot identify an agent with a message. This implies that the agent's report does not alter the audience's belief, hence the reputation remains constant independent of her reporting choice in the anonymous environment. On the other hand, we call an environment non-anonymous when the audience can associate a message with its sender. Formally, write an agent's utility in the anonymous environment as:

$$U_A(i,J,t) = \bar{\pi}(J) - \mathbb{1}(i \notin J)\left[t + c(\pi(\{i\}), \bar{\pi}(J))\right]. \tag{7}$$

We thus have four different environments by varying the restriction on the message space and the anonymity of the agents: non-anonymous with restricted(precise only) communication(NA-R), non-anonymous with unrestricted(potentially vague) communication(NA-UR), anonymous with restricted communication(A-R), and anonymous with unrestricted communication(A-UR). Note that the NA-R environment corresponds to the standard FFH model where an agent can only send precise messages and the audience can identify a message's sender.

# 3 Analysis

Let us first begin with restating the previously known results about the NA-R environment. The main intuition is that some agents lie when they see a small number, while some others always report the truthful message. Thus, when a reported number is small, it is safe to believe that the agent is being honest. The following equilibrium results establish a baseline for comparison with the other environments.

**Lemma 1** (KS LM 1). *In any equilibrium in the NA-R environment, there exists a strict positive probability that agents lie.*

**Lemma 2** (KS PR 4; GKS PR 2). *In any equilibrium in the NA-R environment, no agent underreports.*

**Lemma 3** (KS THM 1; GKS PR 5). *In any equilibrium in the NA-R environment,* *i. there exists a threshold* $1 < l^* < N$ *such that*

$$\forall j \geq l^* \; \exists i \neq j, t \in \mathbb{R} \;\; s.t. \;\; \sigma_{it}^{j} > 0 \;\; and$$

$$\forall j < l^*, i \in \Omega, t \in \mathbb{R} \; \sigma_{it}^{j} = 0;$$

*ii. all agents who observe a value above the threshold report their observed value truthfully.*

Based on the three results, we bring a simple comparison of the behavior of an agent in the NA-R environment to that in the A-R environment. A thought experiment of choosing an agent and comparing her behavior in the two environments easily leads to a conjecture that the absence of the reputation concern should only facilitate more lies. The following lemma shows this is indeed the case in our model.

**Lemma 4.** *If a type* $(i, t)$ *agent lies in an equilibrium in the NA-R environment, then she lies in the A-R environment.*

Using the above lemmas, we now generalize the argument to compare the probability of lying of an arbitrarily chosen agent in the two environments.

**Proposition 1** (NA-R/A-R)**.** *The set of types* $(i, t)$ *of agents who lie in an equilibrium in the NA-R environment is a subset of the set of liar types in any equilibrium in the A-R environment. The expected monetary earning is greater on average in the A-R environment.*

The first step toward understanding vague communication is to clarify its effect on one's internal cost of lying. To achieve this, let us now relax the restriction on the message space and consider the case of the A-UR environment in comparison with the A-R environment. One may immediately notice that the richer message space induces multiplicity in the agent's reporting strategy. The absence of the reputation concern simplifies the problem into a simple comparison between the solutions of a truth-telling-constrained and an unconstrained optimization. That is, for each type $(i, t)$ agent, there exists both $\max_{J:i \in J} U(i, J, t)$ and $\max_{J \in M^{\Omega}} U(i, J, t)$, and the agent reports truthfully only if the two maximums coincides.

We can further add more structure to the constrained optimization. Intuitively, this maximum has the form of a union of the true observation $i$ and some ending sequence $x, x+1, \ldots, N$.

We can solve the optimization problem[3] to find the threshold $x^* = \lceil (N+2) - \sqrt{2N-2i+3} \rceil$. Let us define the optimal vague message for each true observation $i$ as $OVM_i \equiv \{i, x^*, x^*+1, \ldots, N\}$. Then $OVM_i \in \text{argmax}_{J:i\in J}\, U(i,J,t)$ and weakly dominates all truthful messages. This allows us to make the following observations in the A-UR environment.

**Proposition 2.** *In the A-UR environment, all truth-tellers uses the optimal vague messages.*

**Corollary.** *In the A-UR environment,*

i. *no agent's message contains a number less than the true observation.*

ii. *no precise message except $\{N\}$ is truthful.*

Furthermore, because we are now free of the reputation component, there is a clear mapping between an agent's type $(i,t)$ and the resulting behavior. That is, fixing the observation $i$, one's lying behavior is simply a monotone function of $t$ in the A-R environment: one reports $j > i$ when $(j-i) + c(i,j) > t$. Thus, there exists a threshold $t_i^*$ for each observation $i$ below which the agent reports a lie and above which the agent tells the truth. This threshold, however, becomes much higher in the A-UR environment because now the agent is equipped with a set of vague messages that allows her to remain truthful yet increase the expected payoff. This fact leads to the following observations.

**Observation.** *If a type-$t$ agent reports truthfully after observing the state $i$ in the A-R environment, then she also reports truthfully when observing $i$ in the A-UR environment.*

**Observation.** *For each observation $i < N$, there exists a positive mass of agents who lies in the A-R environment yet reports truthfully in the A-UR environment.*

The immediate corollary of the two observations above is the following proposition that states the mass of liars is greater when the communication is restricted in the anonymous environments.

**Proposition 3** (A-R/A-UR)**.** *The set of types (i,t) of agents who lie in an equilibrium in the A-UR environment is a subset of liar types in any equilibrium in the A-R environment. The expected monetary earning is greater on average when the communication is not restricted.*

[3] We can simplify the optimization problem by approximating with a continuous uniform distribution: $\text{argmax}_x \int_i^{i+1} \frac{u}{N-x+2}\,du + \int_x^N \frac{u}{N-x+2}\,du$

A particularly useful element in the analysis of the restricted communication environment is the assumption that there always exists a positive mass of truth-tellers for each observation $i$. While sufficiently reasonable in many contexts, this assumption aids us especially in eliminating the less interesting equilibria with off-path beliefs. Thus, we are able to characterize the equilibria with sharp predictions even in the case of the non-anonymous environment. This, however, is unfortunately not the case when we relax the message space to allow vague messages. We illustrate this point by the following example.

**Example 1.** Consider the case where $\gamma > \frac{N-1}{2}$. Then everyone reports $\{1,2,\dots,N\}$ and $\rho(j) = 1$ when $j = \{1,2,\dots,N\}$ and 0 otherwise is an equilibrium.

This example describes the scenario in which the agents care much about their reputation and are forced to stick to the (exogeneously) given norm of the full vagueness. While the specific example bears potential real-world examples[4], an important implication is that there can exist a plethora of equilibria depending on the combination of $\gamma$ and the off-path beliefs. Consider the following example.

**Example 2.** All agents reports a message in the form of an interval: $[i,N] = \{i, i+1, \dots, N-1, N\}$, and the audience assigns the posterior belief according to the Bayes rule to only those interval-messages, and zero to all other messages. We need the following conditions to hold in order to constitute such equilibrium:

- truth-tellers: $\frac{i+N}{2} + \gamma\rho(J) > U(i, MVT_i, t)$ for all $i$ and an interval message $J$
- liars: $\frac{k+N}{2} - \mathbb{1}(i \neq k)(t + \bar{\pi}(J)) + \gamma\rho(J) > U(i, J', t)$ for all non-interval message $J'$

The conditions hold trivially when $\gamma \to \infty$, and the signaling game becomes isomorphic to that of the NA-R, where each message $i$ is simply replaced with the interval $[i,N]$. The ending sequence carries no information to the audience, and the equilibrium follows that of the NA-R.

**Proposition 4.** *In any equilibrium in the NA-UR environment, there exist some types of agents who use a vague message with positive probability in their mixed strategy.*

[4]For instance, we may imagine a group of politicians all replying by the same vague message to a politically sensitive question because they know any message other than the full vagueness will be interpreted as a lie and cost their reputation. Many similar examples can arise in the situations where the agents value their reputation highly. We appreciate Brian Rogers for suggesting this interpretation of the equilibrium.

Proposition 4 affirms that the use of vague messages arise endogenously in the NA-UR environment.

Also, based on our findings in each environment, we make the following conjectures that the anonymous environment will have a greater mass of liars and a higher expected monetary earning in comparing NA-UR and A-UR and that allowing vague messages will reduce the mass of liars while increasing the expected monetary earning in comparing NA-R and NA-UR.

**Conjecture 5.** *i. The set of types $(i, t)$ of agents who lie in an equilibrium in the NA-UR environment is a subset of the set of liar types in any equilibrium in the A-UR environment. The expected monetary earning is greater on average in the A-UR environment.*

*ii. The set of types $(i, t)$ of agents who lie in an equilibrium in the NA-UR environment is a subset of the set of liar types in any equilibrium in the NA-R environment. The expected monetary earning is greater on average in the NA-UR environment.*

# 4 Hypotheses

lie ↑ earning ↑
→

lie ↓ earning ↑ ↓

| | Non-anonymous & Unobs. | Anonymous & Obs. |
|---|---|---|
| Precise | NA-R | A-R |
| Vague | NA-UR | A-UR |

**Hypothesis 1** (NA-R/A-R)**.** *Under the precise (restricted) communication,*

*i. more agents lie in the anonymous environment: $lie_{NA-R} \leq lie_{A-R}$;*

*ii. agents earn more monetary payoff on average in the anonymous environment: $earning_{NA-R} \leq earning_{A-R}$.*

Hypothesis 1 is a consequence of Proposition 1.

**Hypothesis 2.** *In the A-UR environment,* *i. all truth-tellers in A-UR uses MVT;*

*ii. no message contains a number less than the true observation;*

*iii. no precise message except $\{N\}$ is truthful.*

Hypothesis 2 is a consequence of Proposition 2 and its corollary.

**Hypothesis 3** (A-R/A-UR)**.** *In the anonymous environment,*

*i. more agents lie when the communication is restricted (precise): $lie_{A-R} \geq lie_{A-UR}$;*

*ii. an agent who is truthful in A-R is also truthful in A-UR conditional on the same observation;*

*iii. some agents who lie in A-R reports truthfully in A-UR conditional on the same observation;*

*iv. agents earn more monetary payoff on average when the communication is not restricted (vague): $earning_{A-R} \leq earning_{A-UR}$;*

Hypothesis 3 is a consequence of Proposition 3 and the observations about the behavior in the A-UR environment.

**Hypothesis 4.** *In both the NA-UR and the A-UR environment, people use vague messages.*

Hypothesis 4 is a consequence of Propositions 2 and 4.

**Hypothesis 5.**

*i. Under the vague (unrestricted) communication, agents earn more monetary payoff on average in the anonymous environment: $earning_{NA-UR} \leq earning_{A-UR}$;*

*ii. In the non-anonymous environment, agents earn more monetary payoff on average when the communication is not restricted (vague): $earning_{NA-R} \leq earning_{NA-UR}$;*

Hypothesis 5 is a restatement of Conjecture 5.

## 5 Experiment Design

We vary the experiment treatments along two dimensions: 1) we consider precise or vague messages, and 2) we vary the ability of the experimenter to identify the responses to an individual subject. There are two types of experiment sessions that represent the variation in the anonymity of agents. Within each session, a subject confronts two stages of reporting tasks that represent the availability of vague messages. In each stage, subjects are incentivized to observe a random number and later report the number to the experimenter. The amount they earn depends on the number or numbers they report.

In the anonymous session, the responses are recorded under screen names so that the experimenter cannot map a subject's identity to her response. As the subjects are instructed that the experiment intentionally prohibits such mapping, this treatment should establish the effect of suppressing their reputational concern and emulate the environment where $\gamma \to 0$. In the non-anonymous session, on the other hand, the experimenter knows each subject's response.

A 'stage' is our basic unit of observation. In each stage, subjects first observe the realization of a random integer uniformly distributed between 1 to 10 and later asked to report the number. In the anonymous treatment, the number is generated within the experiment software so that the experimenter knows both the true observation and the report of each subject. In the non-anonymous treatment, subjects use an external website to generate the random number[5] so that the experimenter cannot know the true observation. This design choice drives from the idea that an non-anonymous environment with the true observation known to the experimenter is overly artificial; as result, subjects' discomfort with lying behavior may be exaggerated in such environment.

The observation process is identical and independent in the two tasks. When reporting, however, the available set of messages differs between the two tasks. In the precise stage, the set is restricted to the single-valued messages only. In the vague stage, the subjects are allowed to use both single-valued and set-valued messages. The combination of these between-subject and within-subject variations yields four treatments in our experiment: NA-R[6], NA-UR, A-R, and A-UR. The order of precise and vague treatments within a session is randomized to measure the order effect.

[5]We provided a link to a Google search result for the phrase "random number between 1 to 10."

[6]The NA-R treatment is equivalent to the original experiment of Fischbacher and Föllmi-Heusi (2013).

We recruited 69 student subjects from the subject pool of the Missouri Social Science Experimental Laboratory(MISSEL) at Washington University in St. Louis. The recruitment and experiment management processes are done by the ORSEE system. Students are invited to a virtual Zoom meeting where the experimenter reads the instructions and provides a link to the main experiment web page. The main experiment is implemented using the Qualtrics online survey platform.

Each session lasted for approximately 30 minutes. In all cases, subjects received a $2 show-up fee, thus the total amount they could earn varies from $2 to $12. See Appendix for our experimental instructions and procedural details.

# 6 Results

| | Average Report | | Lie | | | |
|---|---|---|---|---|---|---|
| | Precise | Vague | Precise | Vague | Size | N |
| Identifiable | 6.556 | 8.208 | | | 3.72 | 36 |
| Anonymous | 6.788 | 8.285 | 12 (36.36%) | 6 (18.18%) | 3.06 | 33 |

Table 1: Data summary

We present the basic summary statistics in Table 1. There were 4 anonymous sessions and 7 non-anonymous sessions, with the average size of 8.3 and 5.1 participants in each session, respectively[7]. For the treatments with unrestricted communication, we take the mean of the numbers included in each subject's report when calculating the average report. The average size of vague messages is computed by taking the average number of the numbers included in the messages, conditional on being vague. Overall, we find that allowing vague messages increase the average report. Likewise, anonymity of reports also increase the average report, though the difference is less significant.

**Result 1.** *Under the precise (restricted) communication,*

[7]Participants voluntarily chose time blocks, and we provided the anonymous version of the software in larger sessions.

1. *(not yet tested)*[8]*;*

2. *participants reported higher in the anonymous environment (6.556 in NA-R and 6.788 in A-R).*

While we are not able to compute the likelihood of lying in the non-anonymous treatments yet, we can still test whether the average report is higher in the anonymous treatments as predicted. We observe the average report is slightly higher in the anonymous treatment as predicted (6.788 as opposed to 6.556), the difference is not significant in the one-sided t-test with a p-value of 0.3876.

**Result 2.** *In the A-UR environment,*

1. *44.4% (12 of 27) of truth-tellers reported the optimal vague messages (including the honest 10), 18.5% (5 of 27) used a pair of the true observation and 10, and 18.5% (5 of 27) used a precise message below* $\{10\}$*;*

2. *only 1 out of 33 participant included a number less than the true observation in the report;*

3. *all precise messages below* $\{10\}$ *were truthful.*

This result partially supports Hypothesis 2. Figure 1 summarizes the types of messages used in A-UR treatments. Overall, 45.5% (15 of 33) of participants used a precise message, and 54.5% (18 of 33) used a vague message. Among the 15 precise messages, 40% (6 of 15) were lies and 60% (9 of 15) were not lies. All the liars reported the maximum of 10. Among the 9 precise truth-tellers, 4 people observed 10 and reported so.

The model predicts that all truth-tellers to seek payoff maximization conditional on including the true observation in the report. If we combine both the optimal vague messages (including the honest 10 as the optimal message) and the pair type messages into a broader set of payoff-increasing truthful messages, then we have a majority (62.9%) of truth-tellers maximizing their monetary payoff conditional on being honest. Yet there is still a noticeable size of precise truth-tellers who reported below 10, which contradicts our hypothesis. This may suggest a possibility of other motivation for truth-telling not addressed in our model.

[8] We plan to run a likelihood estimation with additional assumptions on the functional form of the utility and the equilibrium strategies, but the result is not yet available. We are in the process of running more experimental sessions to increase the sample size.

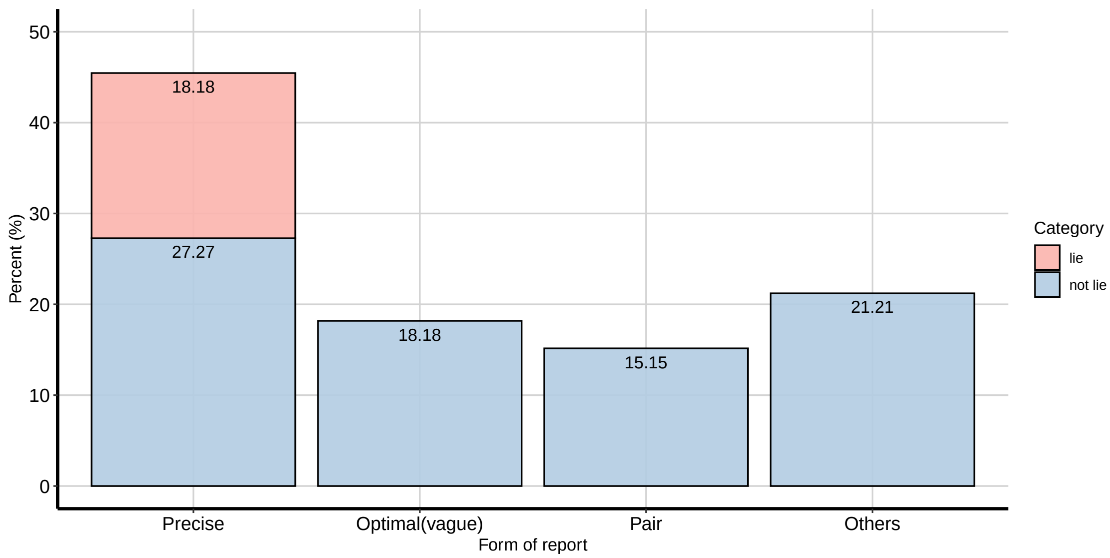


Figure 1: Message types used in A-UR treatments

**Result 3.** *In the anonymous environment,*

1. *more participants lied when the communication is restricted (36.36% in A-R and 18.18% in A-UR);*
2. *all participants who were truthful in A-R remained truthful in A-UR;*
3. *6 out of 12 liars in A-R reported truthfully in A-UR;*
4. *participants reported higher on average when the communication is not restricted (6.788 in A-R and 8.285 in A-UR).*

The result supports Hypothesis 3. The paired t-test result of the proportions of liars in A-R and A-UR is significant with a p-value of 0.006, and the one-sided t-test for the average report also returns a p-value of 0.005. An interesting point is that 5 of the 6 participants who switches from lying to truth-telling reported 10 in A-R, yet they chose a vague message whose expected earning is less than 10. This suggests that the use of vague message exacerbates the internal cost and supports the assumption that the use of vague yet truthful messages would only cost in the reputation.

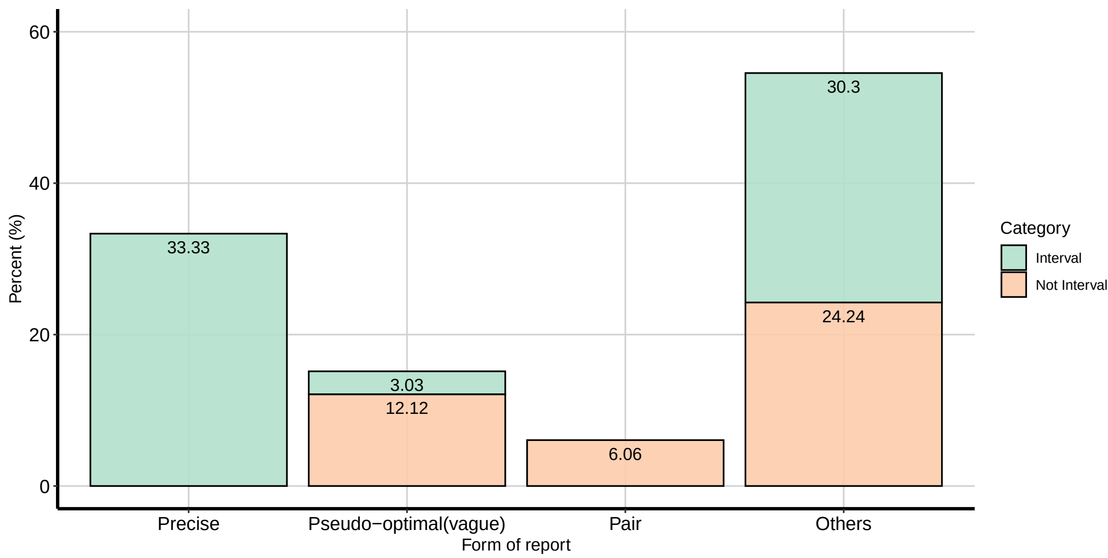


Figure 2: Message types used in NA-UR treatments

**Result 4.** *In the NA-UR environment, 25 out of 36 participants (69.4%) used a vague message. In the A-UR environment, 18 out of 33 participants (54.5%) used a vague message.*

The result supports Hypothesis 4. The majority of the participants used a vague message. It is interesting to see more participants used a vague message in the non-anonymous environment. This difference is possibly due to the blatant liars in the anonymous case who lie by reporting 10 without a reputational concern.

Also, it is worth noting that the size of vague messages is larger in the non-anonymous treatments (3.72 numbers reported on average conditional on being set-valued) compared to the anonymous environment (3.06 numbers reported on average conditional on being set-valued). The difference possibly arises from the impact of vague messages on the external cost of lying. Figure 2 summarizes the types of messages used in NA-UR treatments. Note that, compared to the anonymous treatment as displayed in Figure 1, participants used less of precise, pseudo-optimal(vague)[9], and pair types of messages. On the other hand, we see the vast majority of the messages are now under the 'other' category. Among the 'other' messages, half of the messages take the form of a continuous interval, which was not the case in the

[9]While we do not know the true observation of each participant in the non-anonymous treatments, we categorized the optimal-looking messages using the minimum of the reported numbers in the message as a pseudo-true-observation.

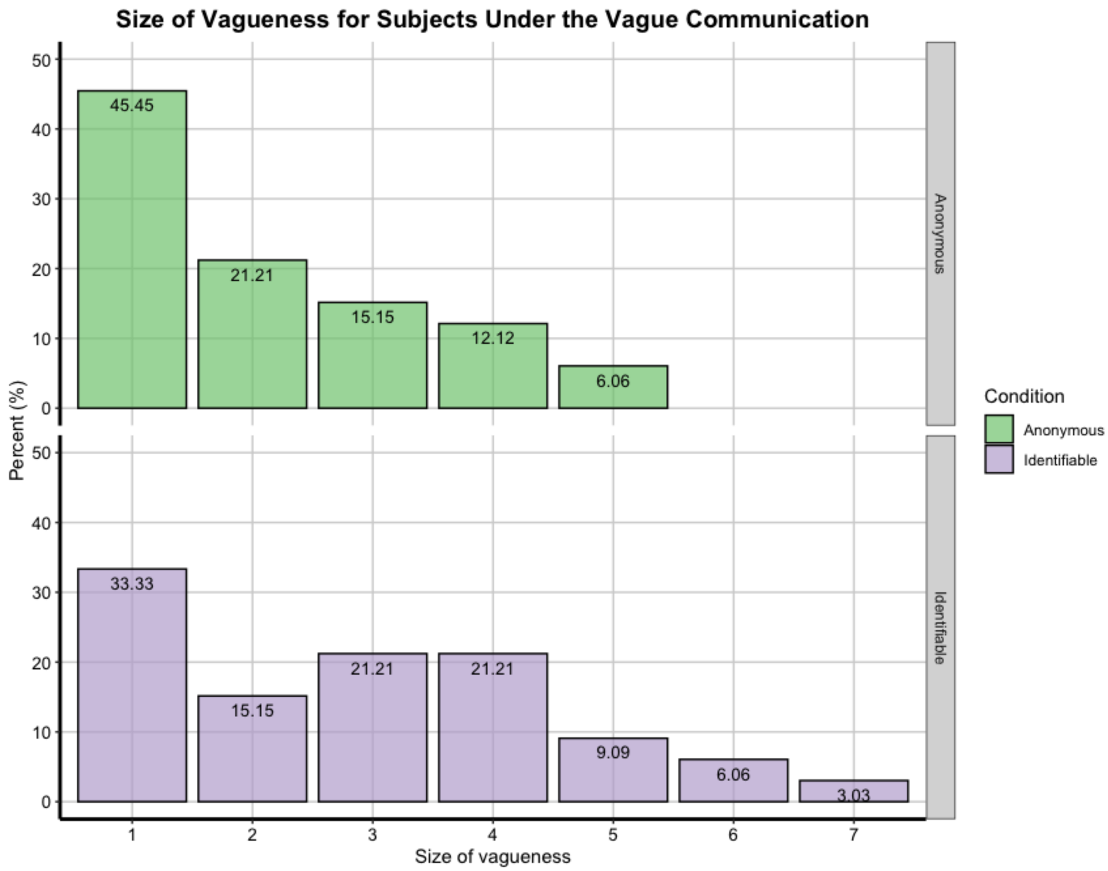


Figure 3: Size of vague messages in A-UR and NA-UR treatments

anonymous treatments.

The distribution of the size of messages also suggests that there is a significant difference between the reporting pattern in the A-UR and the NA-UR treatments. Figure 3 shows that the distribution of messages in the non-anonymous session exhibits a longer tail. It also shows that, while it exhibits a clear decreasing pattern in the anonymous treatment, there is a heap of moderate-sized messages around 3 and 4. Together with the frequent uses of interval messages in the non-anonymous treatment, we interpret this pattern as indirect evidence of the impact of vagueness in the external cost of lying.

**Result 5.**

1. *Under the vague (unrestricted) communication, participants reported lower on average in the anonymous environment (8.208 in NA-UR and 8.285 in A-UR).*

2. *In the non-anonymous environment, participants reported higher on average when the communication is not restricted (6.556 in NA-R and 8.208 in NA-UR).*

The result supports Hypothesis 5. We observe the average report is slightly greater in the anonymous environment. However, the one-sided t-test is inconclusive with a p-value of 0.4154. The average report is greater in the NA-UR than in the NA-R, and the one-sided t-test result is significant with a p-value of 0.0098.

# 7 Concluding Remarks

The empirical findings support the prediction that the use of vague messages endogenously arises from the lying aversion. The prevalence of monetary-payoff maximizing messages in the anonymous sessions support the key conjecture that the vagueness of a message does not affect the internal cost of lying. The difference in the reporting pattern across the NA-UR and the A-UR sessions suggests that the vagueness still plays an important role in the behavioral aspect.

Overall, participants reported much higher on average when vague communication is allowed. This finding sheds new light on the understanding of lying aversion, suggesting that the restricted message space could be a source of the observed aversion for monetary-payoff-maximization in previous experiments. The average reporting in the treatments with restricted communication in our experiment is comparable to that in the previous experiments, claiming approximately two-thirds of the maximum amount available. However, as we allow the use of vague messages, the average report increases to 8.2, reducing the size of the aversion from one-third of the maximum monetary payoff amount to less than one fifth. What more important is, based on our findings from the anonymous treatments, that people lie less as they increase their expected monetary payoff. The decomposition of the aversion toward monetary-payoff-maximization in our experimental design calls for a new perspective of understanding misleading behavior by showing that the observed aversion in many individuals is independent of the consequence of their choice of messages. In other words, as they have the new option of using the vague message to convey their honesty to the audience, they no longer need to yield the monetary payoff as much to provide a credible signal. This finding is analogous to the "warm glow" giving where workers care only whether they exert effort to help others or not, but not the consequential magnitude of the help provided. As long as the message can remain even remotely truthful, most agents exhibit no hesitation in increasing their monetary payoff.

However, the existence of precise truth-tellers in the anonymous environment, though small in the size, suggests that the warm glow argument may not be the whole picture. A possible explanation is the existence of another motivation for truth-telling such as a concern for the self-image of good intention. That is, although the participants might have understood that there is no external observer to judge their behavior, their moral standard may be a combination of the honesty of reporting the true observation and the uprightness of reporting the most accurate message.

The analysis is still in progress and we expect a clearer result as we increase the sample size and as we pin down the functional form of the equilibrium strategy to perform the likelihood estimation of the lying behaviors.

# References


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

Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. "Dynamic psychological games". Journal of Economic Th
144 (1): 1–35.

Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach". Journal of Political Economy 76 (2): 169–217.

Blume, Andreas, et al. 1998. "Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games". The American Economic Review 88 (5): 1323–1340.

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

Deversi, Marvin, Alessandro Ispano, and Peter Schwardmann. 2020. "Spin Doctors: A Model and an Experimental Investigation of Vague Disclosure".

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

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

Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. “Psychological games and sequential rationality”. Games and Economic Behavior 1 (1): 60–79.

Gneezy, Uri. 2005. “Deception: The Role of Consequences”. 95 (1): 384–394.

Gneezy, Uri, Agne Kajackaite, and Joel Sobel. 2018. “Lying Aversion and the Size of the Lie”. American Economic Review 108 (2): 419–453.

Greiner, Ben. 2015. “Subject pool recruitment procedures: organizing experiments with ORSEE”. Journal of the Economic Science Association 1 (1): 114–125.

Hagenbach, Jeanne, and Eduardo Perez-Richet. 2018. “Communication with evidence in the lab”. Games and Economic Behavior 112 (): 139–165.

Khalmetski, Kiryl, and Dirk Sliwka. 2019. “Disguising Lies—Image Concerns and Partial Lying in Cheating Games”. American Economic Journal: Microeconomics 11 (4): 79–110.

Sánchez-Pagés, Santiago, and Marc Vorsatz. 2007. “An experimental study of truth-telling in a sender–receiver game”. Games and Economic Behavior 61 (1): 86–112.

Schmeidler, David. 1973. “Equilibrium points of nonatomic games”. Journal of Statistical Physics 7 (4): 295–300.

Serra-Garcia, Marta. 2018. “Lying in Economics”. Chap. 38 in The Oxford Handbook of Lying, ed. by Jörg Meibauer. Oxford University Press.

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

# A Proofs

*Proof of Lemma 1.* Suppose there exists an equilibrium in the NA-R environment where no agent lies and the audience believes $P(j \text{ is truthful}) = 1$ for any report $j$. The utility of an arbitrary agent with the true observation of $i < N$ and the type $t$ is

$$U(i, j, t) = j - \mathbb{1}(i \neq j)[t + c(i, j)] + \gamma \cdot 1.$$

when she reports some $j$.

As the reputation is constant under the audience's belief that everyone is truth-telling, the agent is better off by reporting $i + 1$ instead of $i$ when

$$U(i, i+1, t) - U(i, i, t) = ((i+1) - i) - (t + c(i, i+1)) > 0.$$

Because we assumed $c(i, i+1) < 1$, there exists some $0 < t < 1 - c(i, i+1)$.

□

*Proof of Lemma 2.* Suppose there exists an equilibrium where an agent lies by reporting a number $j$ below her true observation $i$. Then it must be the case $\rho(j) > \rho(i)$.

It also follows that reporting $j$ must yield a larger utility than reporting $i$:

$$(j - i) + \gamma(\rho(j) - \rho(i)) \geq t + c(i, j).$$

We complete the proof by showing that $\rho(i) = 1$; that is, if any agent choose to report $i$, then it must be the case that the report is truthful.

Take another agent with the intrinsic aversion type $t$ who observes some $i' \neq i$. We claim that $U(i', j, t') > U(i', i, t)$:

$$(j - i) + [c(i', i) - c(i', j)] + \gamma(\rho(j) - \rho(i)) \geq c(i', i) - c(i', j) + t + c(i, j) > 0$$

because of the triangular inequality assumption: $c(i, i') + c(i', j) \geq c(i, j)$. As the choice of this agent is arbitrary, this is the case for all agents who do not observe $i$ never lies by reporting $i$; in turn, this implies $\rho(i) = 1$, a contradiction. □

*Proof of Lemma 3.* Let $L_P^\Omega \subseteq M_P^\Omega$ be the set of messages that liars use to lie with positive probability. Let $l^* = \min L_P^\Omega$. By Lemma 1, $L_P^\Omega$ is nonempty and $l^*$ is well-defined. Also, by the no-underreporting condition, we can deduce that $\rho(1) = 1$ and $1 < l^*$.

We now show $L_P^\Omega = \{l^*, l^*+1, \ldots, N\}$ by contradiction: suppose there exists some elements of $\Omega$ greater than $l^*$ which is not an element of $L_P^\Omega$. Let $n$ be the minimum of such elements, so that $n-1 \in L_P^\Omega$. As $\rho(n) = 1$, we can see that any agent who lies by reporting $n-1$ is strictly better off by reporting $n$ instead:

$$\begin{aligned} U(i,n,t) - U(i,n-1,t) &= (n-(n-1)) - (c(i,n) - c(i,n-1)) + \gamma(\rho(n) - \rho(n-1)) \\ &> c(n-1,n) - (c(i,n) - c(i,n-1)) + \gamma(\rho(n) - \rho(n-1)) \geq 0. \end{aligned}$$

Now it remains to show that the agents who observe $i \in L_P^\Omega$ reports truthfully. Suppose there exists $j, j' \in L_P^\Omega$ such that some agent observing $j$ instead chooses to report $j'$. That is,

$$U(j,j',t) - U(j,j,t) = (j'-j) - (t + c(j,j')) + \gamma(\rho(j') - \rho(j)) \geq 0$$

for any intrinsic aversion type $t$. This implies that any agent who lies by reporting $j$ is strictly better off by reporting $j'$ instead:

$$\begin{aligned} U(i,j',t) - U(i,j,t) &\geq (j'-j) + (c(i,j) - c(i,j')) + \gamma(\rho(j') - \rho(j)) \\ &\geq t + c(j,j') + (c(i,j) - c(i,j')) > 0 \end{aligned}$$

for any intrinsic aversion type $t$ and any true observation $i \neq j$. Therefore, all agents who observe $i \in L_P^\Omega$ reports truthfully. $\square$

*Proof of Lemma 4.* Let $j > i$ be a part of an equilibrium strategy for an $(i,t)$-agent in the NA-R environment:

$$U(i,j,t) = j - [t + c(i,j)] + \gamma\rho(j) \geq U(i,j',t) \quad \forall j' \in M_P^\Omega.$$

We can infer that

$$j - [t + c(i,j)] + \gamma\rho(j) \geq i + \gamma\rho(i),$$

or

$$(j-i)-[t+c(i,j)] \geq \gamma(\rho(i)-\rho(j)).$$

As no agent underreports in the A-R environment, it suffices to show that this agent does not tells the truth in the A-R. Suppose not.

$$U_A(i,i,t) = i > U_A(i,j,t) = j-[t+c(i,j)].$$

This implies $\rho(i) < \rho(j)$, meaning the agent lied in the NA-R both because there were a monetary gain and a reputational gain.

$$(j-i)+[c(i',i)-c(i',j)]+\gamma(\rho(j)-\rho(i)) \geq c(i',i)-c(i',j)+t+c(i,j) > 0$$

because of the triangular inequality assumption: $c(i,i')+c(i',j) \geq c(i,j)$. As the choice of this agent is arbitrary, this is the case for all agents who do not observe $i$ never lies by reporting $i$; in turn, this implies $\rho(i) = 1$, a contradiction. Therefore, the agent must lie in the A-R if she sees the same observation $i$. □

*Proof of Proposition 1.* WTS $P(\text{lie}_{NA-R}) < P(\text{lie}_{A-R})$.

We argue by the law of total probability:

$$P(\text{lie}_{NA-R}) = \sum_{i=1}^{N} P(\text{observe } i) P(\text{lie}_{NA-R}|\text{observe } i);$$
$$P(\text{lie}_{A-R}) = \sum_{i=1}^{N} P(\text{observe } i) P(\text{lie}_{A-R}|\text{observe } i).$$

Because the probability of observing some $i$ is uniform in both environments, it suffices to show that the conditional probability of lying in A-R is greater than or equal to that in NA-R for all true observation $i$.

We learned that there exists some threshold $1 < l^* < N$ in Lemma 3. Thus, conditional on that an agent observing $i \geq l^*$, we have the conditional probability of the agent reporting a lie as

$$P(\text{lie}_{NA-R}|\text{observe } i) = 0;$$

for all intrinsic aversion type $t$, while

$$P(\text{lie}_{A-R}|\text{observe } i) > 0$$

because of the positive probability that the agent has the type $t$ small enough to report a lie.

Now consider the case of $i < l^*$. Let $T_i$ be a subset of $\mathbb{R}$ such that

$$P(\text{lie}_{NA-R}|\text{observe } i, t \in T_i) > 0;$$
$$P(\text{lie}_{NA-R}|\text{observe } i, t \notin T_i) = 0.$$

By Lemma 4, $P(\text{lie}_{A-R}|\text{observe } i, t \in T_i) = 1$; and $P(\text{lie}_{A-R}|\text{observe } i, t \notin T_i) \geq 0$. Thus, regardless of $P(t \in T_i)$, we have

$$P(\text{lie}_{NA-R}|\text{observe } i) \leq P(\text{lie}_{A-R}|\text{observe } i)$$

for all $i < l^*$. Therefore, $P(\text{lie}_{NA-R}) < P(\text{lie}_{A-R})$.

Also, because the monetary payment is a monotone mapping of the reports under the restricted communication, and because any lying takes the form of reporting upward, agents would earn more monetary payoff on average as the probability of lying is greater in the A-R. □

*Proof of Proposition 4.* Suppose there exists an equilibrium where no agent uses a vague message. Let $\rho(j) = 0$ for all $j \in M_V^{\Omega}$. Given that all messages used with positive probability are precise, we know all precise messages are used with positive probability in this equilibrium. We use Lemma 1 and 2 to argue that there exists a positive probability that agents lie upward. Let $l^*$ be the threshold defined in Lemma 3.

Let us first argue that there exists agents who observe $l^*-1$ and lie by reporting $l^*$. Suppose not. Then there must exist some $l > l^*$ such that

$$U(l^*-1, l, t) - U(l^*-1, l^*, t) = (l - l^*) + (c(l^*-1, l^*) - c(l^*-1, l)) + \gamma(\rho(l) - \rho(l^*)) > 0.$$

Note that $c(l^*-1, l^*) - c(l^*-1, l) \leq 0$ because $l > l^*$ and $c$ is increasing in the distance between the two arguments. Also, $c(l^*-1, l^*) - c(l^*-1, l) \leq c(i, l^*) - c(i, l) \leq 0$ for all $i < l^*$ because of the triangular inequality assumption. This implies that for all agents whose true observation is below $l^*$ is better off by reporting $l$ instead of $l^*$. As no agent would lie by reporting $l^*$, this is a contradiction to the definition of threshold.

Now consider the agent who observes $l^*-1$ and lies by reporting $l^*$. The agent receives the utility of

$$U(l^*-1,l^*,t)=l^*-(t+c(l^*-1,l^*))+\gamma\rho(l^*).$$

However, if the agent reports $\{l^*-1,l^*+1\}$ instead, the agent receives

$$U(l^*-1,\{l^*-1,l^*+1\},t)=l^*.$$

That is, the agent is better off by reporting $\{l^*-1,l^*+1\}$ when

$$t>\gamma\rho(l^*)-c(l^*-1,l^*),$$

which happens with a positive probability. Furthermore, the above analysis is valid for all $\rho(\{l^*-1,l^*+1\})\geq 0$. This is a contradiction to the assumption of an equilibrium with no vague messages. Therefore, there exists a positive probability that agents use a vague message in any equilibrium in the NA-UR environment. $\square$

# B Experimental procedure

An experiment session contains the instructions, a preliminary quiz, and two stages of choice tasks. The subjects receive invitations through the Missouri Social Science Experimental Laboratory(MISSEL)'s ORSEE system. Once registered, they receive a link to a Zoom meeting. After subjects join the Zoom meeting room, the experimenter reads the instructions, followed by a preliminary quiz to ensure the understanding of the procedures. In an anonymous session, everyone including the experimenter has her video off, and the experimenter does not explicitly take attendance. In order to assure the subjects about anonymity, subjects are asked to use a screen name. The experimenter emphasizes that the screen name is used solely for data analysis and that she cannot associate the screen name with their true identity. In an non-anonymous session, on the other hand, the experimenter has her video on and asks the subjects to turn their video on as well. Also, the experimenter takes attendance with the experiment roster.

The subjects are allowed to participate in the main experiment only after successfully passing the preliminary quiz. Each subject has three chances to attempt the quiz before turned away. The experimenter provides a link to the main experiment website via the chat window of the Zoom meeting upon the successful completion of the quiz.

At the beginning of the main experiment part, subjects enter their identification information. In the anonymous treatment, subjects are asked to enter their screen name; whereas, in the non-anonymous treatment, subjects are asked their name and student ID.

After entering the identification information, the subjects observe the realization of a random integer uniformly distributed between 1 to 10 and later asked to report the number. In the anonymous treatment, the number is generated within the experiment software. In the non-anonymous treatment, subjects click on a link to a Google search result for the phrase "random number between 1 to 10."

In the reporting screen, subjects are presented with ten boxes labeled from 1 to 10 and report by selecting the numbered boxes. In a precise stage (Figure 4), which allows for precise messaging only, a subject may select only one box at a time, whereas in a vague stage (Figure 5), she may select multiple boxes. Subjects are told that this selection is deemed as a statement that the numbers represented by the selected box include the number they observed.

In the case of a precise message, the subject is paid the equivalent in dollars to the number she selected divided by two. In the case of a vague message, the computer randomly chooses one number from the submitted selection of numbers and pay the equivalent in dollars to the randomly chosen number divided by two.

After subjects complete both precise and vague stages, a confirmation screen reviews their compensation information. In the anonymous treatment, the screen also provides a password. The subjects visit an external website and retrieve their compensation in the form of an Amazon gift card by entering their screen name and the password. This process is to ensure the anonymity of the experiment session. In the non-anonymous treatment, they are asked to enter their email address and taxpayer information to receive the Amazon gift card directly emailed to them.

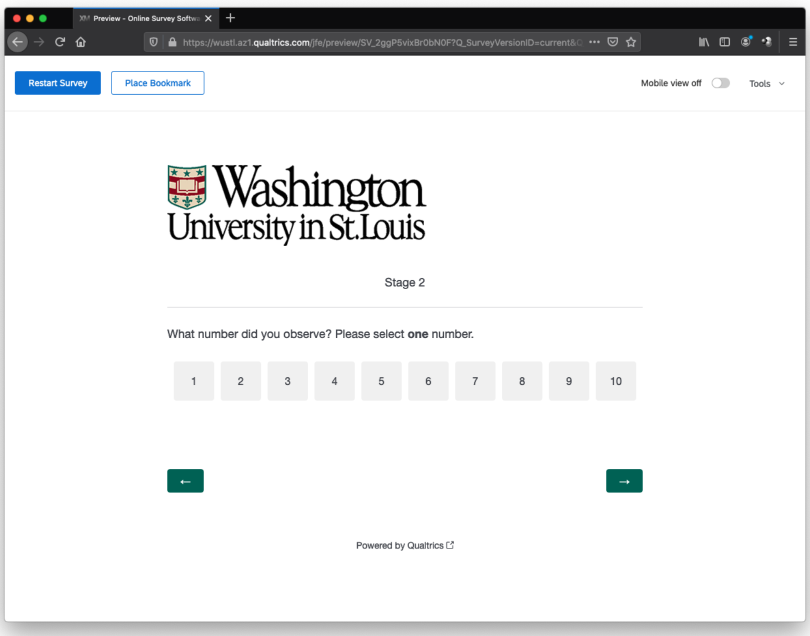


Figure 4: A screenshot of the experiment software displaying the precise-message stage

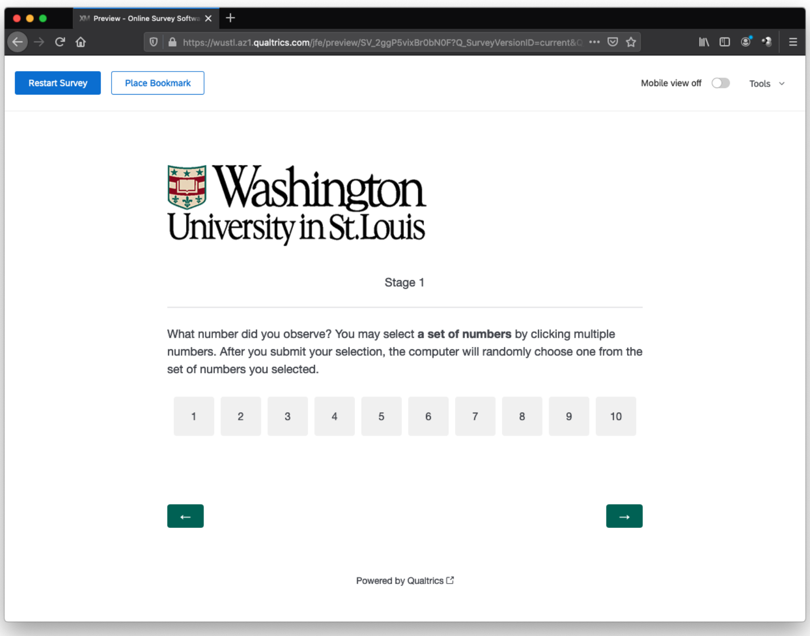


Figure 5: A screenshot of the experiment software displaying the vague-message stage

# C Experimental Instructions

## C.1 Instructions for Anonymous, Observable session (Zoom session; no video)

Welcome to the experiment! Thank you very much for participating today, and I will first walk you through the instructions for the experiment.

We need your full attention during the experiment. If you have trouble with hearing the audio or seeing the shared screen, please let me know. Do not turn the video on during the experiment. If you have any questions during the instructions, please raise hand so that I can unmute you. Your question will be answered out loud, so everyone can hear.

The experiment you will be participating in today is an experiment in individual decision making. At the end of the experiment, you will be paid an Amazon gift card. You will receive the show-up fee of $2 for completing the experiment, with the additional amount that depends on your decisions and on chance. The details of the compensation will be described later. All instructions and descriptions that you will be given in the experiment are accurate and true. In accordance with the policy of this lab, at no point will we attempt to deceive you in any way.

I would like to first point out that we want to ensure this experiment is conducted anonymously, meaning that we cannot connect the responses recorded in this experiment to any particular individual who participated in this research. Qualtrics, the survey platform we are using, provides an option for the researchers to not collect any personal information, such as IP address or geographic location of the participants, for anonymous surveys. Also, your response will be recorded with a SCREEN NAME. You will be asked to choose a screen name that is at least 8 characters in length using letters, numbers, and underscore. This SCREEN NAME is only used in data analysis and distributing your Amazon gift card after the experiment. We cannot and will not attempt to associate SCREEN NAMEs to any particular individual.

I will now describe the main features of the experiment. First, there is a short quiz to ensure your understanding of the procedures. You will be able to repeat the quiz if you make mistakes. You will have three chances to attempt the quiz. If you fail to get all questions correct after three attempts, you may not participate in the main experiment. Even in such a case, please remain connected to the Zoom session until everyone finishes, and you will receive the show-up fee for today's experiment. The main part of the experiment after the preliminary quiz consists of two STAGES. In each STAGE you will observe a number that we ask you to remember and later report to us. The number you report to us determines how much money you will be paid. At the end of the experiment, a confirmation screen will summarize today's experiment and provide the information to retrieve your payment. I will give you more details about each step. At the end of this instructions, we will first provide a link for the preliminary quiz using the chat window. This is a screenshot of the webpage. Please choose a screen name, and make sure you keep this screen name. After you successfully complete the preliminary quiz, you will be reminded of your screen name once again. We recommend you copy and paste the screen name. We cannot recover this information for you, and you will not be able to receive your compensation without the correct screen name. Please wait while everyone else finishes the quiz. Once everyone finishes, we will provide another link for the main experiment using the chat window. This is a screenshot of the first page of the main experiment. Make sure you use the same screen name you used in the preliminary quiz. You may not receive your compensation if the screen names do not match. As previously mentioned, your main task today is to observe a number that we ask you to remember and later report to us. The observation process is identical in both STAGES. At the beginning of both STAGES, the computer will randomly draw a number between 1 to 10. The probabilities are equal across the numbers; that is, each number is chosen with the same probability of one-tenth. We ask you to remember the number and report on the next screen. However, the way you can report differs between STAGES. In one STAGE, you are allowed to select one number after you observe the draw. This is a screenshot of the experiment stage where you can select one number. By one number, we mean that you may click only one number on the screen. We will interpret this selection as a statement that the number you observed is the number you selected, and we will regard this number as your report. The number you report determines how much money you will be paid. You will be paid the equivalent in dollars to the number you report divided by 2. In other words, if you report "1", you receive 50 cents. If you report "2", you receive $1, if you report "3", you receive $1.50 and so on. A confirmation screen after your report will help you review your selection and the corresponding payment. In another STAGE, you are allowed to select a set of numbers after you observe the draw. This is a screenshot of the experiment stage where you can select a set of numbers. By a set of numbers, we mean that you may click multiple numbers on the screen. For instance, you may choose to click on four numbers, one number, two numbers, or even all ten numbers. If you select multiple numbers, we will interpret it as a statement that the number you observed is one of the numbers you selected. If you select a single number, we will interpret it as a statement that the number you observed is the number you selected. After you submit your selection, the computer will randomly choose one number from the set of numbers you selected. We will regard this randomly chosen number as your report. Again, the number you report determines how much money you will be paid. You will be paid the equivalent in dollars to the randomly chosen number from the set of numbers you selected divided by 2. If the randomly chosen number is “1”, you receive 50 cents. If “2”, you receive $1, if “3”, you receive $1.50 and so on. A confirmation screen after your report will help you review your selection and the corresponding payment.

The order of the two STAGES is randomly determined. In other words, it is equally likely that you first participate in the STAGE allowing only a single number and then participate in the STAGE allowing a set of numbers, or first participate in the STAGE allowing a set of numbers and then participate in the STAGE allowing only a single number. In any case, you will play each STAGE only once.

After the completion of both STAGES, a final review screen will provide the information to receive your Amazon gift card code. This is a screenshot of the review screen. This includes your SCREEN NAME you entered, the amount you will receive, and a randomly generated passcode. Because the experiment is anonymous, we have no means to recover this information for you. Please make sure you either print or take the screenshot of this page for your record, because it is very important when you retrieve your compensation.

Please make sure you understand the retrieval process well. You will need to provide the information from the final review screen on a separate webpage on the MISSEL website to retrieve a link to the Word document containing your Amazon gift card code.

This is a screenshot of the retrieval website. You need to provide both the SCREEN NAME and the corresponding amount for verification purposes. The page will provide a unique hyperlink to the Word document associated with the SCREEN NAME-amount combination. The Word document is password-protected, and only the person with the correct password can access the content. Use the passcode from the final review screen to open this file, and you will find your Amazon gift card code inside the document. Please allow us a few hours after the completion of the experiment to validate the data and upload the codes to the website. You will receive an email when the codes are ready for everyone. Please note that this MISSEL webpage only collects the SCREEN NAME and the corresponding amount, and no other personal information. For security reasons, the gift code retrieval webpage will be available only for a week upon the completion of the experiment.

We are sorry for the inconvenience that we are not able to email you with the gift code directly. This payment process is to ensure anonymity in this experiment, and we appreciate your understanding that the anonymity of the reports constitutes a crucial component of our research.

This is the end of the instructions. If you have any questions, please raise hand. Otherwise, I will provide the link via the chat window. Please copy and paste the link to your browser to generate the SCREEN NAME.

## C.2 Instructions for Identifiable, Non-observable session (Zoom)

Welcome to the experiment! Thank you very much for participating today. Before we start, I will go over the roster to take attendance to make sure I have everyone registered for the session. During the attendance please turn your video on.

I will now walk you through the instructions for the experiment. We need your full attention during the experiment. If you have trouble with hearing the audio or seeing the shared screen, please let me know. If you have any questions during the instructions, please use the hand-raising feature of Zoom and your question will be answered out loud, so everyone can hear.

The experiment you will be participating in today is an experiment in individual decision making. At the end of the experiment, you will be paid an Amazon gift card. You will receive the show-up fee of $2 for completing the experiment, with the additional amount that depends on your decisions and on chance. The details of the payment will be described later. All instructions and descriptions that you will be given in the experiment are accurate and true. In accordance with the policy of this lab, at no point will we attempt to deceive you in any way.

This is a screenshot of the first page in the main experiment. At the end of the instructions, I will provide the link to the experiment using the chat window. Please copy and paste the link to your browser. The first screen will ask your identification information – your first and last name and your student ID.

After you enter your information, you will proceed to the next screen and take a short quiz to ensure your understanding of the procedures. You will be able to repeat the quiz if you make mistakes. You will have three chances to attempt the quiz. If you fail to get all questions correct after three attempts, you may not participate in today's experiment. In such case, you will only receive the show-up fee for today's experiment.

The main part of the experiment consists of two STAGES after the preliminary quiz. In each STAGE you will observe a number that we ask you to remember and later report to us. The number you report to us determines how much money you will be paid. At the end of the experiment, a confirmation screen will summarize today's experiment and provide the information to retrieve your payment. I will give you more details about the observation, reporting, and payment processes.

The observation process is identical in both STAGES. At the beginning of both STAGES, we will provide a link to Google page that randomly generates a number between 1 and 10. The probabilities are equal across the numbers; that is, each number is chosen with the same probability of one tenth. We ask you to open the link, remember the number, close the Google page, and report the number on the next screen. This is a screenshot of an example Google page. You will see that this is a search result for 'random number between 1 and 10', and the page displays a randomly generated number that matches the search phrase. Do not click on the 'generate' button on the Google page, because the number you see is already randomly generated. Any additional generation only distorts the statistical properties of the experiment. I will now demonstrate how this Google page works. You will find this link during the experiment, and this is equivalent of opening a google page and typing in "random number between 1 and 10." When you open the link, a new window pops up. As you can see, this number is already randomly generated and you should not generate the number again.

The way you can report differs between STAGES. In one STAGE, you are allowed to select one number after you observe the randomly generated number. This is a screenshot of the experiment stage where you can select one number. By one number, we mean that you may click only one number on the screen. We will interpret this selection as a statement that the number you observed is the number you selected, and we will regard this number as your report. The number you report determines how much money you will be paid. You will be paid the equivalent in dollars to the number you report divided by 2. In other words, if you report "1", you receive 50 cents. If you report "2", you receive $1, if you report "3", you receive $1.50 and so on. A confirmation screen after your report will help you review your selection and the corresponding payment. In another STAGE, you are allowed to select a set of numbers after you observe the randomly generated number. This is a screenshot of the experiment stage where you can select a set of numbers. By a set of numbers, we mean that you may click multiple numbers on the screen. For instance, you may choose to click on four numbers, one number, two numbers, or even all ten numbers. If you select multiple numbers, we will interpret it as a statement that the number you observed is one of the numbers you selected. If you select a single number, we will interpret it as a statement that the number you observed is the number you selected. After you submit your selection, the computer will randomly choose one number from the set of numbers you selected. We will regard this randomly chosen number as your report.

Again, the number you report determines how much money you will be paid. You will be paid the equivalent in dollars to the randomly chosen number from the set of numbers you selected divided by 2. If the randomly chosen number is "1", you receive 50 cents. If "2", you receive $1, if "3", you receive $1.50 and so on. A confirmation screen after your report will help you review your selection and the corresponding payment.

The order of the two STAGES is randomly determined. In other words, it is equally likely that you either participate in the STAGE allowing a single number first and then participate in the STAGE allowing a set of numbers or participate in the STAGE allowing a set of numbers first and then participate in the STAGE allowing a single number. In any case, you will play each STAGE only once.

After the completion of both STAGES, a final review screen will summarize today's experiment. This is a screenshot of the review screen. The last screen will ask your email address to receive the Amazon Gift Code of the amount that corresponds to your responses. We will directly send you Amazon gift code to the email address you provide. Please allow us a few hours after the completion of the experiment to validate the data and send the email. In addition, Washington University in St. Louis recommends student subjects to report their taxpayer identification information for tax purposes. If you are an international student and do not have the taxpayer identification information, please indicate so by entering 'Foreign' in the form. If you do not have or do not wish to provide the identification information, please indicate that you would like to opt out by entering 'Refuse' in the form.

This is the end of instructions. If you have any questions, please raise your hand. Otherwise, I will provide the link via the chat window. Please copy and paste the link to your browser and start the experiment.