Lying Aversion and Vague Communication: An Experimental Study

Keh-Kuan Sun* Guangying Chen† Stella Papadokonstantaki‡§

October 30, 2022

The latest version of the draft is available here.

Abstract

An agent may strategically employ a vague message to mislead the audience's belief about the state of the world, but this may cause the agent to feel guilt or negatively impact how the audience perceives the agent. Using a novel experimental design that allows participants to be vague and at the same time isolates the internal cost of lying from the social identity cost of appearing dishonest, we explore the extent to which these two types of lying costs affect communication. We find that participants exploit vagueness so as to be consistent with the truth, yet at the same time leveraging the imprecision to their own benefit. More participants use vague messages in treatments where the social identity concern is relevant. In addition, we find that social identity concern substantially affects the length and patterns of vague messages used across the treatments.

Keywords: Lying; Vagueness; Communication; Experiments; Behavioral Economics

JEL codes: C91, D90

*Smith Institute for Political Economy and Philosophy, Chapman University. Email: sun.k@wustl.edu

†Olin Business School, Washington University in St. Louis. Email: guangyingchen@wustl.edu

‡Department of Economics, Washington University in St. Louis. Email: p.stella@wustl.edu

§We thank Brian Rogers, Marta Serra-Garcia, Agne Kajackaite, Cynthia Cryder, John Nachbar, Jonathan Weinstein, Marcus Berliant, Sangmok Lee, Paulo Natenzon, John Rehbeck, and seminar participants at Washington University in St. Louis, the ESA Job-Market Candidates Seminar Series, Berlin School of Economics, Korea Information Society Development Institute, Tokyo Institute of Technology, ECONtribute Summer Workshop 2021 on Social Image and Moral Behavior, 2021 North American Summer Meeting of the Econometric Society, 2021 Asian Summer Meeting of the Econometric Society, and 2021 ESA Global Online Around-the-Clock Meetings.

1 Introduction

Communication has been studied extensively in economics over the last decades. In particular, the type of messages sent by an agent in the communication process and its impact on the audience's belief have been important subjects of game theory. Standard economic models predict that a sender will choose a message that yields the largest benefit from the receiver's action, even when such a message is a lie. However, it has been observed that senders often choose to be evasive or use vague messages, instead of telling maximal lies despite them possibly getting a lower benefit. A natural question that follows is why and when a sender uses a vague message over a maximal lie. This issue is crucial for a unified understanding of misleading behavior in many applications, including the public good provision problems (Serra-Garcia, Damme, and Potters 2011), sender-receiver disclosure games (Hagenbach and Perez-Richet 2018), and persuasion game (Deversi, Ispano, and Schwardmann 2021).

The answer to the question requires a careful examination of the behavioral aspects of lying and vague communication. First, as demonstrated by recent development in the literature on lying behavior (Serra-Garcia 2018; Fischbacher and Föllmi-Heusi 2013; Abeler, Nosenzo, and Raymond 2019), people exhibit a non-trivial degree of lying aversion. If an agent intrinsically prefers to be honest, then this internal cost of lying should inhibit the use of lies and encourage the use of vague yet relatively truthful messages. Second, a message not only affects the audience's belief about the state of the world, but also the belief about how honest an agent is. When the agent cares about their social identity as being considered honest, this external concern should impact their choice of message. The nontrivial cost of lying suggests that communication is no longer cheap-talk and the misleading messages may bear strategic significance as a credible signal to the audience about the behavioral types of the agent. 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.

In this paper, we present a model of cheating game in which 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 and their preferences for truth-telling. The monetary payoff is determined solely by the reported message. We study two separable motivations for honesty: the internal motivation for being honest and the external concern (social identity) for being seen as honest. We isolate the two costs by employing an anonymous environment where the audience cannot identify an agent with a message. Based on the model, we hypothesize that people use vague messages to reduce their internal guilt and increase their monetary payoff. The absence of the social identity concern in the anonymous environment should lead to a more straightforward profit-maximizing message choices, while the prediction is more opaque in the non-anonymous environment due to the multiplicity of equilibria by the introduction of a richer message space.

To test hypotheses about vague communication with lying aversion, we compare treatments from an online experiment in which subjects face a variant of Fischbacher and Föllmi-Heusi (2013) 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 increases with the reported number. The basic idea of the experiment is that the discrepancy between the maximum number they could have reported and the actual report should capture one's lying aversion, or their preferences for truth-telling by the same token. We generalize the FFH model by allowing subjects to transmit a set-valued message to understand the effect of vagueness in communication.

We use a 2x2 experiment design to test the predictions and, in particular, to separate the intrinsic cost of lying from the social identity cost of appearing dishonest. First, 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 their response. In the non-anonymous session, on the other hand, the experimenter knows each subject's response. Second, a subject confronts two reporting tasks within each session. The observation process is identical and independent in the two tasks. When the subject reports, however, the set of messages available between the two tasks. In the task with restricted communication, the set is restricted to single-valued messages only. In the task with unrestricted communication, the subjects are allowed to use both single-valued and set-valued messages.

The experiment data shows that most subjects use vague messages and that they report higher on average when vague messages are allowed. However, the way they use vague messages differs across the treatments. First, we observe less lying behavior when vague messages are allowed in the anonymous treatment. As anonymity should suppress the subjects' concern for their social identity, the decrease in lying behavior suggests that a vague yet truthful message reduces one's internal cost. That is, subjects exploit vagueness to be consistent with the truth while leveraging the imprecision to their own benefit. Second, subjects use vague messages in the non-anonymous treatment more often compared to the anonymous counterpart. The increased frequency of vague messages and the relevance of the social identity concern suggest that a message's vagueness affects the message choices through both the internal and external costs of lying. Furthermore, we find that the pattern of the reported messages differs substantially between the two types of sessions. When subjects use vague messages in the anonymous treatments, most do not refrain from the most obvious forms: that is, they report a combination of their true observation and the maximum number (10). However, we find only a much smaller fraction of subjects using vague messages employ such obvious messages in the non-anonymous treatment.

Our paper provides a bridge between the literature of lying behavior to a broader set of studies that involve vague communication. Existing literature on vague communication focuses on the strategic use of vague messages without considering the behavioral aspect of lying aversion. For instance, Serra-Garcia, Damme, and Potters (2011) uses the assumption that a set-valued vague message 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 social identity concern as well. Agranov and Schotter (2012) show that in coordination games with multiple equilibria it might be beneficial for a benevolent sender to use vague communication. Zhang and Bayer (2022) show that in a delegation game the social welfare is higher when messages are intervals. Wood (2022) concludes that the information transmitted is more accurate when the senders have the option to send both precise and vague messages. We find a similarity between our design and that of Deversi, Ispano, and Schwardmann (2021). They study the strategic use of vagueness in voluntary disclosure by allowing subjects to send as a message an interval which contains their type. Our main focus is to explore to which extent the attitude toward lying and misleading behavior affects the use of vague messages in conjunction with their strategic motives. We also provide subjects with more flexibility in their choice of messages by allowing any subset of the state space.

Conceptually, we view the lying behavior as an optimization process in which a rational agent chooses the optimal amount of dishonest act, by balancing its marginal monetary benefit and the non-monetary costs that depend on the probability of being found out as a liar. In a more specific context to the nature of lying behavior, Sobel (2020) provides a comprehensive framework to understand lying and deception in games. He distinguishes between three properties - the form, the interpretation, and the consequence - of a message, and our model closely follows the framework suggested in the paper. We define lying strictly by the formal property of a message: a message is a lie only when it does not represent the true state. The agents in our model face both monetary and psychological consequences of their choice of messages. Furthermore, through the analysis of equilibrium strategies, we model how the interpretation of a message affects the agent's decision.

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), Shalvi et al. (2011) and many more). They argue that individuals certainly express aversion toward lying and take significantly lower payoffs than the theories predicted.

An empirical consensus of the FFH experiments is that people do not lie as much as they could, potentially because lying is costly. 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. 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. As Abeler, Nosenzo, and Raymond (2019) note in their conclusion, however, the FFH paradigm has focused on subjects reporting a single number and excludes lies by omission or vagueness.

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 their monetary payoff but also whether she lies and how others interpret their report. These papers adopt the FFH paradigm and simplify the sender/receiver structure of a signaling game into a cheating game between reporting agents and an observing audience. The idea, with the assumptions of a non-atomic game, minimizes the role of the receiver and allows a sharper focus on that of the sender. We inherit their assumptions of these two motivations for truth-telling, the internal guilt and the external concern for social identity, and extend by allowing vague communication. 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.

We observe that as long as the message can remain even remotely truthful, most agents exhibit no hesitation in increasing their monetary payoff. This is analogous to "moral wiggle room" behavior (Dana et al. 2007), the notion that people increase self-interested behavior when excuses for the selfish actions are available. 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.

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 Ω across agents. Each agent has a private type t that represents their 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 social identity for being seen as honest. Formally, we write a type (i, t) agent's utility for reporting a message J as:

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

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

[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

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 $M_V^\Omega \equiv M^\Omega \setminus M_P^\Omega$ 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 $\{\{1,2\},\{2,3\},\{1,3\},\{1,2,3\}\}$. Note that our definition of lying depends purely on the form of a message, similar to the framework proposed in Sobel (2020). The choice of a message bears its consequences through different components of the agent's utility.

The monetary payoff of an agent maps their 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 $x \in \Omega$, their monetary payoff $\pi(J)$ is simply x.

In addition to their monetary payoff, the agent also has two different motivations for honesty. First, she has the internal motivation for being honest. When their 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 their 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) $c(\pi(\{i\}), \bar{\pi}(J)) + c(\bar{\pi}(J), \bar{\pi}(K)) \geq c(\pi(\{i\}), \bar{\pi}(K))$.

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

She also has the external motivation from the social identity for being seen as honest; that is, their utility depends on the audience's belief on how honest the agent is. We assume the audience is a rational Bayesian who forms their posterior belief based on the agent's 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

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

where σ_{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 ρ in terms of the agent's utility by a parameter γ. The parameter measures the agent's sensitivity toward their social identity by reporting a message J. We further assume that i) γ 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:

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

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

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

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^{Ω}. By restricting the message space to M_P^{Ω}, 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 social identity. 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 social identity remains constant independent of their 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:

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

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.

	Non-anonymous	Anonymous
Restricted	NA-R	A-R
Unrestricted	NA-UR	A-UR

Table 1: Four environments

3 Analysis

This section describes equilibria and the agents' behavior in each of the four environments we defined in the earlier section. We begin by arguing that we should expect to observe vague messages used with positive probabilities in any equilibrium.

Proposition 1. *In any equilibrium under unrestricted communication, there exist some types of agents who use a vague message with positive probability in their mixed strategy.*

Proof. In Appendix. □

The intuition behind the proposition is that there exist agents with moderately high intrinsic lying aversion type t who would prefer to be truthful while increasing their monetary payoff. Therefore, they include higher numbers apart from their true observations in their messages. For these agents, such messages strictly dominate any precise lie.

In comparison to the unrestricted communication environment of our interest, a particularly useful element when analyzing the restricted communication environment is the assumption that there always exists a positive mass of truth-tellers for each observation i. While sufficiently plausible 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: the richer message space induces multiplicity in the agents' reporting strategy. We illustrate this point with the following examples.

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

This example describes the scenario in which the agents care much about their social identity and are forced to stick to the (exogeneously) given norm of the full vagueness. While the specific example bears potential real-world examples[3], an important implication is that there can exist a plethora of equilibria depending on the combination of γ and the off-path beliefs. Consider the following example.

Example 2. All agents reports a message in the form of an interval: $[i,N]=\{i,i+1,\ldots,N-1,N\}$, 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: $\frac{i+N}{2}+\gamma\rho(J) > U(i,\{i,x^*,x^*+1,\ldots,N\},t)$[4] 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.

The multiplicity of equilibria obscures our understanding of the agents' behavior regarding vague communication. We thus first turn our attention to anonymous environments to discern the relationship between one's internal cost of lying and vague messages.

[3]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 social identity. Many similar examples can arise in the situations where the agents value their social identity highly. We appreciate Brian Rogers for suggesting this interpretation of the equilibrium.

[4]This message maximizes the expected payoff conditional on including the true state i. We later define this kind of message as an optimal vague message.

Because we are now free of the social identity component, we can utilize the fact 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 simplifies to 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 vague messages that allow their to remain truthful yet increase the expected payoff. This fact leads to the following observations.

Observation. *If a type (i,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 2 (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.*

Another implication of the anonymous environment worth noting is that the absence of the social identity concern simplifies the problem into a straightforward 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[5] to find the threshold $x^* = \lceil (N+2) - \sqrt{2N - 2i + 3} \rceil$.

Definition 4. The optimal vague message (OVM) for each true observation i defined as

$$OVM_i \equiv \{i, x^*, x^* + 1, \dots, N\}, \tag{8}$$

where x^* the threshold above which including the values maximizes

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 inference in the A-UR environment.

Proposition 3. *In the A-UR environment, all truth-tellers use 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.*

Lastly, we conclude this section by comparing the behavior of an agent in the non-anonymous environment to that in the anonymous environment. A thought experiment of choosing an agent and comparing their behavior in the two environments easily leads to a conjecture that the absence of the social identity concern should only facilitate more lies. The following lemma shows this is indeed the case under restricted communication.

Lemma 1. *If a type (i, t) agent lies in an equilibrium in the NA-R environment, then she lies in the A-R environment.*

Proof. In Appendix. □

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

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

Proposition 4 (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.*

Proof. In Appendix. □

4 Hypotheses

Hypothesis 1. *In both the NA-UR and the A-UR environment, people use vague messages.*

Hypothesis 1 is a consequence of Proposition 1.

Hypothesis 2 (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 2 is a consequence of Proposition 2 and the observations about the behavior in the A-UR environment.

Hypothesis 3. *In the A-UR environment,*

i. all truth-tellers in A-UR uses OVM;

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

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

Hypothesis 3 is a consequence of Proposition 3 and its corollary.

Hypothesis 4 (NA-R/A-R)**.** *Under restricted communication, agents earn more monetary payoff on average in the anonymous environment: $earning_{NA-R} \leq earning_{A-R}$.*

Hypothesis 4 is a consequence of Proposition 4.

5 Experimental 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 their response. As the subjects are instructed that the experiment intentionally prohibits such mapping, this treatment should establish the effect of suppressing their social identity concern and emulate the environment where $\gamma \to 0$. In the non-anonymous (identifiable) 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[6] so that the experimenter cannot know the true observation. The experimenters are not able to identify whether a particular individual has lied or not in any of the treatments. This design choice derives from the idea that an environment where the experimenters can detect cheating at the individual level 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. However, the set of available messages differs between the two tasks. In the restricted stage, the set is restricted to single-valued messages only. In the unrestricted stage, the subjects are allowed to use both single-valued and set-valued messages. In the restricted stage, the subjects are paid the equivalent in dollars to the number they report divided by 2. In the unrestricted stage, when they report a single number, they are paid the equivalent in dollars to the number they report divided by 2 , while when they report multiple numbers, the computer will randomly choose one of the numbers and they are paid the equivalent in dollars to this number divided by 2.

The combination of these between-subject and within-subject variations yields four treatments in our experiment: NA-R[7], NA-UR, A-R, and A-UR. The order of precise and vague treatments within a session is randomized to minimize learning effects.

We recruited 176 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 (Greiner 2015). 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.

We conducted 18 sessions and each session lasted for approximately 30 minutes, including the instructions read over Zoom meeting, a screening quiz to make sure the understanding of the experiment procedures, and two stages of the main experiment task via Qualtrics [8]. In all cases, subjects received a

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

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

[8]The first 11 sessions were conducted from August 2020 to October 2020, while the rest in September 2022.

\$2 show-up fee, thus the total amount they could earn varies from \$2 to \$12. Subjects received \$9.84 on average including the show-up fee. No subject participated in more than one session. See Appendix for our experimental instructions and procedural details.

6 Results

We present the basic summary statistics in Table 2. There were 8 anonymous sessions and 10 non-anonymous sessions, with the average size of 12.3 and 8.2 participants in each session, respectively[9]. 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 length of vague messages is computed by taking the average number of the numbers included in the messages.

	Restricted	Unrestricted			Observations
	Average	Average	Vague (%)	Length	
Anonymous	7.619	8.132	48%	2.103	97
Non-anonymous	7.000	8.282	75%	2.810	79

Table 2: Data summary

We first report that the majority of subjects used vague messages in both treatments.

Result 1. *In the NA-UR environment, 59 out of 79 participants (74.7%) used a vague message. In the A-UR environment, 47 out of 97 participants (48.4%) used a vague message.*

The result supports Hypothesis 1.

In the anonymous treatment, the average reports for the restricted and unrestricted treatments are 7.619 and 8.132, respectively, while the non-anonymous reports are 7 and 8.282, respectively. Those

The results of a logit regression show no significant difference in the probability of lying between the two group of sessions.

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

reports are significantly lower than the profit maximizing reports and are in line with the literature which has highlighted agents' preference for truth-telling. Overall, we find that allowing vague messages increase the average report. Using OLS regressions, we find that there is no order effect on the average report. We also find that the probability of lying is independent of gender both in the restricted and unrestricted treatments. More details about the logit regression outcome are available in Appendix. In addition, we find that the impact of anonymity on average reports is insignificant. Finally, we show that the introduction of non-anonymity significantly affects the choice of vague messages. In the treatment where the social identity concern is relevant, subjects use vague and longer messages more often.

Result 2. *In the anonymous environment,*

1. *more participants lied when the communication is restricted (43.3% in A-R and 30.9% in A-UR);*
2. *almost all participants who were truthful in A-R remained truthful in A-UR*[10]*;*
3. *52.58 % of subjects reported truthfully under both A-R and A-UR, 26.80% of them always lied, while 16% switched from lying to truth-telling when they were allowed to be vague;*
4. *participants reported higher on average when the communication is not restricted (7.619 in A-R and 8.282 in A-UR);*

The comparison between the restricted and the unrestricted communication in anonymous sessions confirms Hypothesis 2. More subjects lied in the restricted than in the unrestricted treatment and their behavior can be classified into three types: truth-tellers (reported truthfully under both treatments, conditional liars (switched from lying to truth-telling when they were allowed to be vague) and liars (reported a lie under both treatments). This is in line with previous work on lying aversion (Gneezy, Kajackaite, and Sobel (2018), Khalmetski and Sliwka (2019), among others). An interesting point is that 15 of the 16 participants who switch from lying to truth-telling reported 10 in A-R, yet they chose a

[10] There are 4 subjects who were truthful in the A-R and lied in the A-UR. Their true observations were much larger in the A-UR than in the A-R treatments. It is likely that the counterfactual would be consistent with our prediction conditional on observing similar numbers in both treatments.

vague and truthful message whose expected earning is less than 10. This suggests that the use of vague message reduces the internal cost when it contains the true state.

Table 3 presents the average reports in the A-R and A-UR treatments for truth-tellers, conditional liars and liars. We find that truth-tellers report on average 1.282 higher when they are allowed to be vague. This result is significant at 1% significance level and is consistent with the model prediction that truth-tellers seek payoff maximization conditional on including the true observation in the report. Conditional liars reported on average 1.156 lower when they were allowed to be vague. This is significant at 1% significance level and supports Hypothesis 2. We interpret this result as that the agents with moderate t prefer to use a vague message and include the true state to reduce the internal cost of lying. Consistent with our expectation about agents with higher t, we did not find a significant impact of the unrestricted communication among the liars in terms of their average reports. In addition, the t-tests on the difference between the subjects' true observation and their report show over-reporting across all three types in the A-UR treatment. This result is also in line with Hypothesis 2, which states that agents earn higher monetary pay-off under unrestricted communication. Likewise, we also find that the true observation has a negative effect on the probability of lying in the restricted treatment, while it has no effect in the unrestricted one. Table 4 presents the estimated coefficients of the logit model of effect of the true observations on the probability of lying. This result is intuitive as an agent who randomly drew a low observation can increase their payoff only by lying when the communication is restricted to precise messages. However, they can always employ a vague yet truthful message in the unrestricted environment; hence the true observation has no effect.

	Restricted	Unrestricted	Difference in means	Observations
Truth-tellers	5.608	6.890	1.282***	51
	(0.416)	(0.293)	(0.464)	
Conditional liars	9.875	8.718	-1.156***	16
	(0.125)	(0.316)	(0.370)	
Liars	10	9.962	0.038	26
		(0.027)	(0.027)	
Test (Report = True)				
Truth-tellers	0	1.953***		51
		(0.267)		
Conditional liars	5.000***	1.531***		16
	(0.791)	(0.432)		
Liars	5.462***	5.076***		26
	(0.521)	(0.543)		

Table 3: Anonymous Treatment

	Observation (Restricted)	Observation (Unrestricted)
Lying dummy	-0.138*	-0.081
	(0.072)	(0.081)

Table 4: Logit model on the probability of lying

Result 3-1. *In the A-UR environment,*

1. (a) *28.4% (19 of 67) of truth-tellers reported the optimal messages (either the OVM or the honest 10), 26.9% (18 of 67) used a pair of the true observation and 10;*

   (b) *23.9% (16 of 67) of truth-tellers used a precise message below* {10}*;*

2. *only 2 out of 97 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 3. Figure 1 summarizes the types of messages used in A-UR treatments. Overall, 51.5% (50 of 97) of participants used a precise message, and 48.5% (47 of 97) used a vague message. Among the 50 precise messages, 54% (27 of 50) were lies and 46% (23 of 50)

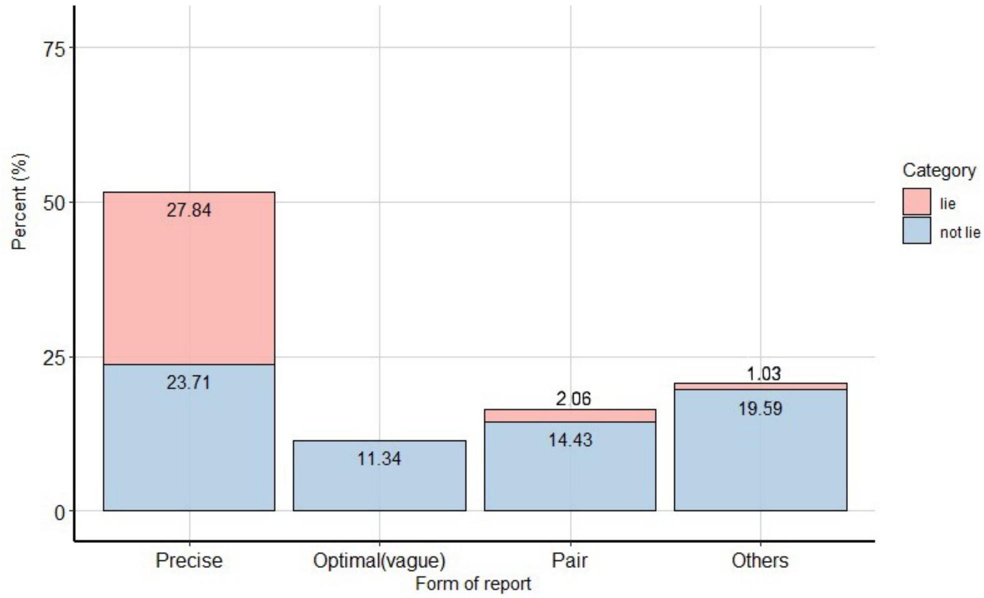

Figure 1: Message types used in A-UR treatments

were not lies. All the liars reported the maximum of 10. Among the 23 precise truth-tellers, 7 people observed 10 and reported so.

The model predicts that all truth-tellers 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 that the majority (55.2%) 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.

Comparing the behaviors across the A-UR and the NA-UR treatments, it is interesting to see that a higher percentage of participants used a vague message in the non-anonymous environment. Overall, 25.3% (20 of 79) of participants used a precise message, and 74.7% (59 of 79) used a vague message. This difference is possibly due to the maximal liars in the anonymous case who lie by reporting 10 without a social identity concern. Also, it is worth noting that the length of vague messages is larger

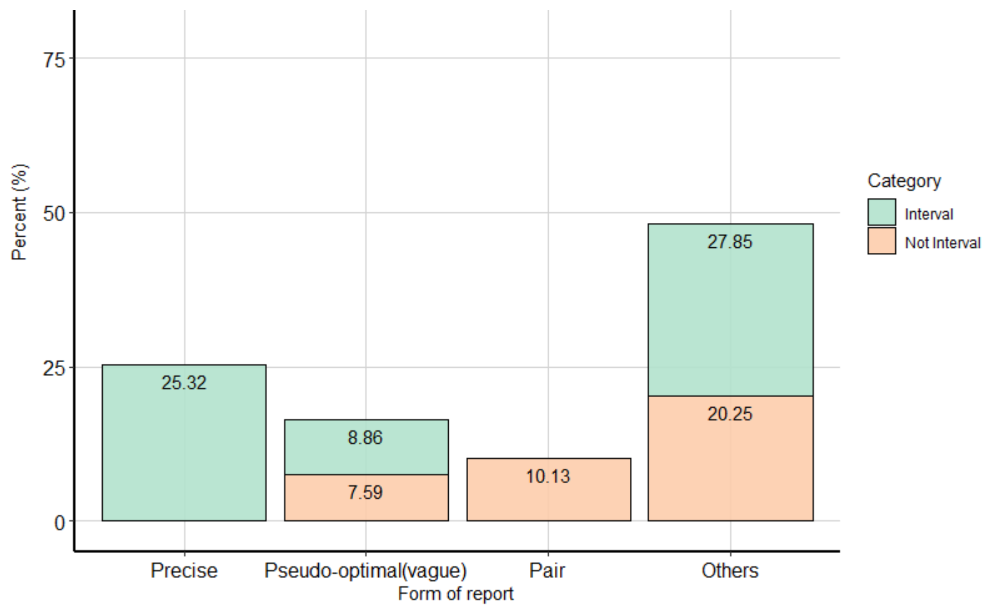

Figure 2: Message types used in NA-UR treatments

in the non-anonymous treatments (2.81 numbers reported on average) compared to the anonymous environment (2.10 numbers reported on average). The difference is significant at 1% significance level and possibly arises from the impact of vague messages on the external cost of lying. In order to further compare the effect of social identity on vague communication, we categorized the vague messages into the following three categories: 'Pseudo-optimal', 'Pair', and 'Others'. While we do not know the true observation of each participant in the non-anonymous treatments, we classified the optimal-looking messages using the minimum of the reported numbers in the message as a pseudo-true-observation. Figure 2 summarizes the types of messages used in NA-UR treatments.

Result 3-2. *In the NA-UR environment,*

1. *22.0% (13 of 59) and 13.6% (8 of 59) of vague messages are pseudo-optimal and pairs, respectively; and*

2. *49.2% (29 of 59) of vague messages take the form of an interval.*

Note that, compared to the anonymous treatment as displayed in Figure 1, participants used less of precise and pair types messages. Conditional on using vague messages, only 36.6% are either pseudo-optimal or pairs in the NA-UR treatment. This ratio differs starkly from the 53.2% in the A-UR treatment. In addition, we see the vast majority of the messages are now under the 'Others' category. Among the 'Others' messages, half of the messages take the form of an interval, which was not the case in the anonymous treatments. In particular, 49.1% of vague messages reported in the NA-UR treatment used intervals, while only 37.2% did in the A-UR. We can infer from the forms of the reported messages that participants understand that the pseudo-optimal and pair messages are obvious and can negatively affect the audience's belief in an equilibrium.

The distribution of the length of messages also highlights that there is a significant difference between the reporting pattern in the A-UR and the NA-UR treatments. Figure 3 reports the distribution of the length of the messages for the A-UR and NA-UR treatments. It also shows that subjects use vague and longer messages more often in the NA-UR treatment. Together with the frequent uses of interval messages in the non-anonymous treatment, we interpret this pattern as an indirect evidence of the impact of vagueness in the external cost of lying.

Result 4. *Comparison between A and NA environment*

1. *The average reports in A-R and NA-R are 7.619 and 7 respectively;*

2. *The average reports in A-UR and NA-UR are 8.132 and 8.232 respectively;*

We do not find statistically significant differences between the average reports of anonymous and non-anonymous treatments. Our results do not reject the null hypotheses both under restricted and unrestricted communication, with the p-values 0.185 and 0.562 respectively. While the result is against our prediction in Hypothesis 4, it is in line with Fries et al. (2021) in which the authors found a similarly perplexing outcome that increasing anonymity of one's report had no effect on average reports and lying behavior. One possible explanation is that the impact of anonymity on people's behavior primarily

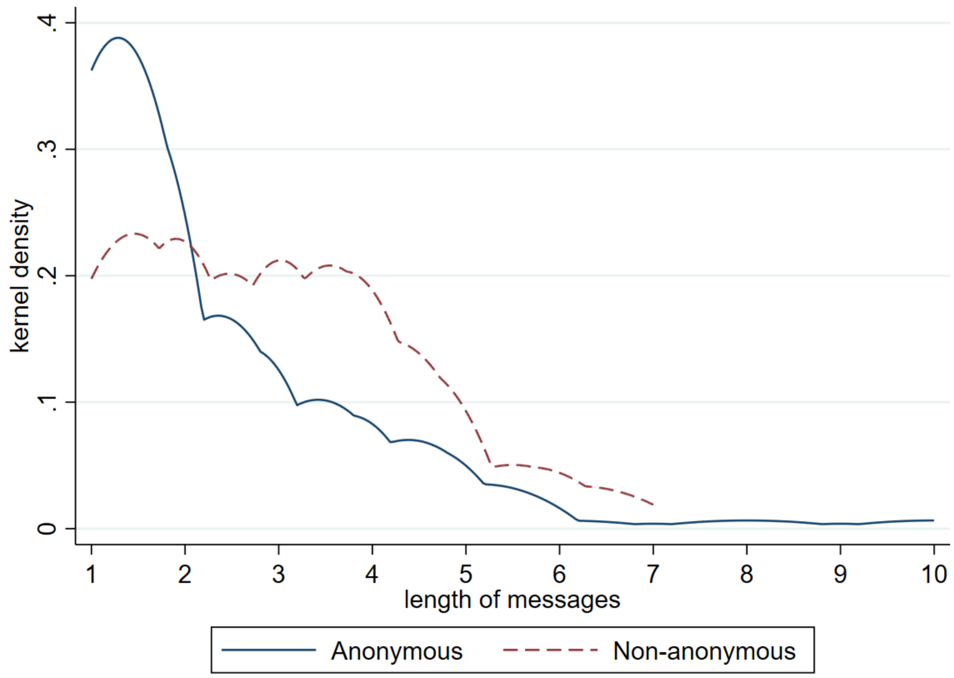

Figure 3: The distribution of the length of vague messages in A-UR and NA-UR treatments

creates only a minuscule difference when collapsed into an average and projected to a unidimensional space, and our sample is far from an infinite one to warrant such a statistical power. The difference we documented across the restricted and unrestricted treatments makes the explanation particularly plausible. Despite the similar magnitude in terms of the average reports, the types of the messages used in the two environments differ substantially in both the forms and the lengths of the messages.

7 Concluding Remarks

The extensive use of vague messages among our subjects, in both anonymous and non-anonymous sessions, provides us with insights to various communication context. We find that subjects exploit vagueness so as to be consistent with the truth, yet at the same time leveraging the imprecision to their own benefit. The prevalence of monetary-payoff maximizing messages in the anonymous sessions under both restricted and unrestricted communication support the key conjecture that a vague and truthful message reduces the internal cost of lying. In addition, we show subjects use vague and longer messages more often when the social identity concern is relevant. 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 seventy percent of the maximum amount available. However, as we allow the use of vague messages, the average report increases to 8.1, reducing the size of the aversion from thirty percent of the maximum monetary payoff amount to less twenty percent. 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 "moral wiggle room" in the sense of Dana et al. 2007 where dictators care only whether they maintain the image of their fairness, but not the consequential magnitude of their partners payoff. In our experiment, most agents exhibit no hesitation in increasing their monetary payoff as long as the message can remain even remotely truthful.

However, the existence of precise truth-tellers in the anonymous environment, though small in the size, suggests that the moral-wiggle-room argument alone may not paint 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.

References

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

Agranov, Marina, and Andrew Schotter. 2012. “Ignorance is bliss: an experimental study of the use of ambiguity and vagueness in the coordination games with asymmetric payoffs”. American Economic Journal: Microeconomics 4 (2): 77–103.

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

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.

Dana, Jason, et al. 2007. “Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness”. Economic Theory 33:67–80.

Deversi, Marvin, Alessandro Ispano, and Peter Schwardmann. 2021. “Spin doctors: An experiment on vague disclosure”. European Economic Review 139 (): 103872.

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.

Fries, Tilman, et al. 2021. “Observability and lying”. Journal of Economic Behavior & Organization 189:132–149.

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.

Shalvi, Shaul, et al. 2011. "Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior". Organizational Behavior and Human Decision Processes 115 (2): 181–190.

Sobel, Joel. 2020. "Lying and Deception in Games". Journal of Political Economy 128, no. 3 (): 907–947.

Wood, Daniel H. 2022. "Communication-Enhancing Vagueness". Games 13 (4): 49.

Zhang, Sookie Xue, and Ralph-Christopher Bayer. 2022. "Delegation based on cheap talk". Theory and Decision: 1–29.

A Proofs

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

Proof. 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 social identity 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 $0 < t < 1 - c(i,i+1)$.

□

Lemma 3 (KS PR 4; GKS PR 2)**.** *In any equilibrium in the NA-R environment, no agent underreports.*

Proof. Suppose there exists an equilibrium where an agent lies by reporting a number j below their 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. $\square$

Lemma 4 (KS THM 1; GKS PR 5). *In any equilibrium in the NA-R environment,*

i. there exists a threshold $1 < l^* < N$ *such that*

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

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

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

Proof. Let $L^\Omega_P \subseteq M^\Omega_P$ be the set of messages that liars use to lie with positive probability. Let $l^* = \min L^\Omega_P$. By Lemma 2, L^Ω_P 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 $L_P^\Omega = \{l^*, l^*+1, \dots, N\}$ by contradiction: suppose there exists some elements of Ω greater than l^* which is not an element of L_P^Ω. 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. $\square$

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 their behavior in the two environments easily leads to a conjecture that the absence of the social identity concern should only facilitate more lies. Applying these observation to the unrestricted communication leads to Proposition 1. We also obtain Lemma 1 from these three lemmas, which we will use to prove

Proposition 4.

Proof of Proposition 1.

Case 1. Anonymous environment: as shown in Proposition 3, all truth-tellers use the optimal vague message.

Case 2. Non-anonymous environment: 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 2 and 3 to argue that there exists a positive probability that agents lie upward. Let l^* be the threshold defined in Lemma 4.

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.

$\square$

Proof of Lemma 1. 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 social identity 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 4. 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 4. 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 1, $P(\text{lie}_{A-R}|\text{observe } i, t \in T_i) = 1$; 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. $\square$

B Gender Difference

B.1 Non-anonymous Environment

In the non-anonymous environment, 28 males and 51 females attended the experiments. We do not find statistically significant differences between males and females neither on their reports nor on the length of vague messages.

Under the unrestricted communication, male and female subjects are almost equally likely to send a precise message. Men are more likely to send a pseudo-optimal vague message, while women are slightly more likely to send a message with contains a pair or an interval.

Table 5: Gender difference on Reports in the Non-anonymous Environment

	Dependent variable:		
	Report_P	Exp_V	vague_size
	(1)	(2)	(3)
Gender difference	0.0553	−0.1587	0.2941
	(0.7291)	(0.2963)	(0.3582)
Constant	6.9804***	8.3379***	2.7059***
	(0.4341)	(0.1764)	(0.2133)
Observations	79	79	79
R^2	0.0001	0.0037	0.0087
Note:			*p<0.1; **p<0.05; ***p<0.01

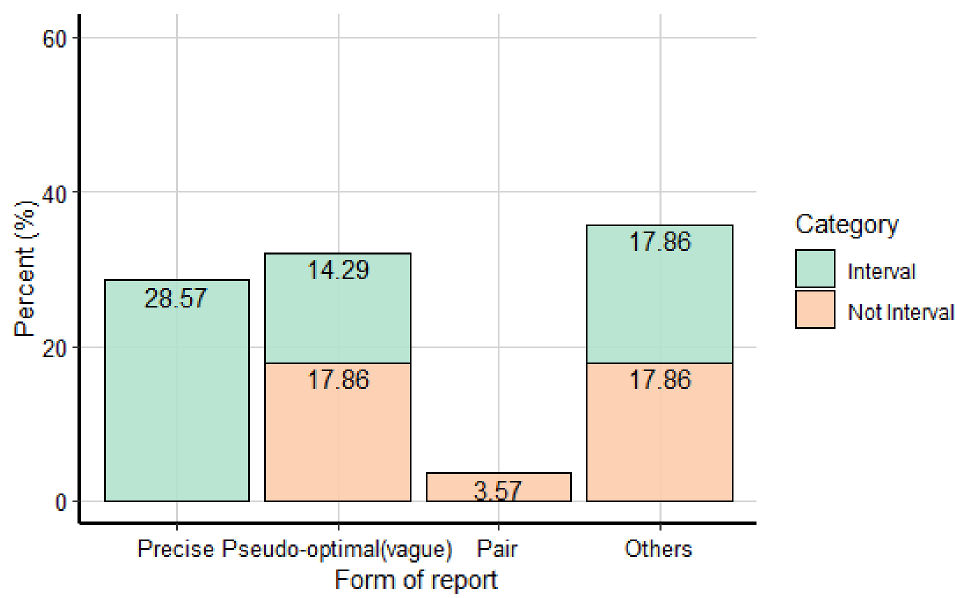

Figure 4: Message types male subjects used in NA-UR treatments

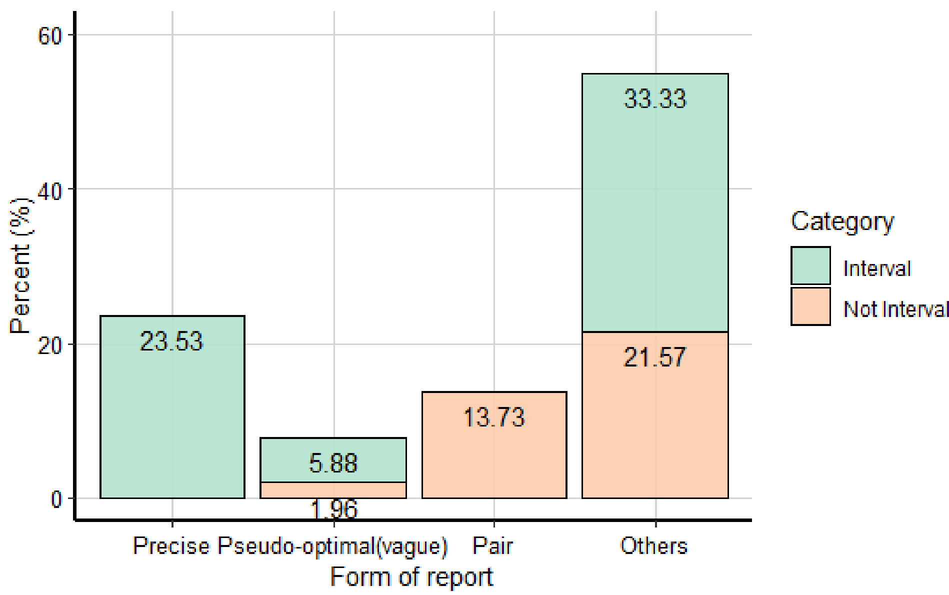

Figure 5: Message types female subjects used in NA-UR treatments

B.2 Anonymous Environment

We used only a subset of anonymous sessions to analyze the gender effect as we could not identify the participants' gender in principle. We ran four single sexed sessions (two male-only sessions and two female-only sessions) to create a gender dummy variable without interfering the anonymous property of the experiment design.

Under the restricted communication, we do not see statistically significant differences between males and females neither on their reports nor on the length of vague messages.

Table 6: Gender difference on Reports in the Anonymous Environment

	Dependent variable:		
	Report_P	Exp_V	vague_size
	(1)	(2)	(3)
Gender difference	0.1034	−0.6061	−0.4867
	(0.7107)	(0.5810)	(0.4461)
Constant	8.000***	8.3290***	2.3143***
	(0.4784)	(0.3911)	(0.3003)
Observations	64	64	64
R^2	0.0003	0.0172	0.0030
Note:			*p<0.1; **p<0.05; ***p<0.01

The results show that, 34.29 % of males and 31.03% of females lied in the non-anonymous treatment. Therefore, we do not find substantial difference in the lying behavior between them. In addition, we do not see significant differences in the choice of type and length of vague messages, as shown by Figures 6 and 7.

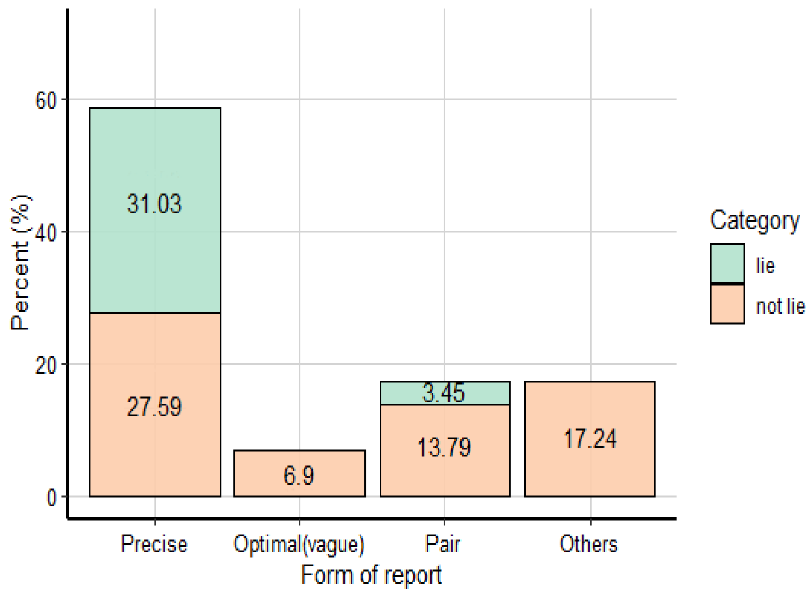

Figure 6: Message types male subjects used in A-UR treatments

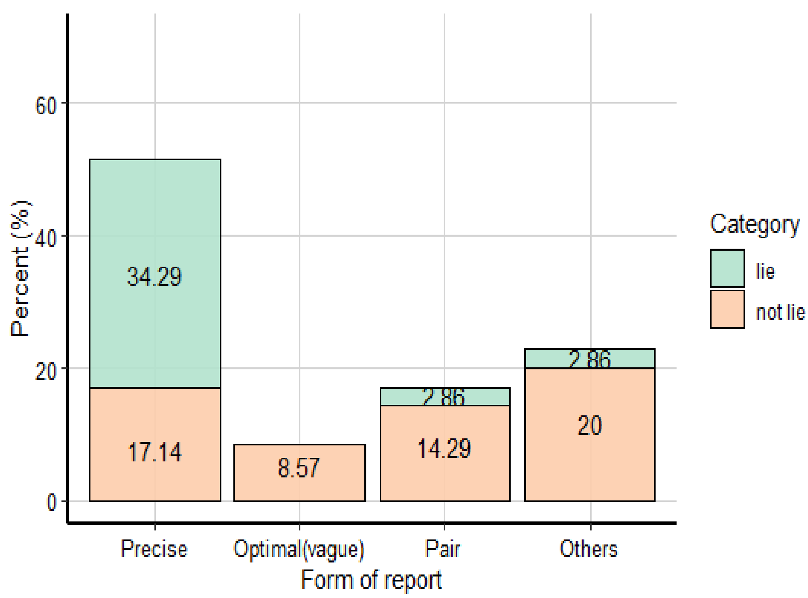

Figure 7: Message types female subjects used in A-UR treatments

C 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 their 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 their 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 8), which allows for precise messaging only, a subject may select only one box at a time, whereas in a vague stage (Figure 9), 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.

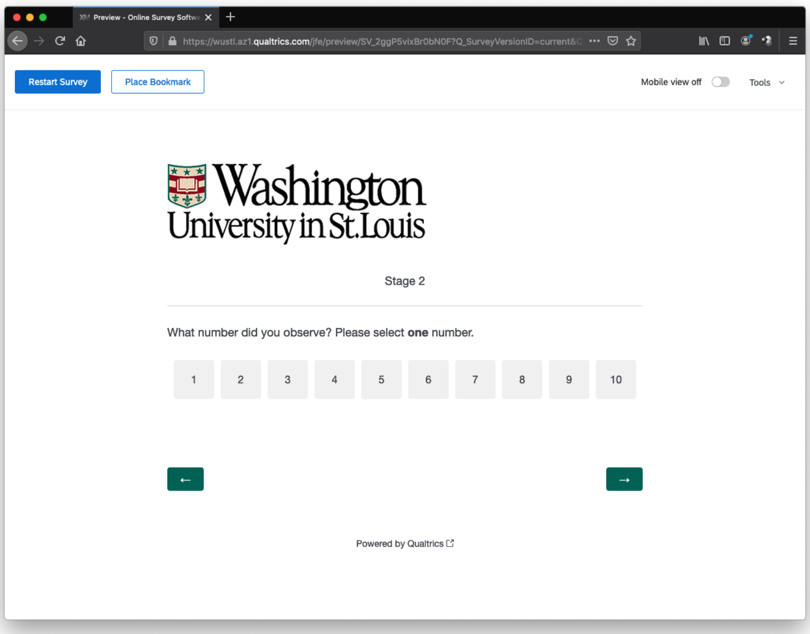

Figure 8: A screenshot of the experiment software displaying the precise-message stage

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

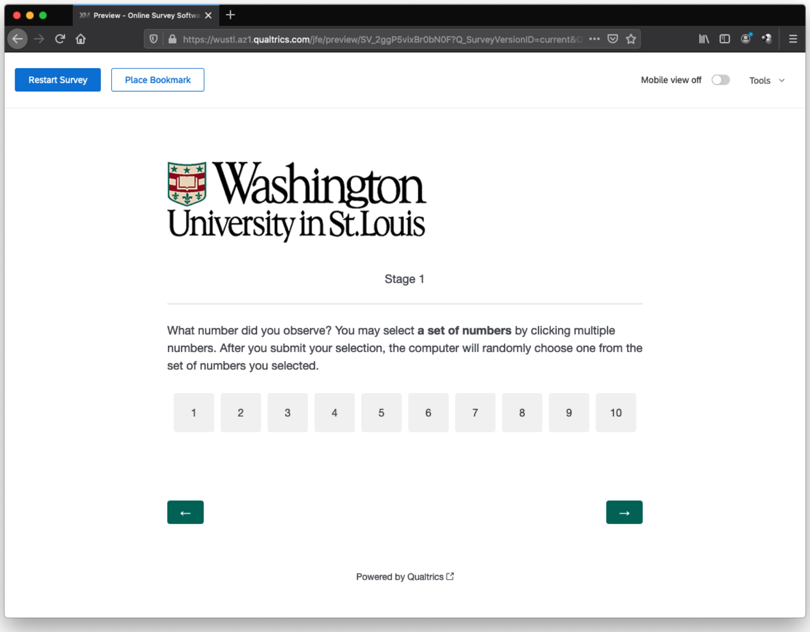

Figure 9: A screenshot of the experiment software displaying the vague-message stage

to receive the Amazon gift card directly emailed to them.

D Experimental Instructions

D.1 Instructions for Anonymous 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.

In the final review screen, you will find a link of a google form. To receive your amazon gift card you have to copy and paste the link in another browser and then fill your personal information. The google form is created by a staff of the department of economics, who won't have access to the data of this experiment.The experimenter only has a list of SCREEN NAMES and the amount associated with them. The experimenter will never be able to access your personal information. This is a screenshot of the google form which you will have to fill to receive your amazon gift card.

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.

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.

D.2 Instructions for Non-Anonymous session (Zoom session; video)

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.