

Pinocchio’s Clock: Using Response Time to Detect Deception in Sender-Receiver Games

Xin Jiang*, Bohan Ye†

September 21, 2026

Abstract

Timestamps that reveal response times (RTs) have become a common feature of contemporary social media platforms and messaging applications. What information does RT reveal when users have conflicting interests, and can they use this information to improve communication outcomes? We conduct a sender–receiver experiment to examine the relationship between RT and message honesty. We find that longer RTs are associated with a higher likelihood of deception, consistent with a cognitive-conflict mechanism under which lying involves greater cognitive struggle. Receivers, however, rely more consistently on senders’ message frequency than on RT when assessing message credibility.

Keywords: lie detection, response time, information acquisition, cognitive conflict

JEL: C92, D01, D91

*University of Southern California, Department of Economics. Email: xjiang.econ@gmail.com

†University of California, Riverside, Department of Economics. Email: bohan.ye@ucr.edu.

1 Introduction

With the widespread adoption of tools such as Slack, email, instant messaging, and social media platforms, response time (RT) is increasingly observable in online communication, yet its interpretation is ambiguous. A slow response may reflect greater effort devoted to acquiring accurate information, but it may also arise from additional deliberation associated with constructing a deceptive message. Moreover, once senders know that their RT is observable, the timing of a message may itself become a strategic choice. Thus, the growing availability of RT raises an important question: when truth discovery and strategic deliberation jointly determine response time, what information does RT convey about honesty, and can receivers make effective use of that information?

A common type of communication involving a conflict of interest in daily life arises when uncovering the truth requires effort from the party holding private information. From car mechanics to social media influencers, individuals often invest time in processing information before delivering a message to serve their own interests. After receiving a mechanic’s maintenance recommendation, the car owner must decide whether to follow it based on the perceived reliability of the recommendation. Did the car mechanic spend enough time investigating the issue? Or did he quickly jump to a costly diagnosis to exaggerate the problem and to make more money? When an influencer posts an eye-catching news item, has the influencer invested sufficient time and effort in verifying its accuracy? We investigate when RT can serve as an informative cue for deception. If variation in RT primarily reflects the effort required to discover the truth, longer RT may be associated with greater trustworthiness. In contrast, if variation primarily reflects post-discovery deliberation over whether to report truthfully, longer RT may instead be associated with deception when lying generates greater cognitive conflict. Importantly, the informativeness of RT may also depend on a sender’s broader communication pattern: when little within-sender variation in honesty exists, RT may contain limited additional information about the truthfulness of an individual message. The same observable cue can therefore have different meanings depending on the process that generates it.

We adopt a sender-receiver game with partially aligned interests, where the sender may acquire private information about the state, and sends a message to the receiver. The receiver, upon receiving the message, updates their beliefs about the true state and takes an action that determines both players’ payoffs. While the sender prefers the receiver to take an action that maximizes the sender’s payoff, the receiver benefits most from choosing the action that aligns with the true state, creating a strategic tension. Unlike conventional games where the sender automatically receives true information, our setting introduces a discovery

process in which the sender must decide whether to invest time in acquiring the truth. In this setting, RT captures both the time spent discovering the truth and the deliberation that precedes message choice. It may therefore reflect both strategic behavior and the cognitive conflict involved in lying. This interaction allows us to examine how RT relates to honesty and its underlying mechanism in communication.

We design a set of treatments to investigate private information and the strategic use of RT. To examine the relationship between RT and sender honesty, we manipulate senders' awareness of RT across two treatments: (1) Sender Unaware (SU), where senders have no incentive to manipulate their RT, and (2) Sender Aware (SA), where senders are informed that their RT will be visible to receivers, and thus have an incentive to manipulate their RT. The SU treatment captures the natural relationship between RT and truth, while the SA treatment allows us to examine whether senders strategically manipulate this relationship when aware of RT disclosure. To study how receivers use RT, we vary RT disclosure in two treatments: (1) Receiver Uninformed (RU), where receivers have no access to RT, and (2) Receiver Informed (RI), where RT is explicitly provided to them. Comparing receiver behavior across these treatments reveals how effectively they incorporate RT into their decision-making.

To vary the composition of RT between information acquisition and deliberation, we manipulate the effort required for truth discovery in the experimental design. In the Effortful experiment, senders need to click several buttons to discover the true state, whereas in the Effortless experiment, the truth is directly available without button clicking. This variation allows us to examine whether differences in the composition of RT affect its interpretation.¹

Our results show that deceptive y messages tend to take longer than honest y messages, consistent with a cognitive-conflict account in which deception requires additional deliberation. Consequently, longer RTs are associated with a higher likelihood of dishonesty and can therefore provide information about message honesty. This relationship is heterogeneous across sender communication patterns and is most pronounced among intermediate-frequency y senders. Importantly, the relationship is strongest when senders are unaware that their RT will be observed and becomes attenuated when RT observability is known. This pattern is consistent with senders adjusting their response timing when it can affect receivers' decisions. At the same time, making RT observable does not significantly change senders' overall honesty rates.

On the receiver side, we find limited evidence that RT is systematically incorporated into

¹<https://www.socialscisceregistry.org/trials/11373>

During the development of the paper, we refined the theoretical framework underlying RT patterns. The experimental design, treatment structure, and core outcome variables remain consistent with the preregistration.

trust decisions. Receivers respond much more consistently to the sender’s frequency of y messages than to variation in RT. Although RT affects trust in some treatment and receiver-type combinations, these effects are heterogeneous and do not form a general pattern of successful deception detection. The results indicate that making a potentially informative process cue observable does not imply that receivers will use it effectively.

Returning to the real-world example, the relationship between RT and honesty suggests that observing how long the mechanic takes to respond may provide some information about the process generating the recommendation, but RT is not a self-interpreting signal. Its meaning depends on whether time is spent acquiring information, deliberating over a deceptive report, or strategically adjusting one’s response when timing is observable. Moreover, our receiver results show that access to RT alone is insufficient for reliably incorporating this information into decisions. The value of RT as a behavioral cue therefore depends not only on whether it contains information, but also on whether the decision maker understands the mechanism that generates it.

In the remaining sections of the paper, Section 2 provides a comprehensive literature review. Section 3 presents the conceptual framework and develops the hypotheses. Section 4 outlines the experimental design. Section 5 presents the results. Section 6 concludes the paper.

2 Literature Review

This research contributes to two strands of literature on communication and information. The first strand of literature relates cues for detecting deception. Previous studies have explored two modes of deceptive communication: written messages and video clips. Regarding written messages, receivers are able to detect lies by analyzing the content and linguistic characteristics. Chen and Houser (2017) finds that promises serve as reliable cues, leading receivers to place greater trust in them, while other factors such as word usage and monetary references are ineffective. On the other hand, video clips provide a more comprehensive set of potential cues, encompassing not only language but also nonverbal elements such as gender, facial expressions, body movements, and hand gestures. Konrad, Lohse and Qari (2014), Dwenger and Lohse (2019), and Serra-Garcia and Gneezy (2021) reveal that receivers display only a slightly better than chance ability in detecting false messages of taxable income. Belot and Van de Ven (2017) observe that receivers in the role of buyers are able to detect the lies of sellers better than chance. Bonnefon, Hopfensitz and De Neys (2013, 2017) demonstrate that trustors in a trust game exhibit limited ability to detect truth based on trustees’ facial pictures. Todorov et al. (2015) also find that trustors easily form first

impressions from faces, although these impressions are unrelated to stable personality traits.

Our study contributes to this literature by examining response time as an indirect cue for deception in strategic communication. Unlike linguistic or visual cues, RT reflects the process through which a message is generated rather than the message content itself. This distinction is particularly relevant when communication involves both information acquisition and strategic deliberation, since the same observed response time may arise from different underlying processes. Our study therefore examines when RT can provide incremental information about message honesty and whether receivers can make effective use of this behavioral cue.

The second body of literature focuses on the value of RT in strategic settings. This emerging literature discusses three main questions when senders have private information: (1) whether their RT carries private information; (2) whether receivers can extract private information from senders' RTs; (3) whether senders manipulate their RT. Frydman and Krajbich (2022) investigate the value of RT in a laboratory social learning game, while Konovalov and Krajbich (2023) examine the value of RT in a laboratory bargaining game. Both studies find that the decision-making processes align with the DDM (Fudenberg, Strack and Strzalecki, 2018; Woodford, 2014), which states that the longer the RT, the smaller the difference between two choices. Cotet and Krajbich (2021) extend these findings to the eBay market with experienced agents. Their findings imply that senders' RT carries private information, even in the field. In addition, Frydman and Krajbich (2022) and Konovalov and Krajbich (2023) demonstrate that receivers can infer private information from senders' RTs. Furthermore, Konovalov and Krajbich (2023) find that individuals attempt to manipulate their RTs when they are aware that their times would be revealed, making it less informative to their counterparts.

Our study extends the investigation of the informational value of RT in the context of lying and truth-telling. Jiang (2013) finds that low-payoff messages take longer to generate than high-payoff messages. Similarly, Shalvi, Eldar and Bereby-Meyer (2012) report that participants sent more high-payoff messages under high-time-pressure conditions than under low-time-pressure conditions. These findings suggest that resisting temptation—i.e., telling the truth when it is costly—requires more deliberation. However, prior work has not distinguished whether high-payoff messages result from honest reporting in an advantageous state or from deception in a disadvantageous state, even though these cases may involve different underlying decision processes.

Our experiment addresses this gap by recording the truth underlying each decision, allowing us to distinguish honest from deceptive high-payoff messages. This design enables us to examine whether RT contains incremental information about message honesty once the underlying state is observed. We further study whether receivers can extract such in-

formation from RT and whether senders strategically adjust their response timing when RT becomes observable.

3 Framework

People generally value seeing themselves as honest, but dishonesty often yields substantial material benefits. Mazar, Amir and Ariely (2008) show that people frequently behave dishonestly enough to profit while maintaining a positive moral self-image. In strategic communication settings, this raises a central question: can RT reveal whether a message is truthful?

We study this question in a sender–receiver environment in which the sender may acquire private information before communicating with the receiver. The state of the world, $\omega \in \{H, L\}$, determines the payoffs of both players, but is initially unknown to the receiver. The sender can choose whether to acquire information about the realized state before sending a message to the receiver.

After the information-acquisition stage, the sender sends one of two possible messages, $m \in \{H, L\}$. Sending $m = H$ is favorable to the sender because it increases the likelihood that the receiver chooses the action associated with a high payoff for the sender. By contrast, sending $m = L$ is less favorable to the sender and is associated with a lower payoff if followed by the receiver.

Upon observing the sender’s message, the receiver chooses an action $a \in \{H, L\}$. The receiver’s payoff depends on whether her action matches the true state of the world. Specifically, the receiver earns a high payoff when $a = \omega$ and a low payoff otherwise. Because the sender benefits when the receiver chooses $a = H$ regardless of the true state, the sender and the receiver have partially aligned but not identical interests, creating incentives for strategic communication.

Because message $m = L$ is associated with the sender’s lowest reward if followed by the receiver, such messages provide relatively little scope for strategic deception. Strategic deception in this environment is therefore most relevant when the sender sends the favorable message $m = H$. For this reason, our analysis focuses on the informativeness of RT conditional on sending $m = H$. However, existing theories generate competing predictions regarding the relationship between RT and honesty.

One possibility is that longer RTs indicate a higher likelihood of honesty. This prediction follows from the moral wiggle-room literature, which shows that individuals may strategically avoid information in order to behave selfishly while preserving a favorable self-image (Dana, Weber and Kuang, 2007; Grossman, 2014). When acquiring information is time-consuming,

short RTs are more likely to reflect deliberate ignorance, whereas long RTs are more likely to reflect information acquisition before communicating.

Under this mechanism, longer RTs are associated with a higher probability that the sender’s message matches the realized state, because truth-telling requires time for information acquisition.

Hypothesis 1. *(Information-Acquisition Mechanism): Conditional on sending $m = H$, longer RTs are associated with a higher likelihood of honesty.*

An opposing prediction arises when senders already know the true state before choosing their message. In this case, deceptive communication requires knowingly overriding a preference for honesty in pursuit of material gain. Prior research suggests that dishonest behavior often involves cognitive conflict between monetary incentives and moral concerns (Greene and Paxton, 2009; Jiang, 2013; Shalvi, Eldar and Bereby-Meyer, 2012). As a consequence, deceptive choices may require more deliberation and therefore longer RTs.

This prediction is also consistent with evidence from decision-making models showing that decisions are slower when the utility difference between alternatives is small (Cotet and Krajbich, 2021; Konovalov and Krajbich, 2019, 2023; Krajbich, 2025; Rubinstein, 2016). If individuals possess a preference for truth-telling (Abeler, Nosenzo and Raymond, 2019), then truthful favorable messages $m = H$ combine both material and moral benefits, producing a relatively clear choice and short RTs. By contrast, deceptive messages $m = H$ involve a tradeoff between monetary gain and honesty concerns, generating greater decision conflict and longer RTs.

Under this mechanism, longer RTs indicate a lower likelihood of honesty.

Hypothesis 2. *(Cognitive-Conflict Mechanism): Conditional on sending $m = H$, longer RTs are associated with a lower likelihood of honesty.*

A growing literature demonstrates that observers can extract private information from RTs. Frydman and Krajbich (2022) show that subjects infer whether another player’s private signal conflicts with prevailing behavior from observed RTs. Similarly, Konovalov and Krajbich (2023) show that bargaining participants infer counterpart valuations from rejection RTs.

These findings suggest that RT may function as an informative behavioral cue in strategic communication environments.

Hypothesis 3. *Receivers infer senders’ private information from observed RTs.*

If senders understand that RT conveys information, they may strategically manipulate their timing behavior. Existing evidence suggests that such manipulation depends on the

strength of strategic incentives. Frydman and Krajbich (2022) find limited evidence of RT manipulation in environments with relatively weak conflicts of interest, whereas Konovalov and Krajbich (2023) show that subjects strategically alter RT when timing directly affects economic outcomes.

Given the conflict of interest between senders and receivers, if receivers infer private information from RT as predicted in Hypothesis 3, observable RT creates incentives for senders to strategically adjust their decision speed to make RT less informative.

Hypothesis 4. *When RT is observable, senders strategically adjust RT to reduce its informativeness.*

The competing hypotheses above impose several requirements on the experimental design.

First, our experiment requires information acquisition to be time-consuming so that RT can partly reflect the process of acquiring information rather than only the deliberation involved in deciding which message to send. This feature is particularly important because the information-acquisition mechanism predicts an opposite direction from the cognitive-conflict mechanism arising during deliberation in the relationship between RT and honesty. The central question is therefore how RT should be interpreted when both mechanisms are present. Although no real-world example fits a laboratory setting exactly, many communication decisions involve both costly information acquisition and strategic or moral conflict. For example, an online influencer deciding whether to share a sensational news item may need time to verify its accuracy, while also facing incentives to post quickly or to post attention-grabbing content. Similar tensions may arise when a salesperson decides whether to disclose product limitations, or when an employee decides how much to verify information before reporting it to a supervisor. By allowing both mechanisms to operate in the same setting, our design examines which mechanism better accounts for the observed relationship between RT and honesty.

In addition, the duration of information acquisition should be flexible so that senders retain substantial control over their observed RT. This feature is necessary for studying whether senders strategically manipulate RT when they know it is observable to receivers. If information acquisition mechanically required a fixed amount of time—for example, exactly 10 seconds to learn the truth—then the sender could influence RT only by choosing whether to acquire information at all. Such a design would substantially restrict the sender’s strategic options and would not allow us to study how senders manipulate RT through the information-acquisition process itself.

Second, receivers must be able to compare RTs within sender and within context. Raw RTs differ substantially across individuals for reasons unrelated to honesty, including reading

speed, motor speed, attention, and general decision style. For this reason, the design allows receivers to observe multiple decisions made by the same sender in comparable situations. This makes it possible for receivers to interpret a sender’s RT relative to that sender’s own behavioral baseline, rather than by comparing raw RTs across heterogeneous individuals.

One possible concern is that observing several decisions from the same sender may itself affect receiver behavior. We view this feature as a necessary part of the design rather than a confound. In many strategic communication settings, receivers rarely evaluate a message in isolation from the sender’s past behavior. Followers often observe an influencer’s posting history before deciding whether to trust a new post; buyers may rely on a salesperson’s prior claims, customer reviews, or reputation before deciding how much weight to place on a recommendation. Importantly, observing multiple decisions by the same sender does not, by itself, imply how RT should be interpreted. Thus, providing sender-specific message histories does not mechanically impose an interpretation of RT. Instead, it allows us to study whether and how receivers use RT when it becomes available.

Third, receivers must have sufficient experience to understand the potential relationship between RT and honesty. To form meaningful interpretations of RT, receivers should first experience the sender’s role. This experience provides direct knowledge of the information-acquisition process and the incentives faced by senders when deciding what message to send. As a result, receiver behavior can be interpreted as informed inference from RT rather than as confusion or speculation about the structure of the task.

The next section presents the detailed experimental design that incorporates these necessary features.

4 Experimental Design

In this section, we describe the laboratory implementation of the sender-receiver game and the experimental conditions. The design was constructed to incorporate the key features motivated by the framework: time-consuming and flexible information acquisition, within-sender RT comparison, and receiver experience with the sender environment before making inference decisions.

4.1 Game

Each subject participated in both the sender and receiver roles. Subjects first made decisions as senders and were later matched with another sender’s messages to make inference decisions as receivers. This procedure ensured that receivers had direct experience with the

information-acquisition process, the message-sending decision, and the incentives faced by senders before interpreting RT.

The true state of the world was determined by the outcomes of six independently rolled fair dice. For each die, outcomes 4, 5, and 6 were defined as “large numbers,” while outcomes 1, 2, and 3 were defined as “small numbers.” Subjects faced the question of whether at least four of the six dice outcomes were large numbers. The state was “yes” if four or more dice showed large numbers and “no” otherwise.

Consistent with the framework, we map the “yes” state to H and the “no” state to L . To match the experimental interface, we use the notation $\omega \in \{yes, no\}$ for states, $m \in \{y, n\}$ for sender messages, and $a \in \{YES, NO\}$ for receiver actions. Given this setting, the prior probability of the “yes” state is $\mu_{yes} = 34.37\%$. This implementation keeps the experimental task intuitive for subjects while remaining consistent with previous sender–receiver studies.

In the sender role, subjects completed 10 independent rounds. At the beginning of each round, Nature first drew the six dice outcomes, thereby determining the true state of the world. The sender then chose whether to acquire information about the realized state before sending a message. Subjects were required to send a message regardless of whether information was acquired. State realizations were independently drawn across rounds and subjects.

In the receiver role, subjects observed all 10 messages sent by the same non-self sender and made a state guess for each message. Receivers did not observe the true states associated with the messages sent by their matched sender. Their own previous state realizations were independently determined and therefore provided no information about the sender’s states.

The receiver’s decision, together with the true state, determined the payoffs of both players as shown in Table 1. The receiver earned \$8 if her action matched the true state and \$4 otherwise. The sender earned \$8 if the receiver chose *YES*, regardless of the true state, and \$4 otherwise.

Table 1: Monetary Payoff.

Truth/Receiver’s choice	<i>YES</i>	<i>NO</i>
<i>yes</i>	(\$8, \$8)	(\$4, \$4)
<i>no</i>	(\$8, \$4)	(\$4, \$8)

Notes: The first item denotes the sender’s payoff, the second item denotes the receiver’s payoff.

4.2 Experiments, Treatments, and Conditions

The baseline condition is Effortful Sender Unaware–Receiver Uninformed (E-SU-RU). This condition implements the core environment of the framework: information acquisition is

time-consuming and flexible, but RT is not made salient to either player. RT was measured as the elapsed time between the moment when the sender had the opportunity to learn the truth and the moment he actually chose a message.

In the Effortful acquisition (E) experiment, senders faced six buttons, each corresponding to one die. Clicking a button revealed the outcome of that die, and senders could choose to click anywhere from zero to all six buttons before sending a message. This feature makes information acquisition time-consuming while preserving sender control over observed RT.² Consequently, the true state could become fully determined after as few as three revealed outcomes, while senders could continue revealing up to all six outcomes.

In the Sender Unaware (SU) treatment, senders were not informed that their RT would later be shown to receivers.³ In the Receiver Uninformed (RU) treatment, receivers observed only the sender’s messages, not RT. This baseline allows us to examine sender behavior when information acquisition is effortful but RT is not explicitly part of the strategic environment.

Starting from the baseline, we introduce three treatment variations.

First, we varied whether receivers observed RT. In the Receiver Informed (RI) treatment, receivers observed both the sender’s message and the RT associated with each message. Comparing RI with RU allows us to study whether receivers use RT to infer the sender’s private information and the honesty of the message.

To facilitate within-sender RT comparisons, receivers observed all 10 messages sent by the same sender. They first made state guesses based only on the messages. After RT information was revealed, their previous guesses were erased and they were asked to make a new set of guesses for the same messages. This procedure avoids mechanically framing the second decision as a revision of the first and allows us to identify the effect of RT on receiver inference.

Second, we varied whether senders knew that RT would be observed. In the Sender Aware (SA) treatment, senders were explicitly informed that their RT would be recorded and shown to receivers. Comparing SA with SU allows us to study whether senders strategically manipulate RT when they know that it may affect receiver beliefs.

Third, in the extended experiment, we varied the effort required for information acquisi-

²Although the button-clicking procedure may introduce some additional cognitive load beyond mechanical search time, the resulting variation in such load appears limited in our data. Once participants acquired sufficient information to determine the state, they revealed all six dice outcomes in more than 90% of fully informed observations. We therefore interpret the Effortful treatment as changing the composition of observed RT primarily by making information acquisition time-consuming, rather than as a clean manipulation of mechanical time alone.

³Not informing participants of the future use of RT data should not be considered deceptive, as no false information was provided. In addition, the consent form and recruitment materials clearly stated the duration of the experiment, making subjects aware that timing information could be recorded.

Table 2: Conditions and Observations.

Experiment	Condition	Information Acquisition	Sender	Receiver	Design	Observations
Main	Baseline (E-SU-RU)	Effortful	Unaware	Uninformed	Within subject	30×10
	Condition 1 (E-SU-RI)	Effortful	Unaware	Informed		30×10
	Condition 2 (E-SA-RU)	Effortful	Aware	Uninformed	Within Subject	32×10
	Condition 3 (E-SA-RI)	Effortful	Aware	Informed		32×10
Extended	Condition 4 (L-SU-RU)	Effortless	Unaware	Uninformed	Within subject	42×10
	Condition 5 (L-SU-RI)	Effortless	Unaware	Informed		42×10

tion. In the Effortless acquisition (L) experiment, all six dice outcomes were automatically displayed to the sender without requiring any clicking. Compared with the E experiment, RT in the L treatment primarily reflects deliberation rather than the time spent acquiring information. The extended experiment therefore provides a useful comparison for understanding how the composition of RT affects receiver inference. Because the purpose of this extension was to examine how removing effortful information acquisition changes the interpretation of RT, rather than to provide an additional test of sender manipulation, we implemented the Effortless experiment only under SU.

We implemented six conditions, summarized in Table 2, as combinations of three treatment variations: sender awareness of RT (SU vs. SA), receiver access to RT (RU vs. RI), and the information-acquisition process (E vs. L). The four conditions in the main experiment (E-SU-RU, E-SU-RI, E-SA-RU, and E-SA-RI) were pre-registered and designed to test the hypotheses developed in the framework. The two conditions in the extended experiment (L-SU-RU and L-SU-RI) were added after the main experiment to provide additional evidence on how receivers interpret RT when information acquisition is effortless and RT primarily reflects deliberation rather than information acquisition.

4.3 Implementation

Comprehension questions were included in each stage to ensure that subjects understood the instructions. Participants could proceed to the experiment only after answering these questions correctly. Each subject completed 10 rounds as a sender and subsequently made 20 receiver decisions: 10 without RT information and 10 with RT information for the same

set of messages. Participants were paid for one randomly selected round in each role. With an effect size of 0.3, an error probability of 0.1, and a power of 0.8, the power calculation indicated a target of 32 observations for each condition.

The experiment was conducted at the Experimental and Behavioral Economics Laboratory at the University of California, Santa Barbara, from May 11, 2023 to May 29, 2024. We recruited 104 undergraduate participants from the UCSB subject pool, with more than 30 participants for each condition, as summarized in Table 2. Participants were required to be at least 18 years old; exact age was not collected. Of the 104 participants, 48 identified as male, 52 as female, and 4 selected “other.” Participants earned an average of \$12 for approximately one hour in the laboratory. The instructions and interfaces can be found in Appendix A.

5 Results

In this section, we begin by presenting the results from Effortful Experiment, which examines informational value of RT in a sender-receiver game involving a time-consuming truth-discovery process. We then discuss the findings from Effortless Experiment, which investigates whether receivers’ interpretation of RT changes in the absence of the mechanical truth-discovery process.

5.1 Results in Effortful Experiment

The main finding of the Effortful Experiment is that, among senders who send intermediate-frequency y messages, faster responses are associated with a higher likelihood of truth, while receivers do not make significant use of this information. We begin by examining senders’ message choices and response times, with particular attention to the relationship between RT and honesty. We then investigate the behavior of receivers. Throughout our analyses, we also examine heterogeneity by sender and receiver type.

5.1.1 Senders’ Message and RT

We begin by describing the composition and RT of senders’ messages. Table 3 reports the number and RT of messages by information-acquisition status and message type. To better describe the results, a message is classified as *fully informed* when the sender revealed enough dice outcomes to determine the state uniquely; *partially informed* when the sender revealed at least one outcome but could not yet determine the state uniquely; and *uninformed* when the sender sent a message without revealing any outcome.

Table 3: Descriptive Statistics of Senders' Messages and Response Times: Effortful Experiment

Treatment	Classification	<i>N</i>	Percent	Mean RT	SD	Median RT [Q1, Q3]
<i>Panel A: Information-acquisition status</i>						
SU	Fully informed	276	92.0%	12.96	9.51	10.00 [8.00, 14.00]
SU	Partially informed	1	0.3%	7.00	—	7.00 [—]
SU	Uninformed	23	7.7%	8.91	11.71	2.00 [0.50, 14.00]
SA	Fully informed	295	92.2%	12.11	7.53	10.00 [8.00, 13.00]
SA	Partially informed	8	2.5%	8.25	5.09	6.50 [4.75, 9.50]
SA	Uninformed	17	5.3%	6.18	6.92	1.00 [1.00, 14.00]
<i>Panel B: Message type</i>						
SU	Honest <i>y</i>	86	28.7%	12.10	9.85	10.00 [9.00, 12.00]
SU	Honest <i>n</i>	138	46.0%	13.48	10.18	10.00 [8.00, 14.00]
SU	Deceptive <i>y</i>	51	17.0%	13.10	6.82	10.00 [9.00, 15.00]
SU	Deceptive <i>n</i>	1	0.3%	9.00	—	9.00 [—]
SU	Not fully informed <i>y</i>	24	8.0%	8.83	11.46	4.50 [0.75, 13.00]
SA	Honest <i>y</i>	87	27.2%	10.85	6.92	10.00 [7.00, 12.00]
SA	Honest <i>n</i>	160	50.0%	12.62	8.32	10.00 [8.00, 13.25]
SA	Deceptive <i>y</i>	45	14.1%	12.67	5.54	12.00 [10.00, 14.00]
SA	Deceptive <i>n</i>	3	0.9%	13.33	2.31	12.00 [12.00, 14.00]
SA	Not fully informed <i>y</i>	25	7.8%	6.84	6.36	5.00 [1.00, 14.00]

Notes: The unit of observation is a sender-round decision. The SU treatment contains 300 observations from 30 senders, and the SA treatment contains 320 observations from 32 senders. A message is classified as fully informed when the sender revealed enough dice outcomes to determine the state uniquely; partially informed when the sender revealed at least one outcome but could not yet determine the state uniquely; and uninformed when the sender sent a message without revealing any outcome. An honest message is a fully informed message consistent with the true state, whereas a deceptive message is a fully informed message inconsistent with the true state. Not fully informed messages include partially informed and uninformed messages. RT is measured in seconds and rounded to the nearest integer at the observation level. Q1 and Q3 denote the first and third quartiles, respectively. Percentages are calculated within condition.

Among fully informed messages, we further define an *honest* message as a message that aligns with the true state of the world, and a *deceptive* message as a message that does not. By contrast, *not fully informed* messages include partially informed and uninformed messages.

Of the 620 sender-round observations in the Effortful Experiment, 571 (92.1%) were sent after the sender had acquired sufficient information to determine the true state. The fully informed rate was nearly identical in the SU and SA treatments (92.0% and 92.2%, respectively), and the difference is not statistically significant (two-sided Fisher’s exact test, $p = 1.000$). This discovery rate is significantly higher than the approximately 55% rates of acquiring information about the other participants’ payoff in Dana, Weber and Kuang (2007) and Grossman (2014) for dictator games. Subjects’ demand for information may be influenced by the game structure. In the dictator game, the dictator directly determines the outcome without sending a message and prefers to be uninformed to maintain innocence regarding potential harm to others, thereby avoiding information. In contrast, in the sender-receiver game, the sender must interact with the receiver by sending a message. The sender may prefer to be informed, generally favoring honesty to achieve their goal and resorting to lying only when absolutely necessary.

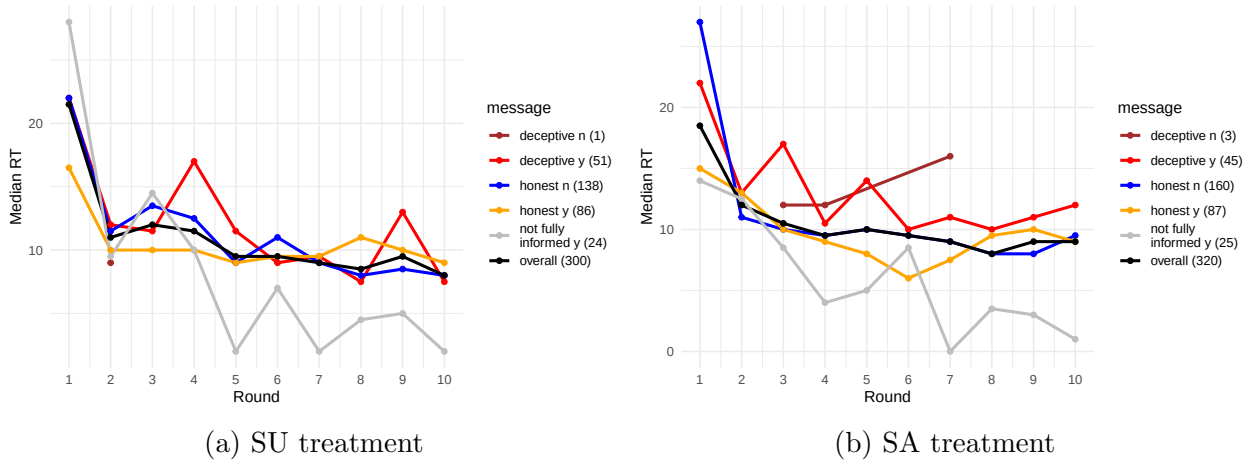
Such preference for certain honesty is different from preference for truth-telling (Abeler, Nosenzo and Raymond, 2019). While preference for truth-telling pertains to an individual’s inclination to be honest once aware of the facts, preference for certain honesty signifies a proactive pursuit of honesty when it serves one’s goals. In our setup, fabricating message y without being informed is unnecessary, as honesty may yield the same payoff. Lying becomes necessary only when the sender discovers a disadvantageous state ($\omega = no$) and seeks to exploit the receiver’s trust for a higher payoff. Consequently, preference for certain honesty may make truth discovery a dominant strategy, motivating the sender to always acquire the truth.

High rates of full information acquisition suggest that the observed variation in RT primarily reflects deliberation after information acquisition, rather than mechanical differences in the acquisition process. Nevertheless, one might still be concerned that the differences in RT are driven by mechanical features of the task, such as the number of button clicks. Our data help alleviate this concern. Among the fully informed messages, participants click all six buttons in more than 90% of cases, leaving relatively little variation in mechanical search effort across observations. This pattern suggests that differences in the number of button clicks are unlikely to be the primary source of the observed RT variation. Taken together, these patterns support our interpretation of RT as primarily capturing post-information deliberation in our setting.

Result 1: The observed response-time variation is primarily associated with deliberation after information acquisition.

This result is consistent with the behavioral literature on RT, which treats decision time as informative about the cognitive process underlying choice. In our setting, this is particularly relevant for y messages. Given the near-perfect honesty rate (99% in the SU treatment, 98% in the SA treatment) of n messages, there is little need to seek an additional cue to assess their accuracy. In contrast, for y messages, where the honesty rate is only 54% (53% in the SU treatment, 55% in the SA treatment), RT is valuable if it correlates with honesty and helps detect deception. We therefore examine whether RT provides additional information about honesty, particularly for y messages, where message content alone is less informative.

Figure 1: Median RT by Message and Round

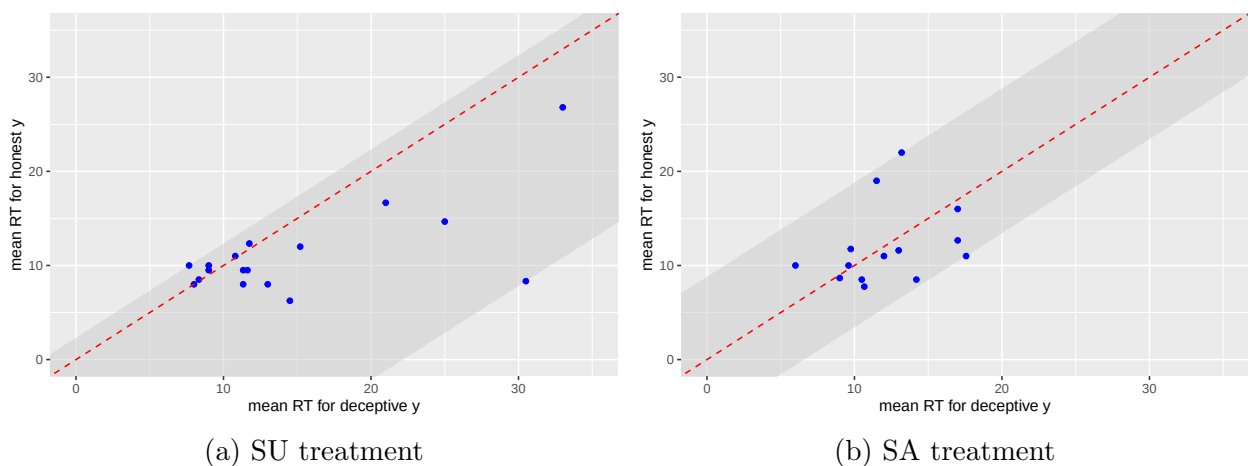


Notes: Each point represents the median RT from all senders for a given message type in a specific round. The number in parentheses indicates the frequency of each message type.

We first examine whether the overall relationship between RT and message honesty is more consistent with Hypothesis 1 or Hypothesis 2. We then examine whether this relationship also holds within individual senders. Figure 1 shows the median population RT by message type and round in both treatments. Overall, the median population RT is highest in round 1 and decreases over subsequent rounds. This pattern suggests that senders spend some time familiarizing themselves with the task in the initial round and then focus on information acquisition and deliberation in later rounds. Not fully informed messages tend to have shorter RTs than the overall median after the initial rounds. The RT for honest n messages is close to the overall median, while deceptive y messages generally take longer than

honest y messages. Because Hypotheses 1 and 2 make opposite predictions about the relationship between RT and honesty, the comparison between honest and deceptive y messages is particularly informative. The aggregate pattern—longer RTs for deceptive y messages than for honest y messages—is consistent with the cognitive-conflict mechanism in Hypothesis 2 and runs against the information-acquisition mechanism in Hypothesis 1. However, this aggregate comparison may partly reflect heterogeneity across senders. To examine whether the same relationship holds within individuals, we focus on senders who sent both honest and deceptive y messages and compare their RTs across the two types of decisions.

Figure 2: The Relationship Between RT and the Type of y Message



Notes: Each data point represents an individual sender who sent both deceptive y and honest y messages. The x-axis represents his mean RT for deceptive y messages, while the y-axis represents their mean RT for honest y messages. The red dashed line represents the 45-degree line. The shaded area represents the range of observed variation around the 45-degree line ($y = x$). The left panel contains 17 data points, and the right panel contains 14.

Figure 2 illustrates the relationship between individual RTs for honest y and deceptive y messages among senders who sent both types of messages. Each data point represents the mean RTs of an individual sender. Points falling on the 45-degree line indicate no difference in RT between the two message types. The left panel contains 17 senders in the SU treatment, while the right panel contains 14 senders in the SA treatment. The majority of data points are located in the lower triangle, implying that, for the same sender, deceptive y messages take longer to send than honest y messages. This within-sender pattern reinforces the aggregate evidence in favor of Hypothesis 2: longer RTs are associated with a lower likelihood of honesty. It also reduces the concern that the aggregate relationship is driven primarily by persistent differences in response speed across individuals.

We then examine whether the difference in RT between honest and deceptive y messages is statistically significant. As shown in Figure 1, RTs in the first round are generally longer

than those in subsequent rounds, likely because participants are still learning the interface and becoming familiar with the task. Including this round could therefore bias the estimated relationship between message type and RT. We accordingly exclude Round 1 from the regression analysis.

We classify senders who send honest messages in all 10 rounds as *honest senders* and those who send at least one deceptive or not fully informed message as *dishonest senders*. Table 4 reports OLS regressions of RT on message type and round for rounds 2–10, with honest y as the omitted message category. Columns 1 and 2 report pooled estimates for the SU and SA treatments, respectively; Columns 3 and 4 add sender fixed effects; and Columns 5 and 6 restrict the sample to dishonest senders. Standard errors are clustered at the sender level.

Table 4: Response-Time Regressions

	<i>Dependent variable: RT</i>					
	all senders		all senders		dishonest senders	
	SU (1)	SA (2)	SU (3)	SA (4)	SU (5)	SA (6)
deceptive n	−5.469*** (1.855)	2.937 (2.169)	−3.713*** (1.158)	2.068 (3.972)	−3.889** (1.533)	1.643 (4.520)
deceptive y	1.059 (1.623)	2.197*** (0.663)	3.428** (1.432)	1.330* (0.737)	3.274** (1.578)	1.406 (0.852)
honest n	1.007 (1.239)	1.419** (0.673)	0.985 (0.991)	1.726** (0.791)	0.664 (1.695)	1.635 (1.254)
uninformed y	−4.024 (4.746)	−3.583 (3.081)	−4.411* (2.627)	−4.048** (2.042)	−4.550* (2.731)	−4.040* (2.132)
round	−0.802*** (0.227)	−0.590*** (0.131)	−0.804*** (0.239)	−0.584*** (0.140)	−0.804*** (0.292)	−0.804*** (0.197)
Constant	16.074*** (2.237)	13.149*** (1.038)	15.060*** (1.506)	11.618*** (0.913)	24.927*** (2.735)	12.980*** (1.370)
Fixed Individual			✓	✓	✓	✓
Observations	270	288	270	288	171	171
R ²	0.076	0.140	0.326	0.426	0.304	0.438

Notes: OLS estimates are reported for rounds 2–10. Standard errors clustered at the sender level are reported in parentheses. Sender fixed effects are included where indicated. A dishonest sender is defined as a sender who sent at least one deceptive or not fully informed message during the experiment. The omitted message category is honest y . All tests are two-sided. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

In the SU treatment, the pooled estimate in Column 1 does not show a statistically significant difference between deceptive and honest y messages. Once sender fixed effects are included, however, deceptive y messages take 3.428 seconds longer than honest y messages (Column 3; $t = 2.39$, $p < 0.05$, $N = 270$). The result remains significant when the sample is restricted to dishonest senders: deceptive y messages take 3.274 seconds longer than honest y messages (Column 5; $t = 2.07$, $p < 0.05$, $N = 171$). These within-sender results indicate

that the aggregate comparison masks substantial heterogeneity in baseline response speed across senders.

The SA treatment exhibits a related but weaker pattern. In the pooled specification, deceptive y messages take 2.197 seconds longer than honest y messages (Column 2; $t = 3.31$, $p < 0.01$, $N = 288$). After sender fixed effects are included, the estimated difference falls to 1.330 seconds and is significant only at the 10% level (Column 4; $t = 1.80$, $p < 0.10$, $N = 288$). Among dishonest senders, the estimated difference remains positive but is no longer statistically significant (Column 6; 1.406 seconds, $t = 1.65$, $p > 0.10$, $N = 171$). Thus, the positive association between RT and deception is more robust in the SU treatment than in the SA treatment once persistent sender-level differences are taken into account.

Taken together, the regression evidence is consistent with the cognitive-conflict mechanism in Hypothesis 2: deceptive y messages tend to require more time than honest y messages. Importantly, this relationship is strongest and most robust when senders are unaware that their RT will be observed, whereas it is attenuated when senders are aware of RT observability. This result is also supported by the robustness check reported in Appendix B.1, which includes round fixed effects.

Result 2: Deceptive y messages are associated with longer RTs than honest y messages, consistent with the cognitive-conflict mechanism. This relationship is robust to sender fixed effects in the SU treatment but is weaker in the SA treatment.

5.1.2 Inference from RT

A difference in RT across message types and an inference based on RT are conceptually distinct. The former asks whether honest and deceptive messages are associated with different response times. The latter reverses the conditioning and asks whether, after observing a particular RT, one can infer whether a message is more likely to be honest or deceptive. Thus, even if deceptive messages take longer on average, RT may still have limited inferential value if honest and deceptive messages have substantially overlapping RT distributions. For binary inference, an RT cue favors honesty when the conditional probability that a y message is honest exceeds 50%, and favors deception when this probability falls below 50%.

For the purpose of inference, we shift our focus to whether the message y is honest or dishonest, where dishonesty includes both deceptive and not fully informed messages. An honest message is one that is true and should be trusted. In contrast, a dishonest message either has a probability of being true no greater than 34.37% (if not fully informed), or is certain to be false (if deceptive), and thus should be disregarded.

The conditional probability that a certain type of y is honest is defined as

$$p(honest|.)=\frac{p(.,honest)}{p(.)}.$$

For example, the probability of slow y being honest is

$$p(honest|slow)=\frac{p(slow|honest)p(honest)}{p(slow|honest)p(honest)+p(slow|dishonest)p(dishonest)}.$$

Result 2 shows that deceptive y messages tend to take longer than honest y messages, particularly in the SU treatment after controlling for sender-level heterogeneity. This establishes an average relationship between RT and honesty, but it does not imply that RT is equally informative for all senders. We therefore examine conditional honesty rates more directly below.

We define a message y as *immediate* if its RT equals zero. In our sample, 6 of 161 y messages (3.7%) in the SU treatment and 2 of 157 y messages (1.3%) in the SA treatment are classified as immediate. Overall, immediate messages account for only 2.5% of all y messages. Response times were recorded in seconds and rounded to the nearest integer, so an observed RT of 0 reflects a response completed in less than 0.5 seconds rather than a literally instantaneous response. Given that the sender needs to click several buttons to reveal the truth before sending an honest y , an immediate y cannot be classified as an honest message under our definition. In fact, the shortest fully informed message has an RT of 3 seconds, supporting this interpretation. Given that immediate messages consistently imply dishonesty, we focus on non-immediate ($RT > 0$) y messages and assess their veracity based on RT.

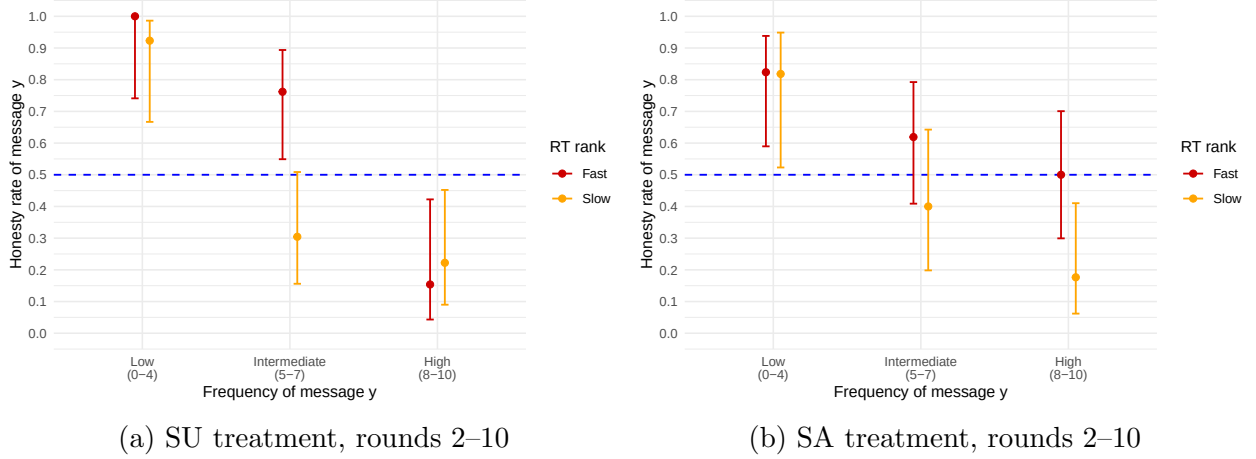
We categorize non-immediate y messages into three groups: Median, where the RT is exactly equal to the individual non-immediate median RT over rounds 2–10; Slow, where the RT exceeds the median RT; and Fast, where the RT is less than the median RT. This message categorization is made at the individual level.

Furthermore, RT may reveal information about a sender’s broader communication pattern. We characterize this pattern by the total number of y messages sent over the ten rounds. Guided by the distribution expected under truthful communication, we classify senders who send 0–4 y messages as low-frequency y senders, those who send 5–7 as intermediate-frequency y senders, and those who send 8–10 as high-frequency y senders.⁴ The resulting

⁴Under truthful communication, the number of y messages, N_y , follows a Binomial(10, 0.3437) distribution. Three y messages is the modal outcome, with $P(N_y = 3) \approx 25.5\%$, but four y messages is nearly as likely, with $P(N_y = 4) \approx 23.4\%$. Moreover, approximately 76.5% of truthful senders are expected to send no more than four y messages. We therefore include four-message senders in the low-frequency category.

distribution of sender-frequency categories is 12 low-frequency, 12 intermediate-frequency, and 6 high-frequency senders in the SU treatment, compared with 17, 9, and 6, respectively, in the SA treatment. The distribution does not differ significantly across treatments ($\chi^2(2) = 1.23$, $p = 0.541$).

Figure 3: Honesty Rates of Fast and Slow y Messages by Sender-Frequency



Notes: Fast and slow RTs are defined at the individual level using rounds 2–10. A message is classified as slow if the sender’s RT exceeds the sender’s own median non-immediate RT and as fast if it falls below that median. Each point represents the mean honesty rate of y messages within the corresponding sender-frequency category, and the error bars indicate 95% Wilson score confidence intervals. The blue dashed line marks the 50% honesty-rate threshold. Sender-frequency categories are based on the total number of y messages sent across all 10 rounds. Senders who sent 0–4 y messages are classified as low-frequency y senders, those who sent 5–7 are classified as intermediate-frequency y senders, and those who sent 8–10 are classified as high-frequency y senders.

Figure 3 depicts the honesty rates of fast and slow y messages across the three sender-frequency categories. Table 5 complements the statistical tests shown in Figure 3. Panel A reports observed honesty rates with Wilson 95% confidence intervals and uses two-sided score tests to assess whether each honesty rate differs from the 50% benchmark. Panel B directly compares fast and slow responses within each condition and sender-frequency category using two-sided Fisher’s exact tests.

The results show that the informativeness of RT is heterogeneous across sender-frequency categories. In the SU treatment, low-frequency y messages are highly likely to be honest regardless of response speed: the honesty rates are 100.0% for fast responses and 92.3% for slow responses. Both rates lie above the 50% benchmark, while the fast–slow difference is not statistically significant (Fisher’s exact test, $p = 1.000$). At the other end of the distribution, high-frequency y messages have low honesty rates for both fast and slow responses (15.4%

At the other end of the distribution, sending eight or more y messages is extremely unlikely under truthful communication, with $P(N_y \geq 8) \approx 0.4\%$, motivating the cutoff for the high-frequency category.

Table 5: Statistical Tests for Fast and Slow y Messages by Sender-Frequency*Panel A: Wilson score tests of honesty rates against 50%*

Condition	Frequency	RT	Honest/N	Honesty rate	95% CI	p-value
SU	low-frequency	Fast	11/11	100.0%	[74.1%, 100.0%]	<.001
SU	low-frequency	Slow	12/13	92.3%	[66.7%, 98.6%]	0.002
SU	intermediate-frequency	Fast	16/21	76.2%	[54.9%, 89.4%]	0.016
SU	intermediate-frequency	Slow	7/23	30.4%	[15.6%, 50.9%]	0.061
SU	high-frequency	Fast	2/13	15.4%	[4.3%, 42.2%]	0.013
SU	high-frequency	Slow	4/18	22.2%	[9.0%, 45.2%]	0.018
SA	low-frequency	Fast	14/17	82.4%	[59.0%, 93.8%]	0.008
SA	low-frequency	Slow	9/11	81.8%	[52.3%, 94.9%]	0.035
SA	intermediate-frequency	Fast	13/21	61.9%	[40.9%, 79.2%]	0.275
SA	intermediate-frequency	Slow	6/15	40.0%	[19.8%, 64.3%]	0.439
SA	high-frequency	Fast	10/20	50.0%	[29.9%, 70.1%]	1.000
SA	high-frequency	Slow	3/17	17.6%	[6.2%, 41.0%]	0.008

Panel B: Fisher's exact tests for fast and slow y messages

Condition	Frequency	Fast honesty	Slow honesty	Slow/Fast odds ratio	OR 95% CI	Fisher p-value
SU	low-frequency	100.0%	92.3%	0.000	[0.000, 46.058]	1.000
SU	intermediate-frequency	76.2%	30.4%	0.144	[0.028, 0.615]	0.003
SU	high-frequency	15.4%	22.2%	1.549	[0.181, 20.134]	1.000
SA	low-frequency	82.4%	81.8%	0.966	[0.091, 13.715]	1.000
SA	intermediate-frequency	61.9%	40.0%	0.421	[0.085, 1.927]	0.311
SA	high-frequency	50.0%	17.6%	0.224	[0.031, 1.172]	0.082

Notes: The unit of observation is a sender-round report. Panel A reports observed honesty rates with Wilson 95% confidence intervals. The reported two-sided score tests evaluate whether each honesty rate differs from 50%. Panel B reports two-sided Fisher's exact tests comparing fast and slow responses within each condition and sender-frequency category. Odds ratios compare slow with fast responses. Reported p -values are not adjusted for multiple comparisons.

and 22.2%, respectively), and the difference between them is again not significant (Fisher’s exact test, $p = 1.000$). Thus, at the two ends of the sender-frequency distribution, the sender’s overall communication frequency already provides a relatively strong indication of honesty, leaving little incremental information to be extracted from RT.

The strongest incremental information from RT appears among intermediate-frequency y senders in the SU treatment. Fast messages are honest 76.2% of the time, whereas slow messages are honest only 30.4% of the time. The difference is statistically significant (two-sided Fisher’s exact test, $p = 0.003$). Hence, when sender frequency alone provides a less decisive signal, response speed can further distinguish honest from dishonest messages.

In the SA treatment, fast messages also have higher honesty rates than slow messages among intermediate- and high-frequency y senders, but neither difference is statistically significant at conventional levels (61.9% versus 40.0%, $p = 0.311$, and 50.0% versus 17.6%, $p = 0.082$, respectively). Overall, RT does not provide a uniform signal across sender-frequency categories. Its incremental informational value depends on the sender’s broader communication pattern and is most pronounced for intermediate-frequency y senders in the SU treatment.

Result 3: RT provides incremental information about message honesty primarily when the sender’s message frequency is itself less informative. In the SU treatment, response speed significantly distinguishes honest from dishonest y messages among intermediate-frequency senders, whereas RT adds little information for low- and high-frequency senders. This distinction is weaker in the SA treatment.

5.1.3 Manipulation of RT

If senders recognize that RT may reveal information about message honesty, making RT observable to receivers may create an incentive to adjust response timing so that honest and deceptive messages become less distinguishable. Hypothesis 4 therefore predicts that the relationship between response time and honesty should be weaker when senders are aware that their RTs are observable.

Comparing the two conditions in Figure 3 and Table 5, the relationship between RT and honesty is less pronounced in the SA treatment than in the SU treatment. Among intermediate-frequency y senders, fast messages are substantially more likely to be honest than slow messages in SU (76.2% versus 30.4%, $p = .003$). Moreover, the honesty rate of fast messages is significantly above 50% ($p = .016$), while that of slow messages is marginally below 50% ($p = .061$). Thus, RT provides useful information about message honesty in SU.

In SA, the corresponding difference is smaller and statistically insignificant (61.9% versus 40.0%, $p = .311$), and neither honesty rate differs significantly from 50%. Although fast and slow messages also exhibit different honesty rates among high-frequency senders in SA (50.0% versus 17.6%), this difference is only marginally significant ($p = .082$); moreover, fast messages are no more informative than chance, although slow messages remain significantly associated with dishonesty ($p = .008$). Taken together, the attenuation of the RT–honesty relationship from SU to SA suggests that senders may make their response times less informative when they know that RT is observable.

We also examine whether awareness of RT observability is accompanied by a broader change in messaging behavior. The overall honesty rate of informed messages is 81.2% in SU and 83.7% in SA. The difference is not statistically significant ($p = .645$, sender-clustered logit). Thus, the main behavioral response appears to operate through response timing rather than through a large shift in the propensity to communicate honestly.

Result 4: The relationship between response time and message honesty is less pronounced when senders are aware that RT is observable.

Overall, the sender–side results show that most senders acquire sufficient information to identify the true state, and that the observed variation in RT mainly reflects post-information deliberation rather than mechanical differences in information acquisition. Among intermediate-frequency senders in the SU treatment, slower responses are more likely to accompany dishonest y messages. At the same time, the informational value of RT appears weaker when senders are aware that their RTs are observable.

The next section turns to receivers and asks whether they actually use these RT differences when making trust decisions.

5.1.4 Receivers’ Behavior

In this section, we examine how receivers use RT information. According to Result 3, an optimal receiver should follow both n and y messages from a sender classified as a low-frequency y -sender, but should always choose *NO* when facing a sender classified as a high-frequency y -sender. When facing an intermediate-frequency y -sender, an optimal receiver should follow a fast y message but choose *NO* following a slow y message. We investigate whether receivers behave in accordance with these predictions and whether their responses vary systematically across receiver types.

One concern is that providing receivers with RT information could act as an experimental demand cue and induce them to change their choices. In the SU treatment, the proportion

of receivers guessing YES was 39% without RT information (RU) and 37% with RT information (RI). In the SA treatment, the corresponding proportions were 42% and 44%. Wilcoxon signed-rank tests indicate that the effect of providing RT information is statistically insignificant in both conditions ($p = 0.42$ in SU, and $p = 0.28$ in SA). Thus, we find no evidence that providing RT information systematically affects receivers’ choices.

Our design does not permit further experimental variation in the relative number or ordering of RU and RI decisions that could fully rule out a demand effect. In particular, once receivers have observed response-time information in RI, subsequently withholding that information in RU would not restore an uninformed baseline. Nevertheless, even if the provision of RT information generated an overall demand effect, it would remain informative to examine whether receivers respond differently to faster versus slower responses.

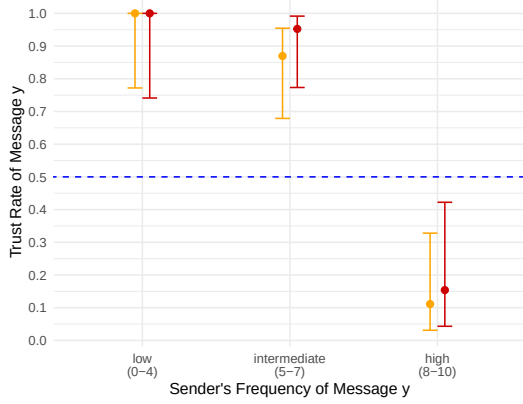
Figure 4 depicts receivers’ trust rates of fast and slow y messages across the four conditions. Overall, receivers tend to trust y messages from low- and intermediate-frequency y senders more than those from high-frequency y senders. However, within each sender-frequency category, trust rates do not differ significantly between fast and slow y messages in any condition. These results suggest that receivers do not systematically use response speed when deciding whether to trust a y message.

To examine these patterns more formally, Table 6 reports OLS regressions of receivers’ trust decisions with receiver individual fixed effects. The regression results show that the sender’s frequency of y messages is strongly associated with receivers’ trust decisions. In particular, receivers place substantially less trust in y messages from high-frequency y senders in both the SU and SA treatments. By contrast, there is little evidence of a common RT effect across conditions: most interactions between RI and RT are not statistically significant.

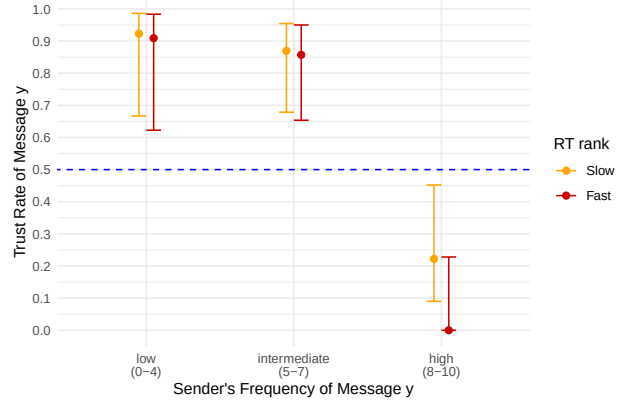
Nevertheless, several significant interaction terms reveal heterogeneity in the use of RT information across treatments and receiver types. Among receivers who previously behaved honestly as senders, the effect of RI in the SU treatment varies significantly with the sender’s RT category when the sender is classified as a high-frequency y -sender. In particular, the positive $RI \times \text{High-frequency}$ coefficient (Column 3, 0.893), together with the negative three-way interactions involving the Fast and Median RT categories (Column 3, -0.893, -0.893), suggests that these receivers place relatively greater trust in slow than in fast y messages from high-frequency y -senders. The corresponding pattern is not evident in the SA treatment. Receivers who previously behaved dishonestly as senders also exhibit several significant RT-related interaction effects, particularly in the SA treatment, although these effects do not form a consistent pattern across RT and sender-frequency categories.

One possible interpretation of this heterogeneity is that receivers’ own prior behavior as senders shapes how they interpret RT. Receivers who previously behaved honestly as senders

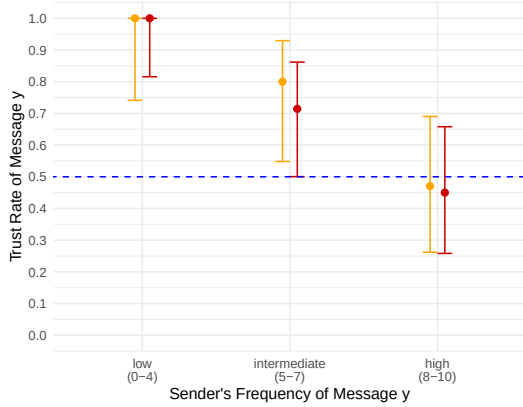
Figure 4: Trust Rates of y Messages in Four Conditions



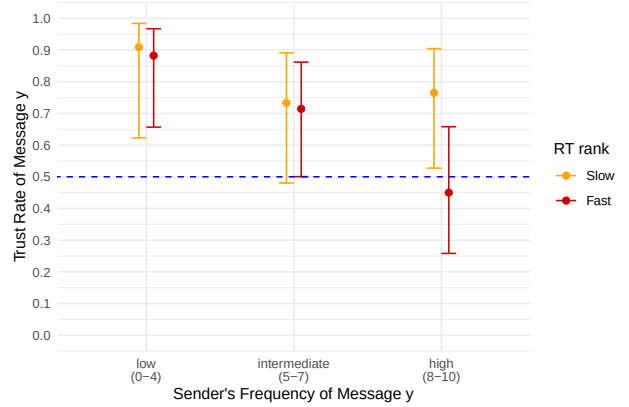
(a) SU-RU condition



(b) SU-RI condition



(c) SA-RU condition



(d) SA-RI condition

Notes: The points represent the trust rates for message y , categorized by the sender's RT rank, conditional on the number of y messages sent by the matched sender. The error bars represent 95% Wilson confidence intervals for the trust rates. The blue line marks the 50% trust-rate benchmark. "Slow" indicates that the sender's RT exceeds their own median RT for non-immediate reports in rounds 2–10, whereas "Fast" indicates that the sender's RT falls below this median.

Table 6: Regression of Receiver Trust on Sender Response Time and Receiver' Prior Behavior as Senders

	<i>Dependent variable: Trust</i>					
	All receivers		Honest receivers		Dishonest receivers	
	SU	SA	SU	SA	SU	SA
<i>Main effects</i>						
Constant	0.854*** (0.026)	1.051*** (0.044)	0.870*** (0.051)	1.015*** (0.029)	0.998*** (0.027)	1.071*** (0.063)
RI	-0.077 (0.078)	-0.091 (0.097)	-0.143 (0.147)	0.000 (0.000)	-0.000 (0.000)	-0.167 (0.182)
Fast	-0.012 (0.020)	0.006 (0.019)	-0.037 (0.045)	0.035 (0.028)	0.008 (0.013)	-0.017 (0.023)
Immediate	-0.073 (0.049)	0.063 (0.043)			-0.088 (0.057)	0.062 (0.063)
Median	0.016 (0.023)	-0.050 (0.074)	0.032 (0.047)	0.015 (0.054)	0.002 (0.027)	-0.125 (0.113)
Intermediate-frequency	0.169*** (0.031)	0.003 (0.073)	0.143** (0.057)	0.069 (0.086)	0.027 (0.037)	-0.680*** (0.072)
High-frequency	-0.781*** (0.056)	-1.089*** (0.073)	-0.870*** (0.051)	0.007 (0.035)	-0.910*** (0.064)	-1.148*** (0.108)
<i>Two-way interactions</i>						
RI×Fast	-0.014 (0.118)	-0.027 (0.131)	0.143 (0.147)	-0.250 (0.165)	-0.143 (0.131)	0.167 (0.182)
RI×Immediate	-0.111 (0.185)	-0.794*** (0.210)			0.071 (0.067)	-0.900*** (0.340)
RI×Median	0.077 (0.078)	0.091 (0.097)	0.143 (0.147)	-0.000 (0.000)	0.000 (0.000)	0.167 (0.182)
RI×Intermediate-frequency	0.077 (0.124)	0.024 (0.117)	0.032 (0.204)	-0.000 (0.000)	0.071 (0.130)	0.042 (0.212)
RI×High-frequency	0.188 (0.201)	0.385* (0.232)	0.893*** (0.147)	0.143 (0.102)	-0.071 (0.067)	0.567 (0.385)
Fast×Intermediate-frequency	0.040 (0.037)	-0.081 (0.111)	0.067 (0.056)	-0.152 (0.160)	0.042 (0.062)	-0.014 (0.156)
Median×Intermediate-frequency	-0.089 (0.085)	0.011 (0.203)	0.039 (0.076)	-0.259 (0.202)	-0.117 (0.119)	0.270 (0.358)
Fast×High-frequency	0.085* (0.046)	-0.080 (0.077)	0.037 (0.045)	-0.037 (0.028)	0.033 (0.036)	-0.102 (0.116)
Median×High-frequency	-0.088 (0.054)	0.137* (0.081)	-0.032 (0.047)	0.107 (0.118)	-0.090 (0.064)	0.210* (0.121)
<i>Three-way interactions</i>						
RI×Fast×Intermediate-frequency	-0.081 (0.197)	0.093 (0.276)	-0.032 (0.204)	0.150 (0.453)	-0.110 (0.280)	0.049 (0.310)
RI×Median×Intermediate-frequency	-0.077 (0.201)	0.226 (0.259)	-0.032 (0.204)	0.167 (0.381)	-0.071 (0.263)	0.292 (0.361)
RI×Fast×High-frequency	-0.251 (0.219)	-0.267 (0.296)	-0.893*** (0.147)	-0.143 (0.353)	0.048 (0.162)	-0.400 (0.444)
RI×Median×High-frequency	-0.188 (0.201)	-0.462* (0.241)	-0.893*** (0.147)	-0.143 (0.102)	0.071 (0.067)	-0.658* (0.393)
Fixed Individual	✓	✓	✓	✓	✓	✓
Observations	290	282	90	120	200	162
R ²	0.757	0.531	0.821	0.304	0.749	0.624

Notes: The sample is restricted to y messages. The dependent variable, Trust, equals one if the receiver follows the sender's y message and zero otherwise. RI is an indicator for decisions in which the sender's response time is revealed. Honest receivers are participants who previously behaved honestly when acting as senders, whereas dishonest receivers are participants who previously behaved dishonestly when acting as senders. The omitted response-time and sender-frequency categories are Slow and Low-frequency y -sender, respectively. All specifications include individual fixed effects. Standard errors, clustered at the individual level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

may associate a longer RT with greater deliberation or effort in formulating a truthful response and may therefore be more likely to use RT when evaluating a message. By contrast, receivers who previously behaved dishonestly as senders may infer less from RT based on their own experience of the decision process. Overall, the sender’s frequency of y messages is a more consistent predictor of trust, whereas receivers’ use of RT information appears limited and context-dependent.

Result 5: Receivers’ trust decisions are more consistently associated with the sender’s frequency of y messages. There is limited evidence of systematic RT use overall.

5.2 Results in Effortless Experiment

In the Effortful (E) experiment, RT comprises both the time potentially required for the mechanical process of truth discovery and the sender’s deliberation over which message to send. Receivers may therefore interpret a longer RT as indicating greater effort in discovering the truth, potentially affecting their perceptions of the sender’s reliability. In the Effortless (L) experiment, the truth-discovery process is absent, so RT primarily captures the sender’s deliberation over the message rather than the time spent acquiring information. Comparing the E and L experiments therefore allows us to examine whether receivers’ interpretation and use of RT depend on its informational composition.

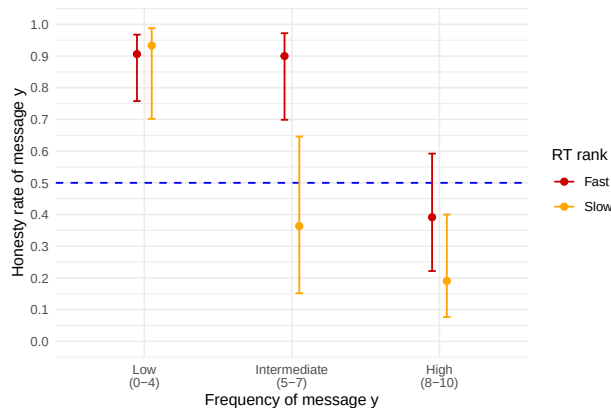
Following the analytical framework used above, we distinguish between senders’ actual behavioral types and the types that receivers can infer from observed message frequencies. Because we cannot observe whether senders attend to the true state, we classify each message as *true* if it matches the state and *false* otherwise. The Effortless (L) experiment also consists of 10 rounds. A sender is classified as *honest* if all 10 messages are true and as *dishonest* otherwise. Based on this criterion, the sample contains 21 honest and 21 dishonest senders.

From the receiver’s perspective, senders are classified according to the observed number of y messages: low-frequency if they send four or fewer y messages, intermediate-frequency if they send between five and seven, and high-frequency if they send eight or more. This classification yields 24 low-frequency, 10 intermediate-frequency, and 8 high-frequency y -senders.

Figure 5 depicts the honesty rates of fast and slow y messages in the Effortless experiment. The relationship closely mirrors that observed in the Effortful experiment (Figure 3). y messages from low-frequency senders are highly likely to be true regardless of whether they

are sent relatively fast or slowly (90.6% versus 93.3%; two-sided Fisher’s exact test, $p = 1.000$), whereas those from high-frequency senders have truth rates below 50% for both RT categories (39.1% versus 19.0%; two-sided Fisher’s exact test, $p = 0.194$). Most importantly, among intermediate-frequency senders, fast y messages are substantially more likely to be true than slow y messages (90.0% versus 36.4%; two-sided Fisher’s exact test, $p = 0.003$). These findings provide evidence that the relationship between RT and honesty is robust to the effort required for information acquisition.

Figure 5: Honesty Rates of Fast and Slow y Messages in Effortless Experiment

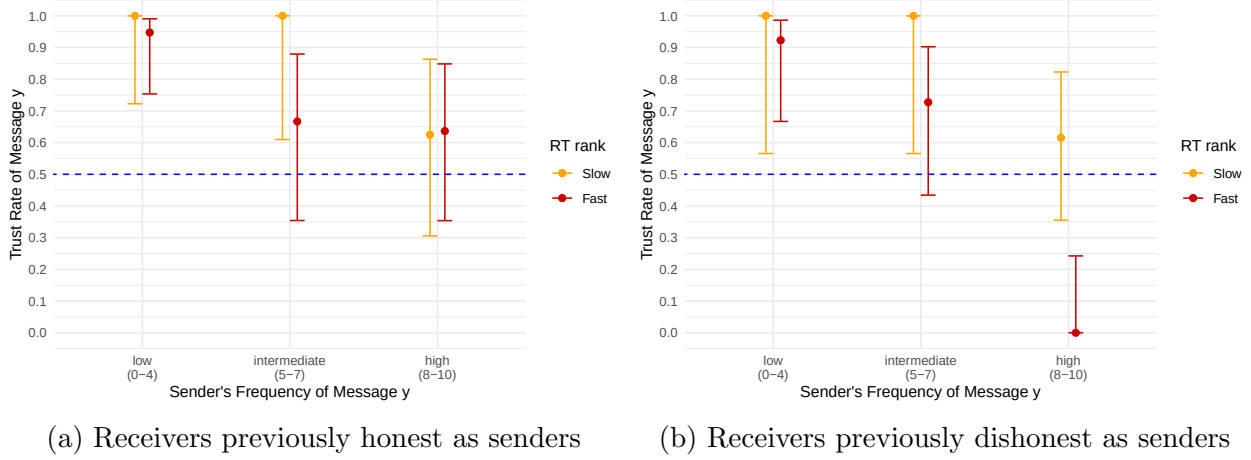


Notes: Fast and slow RTs are defined at the individual level using rounds 2-10. A message is classified as slow if the sender’s RT exceeds the sender’s own median non-immediate RT and as fast if it falls below that median. Each point represents the mean honesty rate of y messages within the corresponding sender-frequency category, and the error bars indicate 95% Wilson score confidence intervals. The blue dashed line marks the 50% honesty-rate threshold. Sender-frequency categories are based on the total number of y messages sent across all 10 rounds.

The main focus of the Effortless experiment is to examine whether receivers can develop an accurate association between truth and RT in the absence of the mechanical process. Figure 6 illustrates that receivers previously honest as senders do not distinguish fast and slow y messages for any classified sender type. In contrast, receivers previously dishonest as senders show a pronounced difference in trust between fast and slow y messages from high-frequency y senders.

Interestingly, receivers previously dishonest as senders placed significantly more trust in slow than in fast y messages from high-frequency senders. An exploratory analysis suggests that they may evaluate the sender’s RT relative to their own experience: messages produced no more slowly than the receiver’s fastest true response were trusted substantially less often than messages exceeding this receiver-specific benchmark. This pattern is consistent with receivers treating very short RTs as implausibly fast, although it does not establish that they explicitly applied such a decision rule. The benchmark definition, descriptive evidence, and statistical analyses are reported in Appendix C.1.

Figure 6: Receivers’ Trust Rates of y Messages in Effortless Experiment



Notes: The figure reports receivers’ mean trust rates for y messages in the RT-information condition, separately by receiver type, sender y -message frequency, and response-time category. The sample is restricted to rounds 2–10 and excludes immediate responses and responses equal to the sender-specific median RT. For each sender, Fast (Slow) denotes an RT below (above) the median of the sender’s positive RTs in rounds 2–10. Senders are classified as low-frequency if they send 0–4 y messages, intermediate-frequency if they send 5–7, and high-frequency if they send 8–10. Error bars represent 95% Wilson confidence intervals. The dashed blue line indicates a trust rate of 50%.

Taken together, the results from the Effortful and Effortless experiments point to a general tendency to interpret longer RTs favorably. In the Effortful experiment, receivers may attribute longer RTs to greater effort in discovering the truth. In the Effortless experiment, where truth discovery requires no mechanical effort, some receivers may instead evaluate the sender’s RT relative to their own truthful-response time. Although the underlying mechanism cannot be conclusively identified, receivers do not appear to recognize that longer deliberation may be associated with deception. When they distinguish between fast and slow messages, they generally place greater trust in the slower message.

Result 5’: When receivers distinguish between fast and slow messages, they tend to place greater trust in slower messages. This pattern is consistent with receivers interpreting longer RTs as reflecting either greater effort in truth discovery or a more plausible amount of deliberation, although the underlying mechanism cannot be conclusively identified.

6 Discussion

Our study investigates the private information conveyed by response time and its strategic use in a sender-receiver game with partially aligned interests. We find that longer RTs are associated with a higher likelihood of deceit, consistent with a cognitive-conflict mechanism

rather than an information-acquisition mechanism. On the receiver side, trust decisions are more consistently associated with the sender’s frequency of y messages than with the sender’s RT.

On the sender side, in the SU treatment of both the Effortful and Effortless experiments, dishonest y messages are associated with slower responses, whereas honest y messages are associated with faster responses among intermediate-frequency senders. In contrast, RT provides little additional information about message honesty for low- and high-frequency senders. In the Effortful experiment, senders acquire sufficient information to identify the true state in more than 92% of rounds. Slow responses are therefore unlikely to reflect greater effort devoted to discovering the truth; rather, they are more likely to capture the additional deliberation and cognitive conflict involved in deciding to lie after the truth has already been discovered.

On the receiver side, the evidence suggests that trust decisions are driven primarily by the sender’s frequency of y messages rather than by response time itself. Providing RT information therefore does not lead receivers to systematically incorporate deliberation time into their trust decisions. We find only limited evidence of RT use in specific subsets of receivers, suggesting that the behavioral interpretation of response time is heterogeneous rather than general. In particular, some patterns are consistent with receivers relying on their own prior experience as a reference point when evaluating others’ response times, but this channel is secondary to the much stronger role played by observed message frequency. Taken together, these findings indicate that response time may contain behaviorally relevant information without being readily or uniformly exploited by receivers.

One question for future research is why receivers use RT selectively and sometimes in the wrong sender group. In our data, RT most clearly predicts honesty among intermediate-frequency senders, yet the strongest evidence of receiver attention to RT appears among high-frequency senders. Future experiments could test whether this mismatch reflects incorrect beliefs about the source of delay or reliance on receivers’ own response times as benchmarks.

A second direction is to examine whether receivers respond more strongly when RT is made more salient or economically consequential. In our setting, RT is an incidental by-product of information acquisition and deliberation, and receivers face no direct reason to treat delay as a costly signal. A money-burning mechanism, for example, could make waiting costly to senders and thereby give response timing a more explicit signaling role. Such a design could reveal whether receivers’ limited use of RT results from difficulty interpreting an incidental behavioral cue or from a more general reluctance to rely on timing information.

Acknowledgment During the preparation of this work the authors used ChatGPT in order to improve the clarity and readability of the writing, including polishing sentence structure and grammar. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

- Abeler, Johannes, Daniele Nosenzo, and Collin Raymond.** 2019. “Preferences for truth-telling.” *Econometrica*, 87(4): 1115–1153.
- Belot, Michele, and Jeroen Van de Ven.** 2017. “How private is private information? The ability to spot deception in an economic game.” *Experimental economics*, 20: 19–43.
- Bonnefon, Jean-François, Astrid Hopfensitz, and Wim De Neys.** 2013. “The modular nature of trustworthiness detection.” *Journal of Experimental Psychology: General*, 142(1): 143.
- Bonnefon, Jean-François, Astrid Hopfensitz, and Wim De Neys.** 2017. “Can we detect cooperators by looking at their face?” *Current Directions in Psychological Science*, 26(3): 276–281.
- Chen, Jingnan, and Daniel Houser.** 2017. “Promises and lies: can observers detect deception in written messages.” *Experimental Economics*, 20: 396–419.
- Cotet, Miruna, and Ian Krajbich.** 2021. “Response times in the wild: eBay sellers take hours longer to reject high offers and accept low offers.” *Available at SSRN 3804578*.
- Dana, Jason, Roberto A Weber, and Jason Xi Kuang.** 2007. “Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness.” *Economic Theory*, 33: 67–80.
- Dwenger, Nadja, and Tim Lohse.** 2019. “Do individuals successfully cover up their lies? Evidence from a compliance experiment.” *Journal of Economic Psychology*, 71: 74–87.
- Frydman, Cary, and Ian Krajbich.** 2022. “Using response times to infer others’ private information: An application to information cascades.” *Management Science*, 68(4): 2970–2986.
- Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki.** 2018. “Speed, accuracy, and the optimal timing of choices.” *American Economic Review*, 108(12): 3651–3684.
- Greene, Joshua D, and Joseph M Paxton.** 2009. “Patterns of neural activity associated with honest and dishonest moral decisions.” *Proceedings of the National Academy of Sciences*, 106(30): 12506–12511.

- Grossman, Zachary.** 2014. “Strategic ignorance and the robustness of social preferences.” *Management Science*, 60(11): 2659–2665.
- Jiang, Ting.** 2013. “Cheating in mind games: The subtlety of rules matters.” *Journal of Economic Behavior & Organization*, 93: 328–336.
- Kononov, Arkady, and Ian Krajbich.** 2019. “Revealed strength of preference: Inference from response times.” *Judgment and Decision making*, 14(4): 381–394.
- Kononov, Arkady, and Ian Krajbich.** 2023. “Decision times reveal private information in strategic settings: Evidence from bargaining experiments.” *The Economic Journal*, 133(656): 3007–3033.
- Konrad, Kai A, Tim Lohse, and Salmai Qari.** 2014. “Deception choice and self-selection—The importance of being earnest.” *Journal of Economic Behavior & Organization*, 107: 25–39.
- Krajbich, Ian.** 2025. “Does Slow Strategizing Indicate Optimal Behavior or Uncertainty?”
- Mazar, Nina, On Amir, and Dan Ariely.** 2008. “The dishonesty of honest people: A theory of self-concept maintenance.” *Journal of marketing research*, 45(6): 633–644.
- Rubinstein, Ariel.** 2016. “A typology of players: Between instinctive and contemplative.” *The Quarterly Journal of Economics*, 131(2): 859–890.
- Serra-Garcia, Marta, and Uri Gneezy.** 2021. “Nonreplicable publications are cited more than replicable ones.” *Science advances*, 7(21): eabd1705.
- Shalvi, Shaul, Ori Eldar, and Yoella Bereby-Meyer.** 2012. “Honesty requires time (and lack of justifications).” *Psychological science*, 23(10): 1264–1270.
- Todorov, Alexander, Christopher Y Olivola, Ron Dotsch, and Peter Mende-Siedlecki.** 2015. “Social attributions from faces: Determinants, consequences, accuracy, and functional significance.” *Annual review of psychology*, 66: 519–545.
- Woodford, Michael.** 2014. “Stochastic choice: An optimizing neuroeconomic model.” *American Economic Review*, 104(5): 495–500.

A Appendix: Instructions and Interfaces

PART I

Part 1 consists of 10 rounds, in which each of you will play the role of a sender. In each round, the computer generates outcomes of 6 dice. The outcome of each die is randomly drawn from the set 1, 2, 3, 4, 5, 6.

We define the outcomes 1, 2, 3 as "small numbers", and 4, 5, 6 as "large numbers".

You can privately learn the outcomes of the 6 dice by clicking the dice buttons on your screen. Each button will uncover the outcome of 1 die, and you can uncover as many as you want. The outcomes of the dice will not change with the total buttons you click to uncover. Then you need to report on the question if there are 4 or more "large numbers" by choosing "Yes" or "No". Your report could be independent from the outcomes of the 6 dice.

Your payment as a sender

Your payment will be based on the Parts 2 and 3 of the experiment. You will be matched with another player, who's a receiver, and his choices will determine your payoff as a sender.

The receiver in Parts 2 and 3 has to answer the same question as in Part 1, without the ability to learn the true outcomes. However, he will receive your report to inform him to guess the answer.

If the receiver chooses **Yes** (there ARE 4 or more "large numbers"), you will get **\$8**, regardless of the true answer.

If the receiver chooses **No** (there ARE NOT 4 or more "large numbers"), you will get **\$4**, regardless of the true answer.

The receiver gets \$ 8 if his guess is right, and \$4 if his guess is wrong.

If the instructions is clear to you, you may go to the next screen to answer comprehension questions.

OK

You will be able to proceed after answering the following questions correctly.

1. If the outcomes of 6 dice are 5,3,3,6,4,1, which statement is true?

A. It is true that there are 4 or more large numbers.
B. It is wrong that there are 4 or more "large numbers".
C. A sender has an obligation to report "No" based on this outcome.

☐ A
☐ B
☐ C

2. How is the sender's payment determined?

A. It depends whether the sender answers the questions correctly, if it is correct, the sender will get \$8, otherwise she will get \$4.
B. Only the receiver's choice will determine the sender' payoff. If the receiver chooses Yes, then the sender will get \$8. If the receiver chooses No, then the sender will get \$4.
C. The sender only sends message to the receiver, the receiver's choice is going to determine the sender' payoff. If the receiver chooses correctly, then the sender will get \$8. If the receiver chooses wrongly, then the sender will get \$4

☐ A
☐ B
☐ C

OK

Display Die 1	Would you report that there are 4 or more numbers are "large numbers"? YES NO
Display Die 2	
Display Die 3	
Display Die 4	
Display Die 5	
Display Die 6	

PART II

Now you are a receiver, who will get 10 reports from a sender in Part 1, one for each round. You'll get reports all at the same time.

It is up to you whether you would like to refer to the reports from the sender to make your guesses or not. You will need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice". Your guess will be either "Yes" (there ARE 4 or more 'large numbers') or "No" (there ARE NOT 4 or more 'large numbers').

Your payment as a receiver

For any question, you will get \$8 if your guess is right, and \$4 if your guess is wrong.

If the instructions is clear to you, you may go to the next screen to answer comprehension questions.

OK

You will be able to proceed after answering the following question correctly.

How is the receiver's payment determined?

☐

A. The receiver will get \$8 if his guess is right, and \$4 if his guess is wrong.

☐

B. If the receiver chooses Yes, he gets \$8. Otherwise, he gets \$4.

☐

C. The receiver must refer to the sender's signal.

OK

Now you are a receiver to review the 10 reports from a sender in Part 1.

You are seeing all 10 reports from the sender. It is up to you whether you would like to refer to the reports from the sender to make your guesses or not. You need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice".

Your payment as a receiver

You will get \$8 if your guess is right, and \$4 if your guess is wrong.

Round	1	2	3	4	5	6	7	8	9	10	
Sender's Report	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No	
Your decision	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes

Submit

PART III

Now, you need to re-make 10 guesses on the questions. But, in addition to the reports from the sender, you will also know the time that the sender used to generate each report.

Your payment as a receiver

Your payment is the same as before. For any question, you will get \$8 if your guess is right, and \$4 if your guess is wrong.

OK

Now, you need to re-make 10 guesses on the questions. But, in addition to the reports from the sender, you will also know the time that the sender used to generate each report.

You are seeing all 10 reports and the response time from the sender. It is up to you whether you would like to refer to the answers from the sender to make your guesses or not. You need to make 10 guesses on "If there ARE 4 or more 'large numbers' in the 6 dice"

Your payment as a receiver

Your payment is the same as before. You will get \$8 if your guess is right, and \$4 if your guess is wrong.

Round	1	2	3	4	5	6	7	8	9	10
Response Time (in sec)	1	0	0	1	0	0	0	0	0	0
Sender's Report	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Your decision	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes	☐ No ☐ Yes

B Appendix: Robustness Checks

B.1 Robustness Check with Round Fixed Effects

The main specifications control for a linear round trend. This approach assumes that changes in response time across rounds are approximately linear. However, learning, adaptation, and fatigue may generate nonlinear changes over the course of the experiment. To allow the effect of each round to vary flexibly, we re-estimate the response-time regressions after replacing the continuous round variable with a full set of round fixed effects.

The main within-sender result in the SU treatment remains robust to the inclusion of round fixed effects. Among all SU senders, deceptive y messages take significantly longer than honest y messages after controlling for sender and round fixed effects (Column 3; $t = 2.50$, $p = 0.013$, $N = 270$, 30 sender clusters). The result remains significant when the sample is restricted to dishonest senders (Column 5; $t = 2.09$, $p = 0.038$, $N = 171$, 19 sender clusters).

The corresponding results in the SA treatment are weaker. Among all SA senders, the difference between deceptive and honest y messages is significant only at the 10% level after controlling for sender and round fixed effects (Column 4; $t = 1.67$, $p = 0.095$, $N = 288$, 32 sender clusters). Among dishonest senders, the estimated difference remains positive but is no longer statistically significant (Column 6; $t = 1.60$, $p = 0.113$, $N = 171$, 19 sender clusters). Thus, the positive relationship between RT and deception remains more robust in the SU treatment than in the SA treatment when both sender and round fixed effects are included.

As an additional analysis reported in the Appendix, we divide the sample of dishonest senders into early and late rounds. The estimated coefficients remain positive in both subsamples but are not statistically significant, likely because splitting the sample substantially reduces the number of observations and, consequently, statistical power. Nevertheless, the full-period estimate remains statistically significant after controlling for both sender and round fixed effects, supporting our central finding that deceptive y messages take longer than honest y messages in the SU treatment.

Table 7: Sender Response Time by Message Type: Round Fixed-Effects Specifications

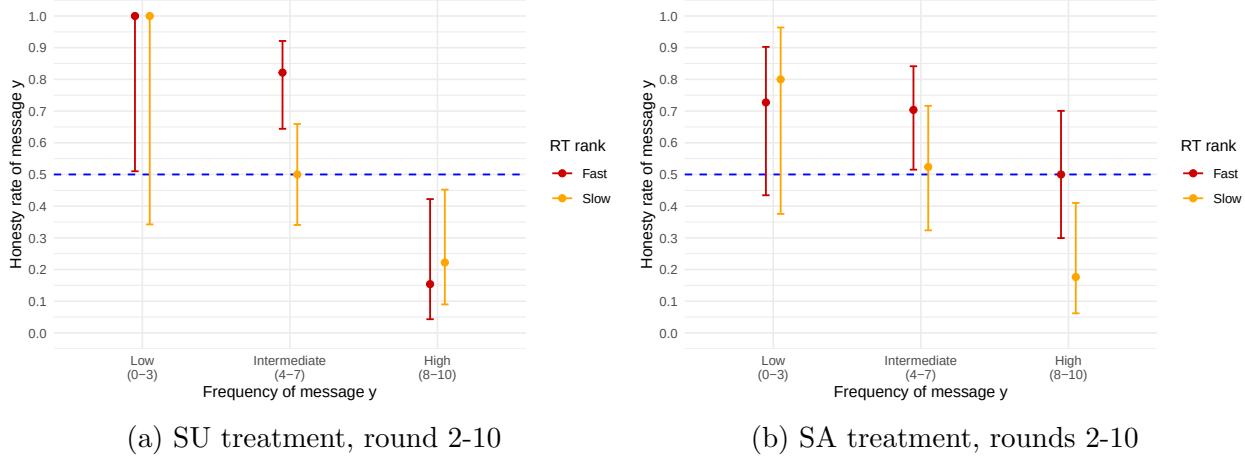
Dependent variable: RT												
	all senders			dishonest senders			dishonest, rounds 2-5			dishonest, rounds 6-10		
	SU	SA	SU	SA	SU	SA	SU	SA	SU	SA	SU	
deceptive n	-4.548* (2.412)	3.304* (1.996)	-2.573 (2.384)	2.480 (3.699)	-4.228 (3.675)	2.340 (3.660)	-2.600 (3.607)	-5.132** (1.954)			6.924*** (2.002)	
deceptive y	1.152 (1.649)	2.286*** (0.805)	3.603** (1.443)	1.383* (0.826)	3.357** (1.604)	1.539 (0.964)	2.372 (3.011)	1.353 (2.574)	2.474 (1.864)	1.252 (1.030)		
honest n	1.037 (1.215)	1.527** (0.732)	0.995 (0.949)	1.868** (0.837)	0.788 (1.476)	2.269* (1.247)	2.185 (5.113)	3.772 (3.580)	-0.721 (1.030)	1.252 (0.858)		
uninformed y	-4.054 (4.838)	-3.544 (3.032)	-4.862 (3.116)	-4.083** (1.894)	-5.126 (3.561)	-3.789* (1.945)	-9.870* (5.875)	-4.796 (3.147)	-5.953 (4.014)	-4.263** (2.082)		
Constant	13.548*** (2.412)	13.650*** (1.355)	12.491*** (1.996)	12.084*** (1.347)	24.266*** (4.669)	13.470*** (1.965)	31.255*** (6.104)	11.631*** (2.631)	18.981*** (3.843)	8.292*** (0.985)		
Fixed Round	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		
Fixed Individual			✓	✓	✓	✓	✓	✓	✓	✓		
Observations	270	288	270	288	171	171	76	76	95	95		
R ²	0.087	0.165	0.338	0.452	0.321	0.479	0.371	0.543	0.528	0.540		

B.2 Robustness Check with Threshold Three

As a robustness check, we redefine the sender-frequency categories using a more restrictive cutoff for low-frequency senders. Specifically, senders who sent 0–3 y messages are classified as low-frequency senders, those who sent 4–7 are classified as intermediate-frequency senders, and those who sent 8–10 are classified as high-frequency senders.

The broad pattern is qualitatively consistent with that obtained using the baseline cutoff, although the distinction around the 50% benchmark becomes less pronounced for intermediate-frequency senders. In the SU treatment, y messages from low-frequency senders are almost always honest, whereas those from high-frequency senders have honesty rates well below 50%, regardless of whether their RTs are fast or slow. Among intermediate-frequency senders, fast y messages continue to exhibit a significantly higher honesty rate than slow y messages (82.1% versus 50.0%, two-sided Fisher’s exact test, $p = 0.015$). However, unlike under the baseline classification, the honesty rate of slow y messages is no longer clearly below 50% (two-sided proportion test against 50%, $p = 1.000$). In the SA treatment, the confidence intervals are generally wider and substantially overlap, providing less evidence that RT systematically predicts honesty. Overall, although the alternative cutoff weakens the sharp separation around the 50% benchmark among intermediate-frequency senders, it does not materially change the broader conclusion that faster y messages tend to be more honest than slower ones in the SU treatment, while the corresponding relationship is less clear in the SA treatment.

Figure 7: Honesty Rates of Fast and Slow y Messages by Sender Type



Notes: Fast and slow RTs are defined at the individual level using rounds 2–10. A message is classified as slow if the sender’s RT exceeds the sender’s own median non-immediate RT and as fast if it falls below that median. Each point represents the mean honesty rate of y messages within the corresponding sender-frequency category, and the error bars indicate 95% Wilson score confidence intervals. The blue dashed line marks the 50% honesty-rate threshold. Sender-frequency categories are based on the total number of y messages sent across all 10 rounds. Senders who sent 0–3 y messages are classified as low-frequency y senders, those who sent 4–7 are classified as intermediate-frequency y senders, and those who sent 8–10 are classified as high-frequency y senders.

C Appendix: More Results

C.1 Using Receivers’ Minimum RTs as Benchmarks for Fast and Slow Responses

One possible interpretation of the earlier Fast–Slow pattern is that receivers who previously behaved dishonestly when acting as senders may use their own response time as a reference point when evaluating others’ messages. To explore this possibility, we compare the RT of each received y message with the receiver’s own response time from the earlier sender role.

Figure 8 shows trust rates for y messages among receivers who previously behaved dishonestly when acting as senders in the Effortless experiment. For each of these receivers, we use the receiver’s own response time from the earlier sender role as an individual benchmark for evaluating the RT of messages received from another sender. Specifically, we choose the benchmark based on that participant’s shortest RT for sending a true y message when previously acting as a sender. If a receiver never sent a true y message, we instead use that participant’s shortest RT for sending a true n message. We then compare the RT of each received y message with this receiver-specific benchmark. We classify a y message as *faster than minimum* if the current sender’s RT is shorter than the receiver’s own benchmark, as *equal to minimum* if the two RTs are equal, and as *slower than minimum* if the current sender’s RT exceeds the receiver’s own benchmark.

The resulting pattern is consistent with the earlier comparison based on the sender-specific Fast–Slow classification. Among these previously dishonest receivers matched with high-frequency senders, none of the messages with RTs below the receiver-specific benchmark were trusted (0/11), compared with 15.4% of messages with RTs equal to the benchmark (2/13) and 58.3% of messages with RTs above the benchmark (14/24). The association between benchmark category and trust is statistically significant (two-sided Fisher’s exact test, $p < 0.001$). Thus, the greater trust previously placed in slow rather than fast messages from high-frequency senders is also evident when RT is evaluated relative to the receiver’s own truthful response time from the earlier sender role.

Figure 8: Receivers’ Trust Rates of y Messages in Effortless Experiment

