CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

University of Nottingham
UK | CHINA | MALAYSIA

Discussion Paper No. 2020-11

Valeria Burdea, Maria Montero & Martin Sefton

May 2020

Communication with Partially Verifiable Information: An Experiment

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Suzanne Robey
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 95 14763
suzanne.robey@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

COMMUNICATION WITH PARTIALLY VERIFIABLE INFORMATION: AN EXPERIMENT

Valeria Burdea*, Maria Montero†, Martin Sefton‡

May 25, 2020§

Abstract

We use laboratory experiments to study communication games with partially verifiable information. In these games, based on Glazer and Rubinstein (2004, 2006), an informed sender sends a two-dimensional message to a receiver, but only one dimension of the message can be verified. We compare a treatment where the receiver chooses which dimension to verify with one where the sender has this verification control. We find significant differences in outcomes across treatments. However, receivers' payoffs do not differ significantly across treatments, suggesting they are not hurt by delegating verification control. We also show that in both treatments the receiver's best reply to senders' observed behavior is close to the optimal commitment strategy identified by Glazer and Rubinstein.

Keywords: communication, partially verifiable messages, verification control, experiment

*Department of Political Science and Pittsburgh Experimental Economics Laboratory, University of Pittsburgh (valeria.burdea@pitt.edu)

†School of Economics, University of Nottingham

‡School of Economics, University of Nottingham

§We are grateful to Johannes Abeler, Jan Potters and Daniel Seidmann for helpful comments. We would also like to thank participants at GAMES 2016, FUR 2016, SING 2016, CED 2017, the Amsterdam Communication Workshop 2018, the Stony Brook International Conference on Game Theory 2019 and seminar participants at University of the Basque Country, Tilburg University, University of Amsterdam, WZB Berlin and New York University at Abu Dhabi. This work was supported by the ESRC Network for Integrated Behavioural Science (NIBS) [grant number ES/K002201/1]. Ethical approval for the experiments was obtained from the Nottingham School of Economics Research Ethics Committee.

1 INTRODUCTION

We report experiments based on partially verifiable information games. In these games an informed sender (he) wants to persuade an uninformed receiver (she) to take a certain action. To do so he sends costless messages about his private information to the receiver, who can then verify some (but not all) of the messages. Thus, the receiver acts on a combination of hard evidence and cheap talk.

Such situations are commonplace, since in practice informed parties often make claims about their private information and it is typically infeasible or prohibitively costly for a receiver to verify all claims. Also, in practice, the verification process varies: in some cases the sender can choose which claims to verify, while in others the receiver can choose what to verify.

For example, consider a job interview where a candidate wishes to persuade an employer to hire him. In some interview formats the candidate may make claims about his various skills, and also be able to demonstrate some (but not all) of these. Which skills will the candidate choose to demonstrate? In other formats the employer may target certain skills and use particular questions to discover the candidate's command of these. As another example, consider a business owner trying to convince an investor to invest in his business. The owner may make claims about the business, but also provide hard evidence to back up these claims (e.g. sales figures, market research reports). Again, which evidence is provided might be determined both by the business owner and the potential investor.

When senders decide which claims to verify, will they disclose their strongest evidence? When receivers decide, will they investigate the strongest claims? Does it matter whether the sender or receiver has control over the verification process? To answer these questions we focus on games that are based on theoretical models of partially verifiable information (Glazer and Rubinstein, 2004, 2006) and we bring experimental evidence to bear on them. As far as we are aware, this is the first experimental study of strategic information transmission in which both cheap talk and hard evidence are present.

In the games we study the state of the world is based on the values of two aspects which are known to the sender, but not to the receiver; in some states of the world the players' interests are aligned while in others they are opposed. In our first game the sender sends a message about the two aspects to the receiver, who then chooses one of the aspects to observe before taking an action (we refer to this as the "Receiver Verifies" game). This game is based on the model introduced in Glazer and Rubinstein (2004), who use a mechanism design approach to identify a receiver's optimal strategy assuming that the receiver is able to commit to a strategy. The game we implement is an extensive-form game in which the receiver cannot commit. Nevertheless, it has a sequential equilibrium where the receiver uses an optimal commitment strategy and attains the optimal payoff of the game with commitment (Glazer and Rubinstein, 2004, section 7). We note that there are many other sequential equilibria leading to a lower receiver payoff, and no obvious reason why the receiver's preferred equilibrium would be played.

The second game we study differs from the first in that the sender decides which aspect will be observed by the receiver. This "Sender Reveals" game is similar to the model introduced in Glazer and Rubinstein (2006), except that in their model there are no messages and the receiver can commit to a strategy. Using a mechanism design approach, they show that the payoff from the receiver's optimal strategy is the same as in the model where the receiver decides which aspect to verify. Again, our laboratory extensive form game has sequential equilibria that result in the receiver attaining this optimal payoff, but many other sequential equilibria resulting in a lower receiver payoff as well.

In summary, the maximum equilibrium payoff for the receiver is the same in both games, but each game has multiple equilibria. Thus, equilibrium theory (absent any refinement criteria) does not offer a sharp prediction about how each game will be played, or which game is better for the receiver. Our experiment allows us to observe how the games are actually played, whether receivers are able to attain their maximal equilibrium payoff, and whether control of the verification process is beneficial to the receiver.

In Sender Reveals, we find that senders typically reveal their stronger aspect and make inflated claims about their weaker aspect. Given this, receiver behavior can be described as a "noisy best response", and the noise results in receivers earning less than their theoretically optimal commitment payoff. In Receiver Verifies senders also usually inflate their weaker aspect. But in response, receivers use a noisy checking strategy, resulting in them being less likely to observe the sender's stronger aspect. Again, receivers earn less than their theoretically optimal commitment payoff. Overall we find that receiver payoffs are similar across treatments, suggesting that the receiver has nothing to lose from delegating verification control.

Our experiment also complements the analysis of optimal receiver strategies in Glazer and Rubinstein (2004, 2006). Whereas they identify optimal commitment strategies for a receiver facing a sender who best responds to the receiver's strategy, we compare the performance of several receiver strategies against the empirically observed sender behavior. Interestingly, in both treatments the best performing strategy is one of the theoretically optimal commitment strategies identified in Glazer and Rubinstein (2004, 2006). Moreover, while there are multiple theoretically optimal commitment strategies, which differ from one another in how they condition on messages, and we find a particular one of these that performs best.

A theoretically optimal commitment strategy is one which, when the sender best responds to it, gives the receiver her highest expected payoff. Thus, intuitively, such a strategy is difficult for a sender to exploit. Our subject senders sometimes "come clean", truthfully admitting that the state is unfavorable. The theoretically optimal commitment strategy that performs best against our subject senders is one which conditions on messages so as to be difficult to exploit, and at the same time is able to take advantage of truthful senders.

The rest of the paper is organized as follows. In Section 2 we discuss the related literature. In Section 3 we describe the partially verifiable information games and their equilibria in detail. Section 4 describes how we implement these games in the lab and presents the experimental design. In Section 5 we report the results of our experiment, and analyze the receiver best response to the empirically observed sender behavior. Section 6 concludes.

2 RELATED LITERATURE

Our experiment studies strategic communication in a setting where an informed sender can send messages about private information and these messages are partially verifiable. To this end, our paper stands at the intersection between the cheap talk and the disclosure research agendas.

A substantial experimental literature has examined cheap talk games based on the model of Crawford and Sobel (1982) (see Blume et al., 2020, for a review), and similarly, previous experiments have studied verifiable message (disclosure) games based on the models of Milgrom (1981) and Grossman (1981) (see e.g. Forsythe et al., 1989, 1999; Hagenbach and Perez-Richet, 2018; Jin et al., 2018; King and Wallin, 1991; Li and Schipper, 2020; Penczynski and Zhang, 2017). The game we use differs from a cheap talk game in that messages can be (partially) backed by evidence, and differs from a verifiable message game in that we force partial disclosure (whereas in verifiable message games senders may remain silent or fully reveal their type). This combination has important theoretical implications. For example, the typical theoretical result in verifiable message games is that senders always reveal their type (the unraveling principle). This does not apply to games with partial disclosure such as ours.

Only a few theoretical studies combine elements of cheap talk games and disclosure games. Among these, Lipman and Seppi (1995) examine the role of competition between senders in a model where information is partially verifiable while Forges and Koessler (2005) characterize the equilibrium set of such games when a communication mediator is present. Though these aspects seem useful in increasing the amount of reliable information the receiver can extract from the sender, in this paper we focus solely on two-person interactions, based on theoretical models introduced by Glazer and Rubinstein (2004, 2006). Carroll and Egorov (2019) provide a theoretical analysis of situations that are similar to those studied in Glazer and Rubinstein (2004). They find that for a specific class of sender payoff functions the receiver can learn the sender's private information fully. The models that we consider in this study do not belong to this class.

Glazer and Rubinstein (2004) analyze a situation where a sender is privately informed about a multi-dimensional state of the world and sends a message about this to a receiver.[1] The receiver then chooses a single dimension of the state to observe, and can thus verify part of the message, before taking one of two actions, Accept or Reject. The sender prefers the receiver to accept independent of the state, whereas the receiver's optimal action depends on the state. The authors

[1] Glazer and Rubinstein (2004) refer to this setting as a persuasion game. This is not to be confused with the Bayesian persuasion games of Kamenica and Gentzkow (2011). These are sender-receiver games where the sender can commit to a message strategy (see Fréchette et al. (2018) and Aristidou et al. (2019) for experiments that allow sender commitment). In contrast, Glazer and Rubinstein (2004) do not allow sender commitment in their model.

identify optimal mechanisms from the receiver's point of view, i.e. mechanisms that maximize the receiver's expected payoff, when the receiver can commit. Glazer and Rubinstein (2006) modify this model by removing messages and the receiver's option to verify and instead allowing the sender to reveal truthfully one dimension of the state. They show that the receiver's optimal mechanism in this case yields the same expected payoff to the receiver. Thus, theoretically the receiver does not suffer by losing verification control.

Glazer and Rubinstein (2004, 2006) also discuss the corresponding extensive form games where the receiver cannot commit to a strategy. In both settings they show that the receiver's payoff from the optimal mechanism can still be achieved in a sequential equilibrium of these games. Our study is designed to test the effect of losing verification control on the receiver's payoff in these extensive form games. In the next section we discuss this setup in more detail.

3 PARTIALLY VERIFIABLE INFORMATION GAMES

We study two partially verifiable information games. In both games the sender's type is determined by the value of two aspects, and the privately-informed sender makes a claim about the values of these aspects. In the "Receiver Verifies game", the receiver then chooses one of the aspects to be checked and, after observing the actual value of that aspect, decides whether to accept or reject. In the "Sender Reveals" game, after the message is sent the sender decides which aspect is observed by the receiver. Hence, the two games differ only in who controls the verification process. We describe the games formally below.

3.1 *The "Receiver verifies" (Rv) game*

The "Receiver verifies" game (henceforth Rv) is based on Glazer and Rubinstein (2004). There are two players: a sender and a receiver. The sender's type depends on the values of two *aspects*. Aspect $i = 1, 2$ is a random variable that can take values in the set $X_i = \{1, ..., 9\}$. The set of possible types is then $X = X_1 \times X_2$. A generic element of X will be denoted as $x = (x_1, x_2)$. The probability of type $x \in X$

is denoted as $p_x = \frac{1}{81}$. The sender's type is "good" if x belongs to the set G, where $G = \{(x_1, x_2) | x_1 + x_2 \geq 11\}$; and "bad" otherwise.[2]

Payoffs depend on type and action, as summarized in Table 1. The receiver wants to take action a ("accept") if the sender's type is good, and action r ("reject") if the sender's type is bad. The sender always wants the receiver to take the action a, irrespective of type. Note that the receiver's utility is 1 if the optimal action has been chosen (i.e., if either $x \in G$ and a has been taken, or $x \in X \backslash G$ and r has been taken) and 0 otherwise. Hence, both types of errors (rejecting a good type and accepting a bad type) are assumed to be equally costly for the receiver, and an expected utility maximizing receiver minimizes the probability of making an error.

Table 1: Payoff Matrix (sender's payoff listed first in each cell)

	Receiver accepts	Receiver rejects
Good type	(1, 1)	(0, 0)
Bad type	(1, 0)	(0, 1)

The timing of the game is as follows. First, the sender sends a message. The set of available messages is denoted by M, and a generic element of M will be denoted by $m = (m_1, m_2)$. A (mixed) strategy for the sender is a function $\sigma : X \to \Delta M$. We assume that the set of messages coincides with the set of types, i.e., $M = X$, but we will keep the notation M to refer to the set of messages for the sake of clarity. We denote the probability that the sender sends message m when the sender's type is x by $\sigma(m|x)$.

After observing the sender's message, the receiver decides which aspect to check and which action to take depending on the message and on the value of the aspect that has been observed. The strategy of the receiver $f = (\pi, d)$ consists of:

- A *checking rule* that determines which aspect to check after receiving the message. The function $\pi_1 : M \to [0, 1]$ denotes the probability of checking aspect 1 as a function of the message received (the receiver must check exactly one aspect, hence $\pi_2 : M \to [0, 1]$ satisfies $\pi_2(m) = 1 - \pi_1(m)$ for all $m \in M$).

[2]Our parametrization is taken from the online experiment of Glazer and Rubinstein at http://gametheory.tau.ac.il/exp5/. In their experiment, the receiver is a computerized player playing an undisclosed strategy. We also conducted sessions with a different parametrization. The results are qualitatively similar to those we report below; for completeness we summarize these sessions in Appendix F.

- A *decision rule* for each aspect that determines the probability of acceptance depending on the message received and the value observed. We denote by $d_k : M \times X_k \to [0, 1]$ the probability of accepting after checking aspect k as a function of the value observed and of the message.

An example strategy for the receiver would be the *fair random strategy* (analogous to the fair random mechanism of Glazer and Rubinstein (2004)) where $\pi_1(m) = \pi_2(m) = 0.5$, and $d_k(m, x_k) = 1$ if and only if $m \in G$ and $x_k = m_k$; that is, the receiver checks one aspect at random and accepts when the sender claims to be of a good type and the observed value coincides with the message.

3.1.1 *Equilibria of Rv*

Our equilibrium concept is sequential equilibrium. A sequential equilibrium consists of strategies and beliefs such that (1) at any information set, the player who has the move is playing a best response, given the beliefs and assuming that players will subsequently stick to their strategies (sequential rationality) (2) beliefs are determined by Bayes' rule and the players' equilibrium strategies (consistency of beliefs)[3]. Note that there are two opportunities for the receiver to update the prior beliefs: after the message is sent, and after the value of an aspect has been observed. In the first case, the updating of the prior may determine which aspect to check; in the second case, the updating of the prior may determine which action to take. See Appendix A for further details.

The Rv game has multiple equilibria. First, there is a family of equilibria which results in the following outcome: hands are accepted if and only if at least one aspect takes a value of at least 7 (henceforth 7+). We will refer to this outcome as the *GR outcome*, since this is also the outcome resulting from an optimal receiver commitment strategy as analyzed in Glazer and Rubinstein (2004) (see Appendix B). It follows that these equilibria achieve the receiver's maximum equilibrium expected payoff.

These equilibria differ in the role of messages. In some of these equilibria, the message simply informs the checking decision; the acceptance decision is indepen-

[3] If all information sets can be reached given the strategies, beliefs are completely determined by Bayes rule. If not, sequential equilibrium requires that the strategies and beliefs are found as the limit of a sequence of fully mixed strategies together with the beliefs that follow from those strategies using Bayes rule.

dent of whether the value observed coincides with the message. For example, the following is an equilibrium. The receiver checks the highest claim (checking at random if both claims are equal) and, conditional on this checking behavior, accepts if and only if the observed value is 7+. The sender sends $(1,9)$ if he has at least one 7+ aspect and $x_2 > x_1$, $(9,1)$ if he has at least one 7+ aspect and $x_1 > x_2$, and randomizes between $(9,1)$ and $(1,9)$ in all other cases. Essentially the message "points to" the 7+ aspect if the sender has one. To see that this is an equilibrium, note that, given the receiver's strategy, hands without a 7+ aspect are rejected irrespective of the sender's message, and hands with two 7+ aspects are accepted irrespective of the sender's message. Hands with only one 7+ aspect will be accepted if the sender reports the 7+ aspect as the higher aspect, and this is what the sender's strategy does. The sender's strategy has been constructed so that the receiver maximizes expected payoff by checking the higher message and accepting if and only if the observed value is 7+. See Appendix C for details.

Other equilibria that also result in the GR outcome use the message to inform both the checking and acceptance decisions. For example, the sender sends a message as shown in Table 2 below. In this table, senders with a good type report it truthfully, while some senders with a bad type inflate their lower aspect. The receiver checks the higher claim ($\pi_1 = 1$ if $m_1 > m_2$; $\pi_1 = 0.5$ if $m_1 = m_2$ and $\pi_1 = 0$ if $m_1 < m_2$) and, conditional on this checking behavior, accepts if and only if a good type is reported and the observed value coincides with the claimed value and is 7+ ($d_i = 1$ if $m_1 + m_2 \geq 11$, $m_i = x_i$ and $x_i \geq 7$; otherwise $d_i = 0$).

Table 2: Example of an equilibrium sender's message strategy in Rv

	9	(2,9)	(2,9)	(3,9)	(4,9)	(5,9)	(6,9)	(7,9)	(8,9)	(9,9)
	8	(3,8)	(4,8)	(3,8)	(4,8)	(5,8)	(6,8)	(7,8)	(8,8)	(9,8)
	7	(4,7)	(5,7)	(6,7)	(4,7)	(5,7)	(6,7)	(7,7)	(8,7)	(9,7)
	6	**(5,6)**	**(6,6)**	**(7,6)**	(4,6)	(5,6)	(6,6)	(7,6)	(8,6)	(9,6)
x_2	5	**(7,5)**	**(6,5)**	(3,5)	(4,5)	(5,5)	(6,5)	(7,5)	(8,5)	(9,5)
	4	**(8,4)**	**(7,4)**	(3,4)	(4,4)	(5,4)	(6,4)	(7,4)	(8,4)	(9,4)
	3	**(8,3)**	(2,3)	(3,3)	(4,3)	(5,3)	**(6,7)**	(7,6)	(8,3)	(9,3)
	2	**(9,2)**	(2,2)	(3,2)	**(4,7)**	**(5,6)**	**(6,6)**	(7,5)	(8,4)	(9,2)
	1	(1,1)	**(2,9)**	**(3,8)**	**(4,8)**	**(5,7)**	**(6,5)**	(7,4)	(8,3)	(9,2)
		1	2	3	4	5	6	7	8	9
						x_1				

Note: Gray-highlighted messages are bad hands profitably lying (since they are accepted). Messages in bold are bad hands lying even though they are rejected.

Given this receiver strategy, senders without a 7+ aspect are rejected irrespective of their message, while senders with a 7+ aspect can be accepted by sending an appropriate message. The gray-highlighted messages in Table 2 are sent by bad types with a 7+ aspect, who best respond to the receiver strategy by inflating the value of the lower aspect, while keeping the claim below the value of the higher aspect, so that the highest claim remains truthful. These types are profitably lying. Good types with a 7+ aspect tell the truth and are accepted.

Table 2 has been constructed so that it is a best response for the receiver to follow the above strategy. Note that not all messaging strategies where the sender reports good types truthfully and inflates bad types with a 7+ aspect would induce a sequential equilibrium. For example, consider message (7,4). If the sender only inflates bad types up to the bare minimum needed to make up a good type, after receiving message (7,4) and verifying that the 7 is correct, the type is equally likely to be (7,4), (7,3), (7,2) and (7,1). Only one out of four is a good type, hence the receiver should reject. In order to have an equilibrium, the sender cannot concentrate the lies on messages that add up exactly to the threshold; one possibility is that (7,3) reports (7,6), (7,2) reports (7,5) and (7,1) reports (7,4) as in the table.

The sender's strategy should also be such that it is optimal for the receiver to always verify the aspect with the highest claimed value. For example, if the message (7,4) were sent only by the (7,4) and (7,1) types, the receiver would want to check the 4 since that would ensure discovering the bad type. The equilibrium would require some types (e.g. (2,4), as in the table) to send (7,4) as well even though it has no benefit for them. All types that send a bolded message in Table 2 are lying not because it is (strictly) profitable to do so, but in order to preserve the optimality of the receiver's strategy.

The equilibrium described also requires the receiver to reject some good types, such as (6,6). If (6,6) is the only type that sends message (6,6), then the receiver would know for sure that this is a good type and should accept instead. In order to have an equilibrium, one needs to assume that either (6,6) reports one of the aspects as 7+ (even though this will be discovered for sure given the strategy of the receiver) or there are bad types that also report (6,6) (as in Table 2), so that a claim of (6,6) is not unambiguously a good type. In both cases, there are sender types that are lying even though they have nothing to gain from doing so.

There are also many sequential equilibria that lead to different outcomes and lower payoffs to the receiver. For example, an equilibrium giving the receiver their lowest equilibrium payoff involves the sender sending a message at random, the receiver checking an aspect at random and accepting if and only if the observed value is at least 6. Types without a 6+ are rejected, types with two 6+ aspects are accepted, while types with one 6+ may or may not be accepted depending on which aspect the receiver observes. In this equilibrium, messages play no role whatsoever, and both players earn a lower payoff than in the GR outcome. There are also other equilibria where messages play a role and in which the sender is better off than in the GR outcome (see Appendix D for details).

3.2 *The "Sender reveals" (Sr) game*

The "Sender reveals" game (henceforth Sr) is based on Glazer and Rubinstein (2006).

The strategy of the sender $g = (\sigma, \rho)$ consists of:

_ A *message rule* that determines the message as a function of type. As in Rv, we denote the message rule as $\sigma : X \to \Delta M$ and the probability of sending message m as a function of type as $\sigma(m|x)$.

_ A *revelation rule* that determines which aspect is to be observed by the receiver depending on the message and on the type. The function $\rho_1 : X \times M \to [0,1]$ denotes the probability of revealing aspect 1 as a function of type and of the message[4]. Since exactly one aspect is revealed, $\rho_2(x, m) = 1 - \rho_1(x, m)$.

The receiver's strategy consists of a decision rule for each aspect, that is, $d_k : M \times X_k \to [0,1]$ for $k = 1, 2$. Given the aspect k to be observed, the receiver strategy determines the probability of acceptance as a function of the message m and the value of the aspect observed x_k.

3.2.1 *Equilibria of Sr*

The Sr game also has multiple sequential equilibria. As in Rv, there is a family of equilibria which results in senders being accepted if and only if at least one aspect takes a value of 7+. Also as in Rv, this is the outcome that results from a

[4]In our experiment the sender first inputs the message and then the aspect to be revealed. Equivalently we could think of the sender first deciding which aspect to reveal and then which message to send. The two decisions are interrelated (for example, we would expect the sender to report the aspect revealed truthfully).

receiver optimal commitment strategy (see Appendix B), hence we refer to it as the GR outcome. In some of these equilibria, the sender displays the higher of the two aspects and sends an uninformative message (e.g., sends a message at random, or always sends the same message). For example, $\rho_1 = 1$ if $x_1 > x_2$, $\rho_1 = 0$ if $x_1 < x_2$ and $\rho_1 = 0.5$ if $x_1 = x_2$, and $\sigma(m|x) = \frac{1}{81}$ for all $m \in M$, $x \in X$. The receiver then best responds by setting $d_i(m, x_i) = 1$ if $x_i \geq 7$ and $d_i(m, x_i) = 0$ if $x_i < 7$ for $i = 1, 2$ irrespective of m. Note that in this equilibrium the receiver's strategy depends only on the value observed and not on the message.

There are also equilibria that lead to the GR outcome but differ in the role of messages. Again analogously to Rv, consider the following strategy combination. The sender shows the higher of the two aspects and uses the messaging strategy in Table 2. The receiver accepts if and only if a good type is reported, the observed value coincides with the reported value and the observed value is 7+ ($d_i(m, x_i) = 1$ if $x_i \geq 7$, $m_1 + m_2 \geq 11$ and $x_i = m_i$; $d_i(m, x_i) = 0$ otherwise for $i = 1, 2$). This constitutes an equilibrium where the receiver conditions the acceptance decision not only on the aspect observed but also on the message.

Again similarly to Rv, not all messaging strategies where the sender reports good types truthfully and inflates bad types with a 7+ aspect would induce a sequential equilibrium. In order for the receiver's strategy to be a best response to the sender's strategy, the sender cannot concentrate the lies on messages that add up to the bare minimum required to make up a good type, and, if all good types tell the truth, there must be bad types that send messages $(6, 6)$, $(6, 5)$ and $(5, 6)$ even though there is no strict gain from doing so.[5]

There are also equilibria resulting in different outcomes and a lower payoff for the receiver. Some of these give a lower payoff for the sender. For example, the following equilibrium gives the sender a 0 payoff. Senders with good types reveal their lower aspect, while senders with bad types reveal their higher aspect (if both aspects are equal, senders reveal an aspect at random), and all types send a message at random. Given this sender strategy, the receiver best responds by rejecting irrespective of the value of the aspect observed. Given that the receiver is rejecting whatever she observes, the sender has no incentive to deviate.

[5]Because the sender chooses which aspect to reveal, bold messages in Table 2 other than (6,6), (6,5) and (5,6) could be changed to truthful messages without affecting the receiver's best response.

Other equilibria result in a higher payoff for the sender than the GR outcome. Appendix D describes an equilibrium which results in types being accepted if and only if the sender has a 6+ aspect.

Given the multiplicity of equilibria in both Sr and Rv, an obvious question is which of these, if any, would be observed in actual game play. What messages will be sent? For example, will senders with one high aspect (7+) inflate the lower aspect while keeping the highest claim truthful? What evidence will be observed and how will receivers respond to this? For example, will the receiver observe the sender's high aspect and accept if and only if it is 7+? How will receivers condition their decisions on messages and evidence? Will receivers be able to attain a payoff as high as that realized in the receiver's best equilibrium? Will the receiver be better off by being able to choose which aspect is verified?

The next section describes our experiment, designed to answer these questions.

4 EXPERIMENTAL DESIGN

The experiment was conducted in the CeDEx laboratory at the University of Nottingham, UK. There were 192 subjects, recruited from a university-wide pool of undergraduate and graduate students using ORSEE (Greiner, 2015). The experiment was programmed in z-Tree (Fischbacher, 2007).

Our experiment varies the game (Sr and Rv) across sessions. 4 sessions of each treatment were conducted with 24 subjects per session, and each session was divided into two matching groups. This gives us 8 independent observations per treatment.[6]

Upon arrival at a session, subjects were randomly allocated a seat number and given a set of instructions, which were then read aloud by the experimenter.[7] The decision-making part of the experiment consisted of 30 periods, where subjects were randomly matched in each period to play the relevant game. Subjects were re-

[6]Recall that we also conducted treatments with a lower threshold determining a good hand which resulted in qualitatively similar findings to those reported below. We present a summary of these additional treatments in Appendix F.

[7]See Appendix G for a copy of the instructions.

matched within their matching group at the beginning of each period, but retained the same role (sender or receiver) during the entire session.

At the beginning of each period, each sender is dealt two cards (blue and orange). Each of the cards is equally likely to be any integer value between 1 and 9, and all draws are independent across colors and senders.[8] A hand is defined as "good" if the sum of the values of the two cards is at least 11. Having observed the two cards, the sender sends a message to the receiver of the form "The value of the orange card is __; The value of the blue card is __". Next, the receiver chooses one of the cards to observe (Rv treatment) or the sender chooses one of the cards for the receiver to observe (Sr treatment). The receiver then accepts or rejects. The sender earns 1 point if the receiver accepts and 0 if the receiver rejects; the receiver earns 1 point if she accepts a good hand or rejects a bad hand, and 0 otherwise.

At the end of each period, a summary screen displayed the true values of the two cards, the message sent by the sender, the card chosen to be observed (by the receiver or by the sender, depending on treatment), the receiver's decision, and the point-earnings of the two subjects.

At the end of the experiment subjects received their accumulated earnings for the 30 periods (1 point = £0.50) plus a £3 participation fee. Each session lasted around 90 minutes, and average earnings were £12.92.

5 RESULTS

We report our results in three subsections. First we analyze subjects' decisions in the order they were sequenced in the experiment: message, revelation/verification, acceptance. Next, we look at the implications for game outcomes and payoffs. In the final subsection, we examine the receiver's best response given the empirically observed sender behavior.

Unless otherwise specified, all statistical tests are two-tailed signed-rank tests taking the matching group as the unit of observation. (Recall, for comparisons

[8] For the Rv treatment the card draws were randomized using the random number generator during the session. To enhance comparability across treatments, we then used these realizations in the corresponding sessions of the Sr treatment. This allows us to perform the statistical comparisons on paired observations.

between treatments, our design allows us to perform statistical comparisons on paired observations.) Also, unless otherwise specified, tests are based on the data pooled from all periods. (We also examined whether there were any dynamic trends in behavior and how these might affect our results. We find only mild evidence of learning effects, and the results based on the pooled data are robust to these - see Appendix E for details.)

5.1 *Decisions*

5.1.1 *Messages*

Figure 1 presents the distribution of the realized and the reported hands. The middle panel (b) displays the 1,440 random hands dealt in each treatment. Recall that the random draws are identical across treatments. As each combination of values was equally likely, the distribution of the realized draws is approximately uniform. Panel (a) presents the distribution of the values reported by senders in the Sr treatment; panel (c) presents the distribution of reports in the Rv treatment.

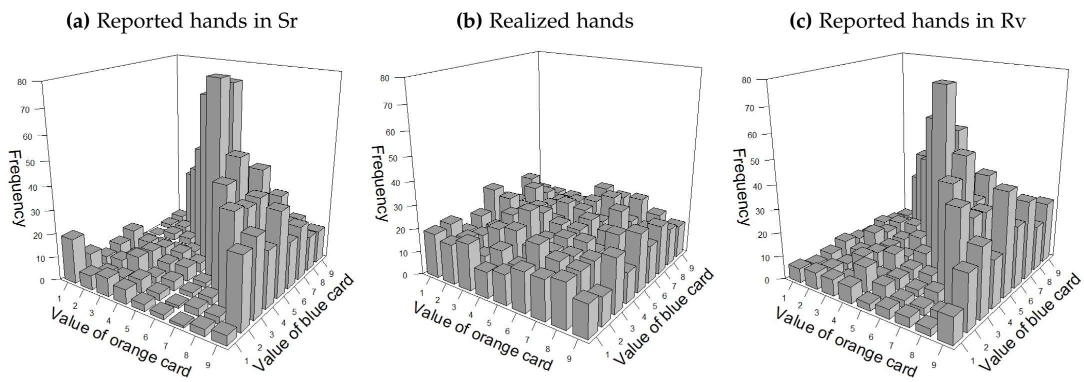


Figure 1: Distribution of realized vs. reported hands in Sr and Rv

The realized and the reported distributions clearly differ and there is a shift in the distribution of reported values towards the area where these add up to at least 11. Table 3 reports the truthtelling rate in both treatments. When senders have a good hand they tend to report it truthfully, although slightly and significantly more often in Rv. Notably, senders tend to lie about bad hands; less than 20% of bad hands are truthfully reported in both treatments.[9] In Sr this is especially the case

[9]Interestingly, very few subjects can be characterized as truth-tellers (never misreporting). Out of 48 senders per treatment, we observed only 2 truth-tellers in the Rv treatment and 1 in the Sr

when they have a 7+ card; less than 8% of senders with a 7+ card tell the truth compared with more than 24% of senders without a 7+ card (p-value=0.008).

Table 3: Truth-telling rates in Sr and Rv

Type of hand		obs.	Truth-telling rate Sr	Truth-telling rate Rv	p-value
Good hands	High card < 7	51	0.863	0.941	0.313
	High card ≥ 7	580	0.847	0.916	0.016
	All good hands	631	0.848	0.918	0.016
Bad hands	High card < 7	591	0.245	0.201	0.461
	High card ≥ 7	218	0.073	0.170	0.188
	All bad hands	809	0.199	0.193	0.813
All hands		1440	0.483	0.510	0.461

Table 4 focuses on the reporting behavior of senders with bad hands with a 7+ card. Recall there are equilibria resulting in the GR outcome (i.e., all hands with a 7+ card are accepted and all other hands are rejected) where these senders report a good hand by inflating the reported value of the lower card while keeping the higher claim truthful. Table 4 shows that most reports are consistent with these equilibria in both treatments, and this behavior is more prevalent in Sr than Rv (79% vs 61%).

Table 4: Proportions of different types of reports (bad hands with a 7+ card)

	Sr	Rv
Good hand reported; higher message is truthful	0.789	0.606
Good hand reported; lower message is truthful	0.018	0.069
Good hand reported; one out of two equal messages is truthful	0.092	0.078
Good hand reported; neither message is truthful	0.005	0.055
Bad hand reported	0.096	0.193

Result 1 *Senders with a bad hand and a 7+ card usually inflate the value of the lower card while keeping the higher claim truthful. This happens significantly more often in Sr than in Rv ($p - value = 0.031$).*

treatment. Thus, our setting does not induce a very strong norm of honesty (cf. Abeler et al. (2019)). This could be due to the conflict of interests which may crowd out lying aversion (Cabrales et al., 2020; Minozzi and Woon, 2013), or it could reflect a different norm induced by the framing of our game.

5.1.2 *Revelation/verification*

When the sender reports a bad hand, the verification strategy is immaterial since the sender is almost certainly telling the truth (indeed, more than 98% of all reported bad hands are actually bad hands in both treatments). In what follows we focus on cases where the sender reports a good hand. We also focus on cases where, for Rv, the two reported values are different and, for Sr, the two card values are different.

In Sr, senders nearly always reveal the higher of the two cards: 96.23% of cases. In Rv, receivers check the higher of the two reports only in 64.86% of the cases. Thus, it appears that revealing the higher card is a more compelling strategy for senders than verifying the higher claim is for receivers.

Result 2 *Senders reveal the higher card significantly more often than receivers check the higher claim ($p - value = 0.008$).*

It is informative to look at this behavior at the individual level. Figure 2 shows the individual propensity to reveal the higher card (Sr) or check the higher claim (Rv). Not surprisingly given the high aggregate frequency of revealing the higher of the two cards, in Sr 60% of senders always reveal the higher card. In contrast, none of the receivers in Rv checks the higher claim all the time. Instead, receivers' behavior is more like a random auditing strategy, though most receivers check the higher claim more often than not.

An implication of this different revelation/verification behavior is that the sender's high card is almost always observed in Sr, while in Rv the sender's low card is observed about a third of the time.

5.1.3 *Acceptance*

When the sender reports a bad hand, the receiver nearly always rejects (over 96% of cases in both treatments). In what follows, we focus on cases where a good hand is reported.

One important factor that influences the receiver's decision is the value of the observed card. The upper panel of Figure 3 presents the acceptance rate conditional

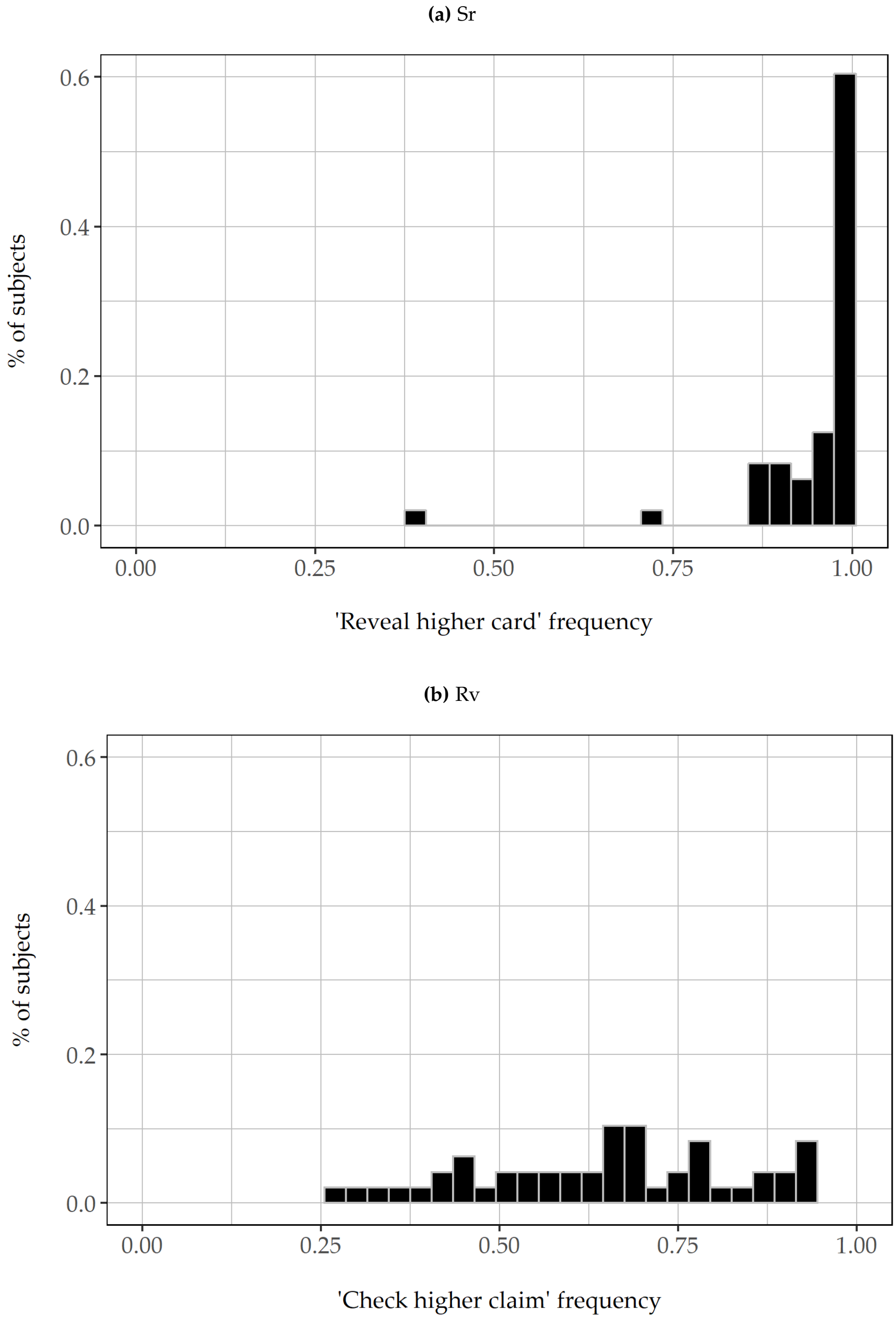


Figure 2: Individual propensity to reveal the higher card (Sr) or check the higher claim (Rv)

on the value observed by the receiver. Acceptance rates increase with the observed value and are generally higher in Rv than Sr.

Result 3 *Conditional on the observed value, the acceptance rate is higher in Rv than Sr. This difference is significant for observed values 3-7 ($p - value < 0.047$ for each case).*

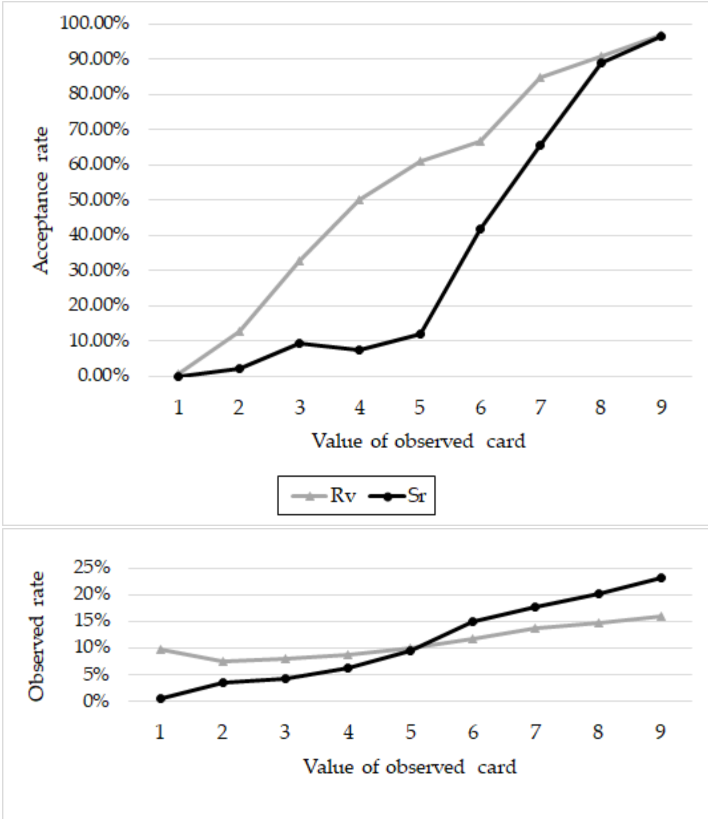


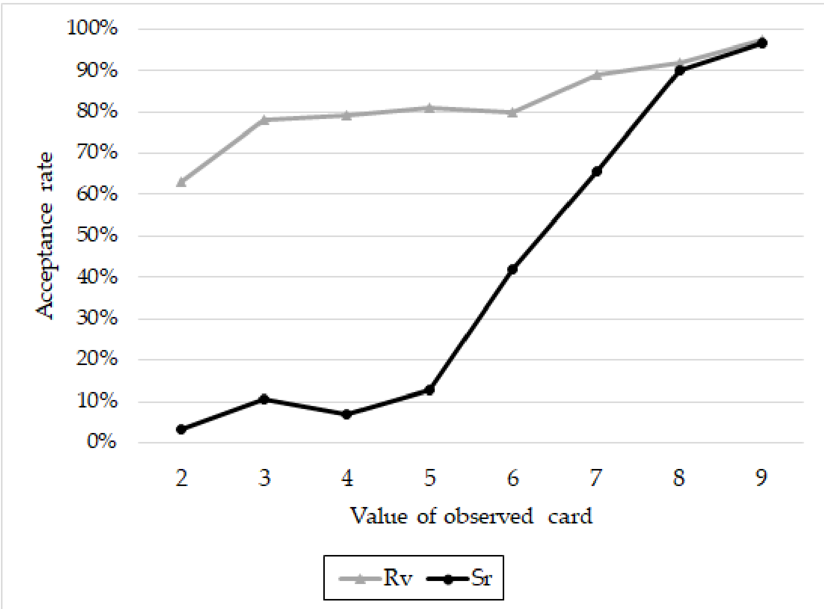
Figure 3: Acceptance rates conditional on the value of the observed card and corresponding relative frequencies of each observed value (reported good hands only; 1245 observations for Sr, 1242 observations for Rv)

The lower panel in Figure 3 depicts the relative frequency of each value observed by receivers. Even though the same random draws are used for both treatments, differences in the distributions of observed values arise from different behavior across treatments. Specifically, the observed card in Sr is almost certain to be the highest of the two, while this is not the case in Rv. Consequently, the distribution of observed cards in Sr is very similar to the distribution of the maximum value of the two cards, while the distribution in Rv is comparatively flat.

An implication of this is that, for a given value of the observed card, the hand is more likely to be good in Rv than in Sr, hence the higher propensity to accept in Rv may be justified (we will return to this point when we discuss Figure 4a below). It is also worth noting that overall acceptance rates conditional on the sender claiming to have a good hand are very similar (60% in Sr and 62% in Rv). The reason for this is that, while the receiver is more likely to accept for a fixed value of the observed card in Rv, she is also more likely to observe lower cards in Rv (and lower cards are less likely to be accepted). We will also return to this later in our discussion of payoffs.

Another factor that influences receivers' decisions is how the observed card compares with the message. In Sr, senders hardly ever misreport the card they reveal (less than 4% of the time), but in Rv a misreport is observed in 30% of cases. Receivers typically reject when observing a misreport (93% of the time in Sr and 95% in Rv). Rejecting after observing a misreport is optimal for the receiver since only about 10% of such hands are actually good hands. Figure 4a shows the acceptance rates conditional on the value of the observed card given that no misreport is observed.

(a)



(b)

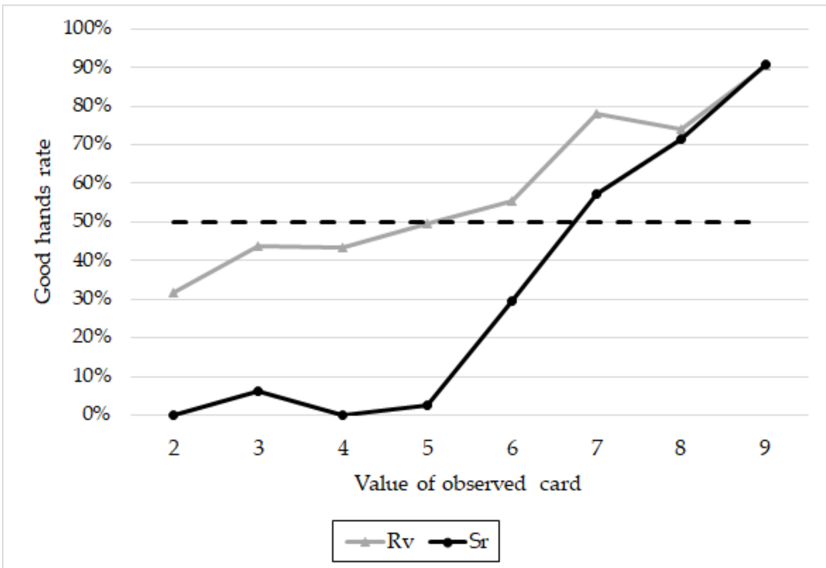


Figure 4: Acceptance rates (a) and proportion of good hands (b) given that a good hand was claimed and no misreport was observed (1199 observations for Sr, 862 observations for Rv)

As the figure shows, the receiver is very likely to accept in Rv if no misreport is observed. This behavior is not as costly as one may think since a) some bad hands have been weeded out when a misreport is discovered (the sender's messaging behavior facilitates this) and b) the receiver's verification strategy implies that the card observed is not always the higher of the two. This is confirmed in Figure 4b which shows the proportion of good hands when a good hand is reported and no misreport is observed. The figure includes a dotted line at 50% - above this line, it is optimal to accept, while below this line it is optimal to reject.[10]

In Sr, receivers are clearly better off rejecting values of 6 and less, and this is what they usually do, though for the value of 6 they accept more than 40% of the time. For values of 8 or 9, they are clearly better off accepting and this is what they almost always do. When a 7 is observed, accepting would give a slightly higher average payoff than rejecting, and receivers do so about two thirds of the time. In Rv, the probability of a good hand remains close to 0.5 even for values as low as 3, and so even though the receiver is too lenient for low observed values, the cost of an acceptance decision is small in terms of payoffs.

To further investigate factors influencing receivers' acceptance decisions we conduct a probit analysis. We include as explanatory variables the value of the observed card, the value of the unverified claim, and a dummy for whether a misreport is observed. We also include four control variables. We include a color dummy to test whether receivers condition on the color of the observed card. We include a good hand dummy to check whether there is any other information apart from the included variables that may help (if they are more likely to accept a good hand) or harm (if they are less likely to accept a good hand) receivers. We include a period trend to check for potential learning effects. Finally, we include a gender dummy. In Table 5 we present the results based on all periods (first two columns) and after excluding the first five periods (third and fourth columns).

We begin by discussing the regressions based on all periods. The value of the observed card has a significantly positive effect on the acceptance probability, but the effect in Sr is more than twice that in Rv. In addition, the value of the unverified claim has a significant, though smaller positive effect on the acceptance probabil-

[10]Optimality here refers to strategies that do not condition on other features of the message. In principle, if the message is informative (for example, by always inflating bad hands up to the bare minimum) it would be possible for the receiver to use the message to improve the accuracy of the acceptance decision. Below, we check whether there is any evidence receivers are able to distinguish between good and bad hands.

Table 5: Probit analysis of acceptance decision

	Dependent variable: Acceptance Decision			
	(all periods)		(excluding first 5 periods)	
	Rv	Sr	Rv	Sr
Value of observed card	0.110***	0.256***	0.131***	0.286***
	(0.025)	(0.023)	(0.027)	(0.024)
Value of unverified claim	0.059***	0.048***	0.078***	0.049*
	(0.021)	(0.018)	(0.024)	(0.026)
Misreport observed	−0.722***	−0.335	−0.742***	−0.370
	(0.039)	(0.303)	(0.033)	(0.465)
Observed card = orange	0.040	0.020	0.053	0.057
	(0.043)	(0.039)	(0.034)	(0.043)
Hand = good	0.049	0.065*	0.005	0.041
	(0.048)	(0.032)	(0.038)	(0.043)
Period	−0.006***	−0.005***	0.000	−0.000
	(0.001)	(0.002)	(0.003)	(0.002)
Female	−0.027	−0.035	−0.044	−0.041
	(0.039)	(0.063)	(0.036)	(0.061)
Observations	1,242	1,245	1,050	1,044

Notes: The table presents marginal effects; *p<0.1; **p<0.05; ***p<0.01; standard errors in parentheses are clustered at the matching group level. The analysis excludes cases where the sum of the two reports was less than 11 (claimed bad hands).

ity in both treatments. Observing a misreport substantially reduces the receiver's acceptance probability; the estimated effect in Sr is not significant. In terms of the control variables, the regressions give no evidence of a color bias or gender effect. In Rv, it appears that after controlling for other variables, bad hands are as likely to be accepted as good hands. In Sr, receivers are slightly more likely to accept good hands; however, this effect is small and only marginally significant. The period variable has a small but statistically significant coefficient in both treatments.

The significant period effect is essentially picking up an effect that takes place over the first few periods. The last two columns of Table 5 report the same regression dropping the first five periods. We find that the period variable, and the good hand dummy in Sr, are no longer significant, while the significance and approximate size of the other coefficients remain unchanged.

5.2 *Outcomes and payoffs*

What are the implications of this behavior for outcomes and payoffs? Figure 5 shows the observed acceptance rates conditional on the value of the higher card, for good and bad hands separately. The dashed lines represent the GR outcome, where a hand is accepted if and only if it contains a 7+ card. In Sr, hands with a 7+ card are usually accepted, and hands without a 7+ card are usually rejected. Thus, the outcome of most hands in Sr is consistent with the GR outcome. Outcomes appear to conform less well with the GR outcome in Rv.

(a) Bad hands **(b)** Good hands

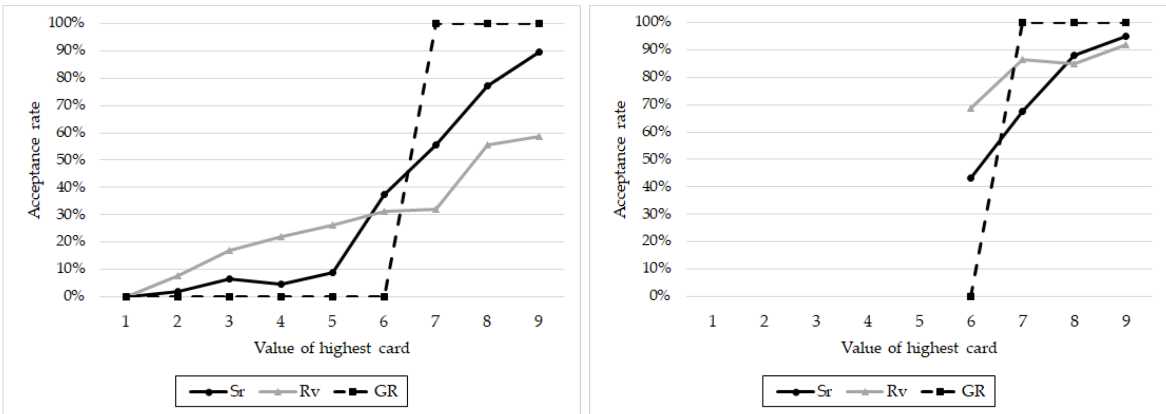


Figure 5: Acceptance rates for a given value of the highest card

In Table 6, we present the acceptance rates, conditioning on whether the hand is good or bad and on whether the hand is accepted or rejected in the GR outcome. Hands that are rejected in the GR outcome (highest card < 7) are more likely to be accepted in Rv than Sr. Bad hands that are accepted in the GR outcome (highest card ≥ 7) are more likely to be accepted in Sr than Rv.

Table 6: Acceptance rates for good and bad hands conditional on the value of the highest card

Type of hand		obs.	Acceptance rate: Sr	Acceptance rate: Rv	p-value
Highest card < 7	Bad hands	591	0.139	0.228	0.039
	Good Hands	51	0.431	0.686	0.016
	All hands	642	0.162	0.265	0.008
Highest card ≥ 7	Bad Hands	218	0.679	0.440	0.039
	Good hands	580	0.869	0.884	0.469
	All hands	798	0.817	0.763	0.461

Turning to payoffs, note that the sender's average payoff is simply the acceptance rate. Although some types of senders get a higher payoff in Sr than Rv (e.g. senders with a bad hand and a 7+ card), other types prefer Rv to Sr (e.g. senders with a bad hand and no 7+ card). Overall, it turns out that, averaging over types, the sender's payoff does not significantly differ between treatments (see Table 7). The receiver's average payoff is a weighted average of her payoffs from good and bad hands, where her average payoff from good hands is the acceptance rate for good hands and her average payoff from bad hands is one minus the acceptance rate for bad hands. The receiver's average payoff also does not differ significantly between treatments (see Table 7).

Table 7: Average payoffs comparison across treatments

	Sr	Rv	$p-value$
Receiver	0.767	0.782	0.641
Sender	0.525	0.541	0.469

Result 4 *Average payoffs do not differ significantly between treatments.*

5.3 *Best-response analysis*

Given that the sender always wants the receiver to accept, how should the receiver respond to different messages and evidence? We answer this question by identifying the optimal strategy for the receiver given the observed sender's behavior. We begin with Sr.

5.3.1 *Sender reveals*

In the Sr game, the receiver only has an acceptance decision to make. The information available to the receiver when making this decision consists of the revealed value and the sender's message. How should the receiver use this information optimally? To identify an optimal benchmark, we allow the sender to condition on the revealed value and the claim about the other card.[11] Table 8 presents the proportion of good hands conditional on the revealed value and the unverified claim.

[11] Thus, we ignore the claim about the revealed value. This does not affect the conclusions of our analysis because misreports of the revealed value are rare (less than 5% of cases) and, when a misreport is revealed, the optimal decision would be reached in over 95% of cases without using this information.

If this proportion is greater than 50%, it is optimal for the receiver to accept, and these cells are highlighted (when the proportion of good hands is equal to 50%, any decision is a best response). We refer to the strategy of making the optimal decision for each combination of revealed value and unverified claim as the empirical best response. The resulting expected payoff is 0.827, and we refer to this as the empirical optimum.

Table 8: Proportion of good hands given revealed value and unverified claim

		Revealed Value								
		1	2	3	4	5	6	7	8	9
Unverified Claim	1	0%	0%	0%	0%	11%	0%	50%	33 %	0%
	2	0%	0%	0%	0%	0%	0%	0%	0%	78%
	3	0%	0%	0%	0%	0%	0%	0%	55%	88%
	4	0%	0%	0%	0%	0%	0%	49%	52%	75%
	5	0%	0%	0%	0%	0%	27%	43%	74%	89%
	6	0%	NA	0%	0%	0%	32%	74%	86%	98%
	7	0%	0%	0%	0%	4%	50%	63%	88%	100%
	8	NA	0%	10%	0%	10%	20%	67%	75%	100%
	9	0%	3%	0%	0%	0%	0%	50%	100%	100%

Note: Highlighted cells represent the cases where accepting yields an expected payoff above 50%.

Next, we evaluate the performance of some alternative strategies. Table 9 presents the alternatives. We categorize these based on the amount of information used to inform the receiver's strategy. For the strategies in which the receiver uses a threshold acceptance rule, conditioning on the value of the observed card, we present the threshold that gives rise to the highest payoff and the two closest thresholds.

First, note that if the receiver ignores messages and evidence, the best she can do is to always reject since the prior probability of a good hand is below 50%. The receiver can do substantially better by using evidence to inform the acceptance decision; the optimal threshold rule is to accept if the revealed card is 7+. This is an optimal commitment strategy and would result in the GR outcome if the sender best responds to it. The receiver can do slightly better by considering both the message and the evidence; the highest payoff in this class of strategies is achieved by accepting when a 7+ is revealed and the hand is claimed to be good. This is also an optimal commitment strategy. If the sender best responds to the receiver, these optimal commitment strategies do equally well. However, because some senders own up to having a bad hand, the strategy that considers both messages and evidence does slightly better. In fact, the strategy of accepting if and only if a good

Table 9: Receiver's expected payoff given observed sender behavior in Sr

Strategy	Payoff
Empirical Best Response	0.827
Ignore message and evidence	
- - - accept or reject at random	0.500
- - - always reject	0.562
Ignore message	
accept iff	
- - - $x_k \geq 8$	0.794
- - - $x_k \geq 7$	0.810
- - - $x_k \geq 6$	0.750
Use message for acceptance decision	
accept iff reported sum ≥ 11 &	
- - - $x_k \geq 8$	0.800
- - - $x_k \geq 7$	0.822
- - - $x_k \geq 6$	0.768

hand is claimed and a 7+ is revealed prescribes very similar decisions to those of the empirical best response, and as a result achieves a very similar payoff.

Result 5 *Given the sender's observed behavior in Sr, following the optimal commitment strategy of accepting if and only if a good hand is claimed and a 7+ card is revealed gives the receiver 99.40% of the empirical optimum.*

5.3.2 *Receiver verifies*

For Rv, the receiver must decide which card to verify and whether to accept or reject. We allow the verification strategy to depend on the message, and the acceptance decision to depend on the message and on whether the verified card was equal to the claim or misreported.

Given the actual sender behavior in the Rv treatment, should the receiver check the card corresponding to the highest or to the lowest claim? Moreover, should she pay attention to the unverified claim? Lastly, which values should the receiver accept? First, if the sender reports a bad hand, the hand is almost certainly bad (only 1 out of 198 reported bad hands is good). In this case, it is optimal to reject independent of which card is checked and whether the check reveals a truthful claim or a lie.

What about the hands that are reported as good? In Table 10 we present the payoffs corresponding to each checking and acceptance decision for all messages representing reported good hands. We see that checking the highest claim and accepting if this is true is the better choice for most messages since it leads to a higher expected payoff. Why is it better for the receiver to accept only if the checked claim turns out to be true? This is because the vast majority of discovered misreports (i.e. 89.76%) represent bad hands.

Table 10: Receiver's best response to sender messages in Rv

Message[12]	Absolute frequency of message	Absolute frequency of good hands given message	Expected payoff from: check highest claim & accept if true[13]	Expected payoff from: check lowest claim & accept if true[14]	Expected payoff from: check either claim & always reject[15]	Expected payoff from: check either claim & always accept[16]
(9,9)	55	32	**0.773**	**0.773**	0.418	0.582
(9,8)	40	35	**0.900**	**0.900**	0.125	0.875
(9,7)	45	38	**0.978**	0.933	0.156	0.844
(9,6)	36	34	**0.972**	0.944	0.056	0.944
(9,5)	55	45	0.909	**0.945**	0.182	0.818
(9,4)	29	24	**0.966**	0.862	0.172	0.828
(9,3)	55	35	0.800	**0.836**	0.364	0.636
(9,2)	73	23	**0.932**	0.438	0.685	0.315
(8,8)	31	22	**0.726**	**0.726**	0.290	0.710
(8,7)	40	29	**0.850**	**0.850**	0.275	0.725
(8,6)	51	40	**0.941**	0.843	0.216	0.784
(8,5)	52	37	**0.885**	0.827	0.288	0.712
(8,4)	69	27	**0.739**	0.681	0.609	0.391
(8,3)	117	32	**0.786**	0.496	0.726	0.274
(7,7)	47	24	**0.872**	**0.872**	0.489	0.511
(7,6)	41	31	0.902	**0.927**	0.244	0.756
(7,5)	81	30	0.691	**0.704**	0.630	0.370
(7,4)	129	38	**0.798**	0.527	0.705	0.295
(6,6)	56	25	**0.696**	**0.696**	0.554	0.446
(6,5)	140	30	0.586	0.621	**0.786**	0.214

Note: Highlighted cells represent receiver's optimal payoff for a given message.

[12]Each message gives the highest claim first and ignores whether this claim refers to the blue or orange card.

[13]Computed by counting the instances in which the highest claim is untrue and the hand is bad and those where the highest claim is true and the hand is good. This is then divided by the frequency of the corresponding message.

[14]Computed by counting the instances in which the lowest claim is untrue and the hand is bad and those where the lowest claim is true and the hand is good. This is then divided by the frequency of the corresponding message.

[15]Computed by counting the number of bad hands and then dividing by the frequency of the corresponding message.

[16]Computed by counting the number of good hands and then dividing by the frequency of the corresponding message.

Next, the receiver should check the highest claim as this claim is more likely to be false. Since a misreport is very often a bad hand, checking the highest claim allows the receiver to take as many bad hands out of the sample as possible by rejecting. This also increases the probability that the hand is good conditional on the claim being true and reduces the number of errors in case of acceptance. In cases where this probability is less than 50% for both messages, it is optimal to reject all hands and which message is checked is immaterial. This is the case of the (6,5) message when the receiver is better off rejecting even if the observed value is as reported.

The receiver's expected payoff from best responding to each message is 0.853. How does this payoff compare to what the receiver could get from other possible strategies? Table 11 presents some alternatives. Again, we categorize these based on the amount of information used to inform the receiver's strategy.[17]

We start by recalling that there are more bad hands than good hands, hence the receiver can obtain more than 50% by rejecting, regardless of messages or evidence. By checking at random and accepting if the observed card is high enough, the receiver does even better. The optimal threshold is to accept if the observed card is 6+ (leading to a payoff of 0.752). The receiver does even better using the "fair random mechanism" of checking at random and accepting if a good hand is claimed and no misreport is observed (0.792).

Using messages to inform the checking decision is potentially beneficial. Table 11 shows that checking the high claim is better than checking the low claim (or checking at random). The best performing strategy in this category checks the high claim and accepts if a 7+ is observed (0.817). Interestingly, this is an optimal commitment strategy: if the sender best responds it leads to the GR outcome.

The final category of strategies use messages for both checking and acceptance decisions. It turns out that it pays to use messages in this way. The best performing strategy is to check the higher claim and accept if and only if no misreport is observed, a good hand is claimed, and the observed value is 7+, and this gives a payoff of 0.848. This is also an optimal commitment strategy. Note that although there are multiple optimal commitment strategies, all leading to the GR outcome if the sender best responds, they do not perform equivalently given senders' observed

[17]For the strategies in which the receiver uses a threshold acceptance rule based on the value of the observed card we present the threshold that gives the highest payoff and the two closest thresholds.

Table 11: Receiver's expected payoff given observed sender behavior in Rv

Strategy	Payoff
Empirical Best Response	0.853
Ignore message	
accept or reject at random	0.500
always reject	0.562
check at random & accept iff:	
- - - $x_k \geq 7$	0.740
- - - $x_k \geq 6$	0.752
- - - $x_k \geq 5$	0.740
Use message for acceptance decision only	
check at random & accept iff:	
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ (fair random)	0.792
Use message for checking decision only	
check lower claim & accept iff:	
- - - $x_k \geq 6$	0.731
- - - $x_k \geq 5$	0.758
- - - $x_k \geq 4$	0.730
check higher claim & accept iff:	
- - - $x_k \geq 8$	0.796
- - - $x_k \geq 7$	0.817
- - - $x_k \geq 6$	0.772
Use message for checking and acceptance decision	
check lower claim & accept iff:	
- - - $m_k = x_k$ & $m_i + m_j \geq 11$	0.753
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 5$	0.765
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 4$	0.766
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 3$	0.765
check higher claim & accept iff:	
- - - $m_k = x_k$ & $m_i + m_j \geq 11$	0.831
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 8$	0.802
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 7$	0.848
- - - $m_k = x_k$ & $m_i + m_j \geq 11$ & $x_k \geq 6$	0.831

behavior. In fact, using the message to inform acceptance pays off because it allows the receiver to take advantage of cases where the sender owns up to having a bad hand or misreports the high card.

Result 6 *Given the sender's observed behavior in Rv, following an optimal commitment strategy of checking the higher claim and accepting if and only if no misreport is observed, a good hand is claimed, and the observed value is 7+, gives the receiver 99.41% of the empirical optimum.*

6 CONCLUSION

Sender-receiver games in which a sender's cheap talk claims are partially backed by evidence reflect many natural environments. In buyer-seller interactions a buyer is often exposed to a sales pitch or advertisements but can also test products (e.g. test-drive a car or download a sample of software). In lobbying environments, policy makers listen to claims of lobbyists but can also investigate claims. Such settings have been analysed in a growing theoretical literature, but they have attracted little experimental research. In this paper we introduce an experimental approach focussing on two theoretical models of Glazer and Rubinstein (2004, 2006).

While Glazer and Rubinstein focus on optimal mechanisms where the receiver can commit to a verification/acceptance rule, we focus on non-cooperative games in which the receiver cannot commit. We study two games in which a sender has private information about two aspects. These aspects determine whether the sender's type is good or bad, where the receiver's optimal action is to accept good types and reject bad types and all sender types want the receiver to accept. The sender makes claims about both aspects and the receiver, before making a decision, then gets evidence about one of them. The games differ in the control the receiver has over the available information. In the "Sender reveals" game the sender chooses which cheap talk claim to back up with evidence, while in the "Receiver verifies" game the receiver chooses which cheap talk claim to verify. Both games have multiple equilibria, including, but not limited to, equilibria that result in the same outcome as the optimal commitment strategy.

We find that in the "Sender reveals" game, senders almost always reveal their strongest aspect. When the sender's type is bad, they usually accompany this with an inflated claim about the weaker aspect. The receiver's acceptance behavior depends on both the observed evidence and the claim about the unverified aspect. Receivers almost always reject when senders claim to be a bad type. When senders claim to be a good type, acceptance behavior can be described as a noisy best response: they are more likely to accept the higher the observed aspect, while the best response would be a threshold strategy.

In "Receiver verifies", senders also usually misreport when their type is bad. Receivers respond to the message with a "random auditing" strategy. They are more likely to check the higher claim but check the lower claim about a third of the time. Moreover, receivers almost always reject if they uncover a misreport, and usually accept when the evidence is consistent with the sender's claim.

There are some notable differences between the two treatments. For example, the receiver is less likely to observe the strongest aspect in "Receiver verifies", and acceptance behavior is more sensitive to the value observed in "Sender reveals". This translates into some differences in outcomes; for example when the sender's type is bad but the sender has one sufficiently strong aspect, the sender would fare better in "Sender reveals" (and the receiver would fare better in "Receiver verifies"). Conversely, when the sender's type is bad and neither aspect is sufficiently strong, the sender would fare better in "Receiver verifies" (and the receiver would fare better in "Sender reveals"). Overall, averaging over all possible sender types, there are no significant differences in payoffs between treatments.

Given the observed senders' behavior, we find a simple receiver strategy for each game that would give the receiver more than 99% of her best possible payoff. For "Sender reveals", the strategy involves accepting if and only if the sender claims a good type and the aspect revealed is sufficiently high. For "Receiver verifies", this strategy involves checking the high claim, and accepting if and only if the observed aspect is sufficiently high, no misreport is detected, and the sender claims to be a good type. Interestingly, these strategies are optimal commitment strategies in a situation where the receiver can commit to verification/acceptance rules as shown by Glazer and Rubinstein (2004, 2006). Thus, the strategies they identify are not only theoretically optimal commitment strategies, they also provide good recommendations for how a receiver should play our games.

Abeler, J., Nosenzo, D., and Raymond, C. (2019). Preferences for truth-telling. *Econometrica*, 87(4):1115–1153.

Aristidou, A., Coricelli, G., and Vostroknutov, A. (2019). Incentives or persuasion? an experimental investigation. *Mimeo*.

Blume, A., Lai, E. K., and Lim, W. (2020). Strategic information transmission: A survey of experiments and theoretical foundations. In Capra, C. M., Croson, R., Rigdon, M., and Rosenblat, T., editors, *Handbook of Experimental Game Theory*. Edward Elgar Publishing, Cheltenham, UK and Northampton, MA, USA.

Cabrales, A., Feri, F., Gottardi, P., and Meléndez-Jiménez, M. A. (2020). Can there be a market for cheap-talk information? An experimental investigation. *Games and Economic Behavior*, 121:368–381.

Carroll, G. and Egorov, G. (2019). Strategic Communication with Minimal Verification. *Econometrica*, 87(6):1867–1892.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Forges, F. and Koessler, F. (2005). Communication equilibria with partially verifiable types. *Journal of Mathematical Economics*, 41(7):793–811.

Forsythe, R., Isaac, R. M., and Palfrey, T. R. (1989). Theories and tests of "blind bidding" in sealed-bid auctions. *The Rand Journal of Economics*, 20(2):214–238.

Forsythe, R., Lundholm, R., and Rietz, T. (1999). Cheap talk, fraud, and adverse selection in financial markets: Some experimental evidence. *The Review of Financial Studies*, 12(3):481–518.

Fréchette, G., Lizzeri, A., and Perego, J. (2018). Rules and commitment in communication: An experimental analysis. *Mimeo*.

Glazer, J. and Rubinstein, A. (2004). On optimal rules of persuasion. *Econometrica*, 72(6):1715–1736.

Glazer, J. and Rubinstein, A. (2006). A study in the pragmatics of persuasion. *Theoretical Economics*, 1:395–410.

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

Grossman, S. J. (1981). The Informational Role of Warranties and Private Disclosure about Product Quality. *The Journal of Law and Economics*, 24(3):461–483.

Hagenbach, J. and Perez-Richet, E. (2018). Communication with evidence in the lab. *Games and Economic Behavior*, 112:139–165.

Jin, G. Z., Luca, M., and Martin, D. (2018). Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure. *NBER Working Paper No. 21099*.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

King, R. R. and Wallin, D. E. (1991). Market-induced information disclosures: An experimental markets investigation. *Contemporary Accounting Research*, 8(1):170–197.

Li, Y. X. and Schipper, B. C. (2020). Strategic reasoning in persuasion games: An experiment. *Games and Economic Behavior*, 121:329–367.

Lipman, B. L. and Seppi, D. J. (1995). Robust inference in communication games with partial provability. *Journal of Economic Theory*, 66(2):370–405.

Milgrom, P. R. (1981). Good News and Bad News: Representation Theorems and Applications. *The Bell Journal of Economics*, 12(2):380–391.

Minozzi, W. and Woon, J. (2013). Lying aversion, lobbying, and context in a strategic communication experiment. *Journal of Theoretical Politics*, 25(3):309–337.

Penczynski, S. P. and Zhang, S. (2017). Disclosure of Verifiable Information under Competition: An Experimental Study. *Mimeo*.

APPENDIX

A SEQUENTIAL EQUILIBRIUM DEFINITION IN RV AND SR

A.1 *The Rv game*

An equilibrium consists of a sender's message strategy $\sigma(m|x)$, a receiver's checking rule $\pi_1(m)$, a receiver's decision rule for each aspect, $d_1(m, y_1)$ and $d_2(m, y_2)$, where y_k is the observed value of aspect $k = 1, 2$, receiver's beliefs after each message, $b(x|m)$, and receiver's beliefs after having checked an aspect, $b_1(x|m, y_1)$ and $b_2(x|m, y_2)$, satisfying the conditions (i)-(v) below.

In what follows, denote by $b_k(y_k|m) = \sum_{x:x_k=y_k} b(x|m)$ the belief probability the receiver assigns to observing value y_k if she checks aspect k after receiving message m.

Also, denote by $g_k(m, y_k) = \sum_{x \in G} b_k(x|m, y_k)$ the belief probability that the receiver assigns to the sender being a good type after receiving message m, checked aspect k and observed value y_k.

(i) (Sender sequential rationality) For all $x \in X$, $m \in M$, $\sigma(m|x) > 0 \Rightarrow m \in \arg\max_{m' \in M} U_S(m'|x)$, where $U_S(m'|x) = \pi_1(m')d_1(m', x_1) + [1 - \pi_1(m')]d_2(m', x_2)$ is the probability that type x is accepted if he sends message m', given the receiver's strategy.

(ii) *(Sequential rationality of the receiver's decision rule)* For all $m \in M$, $y_k \in X_k$ and $k = 1, 2$, the receiver sets $d_k(m, y_k) = 1$ if $g_k(m, y_k) > 0.5$, and $d_k(m, y_k) = 0$ if $g_k(m, y_k) < 0.5$.

(iii) *(Sequential rationality of the receiver's checking rule)*

For all $m \in M$, $\pi_1(m) = 1$ if

$$\begin{aligned} &\sum_{y_1} b_1(y_1|m)[d_1(m, y_1)g_1(m, y_1) + [1 - d_1(m, y_1)][1 - g_1(m, y_1)]] \\ > &\sum_{y_2} b_2(y_2|m)[d_2(m, y_2)g_2(m, y_2) + [1 - d_2(m, y_2)][1 - g_2(m, y_2)]] \end{aligned}$$

and $\pi_1(m) = 0$ if

$$\begin{aligned} & \sum_{y_1} b_1(y_1|m)[d_1(m,y_1)g_1(m,y_1) + [1-d_1(m,y_1)][1-g_1(m,y_1)]] \\ < \; & \sum_{y_2} b_2(y_2|m)[d_2(m,y_2)g_2(m,y_2) + [1-d_2(m,y_2)][1-g_2(m,y_2)]] \end{aligned}$$

(iv) *(Consistency of receiver beliefs after receiving the message)*

For all $x \in X$, $m \in M$,

$$b(x|m) = \frac{\sigma(m|x)p_x}{\sum_{z\in X}\sigma(m|z)p_z} \text{ whenever } \sum_{z\in X}\sigma(m|z)p_z > 0.$$

(v) *(Consistency of receiver beliefs after having checked an aspect)*

For each $k = 1,2$ it holds that $b_k(x|m, y_k) = 0$ for all x such that $x_k \neq y_k$, and

$$b_k(x|m,y_k) = \frac{\sigma(m|x)p_x}{\sum_{z:z_k=y_k}\sigma(m|z)p_z} \text{ whenever } \sum_{z:z_k=y_k}\sigma(m|z)p_z > 0 \text{ and } x_k = y_k.$$

Condition (i) requires that the sender only sends messages that maximize the probability that the receiver accepts. Condition (ii) requires that the receiver takes the action that maximizes the probability of taking the right decision (i.e., accepting a good type or rejecting a bad type) given the message, the aspect checked, the value observed and the beliefs. In order to do this, the receiver should accept if she believes that the type is more likely to be good than bad, and reject if she believes the type is more likely to be bad. Condition (iii) requires that, given the receiver's beliefs after receiving the message, the receiver checks the aspect that maximizes the probability of taking the right decision.

Conditions (iv) and (v) state that the receiver's beliefs must be determined by Bayes rule whenever possible, given the prior and the players' strategies. Condition (iv) requires that, for messages observed on the equilibrium path, the belief probability assigned by the receiver to type x after observing message m is derived as the probability that the sender is of type x and sends message m divided by the total probability of message m being sent. If $\sum_{z\in X}\sigma(m|z)p_z = 0$, no sender type ever sends message m, so the receiver's beliefs are not constrained by Bayes rule. The sequential equilibrium refinement does not bite either, since it is possible to support any beliefs as the limit of a sequence. Condition (v) requires that the receiver rules out sender types with $x_k \neq y_k$ after observing y_k; for other types, the belief probability is the ratio of the probability that the sender is of type x and sends message m divided by the overall probability that the sender is of a type with $z_k = y_k$ and sends message m. If message m is never sent by a type with $z_k = y_k$, the receiver's beliefs are not constrained except by the value y_k itself (i.e., the receiver may have any beliefs as long as the total probability of 1 is distributed among types with $z_k = y_k$).

A.2 *The Sr game*

An equilibrium consists of a sender's message strategy $\sigma(m|x)$, a sender's revelation rule $\rho_1(x, m)$ (with $\rho_2(x, m) := 1 - \rho_1(x, m)$), a receiver's decision rule for each aspect revealed, $d_1(m, y_1)$ and $d_2(m, y_2)$, where y_k is the observed value of aspect $k = 1, 2$, and receiver's beliefs after receving the message and observing the actual value of an aspect, $b_1(x|m, y_1)$ and $b_2(x|m, y_2)$, satisfying the conditions (i)-(iii) below.

Let $\sigma(m, k|x) = \sigma(m|x)\rho_k(x, m)$ denote the probability that type x sends message m and reveals aspect k. Analogously to Rv, denote by $g_k(m, y_k) = \sum_{x \in G} b_k(x|m, y_k)$ the belief probability that the receiver assigns to the sender being a good type given that the sender sends message m, reveals aspect k and the observed value is y_k.

(i) Sender sequential rationality

For any type $x \in X$, any message $m \in M$ and any aspect $k = 1, 2$, $\sigma(m, k|x) > 0$ implies $d_k(m, x_k) \geq d_j(m', x_j)$ for all $m' \in M$, $j = 1, 2$.

(ii) *Receiver sequential rationality*

For all $m \in M$, $y_k \in X_k$ and $k = 1, 2$, the receiver sets $d_k(m, y_k) = 1$ if $g_k(m, y_k) > 0.5$, and $d_k(m, y_k) = 0$ if $g_k(m, y_k) < 0.5$.

(iii) *Consistency of receiver beliefs*

For each $k = 1, 2$ it holds that $b_k(x|m, y_k) = 0$ for all x such that $x_k \neq y_k$, and

$$b_k(x|m, y_k) = \frac{\sigma(m, k|x)p_x}{\sum_{z:z_k=y_k} \sigma(m, k|z)p_z} \text{ whenever } \sum_{z:z_k=y_k} \sigma(m, k|z)p_z > 0 \text{ and } x_k = y_k.$$

Condition (i) states that the sender strategy maximizes the probability of acceptance. If a combination of message and aspect being revealed has positive probability, it must be the case that the sender cannot do better by sending a different message and/or revealing a different aspect.

Condition (ii) is identical to the corresponding condition in Rv.

Condition (iii) is analogous to the corresponding condition in Rv, but not identical. Given that the sender decided to reveal aspect k and that the value of aspect k is y_k, the receiver must rule out all types x with $x_k \neq y_k$. For sender types with $x_k = y_k$, the probability that the sender is of type x equals the probability that the sender sends message *m and reveals aspect* k, divided by the total probability that a sender has value y_k of aspect k and reveals aspect k. When the sender reveals, two senders with the same value of aspect k may send the same message but have different probabilities of revealing aspect k; when the receiver chooses which aspect is observed, any two senders that send the same message must induce the same probability of observing aspect k, since the receiver has no way to distinguish the two cases.

B OPTIMAL COMMITMENT STRATEGIES IN RV AND SR GAMES

B.1 *Receiver Verifies*

We will apply Glazer and Rubinstein (2004) proposition 0 (the L-principle). Glazer and Rubinstein define an L-set to be a set of three types, where $x \in G$, $y \in X \backslash G$ and $z \in X \backslash G$, such that $x_1 = y_1$ and $x_2 = z_2$, that is, each of the bad types differs from the good type in the value of exactly one aspect. The idea of the L-principle is that, since the receiver can only verify one aspect and the sender best replies to the strategy of the receiver, the receiver must make a mistake for at least one of these three types. Either the good type x is rejected, in which case the receiver is making an error for this type, or the good type x is accepted after the receiver checks aspect 1 (in which case the bad type y must be accepted as well, since a sender of type y would have the option of pooling with type x), or the good type x is accepted after the receiver checks aspect 2 (in which case the bad type z must be accepted as well, since a sender of type z would then be able to pool with type x). A consequence of proposition 0 is that, when the prior probability distribution is uniform as in our case, ”an optimal mechanism can be found by using a technique that relies on the L-principle: finding a mechanism that induces H mistakes, and finding H disjoint L-sets” (Glazer and Rubinstein, 2004, p. 1721). The number of disjoint L-sets sets a lower bound on the receiver’s mistake probability. Thus, the L-principle ensures that a mechanism leading to the same number of mistakes as the number of disjoint L-sets is optimal for the receiver.[18]

Table B1 shows the set of possible types in our game, and indicates the good types by the letter G. We mark 15 disjoint L-sets (the three elements of each L are indicated by the same number; for example, types $(9, 2)$, $(9, 1)$ and $(1, 2)$ constitute and L-set). Thus, by the L-principle, there is no commitment strategy that yields fewer than 15 mistakes when the sender best responds. A receiver commitment strategy that, when the sender best responds to it, results in all types with 7+ being accepted and all other types being rejected implies 15 mistakes for the receiver, and hence is an optimal commitment strategy. An example of such a strategy is

[18]The reference to the ”number of mistakes” suggests a deterministic strategy on the part of the receiver. Glazer and Rubinstein point out on p. 1721 that when the prior distribution of types is uniform, the optimal mechanism does not require randomization when the sender’s aim is to persuade the receiver that the average [or, equivalently, the sum] of the two aspects is above a certain threshold.

checking the higher claim and accepting if and only if the aspect observed is 7+. Our example is analogous to Example 2 in Glazer and Rubinstein (2004) and the table below is analogous to Figure 2 in Glazer and Rubinstein (2004). Note that our notation differs slightly from theirs: we denote the set of good types by G, while they denote it by A.

Table B1

	9	15	G15	G	G	G	G	G	G	G
	8	13	14	G13	G14	G	G	G	G	G
	7	10	11	12	G10	G11	G12	G	G	G
	6	9	8	6		G9	G8	G6	G	G
x_2	5	5	7				G7	G5	G	G
	4	3	4					G4	G3	G
	3	2					12	6	G2	G
	2	1			10	9	8	5	3	G1
	1		15	13	14	11	7	4	2	1
		1	2	3	4	5	6	7	8	9
						x_1				

B.2 *Sender Reveals*

Glazer and Rubinstein (2006) show that the same L-principle technique can be applied to the game where the sender chooses which aspect to reveal (see Lemma 2 on p. 400 of their paper; note also that their Proposition 1 shows that there is an optimal commitment strategy that is deterministic, so there is no loss in focusing on deterministic commitment strategies). Therefore, as per the above analysis for the Rv game, a commitment strategy that, when the sender best responds, induces 15 mistakes is an optimal commitment strategy for the receiver. An example of such a strategy is to ignore the message and accept if and only if a 7+ aspect is revealed. This strategy results in hands with a 7+ aspect being accepted and other hands being rejected, just as the strategy we presented for Rv.

C PROOF THAT THE GR OUTCOME CAN BE SUPPORTED AS A SEQUENTIAL EQUILIBRIUM

There are multiple sequential equilibria resulting in the GR outcome in both games. In this section we describe two equilibria that differ in the role of messages for each game.

C.1 *Receiver verifies*

We start by presenting a sequential equilibrium in which the message informs the checking decision (by pointing the receiver to a 7+ aspect if the sender has one), but the acceptance decision is independent of whether the value observed coincides with the message.

C.1.1 *A sequential equilibrium leading to the GR outcome where the message is used for the checking decision only*

Let the sender's message strategy be as in Table C1 below, where the first entry in a cell is the reported value of x_1, and the second entry is the reported value of x_2; for empty cells the sender randomizes between messages (1,9) and (9,1).

Table C1

x_2 \ x_1	1	2	3	4	5	6	7	8	9
9	1,9	1,9	1,9	1,9	1,9	1,9	1,9	1,9	
8	1,9	1,9	1,9	1,9	1,9	1,9	1,9		9,1
7	1,9	1,9	1,9	1,9	1,9	1,9		9,1	9,1
6							9,1	9,1	9,1
5							9,1	9,1	9,1
4							9,1	9,1	9,1
3							9,1	9,1	9,1
2							9,1	9,1	9,1
1							9,1	9,1	9,1

In words, the sender's strategy only uses two messages. If both aspects are equal, or if the sender has no 7+ aspect, the sender sends one of the two messages at random. Otherwise the sender reports the higher value as 9 and the lower value as 1.

We now construct a strategy for the receiver such that the sender and receiver strategies, together with appropriate beliefs, constitute a sequential equilibrium.

The receiver's checking strategy is to check the higher claim, checking at random if both reports are equal. The acceptance strategy for messages $(9,1)$ and $(1,9)$ is as follows: conditional on having checked the higher claim, accept if and only if a 7+ value is observed. Conditional on having checked the lower claim (i.e., conditional on the receiver having deviated from their own checking strategy), accept if and only if a 5+ value is observed. For messages other than $(9,1)$ and $(1,9)$, the acceptance strategy is to accept if and only if a 7+ is observed irrespective of what claim was checked. We now check the optimality of the receiver's acceptance strategy, starting from subgames on the equilibrium path.

Suppose the message was $(9,1)$ (the case $(1,9)$ is analogous) and, having checked the higher claim (i.e. the first aspect), the receiver observes a value of 7. Given the messaging strategy it is not possible for the value of the second aspect to be 8 or 9 (those sender types send message $(1,9)$). The second aspect may be any value from 1 to 7; those values are equally likely except 7 itself, which is only half as likely given that types of the form $(7,x_2)$ with $x_2 < 7$ send message $(9,1)$ while type $(7,7)$ randomizes between $(9,1)$ and $(1,9)$. The probability of a good type is then $\frac{3.5}{6.5} = \frac{7}{13} > 0.5$, hence it is optimal for the receiver to accept, which is what the acceptance strategy specifies. An analogous reasoning applies if the value observed is 8 or 9 (the corresponding probabilities of a good type are $\frac{11}{15}$ and $\frac{15}{17}$).

Now suppose the message was $(9,1)$ and, having checked the first aspect as the checking rule specifies, the receiver observes a value of 6. Given the sender's messaging strategy, the second aspect may be 1, 2, 3, 4, 5 or 6, and all these values have the same probability because types of the form $(6,x_2)$ with $x_2 \leq 6$ send message $(9,1)$ with probability 0.5, while types of the form $(6,x_2)$ with $x_2 > 6$ never send message $(9,1)$. The probability that the type is good is then equal to $\frac{2}{6}$, and it is optimal for the receiver to reject, which is what the acceptance strategy specifies. An analogous reasoning applies if the value observed is 5 or less (the corresponding probabilities of a good type are $\frac{1}{6}$ if 5 is observed and 0 if 4 or less is observed).

Still dealing with subgames on the equilibrium path, let us look at the optimality of the receiver's checking strategy, conditional on having received message $(9, 1)$ or $(1, 9)$. The receiver's checking strategy is optimal if it minimizes the overall probability of error. The overall probability of error is the probability of error conditional on the observed value, weighted by the probability of observing that particular value. The tables below show these probabilities conditional on message $(9, 1)$ or $(1, 9)$ being received, depending on whether the receiver checks the higher claim (Table C2) or the lower claim (Table C3).

In Table C2 below, the first column contains each possible value that may be observed. The second column contains the probability of observing each value; this probability depends on the sender's messaging strategy and on the fact that the receiver is checking the higher claim. The third column contains the probability of a good type conditional on the observed value. The fourth column gives the probability of error that results from the receiver's acceptance strategy for each possible observed value. The overall probability of error if the receiver checks the aspect reported as 9 and takes the optimal acceptance decision can then be calculated as $\frac{6}{81} \cdot \frac{1}{6} + \frac{6}{81} \cdot \frac{1}{3} + \frac{13}{81} \cdot \frac{6}{13} + \frac{15}{81} \cdot \frac{4}{15} + \frac{17}{81} \cdot \frac{2}{17} = \frac{5}{27}$.

Table C2: Probabilities after receiving (1,9) or (9,1) and checking the higher claim

y_i	Prob(y_i)	Prob($GoodHand \mid y_i$)	Prob($Error \mid y_i$)
1	$\frac{6}{81}$	0	0
2	$\frac{6}{81}$	0	0
3	$\frac{6}{81}$	0	0
4	$\frac{6}{81}$	0	0
5	$\frac{6}{81}$	$1/6$	$1/6$
6	$\frac{6}{81}$	$1/3$	$1/3$
7	$\frac{13}{81}$	$7/13$	$6/13$
8	$\frac{15}{81}$	$11/15$	$4/15$
9	$\frac{17}{81}$	$15/17$	$2/17$

If the receiver checks the lower claim after receiving message $(9, 1)$ or $(1, 9)$, the relevant probabilities can be found in Table C3 below. The table also illustrates that accepting if the observed aspect is 5+ is optimal in this situation, precisely what the receiver's strategy specifies. The overall error probability if the receiver checks the lower claim and takes the optimal acceptance decision is $\frac{7}{27}$, which is above $\frac{5}{27}$. This shows the optimality of the checking strategy of the receiver conditional on having received message $(9,1)$ or $(1,9)$: the receiver is less likely to make an error by checking the higher claim.

Table C3: Probabilities after receiving (1,9) or (9,1) and checking the lower claim

y_i	Prob(y_i)	Prob($GoodType \mid y_i$)	Prob($Error \mid y_i$)
1	$\frac{12}{81}$	0	0
2	$\frac{12}{81}$	1/6	1/6
3	$\frac{12}{81}$	1/3	1/3
4	$\frac{12}{81}$	0.5	0.5
5	$\frac{12}{81}$	7/12	5/12
6	$\frac{12}{81}$	2/3	1/3
7	$\frac{5}{81}$	1	0
8	$\frac{3}{81}$	1	0
9	$\frac{1}{81}$	1	0

We have established the optimality of the receiver's checking and acceptance strategies conditional on message $(9,1)$ or $(1,9)$ being received. Receiver beliefs about the probability of a good type or about the probability of observing each value follow directly from the strategies and Bayes rule.

There are other information sets that are not reached given the strategy of the sender, namely information sets where messages other than $(9,1)$ and $(1,9)$ are used. We can construct a sequence of fully mixed strategies for the sender that converge to the strategy played, and that would induce beliefs that would make it optimal for the receiver to check the higher claim and accept if and only if the observed value is 7+, also for other messages.[19]

For example, consider the following fully mixed strategy: senders with a 7+ aspect send the message prescribed by Table C1 with probability $1-\varepsilon$ and randomize between all 81 messages with the remaining probability; senders with no 7+ aspect send the message prescribed by Table C1 with probability $1-2\varepsilon$ and randomize between all 81 messages with the remaining probability. Receiver's beliefs for messages other than $(9,1)$ and $(1,9)$ are constructed from this fully mixed strategy using Bayes rule.

[19] In principle we would also need to specify a fully mixed strategy for the receiver but the details of this strategy are of no consequence. For example, let the receiver check the higher claim with probability $1-\varepsilon$ and then take the optimal acceptance decision with probability $1-\varepsilon$.

This sender strategy clearly converges to the strategy specified earlier as $\varepsilon \to 0$. Furthermore, if a message other than $(9,1)$ or $(1,9)$ is received, it does not matter which of the two reports is checked, hence the receiver may as well check the higher claim if the two reports are different, and check at random if they are equal (the probability of observing each of the 9 possible values of an aspect and the probability of a good type conditional on the value observed do not depend on which claim is checked). As for the acceptance strategy, accepting if and only if the value observed is $7+$ is still optimal. In particular, if a value of 6 is observed, given that types without a $7+$ aspect are disproportionately likely to deviate, the type is slighly more likely to be bad than good. This is crucial since otherwise the receiver would accept 6 given a message other than $(9,1)$ and $(1,9)$, and the sender would then have an incentive to deviate.[20]

C.1.2 *A sequential equilibrium leading to the GR outcome where the receiver uses the message to inform both the checking and the acceptance strategy*

The sender's messaging strategy is in Table 2.

The receiver's strategy is to check the higher claim (checking at random if both reports are equal) and accept if and only if a good type is reported, the observed value coincides with the claim and the observed value is 7+.[21]

The sender's strategy is optimal since, given the receiver's strategy, it results in all sender types with a 7+ being accepted; this is the best the sender can do given that the receiver's strategy conditions acceptance on observing a 7+ aspect, so senders with no 7+ aspect cannot be accepted.

As for the receiver strategy, let us begin by the subgames in which the sender is sticking to the messaging strategy in Table 2. In some of the cells (for example $(4,4)$) the sender is reporting a bad type. All bad type reports in Table 2 are sent by senders with bad types, hence it is optimal for the receiver to (check the higher claim and) reject, which is what the strategy specifies. In other cells (for example $(9,5)$) the sender has a good type and is reporting it truthfully; no other sender type

[20] If a value of 6 is observed, the other aspect may be any value between 1 and 9 but the distribution is not uniform. Each value between 1 to 6 is twice as likely to occur as each value between 7 and 9, hence the probability of a good type conditional on observing a value of 6 would be $7/15$. If a 7 is observed, we know a sender type with a 7+ aspect has deviated, and the other aspect is equally likely to be any value between 1 and 9; the probability of a good type is then $6/9$.

[21] The 7+ threshold is relaxed at some information sets that are not reached in equilibrium, see below.

is sending the same message, so it is optimal for the receiver to (check the higher claim and) accept.[22] Finally, there are messages such that the sender is reporting a good type, and these messages are sent by several sender types. The types that send those messages form an L-set (cf Table B1). For example, if the message received is (9,2), the receiver's prescribed strategy is to check the higher claim and, if the value observed is indeed a 9, accept. Conditional on the message and on observing a 9, the type is equally likely to be (9,2) and (9,1), so it is optimal for the receiver to accept; if a 1 is observed, the type is sure to be a bad type, hence the receiver should reject which is what the strategy specifies. Could the receiver have done better by checking the lower claim upon receiving message (9,2)? Conditional on observing a value of 2, the type is equally likely to be (9,2) and (1,2); it is (weakly) optimal to reject as the strategy specifies. If a 1 is observed, the type is sure to be (9,1) given the sender's strategy, and it is optimal to reject as the strategy specifies. Overall, the probability of error is $1/3$ irrespective of which message is checked (the receiver erroneously accepts type $(9,1)$ if the higher claim is checked and erroneously rejects type $(9,2)$ if the lower claim is checked), so it is optimal to check the higher message. Something analogous happens for messages like $(6,6)$, where the sender reports a good type but the receiver rejects. After checking either message, if a 6 is observed the type has an equal probability of being good or bad and it is weakly optimal to reject; if a 2 is observed instead, the type is sure to be a bad type and it is optimal to reject.

In all the subgames above, the sender is sticking to their prescribed strategy and the receiver's beliefs follow by Bayes rule.

We now turn to the optimality of the receiver strategy for combinations of values and messages that cannot be observed given the sender strategy (this involves all messages that are never sent by the sender in equilibrium as well as cases such as receiving message $(9,2)$ and, having checked the first aspect, observing a value of 5). The definition of sequential equilibrium requires the receiver to have beliefs that make it optimal for the receiver to check the higher claim and (irrespective of what claim was checked) accept if and only if a good type is reported, the value observed coincides with the claim, and the value observed is $7+$.

[22]Similarly to the previously described equilibrium, there is some relaxation of the acceptance threshold at information sets where the receiver deviates from their own checking strategy. Since there are messages that are only sent by good types, the receiver's acceptance strategy if the lower claim is checked is to accept if one of those messages are received and the value observed coincides with the message. The receiver cannot gain from checking the lower claim for these messages.

Let the receiver beliefs place probability 1 on the type being bad for all those situations that have 0 probability given the sender's strategy. In order to have a sequential equilibrium, these beliefs must be obtainable as the limit of a sequence of beliefs, which themselves are derived (by Bayes rule) from a sequence of fully mixed sender strategies that converge to the strategy in Table 2.

The auxiliary sequence of fully mixed strategies for the sender is as follows. Senders with a good type send the message prescribed by Table 2 with probability $1 - \varepsilon^2$ and randomize between all 81 messages with the remaining probability; senders with a bad type send the message prescribed by Table 2 with probability $1 - \varepsilon$ and randomize between all 81 messages with the remaining probability. Receiver beliefs are the limit when $\varepsilon \to 0$ of the beliefs that follow from this sequence by Bayes rule. At any subgame that cannot be reached given the sender strategy, the receiver is certain that the type is bad.

For example, suppose the receiver gets message $(7,4)$ and, upon checking the higher claim, observes a 9. The receiver then believes that the type is certain to be $(9,1)$, and rejects. This belief can be constructed as the limit when $\varepsilon \to 0$ of $\frac{\frac{1}{81}\varepsilon}{\frac{1}{81}\varepsilon + \frac{8}{81}\varepsilon^2} = \frac{1}{1+\varepsilon}$. Hence, we can construct beliefs that justify the receiver rejecting when the observed value does not coincide with the message, even if the observed value is 7+. Analogously, if the receiver gets a message that is not in Table 2 such as $(8,2)$, the receiver believes that the message comes from a bad type, even if the observed value coincides with the claim. Note that all messages not used in Table 2 correspond to reported bad types. The beliefs we have constructed make it optimal for the receiver to reject when a bad type is reported, irrespective of the observed value or of whether it coincides with the claim.

The checking strategy is also optimal off the equilibrium path. If a message off the equilibrium path is observed, the receiver is indifferent between checking the higher and the lower claim, so may as well check the higher claim.

C.2 *Sender reveals*

C.2.1 *A sequential equilibrium leading to the GR outcome where the acceptance decision does not depend on the message*

The sender sends one of the 81 possible messages at random, and reveals the aspect with the higher value (revealing one aspect at random if both aspects have the same value). The receiver accepts if and only if the observed value is 7+ irrespective of the message.

This strategy combination leads to the GR outcome: all types with a 7+ are accepted, and all other types are rejected.

To see that this is a sequential equilibrium, note that the strategy of the sender is a best response: all types with a 7+ are accepted, while other types cannot be accepted given the receiver's strategy. As for the strategy of the receiver, it is a best response because, conditional on a 7+ value being observed, the type is more likely to be good than bad (and this is true irrespective of the message); conditional on a value under 7 being observed, the type is more likely to be bad than good (again, irrespective of the message). For example, if a 7 is observed, the other aspect may be any value between 1 and 7, with 7 itself being only half as likely; this results in a probability of $\frac{7}{13} > 0.5$ that the type is good, so it is optimal to accept as the strategy specifies. If a 6 is observed, the other aspect may be any value between 1 and 6, with 6 itself being half as likely; this results in a probability of $\frac{3}{11} < 0.5$ that the type is good, so it is optimal to reject as the strategy specifies.

Note also that we have constructed the strategy of the sender in such a way that all combinations of messages and values are observed in equilibrium, so the receiver never knowingly encounters an off-equilibrium information set and the sequential equilibrium requirement does not bite. The sender's messaging strategy is already fully mixed, and a fully mixed revelation strategy can be constructed in such a way that the sender reveals the higher value with probability $1 - \varepsilon$ and the lower value with probability ε. The receiver's beliefs are such that the receiver places probability 1 on the higher of the two values being revealed.

C.2.2 *A sequential equilibrium leading to the GR outcome where the acceptance decision depends on the message*

The sender follows the message strategy in Table 2 and reveals the aspect with the higher value (revealing an aspect at random if both aspects have the same value). The receiver accepts if and only if a good type is reported and a 7+ aspect is revealed.

The sender's strategy is a best response to the receiver's strategy since all types with a 7+ are reporting a good type and revealing a 7+ aspect, ensuring acceptance. Other types cannot be accepted given the receiver's strategy.

Given the sender's strategy, there are combinations of messages and values observed that are never observed if the sender sticks to the strategy described. In order to have a sequential equilibrium, we need to construct an auxiliary sequence of strategies and beliefs as explained earlier.

Take the following sequence of fully mixed strategies for the sender. The sender follows the messaging strategy in Table 2 with probability $1 - \varepsilon - \varepsilon^2$. Senders with good types and a 7+ aspect send one of the 36 good type messages at random with probability ε; with probability ε^2 they send one of the 45 bad type messages at random. All other senders send one of the 45 bad type messages at random with probability ε and one of the 36 good type messages at random with probability ε^2. As for the revelation strategy, all senders reveal the higher aspect with probability $1 - \varepsilon$ (and reveal one aspect at random if both values are equal).

A sequence of receiver beliefs is constructed from the sequence of sender beliefs using Bayes rule. The beliefs that we specify for the receiver are the limit of this sequence of beliefs, and the receiver strategy we have specified must be optimal given these beliefs.

For combinations of values and messages that are possible given Table 2, the receiver beliefs are derived from Table 2 itself (recall that $\varepsilon \to 0$, so for example if $(9, 2)$ is sent and a 9 is displayed, the receiver believes the type is equally likely to be $(9, 1)$ and $(9, 2)$); we have already established that the receiver acceptance strategy is optimal in these cases (see our earlier discussion for Rv, where the sender also uses the message strategy in Table 2).

As for other cases, receiver beliefs constructed as the limit of the sequence are such that, irrespective of what value is observed, the receiver believes that the aspect being observed is the higher of the two. This means that, using the information of the value observed only, a value of 6 or less suggests the type is more likely to be bad than good. The message does not change this conclusion since it contains no additional information as to whether the type is good or bad.

For types with a 7+ however, the receiver strategy is such that they are accepted if the reported type is good but rejected if the reported type is bad. Bad types with a 7+ aspect are disproportionately more likely to report a bad type in the sequence we constructed, and this justifies receiver's beliefs that a bad type message makes the type more likely to be bad than good even if a 7+ value is observed.

D OTHER EQUILIBRIUM EXAMPLES

D.1 *Receiver verifies*

D.1.1 *Receiver's worst equilibrium*

The receiver achieves their lowest possible equilibrium payoff when the message is uninformative (e.g. all sender types send each of the possible messages with equal probability, or all sender types send the same message). In such an equilibrium the receiver does not condition the checking decision on the message (e.g. the receiver always checks the same aspect, or checks each of the two aspects with equal probability), and, upon observing one of the aspects, takes the optimal acceptance decision given that the other aspect is equally likely to be any value between 1 and 9. In our game the optimal acceptance decision is to accept if and only if the observed aspect is 6 (since in 5/9 of cases the sum is at least 11) or higher. This equilibrium gives rise to a payoff of $\frac{61}{81}$ for the receiver and $\frac{36}{81}$ for the sender.

It is not possible for the receiver to obtain a lower expected payoff in equilibrium. This is because, given any sender message, the receiver can always ignore the message, check the first aspect and accept if and only if the observed value is 6+. If, in equilibrium, the receiver does something different conditional on the received message, the receiver must be at least as well off as if she followed the above strategy.

D.1.2 *Sender's best equilibrium*

Table D1 below depicts the sender's message strategy in an equilibrium which gives the highest possible equilibrium payoff to the sender. The receiver's checking strategy is as follows. If message $(6, 1)$ or $(1, 6)$ is received, the receiver checks the claim of 6 (and checks at random if any other message is received). The acceptance strategy is to accept if and only if the observed value is 6+, irrespective of which claim was checked.

Table D1

x_2									
9	1,6	1,6	1,6	1,6	1,6	1,6	6,1	6,1	6,1
8	1,6	1,6	1,6	1,6	1,6	6,1	6,1	6,1	6,1
7	1,6	1,6	1,6	1,6	1,6	6,1	1,6	1,6	1,6
6	1,6	1,6	1,6	1,6	1,6	6,1	1,6	1,6	1,6
5	6,1	6,1	6,1	6,1	1,6	6,1	6,1	6,1	6,1
4	6,1	6,1	6,1	6,1	1,6	6,1	6,1	6,1	6,1
3	1,6	1,6	1,6	1,6	1,6	6,1	6,1	6,1	6,1
2	1,6	1,6	1,6	1,6	1,6	6,1	6,1	6,1	6,1
1	1,6	1,6	1,6	1,6	6,1	6,1	6,1	6,1	6,1
	1	2	3	4	5	6	7	8	9
					x_1				

Note that in this equilibrium the sender uses only two messages, (6,1) and (1,6). These messages are used, together with the checking strategy, to ensure that the receiver observes a 6+ aspect if there is one. The acceptance strategy then results in all senders with a 6+ aspect being accepted. The lower acceptance threshold (compared to the equilibrium discussed earlier that results in the GR outcome) is made possible by the strategy of the sender, which does not necessarily point to the higher of the two aspects when both aspects are 6+.

The sender is clearly playing a best response to the receiver's strategy, since the receiver accepts only after observing a 6+ aspect and the sender's strategy leads to all types with a 6+ being accepted.

The receiver's strategy can be divided into acceptance strategy and checking strategy. It is tedious but straightforward to check that the receiver is playing a best response to the sender. For example, suppose the message is $(6,1)$, the receiver checks the first aspect (as required by the strategy) and observes a value of 6. Given the sender's strategy in the table, this means that the type is equally likely to be any of the types where $x_1 = 6$ and $x_2 < 9$; since four out of eight such types are good types it is indeed optimal for the receiver to accept. As for the checking strategy, conditional on message $(6,1)$ or $(1,6)$ being received and on the receiver's acceptance strategy, the receiver is making 20 errors (all consisting of accepting bad types). Can the receiver do better by checking the other aspect? Suppose the message is $(6,1)$. If the receiver checks the first aspect and accepts if the observed value is $6+$, the receiver is making 10 errors (accepting $(6,4)$, $(6,3)$, $(6,2)$, $(6,1)$, $(7,3)$, $(7,2)$, $(7,1)$, $(8,2)$, $(8,1)$ and $(9,1)$). If the receiver checks the second aspect instead, it is still optimal to accept if the observed aspect is $6+$ and to reject otherwise, and the receiver would still make 10 errors (rejecting the good types that send the message $(6,1)$ and have $x_2 < 6$, namely $(6,5)$, $(7,5)$, $(8,5)$, $(9,5)$, $(7,4)$, $(8,4)$, $(9,4)$, $(8,3)$, $(9,3)$ and $(9,2)$). Analogously, it can be checked that if $(1,6)$ is sent but the receiver checks the first aspect, the receiver would still make at least 10 errors.

As for messages other than $(6,1)$ and $(1,6)$, we have specified that the receiver checks a claim at random and accepts if and only if a 6+ value is observed. Is this strategy part of a sequential equilibrium? Consider the following sequence of fully mixed strategies by the sender. The sender plays the messaging strategy described above with probability $1-\varepsilon$, and sends one of the 81 messages at random with probability ε. If the receiver gets a message other than $(1,6)$ or $(6,1)$, the message is not informative and each of the 81 possible types is equally likely. It would then be optimal for the receiver to check either aspect and accept if and only if a 6+ value is observed.

In this equilibrium, all good types and 20 bad types are accepted. The only way the sender could obtain a higher payoff would be if another bad type was accepted, but this would bring the receiver's payoff below $61/81$, which is the lower bound for the receiver's equilibrium payoff (see previous subsection).

D.2 *Sender reveals*

D.2.1 *Receiver's worst equilibrium*

We now describe an equilibrium that attains the receiver's lowest equilibrium payoff. In this equilibrium the sender sends a random message and reveals the highest aspect for bad types and the lower one for good types (when the two aspects are equal, one is revealed at random). The receiver rejects irrespective of the observed value and of the message.

To see that this is an equilibrium, we note that given the sender's strategy, the probability of the type being good is lower than 0.5 for any combination of message and value observed. For example, suppose the sender reveals that the first aspect is a 9. Given the sender's strategy, there are only two types that reveal the first aspect when its value is a 9: (9,1) and (9,9). Because (9,9) reveals the first aspect with probability 0.5, while (9,1) always reveals the 9, the probability of a good type conditional on observing the first aspect to be a 9 is $\frac{1}{3}$ which is less than 0.5. Hence, the receiver best replies by always rejecting.

This equilibrium gives rise to a payoff of $\frac{45}{81}$ for the receiver and 0 for the sender. This is the receiver's lowest possible equilibrium payoff since the receiver always has the option to reject for any given message and evidence, and this would lead to a payoff of $\frac{45}{81}$.

D.2.2 *Sender's best equilibrium*

In this equilibrium, the sender sends a random message and reveals the lower aspect for types with two 6+ aspects and the higher aspect for all other types. The receiver's strategy is to accept if and only if the observed aspect is a 6+.

To see that this is an equilibrium, we note that given the sender's strategy, the probability of the type being good conditional on the observed aspect is higher than 0.5 as long as the observed aspect is a 6+, and equal to 0 otherwise. This makes it optimal for the receiver to accept if and only if she observes a 6+. In this equilibrium, the receiver obtains a payoff of $\frac{51}{81}$ while the sender gets $\frac{56}{81}$.

This is the best equilibrium for the sender because values of 1 to 5 of either aspect must be rejected in any sequential equilibrium, so the best the sender can achieve is to be accepted if he has a 6+ aspect. [23]

[23] The proof is recursive. Start by noting that a value of 1 of either aspect must be rejected since it is certain to be a bad type. Types with an aspect above 1 will then display the other aspect if the other aspect has a positive probability of acceptance. This can then be used to prove (by contradiction) that a value of 2 of either aspect must be rejected with certainty, and so on. The recursion continues up to the value of 5.

E DYNAMICS IN T=11

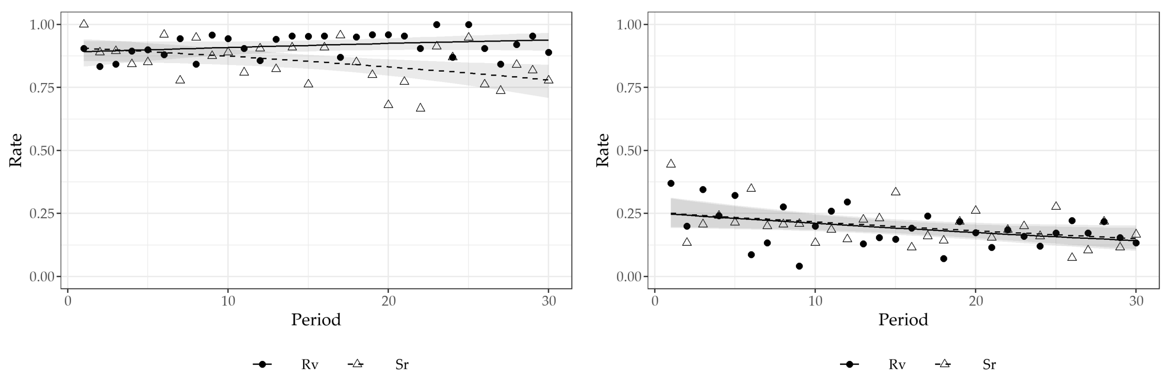
In this section we investigate whether subjects' decisions change across periods, and show that the results reported in the main text are robust to period effects.

E.1 *Messages*

We first look at the rate of truthful reporting. Figure E1 shows the average truth-telling rate for good and bad hands in each treatment. Panel (a) suggests a small downward trend in truth-telling for good hands in the Sr, but not the Rv, treatment, while panel (b) suggests a small downward trend in truth-telling for bad hands in both treatments.

Figure E1: Truth-telling rates across periods

(a) Good hands **(b)** Bad hands



Note: The lines represent predicted rates from probit regressions (standard errors clustered at the matching group level).

Table E1 presents the marginal effects from a probit regression of whether the sender's message is truthful on the Sr treatment dummy, and variables for the interaction between period and each treatment.[24] This analysis is performed separately for good and bad hands. For good hands, the results show that the probability that the sender tells the truth is stable in the Rv treatment, but decreases significantly ($p < 0.05$) in the Sr treatment. For bad hands, the probability that the sender tells

[24] This specification allows us to directly observe if the two treatments have significant, and potentially different period trends. In addition, using the Wald test, we can check if a treatment effect is present while controlling for period effects by testing the joint hypothesis that the coefficient on the treatment dummy is equal to 0 while the coefficients of the two interaction terms are not different from each other. We report the associated χ^2 statistic in all regression tables.

the truth decreases significantly over periods ($p < 0.05$) in both treatments. Note, however, that the estimated effects are quite small, about 0.4 percentage points per period in each case where the decrease is significant.

Table E1: Probit analysis of truth-telling rate

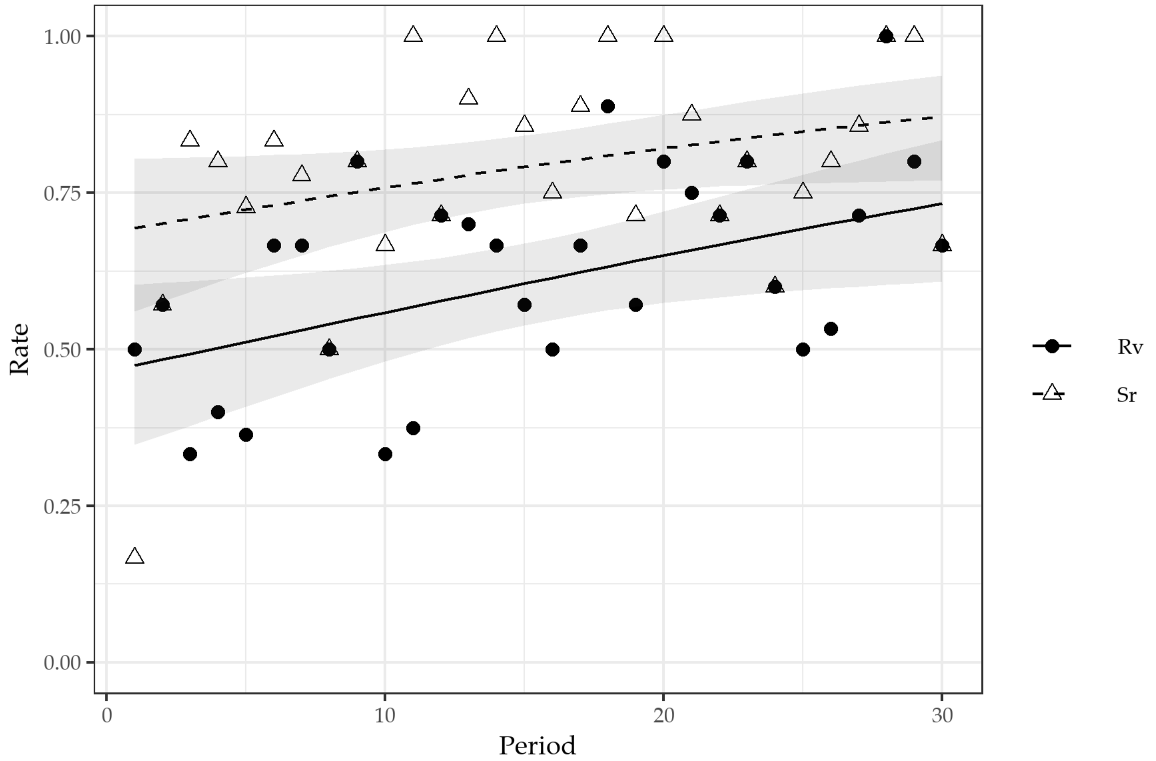
	Dependent variable:	
	Truth-telling decision	
	(Good hands)	(Bad hands)
(1) Treatment = Sr	0.021	0.003
	(0.041)	(0.068)
(2) Period x (Treatment = Sr)	-0.004^{**}	-0.003^{***}
	(0.002)	(0.001)
(3) Period x (Treatment = Rv)	0.002	-0.004^{**}
	(0.002)	(0.002)
Observations	1,262	1,618
χ^2 statistic for (1) = 0 & (2) = (3)	9.146^{**} (df=2)	0.025 (df=2)

Note: The table presents marginal effects; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

How do these dynamics affect the robustness of Result 1? Recall, Result 1 states that senders with a bad hand and a 7+ card usually inflate the value of the lower card while keeping the higher claim truthful, but are significantly more likely to do this in Sr compared to Rv. Focusing on bad hands with one 7+ card, Figure E2 shows the rates of reporting a good hand while keeping the higher message truthful in the two treatments across periods.

The rates show a similar increasing trend in both treatments, but, except for the very first period, the average rate is at least as high in Sr as Rv in every period. Table E2 presents the results from a probit analysis of the senders' likelihood of reporting a bad hand with one 7+ card as good while keeping the higher claim truthful. Although the treatment dummy is not individually significant, a test for the absence of treatment effects rejects this hypothesis (see χ^2 statistic in Table E2).

Figure E2: Rates of reporting a good hand while keeping the higher message truthful across periods. (bad hands with one 7+ card.)



Note: The lines represent predicted rates from probit regressions (standard errors clustered at the matching group level).

Table E2: Probit analysis of the rate of reporting a good hand while keeping the higher claim truthful

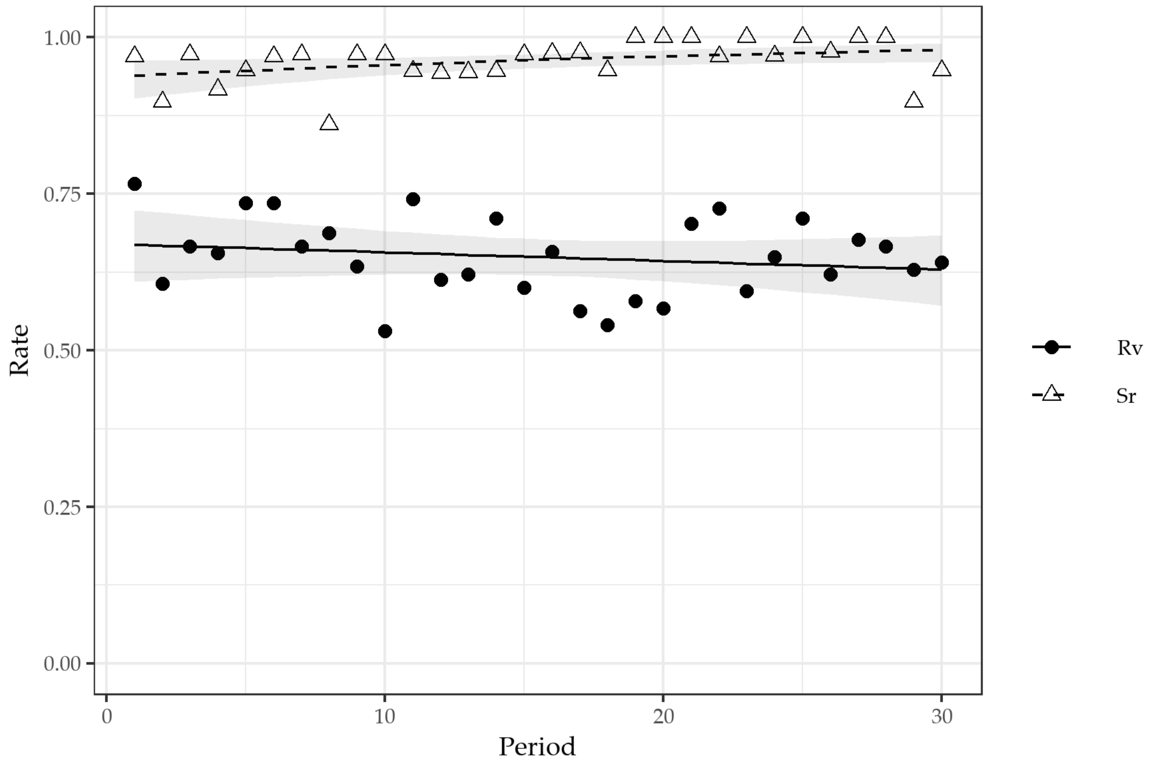
	Dependent variable:
	Good hand claimed & higher claim truthful
(1) Treatment = Sr	0.195
	(0.135)
(2) Period x (Treatment = Sr)	0.008
	(0.005)
(3) Period x (Treatment = Rv)	0.008*
	(0.005)
Observations	436
χ^2 statistic for (1) = 0 & (2) = (3)	7.877** (df=2)

Note: The table presents marginal effects; data includes only bad hands with a 7+ card; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

E.2 *Revelation/verification*

Our main result concerning subjects' revelation/verification decision is that senders are more likely to reveal their higher card in Sr than receivers are to check the higher claim in Rv. Figure E3 shows the average rates of revealing the higher card in Sr and verifying the higher claim in Rv across periods. The rate in Sr is above 85% in every period and there is a slight increasing trend, while that in Rv is below 80% in every period and there is a slight decreasing trend. A probit regression (see Table E3) shows that the trend in Sr is significant while that in Rv is insignificant and overall the treatment difference remains high and strongly significant after controlling for period effects.

Figure E3: Rates of revealing the higher card / verifying the higher claim across periods. (Hands with non-equal cards/reports).



Note: The lines represent predicted rates from probit regressions (standard errors clustered at the matching group level).

Table E3: Probit analysis of the rate of revealing/verifying the higher card/claim

	Dependent variable:
	Reveal/verify higher card/claim
(1) Treatment = Sr	0.239***
	(0.049)
(2) Period x (Treatment = Sr)	0.004**
	(0.001)
(3) Period x (Treatment = Rv)	−0.001
	(0.002)
Observations	2,167
χ^2 statistic for (1) = 0 & (2) = (3)	72.16*** (df=2)

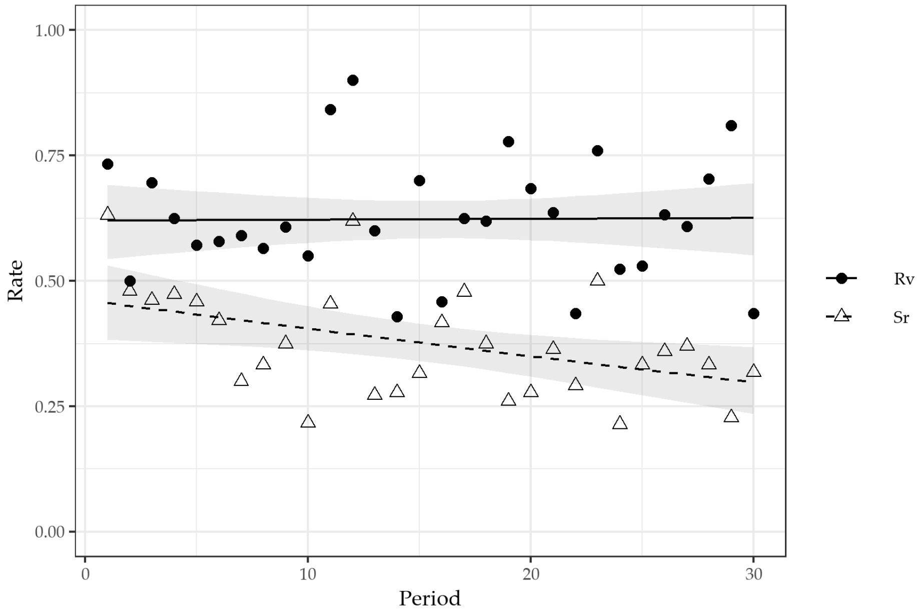
Note: The table presents marginal effects; data for Sr excludes hands where the two cards were equal, while for Rv, it excludes hands where the two reports were equal; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

E.3 *Acceptance*

Recall that our main result regarding receiver's acceptance decision is that the acceptance rate conditional on the observed value is higher in Rv than Sr, significantly so for observed values 3-7 (Result 3). We now check the robustness of this result across periods.

Figure E4 shows the acceptance rates in each treatment across periods. We focus on cases where a good hand was reported and the observed value was between 3 and 7. The acceptance rates in Rv are consistently above those in Sr. A decreasing trend in the acceptance rate in the Sr treatment leads to an increase in the difference between treatments as subjects gain more experience.

Figure E4: Acceptance rates across periods (claimed good hands only, observed values 3-7)



Note: The lines represent predicted rates from probit regressions (standard errors clustered at the matching group level).

Table E4 presents the marginal effects from a probit regression of the acceptance decision on the Sr treatment dummy, interaction terms between period and each treatment, and the value of the observed card. The results suggest that the treatment difference is large and highly significant even after controlling for period effects.

Table E4: Probit analysis of the acceptance decision (claimed good hands only, observed values 3-7)

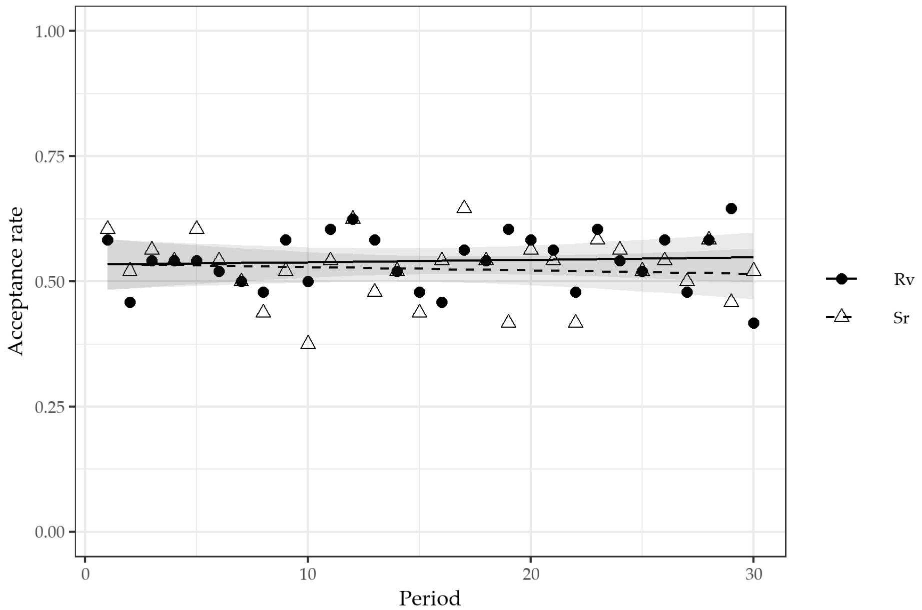
	Dependent variable:
	Acceptance decision
(1) Treatment = Sr	-0.278^{***}
	(0.062)
(2) Period x (Treatment = Sr)	-0.006^{***}
	(0.002)
(3) Period x (Treatment = Rv)	-0.001
	(0.002)
(4) Value of observed card	0.172^{***}
	(0.019)
Observations	1,304
χ^2 statistic for (1) = 0 & (2) = (3)	69.677^{***} (df=2)

Note: The table presents marginal effects; the data excludes cases where the reported values add up to less than 11 and the observed value is less than 3 or greater than 7; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

E.4 *Outcomes and payoffs*

Figure E5 shows the acceptance rate in each treatment across periods. This is also the sender's average payoff. The results of a probit analysis reported in Table E5 confirm the lack of significant period or treatment effects.

Figure E5: Acceptance rate (sender's average payoff) across periods



Note: The lines represent predicted averages from probit regressions with standard errors clustered at the matching group level.

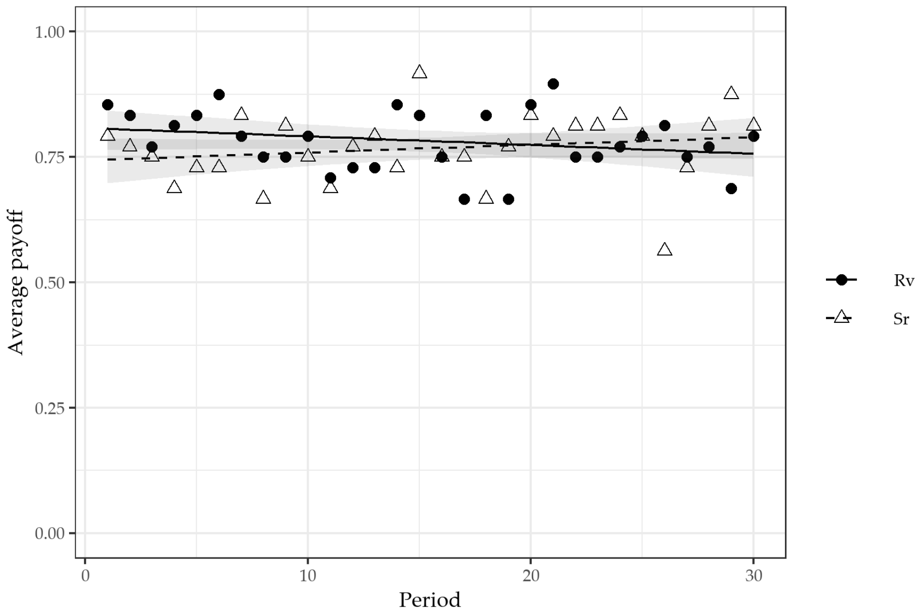
Table E5: Probit analysis of acceptance rate (sender average payoff)

	Dependent variable:
	Acceptance decision (sender payoff)
(1) Treatment = Sr	0.002
	(0.032)
(2) Period x (Treatment = Sr)	−0.001
	(0.001)
(3) Period x (Treatment = Rv)	0.000
	(0.001)
Observations	2,880
χ^2 statistic for (1) = 0 & (2) = (3)	1.242 (df=2)

Note: The table presents marginal effects; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

Figure E6 shows the receiver's average payoff across periods. It appears that there is a slight increasing trend in Sr and a slight decreasing trend in Rv. The probit analysis in Table E6 suggests that the trend in Sr is significant leading to a significantly higher receiver payoff in Rv in early periods and a significantly higher receiver payoff in Sr in later periods. However, further analysis shows that this is driven by behavior in the first 5 periods. When we exclude the first 5 periods the significant treatment differences disappear (see Table E6).

Figure E6: Receiver's average payoff across periods



Note: The lines represent predicted averages from probit regressions with standard errors clustered at the matching group level.

Table E6: Probit analysis of receiver average payoff

	Dependent variable:	
	Receiver payoff	
	(all periods)	(excluding first 5 periods)
(1) Treatment = Sr	-0.065^{**}	-0.042
	(0.031)	(0.047)
(2) Period x (Treatment = Sr)	0.002	-0.002
	(0.001)	(0.002)
(3) Period x (Treatment = Rv)	-0.002	-0.001
	(0.001)	(0.002)
Observations	2,880	2,400
χ^2 statistic for (1) = 0 & (2) = (3)	5.316^{*} (df=2)	0.753 (df=2)

Note: The table presents marginal effects; standard errors in parentheses are clustered at the matching group level; *p<0.1; **p<0.05; ***p<0.01.

F A DIFFERENT PARAMETRIZATION

To examine the robustness of these results to the parametrization, we also conducted sessions of the same two games but with a different threshold (T) defining good hands. In this implementation of the Sr and Rv games the two cards have to add up to at least 9 to represent a good hand. Lowering the good hand threshold does not change the multiplicity of equilibria in Sr or Rv. Moreover, there are still multiple equilibria in both games resulting in a GR outcome (the outcome of an optimal commitment strategy as analyzed by Glazer and Rubinstein (2004, 2006)). The GR outcome maintains the same structure as for T=11 but now, more hands are accepted, reflecting the different distribution of good hands: hands with at least one 6+ card are accepted and the rest are rejected.

We used the same experimental design and procedures. There were 180 subjects in total. One session in the Rv treatment had 20 instead of 24 subjects. The corresponding session in Sr as well as one other also had only 20 subjects. In the following analysis, for the sessions in which the Rv treatment had more subjects than the Sr one, we drop the equivalent extra observations in Rv to maintain comparability between the underlying draws. The main analysis is therefore performed on 8 independent observations, but now 4 of them comprise of 10 individuals rather than 12.

F.1 *Decisions*

The results from the T=9 treatments are consistent with those from the T=11 treatments. First, we note that senders' reporting strategies are in line with Result 1: for bad hands with a 6+ card, senders in Sr report a good hand while keeping the highest of the two claims truthful in 85% of the cases, while in Rv this happens only in 53% of cases which is significantly lower ($p - value = 0.016$). Second, consistent with Result 2, senders reveal the higher card more often in Sr (98.56%) than receivers verify the higher claim in Rv (82.64%), and this difference is statistically significant ($p - value < 0.001$). Third, in line with Result 3, acceptance rates conditional on the observed value are significantly higher in Rv than in Sr for intermediate observed values ($p - value = 0.008$ for each value from 3 to 6).

F.2 *Outcomes and payoffs*

As for Result 4, Table F1 shows that sender payoffs do not vary significantly across treatments. Receiver payoffs however, are marginally significantly higher in Rv ($p = 0.086$).

Table F1: Average payoffs when T=9

	Sr	Rv	p-value
Receiver	0.830	0.871	0.086
Sender	0.678	0.701	0.266

To investigate this further, Table F2 reports acceptance rates for good and bad hands, distinguishing between hands that are accepted in the GR outcome and hands that are rejected in the GR outcome. Note that, as in T = 11, the receiver prefers Sr for bad hands with no high card, and Rv in all other cases. However, in T = 11, 41% of hands were in the first category, while in T = 9 there are only 28% such hands. As a result, the receiver is marginally better-off in Rv when T=9.

Table F2: Acceptance rates for good and bad hands conditional on the value of the highest card (T=9)

			Acceptance rate		
Type of hand		obs.	Sr	Rv	p-value
Highest card < 6	Bad hands	365	0.173	0.211	0.156
	Good Hands	57	0.509	0.754	0.031
	All hands	422	0.218	0.284	0.039
Highest card ≥ 6	Bad Hands	86	0.721	0.430	0.031
	Good hands	812	0.911	0.948	0.234
	All hands	898	0.893	0.899	0.461

F.3 *Best-response analysis*

The results from analyzing the receiver's best response to the observed sender behavior in T = 9 are also consistent with the results reported for T = 11.

Recall that, for Sr, we restrict attention to strategies that condition on the observed value and the claim about the other card. If the receiver makes the optimal decision for each combination of observed value and unverified claim, the receiver's expected payoff is 0.896, which we refer to as the empirical optimum. Analogously to Result 5, following the optimal commitment strategy that accepts if and only if a good hand is reported and a 6+ value is observed would give the receiver 99.88% of the empirical optimum.

Similarly, for Rv, behaving optimally for all senders' messages gives the receiver an average payoff of 0.882. Analogously to Result 6, following the optimal commitment strategy of checking the higher message and accepting if and only if no misreport is observed, a good hand is claimed and the observed value is 6+, gives the receiver 99.77% of the empirical optimum.

G INSTRUCTIONS

G.1 *Rv, T=11*

INSTRUCTIONS

Welcome and thank you for participating in this experiment. Throughout the whole experiment you are kindly asked to remain seated and refrain from communication with the other participants. Mobile phones and other electronic devices should be switched off. If there are any questions please raise your hand and an experimenter will come to answer your questions in private.

Payment: This experiment consists of 30 rounds. In each round you can earn points. At the end of the experiment you will be paid according to your accumulated point-earnings from all rounds. You will be paid in private and in cash with **£0.50 for each point earned**. Additionally, you will receive a **participation fee of £3.**

All your decisions are anonymous, so your identity will be kept secret at all times.

At the beginning of each round you will be randomly matched with another participant (i.e. the person you are paired with will change from round to round). One of you will have the role of Person A and the other the role of Person B. Your role will be assigned at the beginning of the first round and you will keep this role for all 30 rounds.

Each round consists of 2 stages which are described below.

Stage 1: *Person A observes two cards and sends a message*

In this stage, the computer will randomly select two cards, one **orange** and one **blue**, each carrying a value between 1 and 9. Each combination of values on these 2 cards is equally probable. At this stage, only **Person A** will be able to observe these values.

The hand is "**GOOD"** if the sum of the values on the two cards is **at least 11**. The hand is "**BAD**" if the sum of the values is **10 or below**.

This is an example of a BAD hand:

After observing the randomly drawn cards, Person A will send a message to Person B of the following form:

The value of the **orange card** is: ____

The value of the **blue card** is: ____

where Person A fills each blank box with a number between 1 and 9.

Stage 2: *Person B selects one of the cards to observe and makes a decision*

After observing Person A's message, Person B will **select one of the two cards** (**orange** or **blue**) and the computer will reveal its value.

After observing the value on the selected card Person B will **decide between "Accept" and "Reject"**.

End of the Round

At the end of the round both Person A and Person B will receive a summary of the round including:

- The cards that were randomly dealt;
- Person A's message;
- Person B's choice regarding which card to observe;
- Person B's decision to accept or reject;
- Person A and Person B's point-earnings for the round.

How your point earnings are determined:

Person A earns 1 point if B accepts and 0 points if B rejects.

Person B earns 1 point if A has a good hand and B accepts or if A has a bad hand and B rejects. Person B earns 0 points otherwise.

This is summarised in the Table below:

	B **Accepts**	B **Rejects**
A has a **GOOD** hand	Person A receives 1 point, Person B receives 1 point	Person A receives 0 points, Person B receives 0 points
A has a **BAD** hand	Person A receives 1 point, Person B receives 0 points	Person A receives 0 points, Person B receives 1 point

Preliminary questions: Before the 30 rounds begin, you will be asked to answer a few questions regarding your understanding of the instructions. The rounds will begin only after all participants have answered these questions correctly.

Final questionnaire: After the 30 rounds, you will be asked to fill in a short questionnaire. You will then be paid your earnings in private and in cash.

G.2 *Sr, T=11*

INSTRUCTIONS

Welcome and thank you for participating in this experiment. Throughout the whole experiment you are kindly asked to remain seated and refrain from communication with the other participants. Mobile phones and other electronic devices should be switched off. If there are any questions please raise your hand and an experimenter will come to answer your questions in private.

Payment: This experiment consists of 30 rounds. In each round you can earn points. At the end of the experiment you will be paid according to your accumulated point-earnings from all rounds. You will be paid in private and in cash with **£0.50 for each point earned**. Additionally, you will receive a **participation fee of £3.**

All your decisions are anonymous, so your identity will be kept secret at all times.

At the beginning of each round you will be randomly matched with another participant (i.e. the person you are paired with will change from round to round). One of you will have the role of Person A and the other the role of Person B. Your role will be assigned at the beginning of the first round and you will keep this role for all 30 rounds.

Each round consists of 3 stages which are described below.

Stage 1: *Person A observes two cards and sends a message*

In this stage, the computer will randomly select two cards, one **orange** and one **blue**, each carrying a value between 1 and 9. Each combination of values on these 2 cards is equally probable. At this stage, only **Person A** will be able to observe these values.

The hand is "**GOOD"** if the sum of the values on the two cards is **at least 11**. The hand is "**BAD**" if the sum of the values is **10 or below**.

This is an example of a BAD hand:

After observing the randomly drawn cards, Person A will send a message to Person B of the following form:

The value of the **orange card** is:	
The value of the **blue card** is:	

where Person A fills each blank box with a number between 1 and 9.

Stage 2: *Person A selects one of the cards for Person B to observe*

After Person B observes Person A's message, Person A will **select one of the two cards** (**orange** or **blue**) for Person B to observe its value in the next stage.

Stage 3: *Person B observes the value of the card and makes a decision*

After observing the value of the selected card, Person B will **decide between "Accept" and "Reject"**.

End of the Round

At the end of the round both Person A and Person B will receive a summary of the round including:

- The cards that were randomly dealt;
- Person A's message;
- Person A's choice regarding which card to be observed by Person B;
- Person B's decision to accept or reject;
- Person A and Person B's point-earnings for the round.

How your point earnings are determined:

Person A earns 1 point if B accepts and 0 points if B rejects.

Person B earns 1 point if A has a good hand and B accepts or if A has a bad hand and B rejects. Person B earns 0 points otherwise.

This is summarised in the Table below:

	B **Accepts**	B **Rejects**
A has a **GOOD** hand	Person A receives 1 point, Person B receives 1 point	Person A receives 0 points, Person B receives 0 points
A has a **BAD** hand	Person A receives 1 point, Person B receives 0 points	Person A receives 0 points, Person B receives 1 point

Preliminary questions: Before the 30 rounds begin, you will be asked to answer a few questions regarding your understanding of the instructions. The rounds will begin only after all participants have answered these questions correctly.

Final questionnaire: After the 30 rounds, you will be asked to fill in a short questionnaire. You will then be paid your earnings in private and in cash.