What makes communication informative? Cheap talk with multiple senders

Nejat Anbarci
Department of Economics and Finance
Durham University
Durham DH1 3LB, United Kingdom
nejat.anbarci@durham.ac.uk

Nick Feltovich*
Department of Economics
Monash Business School
Monash University
Clayton VIC 3800, Australia
nicholas.feltovich@monash.edu

Mehmet Y. Gürdal
Department of Economics
Boğaziçi University
Bebek, Istanbul, Turkey TR–34342
mehmet.gurdal@boun.edu.tr

Abstract

We investigate communication and decision making in a signalling game that is a simplified version of many real–life settings. There are two senders and one receiver. Each sender knows both senders' qualities, while the receiver is completely uninformed. After a round of communication, the receiver chooses to match with one of the senders. The receiver prefers to match with the higher–quality sender, while each sender simply prefers to be matched. We vary the form of communication. In our *Byte* treatment, a sender sends a single number, indicating her quality. In our *Rich* treatment, a sender sends a single free–form text message. In our *Chat* treatment, the receiver can chat with each sender individually, in simultaneous two–way conversations. Our main result is that receivers in the Chat treatment perform substantially better than chance in picking the higher–quality sender, while receivers in the other treatments fare no better than chance. Additional results suggest that communication in the Chat treatment allows the receiver to extract a body of information that is more likely to be (i) precise and (ii) consistent in its implication for decision making. Receivers in all treatments fare well when both (i) and (ii) are present, and poorly otherwise.

Journal of Economic Literature classifications: D83, C92, D91, D82

Keywords: signalling; lying; experiment; messages; private information

*Corresponding author. Much of this work took place when Anbarci was at Deakin University. Financial support from Boğaziçi University (BAP 7921), Deakin University, and Monash University is gratefully acknowledged. We thank Timo Heinrich, Emin Karagözoğlu, Aaron Nicholas, Birendra Rai and participants at several conferences and seminars for helpful comments and suggestions.

1 Introduction

Deception is the subject matter of a wide range of scientific disciplines, including political science (e.g., Cliffe et al., 2000), psychology (DePaulo et al., 1996; Newman et al., 2003), sociology (Barnes, 1994), biology (see below), neuroscience (Harada et al., 2009; Vartanian et al., 2012), and linguistics (Meibauer, 2011), as well as business and economics. The ability to deceive is evolutionarily advantageous (Dawkins and Krebs, 1978; Wright, 1995), and deception is indeed widespread in nature. Natural selection also favours individuals who accurately spot attempts at deceit (Dawkins, 1978; Trivers, 1985). Strategies to deceive others and at the same time to spot deception by others are ever evolving, and these abilities may have shaped the development of the human brain (Cosmides and Tooby, 1992).

Human economic and social relationships are replete with settings where deception, detection of deception, or both are important. The extent of lying in society can be illustrated by estimates of its aggregate costs, and while clearly difficult to ascertain precisely, even conservative estimates are staggering. As an example, Walczyk et al., 2005, citing earlier work by Berry, 2003, and Lipman and McGraw, 1988, estimate costs to business from worker dishonesty in the United States alone of between \$6 billion and \$200 billion per year.[1]

A natural setting for studying both deception and the detection of deception involves markets for experts' services. As a concrete example, suppose a homeowner needs to choose an electrician to do some rewiring. This interaction is characterised by *asymmetric information*: the electrician will typically have better information about his competence, and perhaps the scope of work that will be required, than the homeowner. There are also *imperfectly aligned preferences*: the homeowner prefers to hire a competent electrician, while the electrician prefers to be hired regardless of his level of competence. Asymmetric information implies that communication between the electrician and the homeowner (whom, for generality, we will often henceforth call "sender" and "receiver") has the potential to reveal the relevant information, aiding the homeowner's decision. However, imperfectly aligned preferences imply that the electrician has incentives to lie to secure a better outcome for himself at the homeowner's expense (hiring that expert rather than a more competent rival, purchasing unnecessary services, paying excessive prices, etc.), and a sensible receiver will keep this in mind when acting on the sender's messages. The result is likely to be limited transmission of information and limited exchange, causing inefficiencies.[2]

Settings like this have received attention in the behavioural economics literature and elsewhere, but much

[1]Of course, worker dishonesty encompasses some kinds of misbehaviour that are distinct from lying, such as workplace theft. However, there are also many kinds of lying in business settings outside the workplace – such as companies lying to their customers (Boush et al., 2015; Darke and Ritchie, 2007) and customers lying to the companies (Cowley and Anthony, 2019), and of course lying in other settings (e.g., by suspects or witnesses in court) – none of which are counted in this estimate. We also acknowledge the moral and ethical costs associated with lying, beyond the economic costs. These are well–known and we do not discuss them here, though we note that some recent work has argued that in some settings, liars may be perceived as having talent in areas requiring sales skills, such as advertising, investment banking, and of course sales (Gunia and Levine, 2019).

[2]If both parties expected their interaction to be repeated, these inefficiencies might be reduced. However, infrequent or even one–shot interactions are very common in these settings. Reputational concerns may also reduce these inefficiencies, if feedback from previous customers is available. However, the incentives for customers to provide truthful post–interaction feedback – good or bad – are often weak, so feedback need not be credible.

of this work has involved one sender interacting with one receiver. (See Section 2 for examples; we note some exceptions below.) In this paper, we consider a small but important modification: one receiver interacting with *two* senders; in the above example, consider a homeowner (receiver) choosing between two electricians (senders). Both senders are presumed to know each others' ability as well as their own, while receivers know nothing beyond the underlying population distribution. The receiver's objective is to choose the higher–ability sender, while each sender wants only to be chosen. The combination of these mis–aligned preferences and the asymmetry of information gives senders both the means and the motive to lie about their quality. We investigate this setting with a laboratory experiment, allowing us a high degree of control over features of the environment like senders' and receivers' preferences, the information they are given, the way communication occurs between them, and the nature and extent of their interaction.

The nature of individuals' preferences in our setting stands in contrast with other recent experimental studies of multiple–sender settings (e.g., Lai et al., 2015; Vespa and Wilson, 2016). In those previous studies, senders' preferences were biased relative to receivers', but to a limited extent, and in such a way that a clever receiver could infer a lot of information from the senders' messages in equilibrium. In our setting, senders' biases are extreme and diametrically opposed (each sender wants the receiver to maximally over–estimate her own ability and maximally under–state the rival sender's ability). Hence as long as agents have standard preferences (maximising their own monetary earnings), the only theoretical solutions are "babbling" equilibria where messages are completely ignored by the receiver, who then can do no better than randomly choose between the senders. This setting may seem contrived, but it is perhaps only a slight exaggeration of many real–life situations: in addition to the example above of tradespeople seeking to perform work for a homeowner, consider two job candidates in a specialised field being interviewed by a panel of non–specialists; two fund advisors trying to manage an investor's portfolio; two agents wishing to represent an actor or athlete; witnesses for opposing sides in a trial trying to convince a judge or jury; two political candidates competing to get elected; and so on.

In our experiment, we vary the manner in which senders communicate with the receiver. In our "Byte" treatment, messages are tightly controlled; each sender can only send a single numerical message indicating his own quality. In our "Rich" treatment, each sender can similarly send only a single message, but is able to use free text, with almost no restrictions on content. In our "Chat" treatment, the receiver can engage in separate private conversations with each sender concurrently, with both sender and receiver able to send multiple messages to the other – again with no structure imposed. The Rich and Chat treatments thus correspond to real–world styles of communication (e.g., many forms of advertising are similar in nature to our Rich treatment, while pairwise email conversations between customers and suppliers are similar to our Chat treatment). Our Byte treatment is a typical implementation of communication in lab experiments and game–theoretic models.

Standard game theory is silent regarding differences across our treatments, as each treatment admits only the babbling equilibria described above, thus implying no treatment effects. However, previous research (see Section 2) suggests these treatments may lead to different outcomes. Our Rich and Byte treatments are similar in that both consist of monologues from senders to the receiver, but differ in the form these monologues take: highly structured in the Byte treatment versus unstructured in the Rich treatment. Some evidence suggests that messages are more likely to be truthful when they are in natural language (as in the Rich treatment but not the Byte treatment). Even if they are not, receivers in the Rich treatment may be able to use their skills in interpreting natural language, to detect and exploit any useful cues in the senders' messages, improving their judgements.

By the same token, our Chat and Rich treatments both use natural language, but they differ in whether the conversations are dialogues (in Chat) or monologues (in Rich). Receivers in the Rich treatment are therefore passive during the message stage, while they can be active in the Chat treatment – allowing them potentially to influence the type, sequence and extent of information conveyed by each sender. While they cannot force senders to reveal information, they can ask for clarification if a sender's message is imprecise, they can confront a sender whose message contradicts something said by the other sender, and they can press a sender whom they believe to be lying. To the extent that hiding information (either by outright lying or by remaining silent) is unpleasant or cognitively taxing for senders, this potential for interrogation in the Chat treatment may improve receivers' outcomes beyond any effect from introducing natural language on its own (i.e., moving from Byte to Rich).

In the experiment, we observe significant departures from the standard–theory predictions, consistent with the intuition described above. Lying about quality, while rampant, is not universal, and varies in systematic ways. Higher–quality senders are mostly truthful, while lower–quality senders typically over–state their own quality and under–state the rivals'. Also, truthfulness broadly increases as we move from the Byte treatment to the Rich treatment and thence to Chat.

Our main result is that the form of communication matters. In the Byte and Rich treatments, receivers are on average unable to exploit the information they receive; they are no more likely than a coin toss to pick the better sender. In the Chat treatment, they are much more successful, choosing correctly almost two–thirds of the time, with a corresponding increase in payoffs. Thus the important distinction is not between structured and unstructured communication, as much previous research has emphasised, but rather between one–way and two–way unstructured communication. A disaggregated analysis of the data suggests that receivers' success in the Chat treatment is due to their improved ability to extract more and better information from the senders. Namely, the body of communication from pairs of senders to receivers is more likely in the Chat treatment to be both *jointly consistent* (not contradicting each other regarding the receiver's "better choice"), and *precise* in senders' claims of own and rival quality. Importantly, both of these criteria are necessary: regardless of the treatment, receivers fare well when both are present, but poorly when only one (or neither) is.

2 Literature review

The literature on lying in business, economics and related areas contains several strands – both theoretical and empirical – and has exploded in recent years. Some of the early behavioural economics literature was arguably a by–product of studies focussed on testing how behaviour changes as a result of communication, with the truth value of the communication itself – either in actuality or in the perception of its recipients – as a secondary question. For example, Duffy and Feltovich's (2002, 2005) studies of pre–play messages focus on their efficacy for improving outcomes, though they also note the truthfulness of pre–play messages depends on the game's structure, and in particular how strong monetary incentives are to be truthful.[3] A subset of this literature, of particular relevance to us, involves *comparisons* of two or more forms of communication. Sally (1995) finds from a meta–analysis of prisoners' dilemma experiments that verbal statements of intended play are more likely to be true and more likely to lead to increased cooperation, compared to written statements (Balliet, 2010, reports a similar result in a later meta–analysis), and statements framed as promises are even more effective. Bochet et al. (2006) find in a controlled experiment that computerised chat (similar to our Chat treatment) increases contributions to public goods almost as much as face–to–face communication, while structured communication (like our Byte treatment) has little effect.[4]

A second strand looks at individuals' willingness to lie about a piece of information they know. Often in these studies, there is no active second player, though sometimes others are affected by the message sent. Experiments of this type include those summarised by Ariely (2012), where subjects perform a real–effort task, then report their performance (which determines their payment) to the researcher – so that lying is possible and incentivised – and where monitoring by the researcher may be impossible. The "die roll" experiments of Fischbacher and Föllmi-Heusi (2013) and others are another example; there, subjects privately roll dice and report the outcomes to the experimenter, which determine their payment. (Batson et al., 1997, and Batson et al., 1999, apply psychology–experiment methodology to a similar setting.) There are also experiments of this type where the researcher – as well as the decision maker – knows the true state (and hence can determine lying directly rather than statistically); see, for example, Gneezy (2005). Many questions have been studied within these settings (see Abeler et al., 2019, for a thorough review); one common finding relevant to us is that attitudes toward lies that affect others' well–being appear to be sensitive to how the others are viewed. Broadly, subjects are *more* averse to lying that harms friends, family members, innocent bystanders, or accomplices compared to "victimless lying" (that only affects the experimenter or her funders), but *less* averse when the lying harms rivals or competitors, while the ordering is reversed when lying benefits others rather than harming them (see Kocher et al., 2018 for an example).

Conrads and Lotz (2015) combine these two strands of the literature, with subjects self–reporting sets of four coin tosses via either face–to–face communication, phone communication, a computer interface in the lab, or a computer interface remotely. Monetary incentives favour over–stating the number of "tails" reported, but the level of such over–statement was similar across all of the treatments. Relatedly, Conrads and Reggiani (2017) found no systematic differences across the same communication treatments (plus one more, computerised chat) in how likely subjects are to keep their promises, though they found that they were more likely to *make* promises the more the communication mode reduced social distance (most in face–to–face, least when

[3]Indeed, the messages themselves are not even analysed in some relevant studies. For example, Azfar and Nelson's (2007) experimental investigation of voter reactions to corruption had candidates for political office giving 15–second speeches to the voters, but the content of the speeches is not discussed in the paper.

[4]Other studies have confirmed the efficacy of face–to–face communication in improving outcomes in social dilemmas; see, for example, Frohlich and Oppenheimer (1998) and Brosig et al. (2003). However, it is difficult to disentangle the effects of reducing social distance from those of losing anonymity of subjects (leading to the possibility of rewards or punishments outside the lab).

remotely via computer, etc.).

A third strand of the literature emphasises lie detection. This literature arose earlier in psychology (Bond and DePaulo, 2006, 2008; Buller and Burgoon, 1996; DePaulo et al., 1985; Ekman and O'Sullivan, 1991; Hartwig et al., 2004; Vrij, 2008). A typical finding is that laymen are poor lie detectors: often not significantly better than chance. Most of these studies are like our Byte and Rich treatments, in that they do not provide opportunities for interaction between potential deceivers and those who try to spot deception. Since relevant questions may well uncover inconsistencies in senders' stories, the cognitive cost to fabricate a consistent story is likely to be higher when there is interaction (DePaulo et al., 2003; Zuckerman et al., 1981), as in our Chat treatment. Von Hippel et al. (2015) assign subjects to groups for a task, and afterward compare their ability to detect a saboteur in the group before and after "interrogation" (like our Chat treatment, but face–to–face instead of via computer); they report that subjects guess correctly much more than chance would imply (about two–thirds of the time, compared to at most one–fourth from random guessing), and interrogation improves guesses (to about 70 percent) but only insignificantly. Minson et al. (2018) show that when subjects are asked questions about undesirable workplace behaviour (taking sick days when well, gossipping about co–workers, etc.), more admit their behaviour when the question is posed with a negative frame ("You have occasionally gossipped about a co–worker, right?") than a positive frame ("You haven't gossipped about a co–worker, right?"), and more in either case than when neither frame is used ("What can you tell us about your interactions with co–workers?"), again suggesting a role for two–way communication in discerning lies. On the other hand, Dwenger and Lohse (2019) find that subjects watching video–recorded statements in a tax–compliance experiment (so that questioning by the subjects is impossible) cannot distinguish truthful statements from lies.

The corresponding literature in economics is fairly recent (Belot et al., 2012; Iyer et al., 2015; Konrad et al., 2014). A noteworthy recent study by Belot and van de Ven (2017), involving deception in buyer–seller interactions, finds that sellers' accuracy of detection exceeds the theoretically predicted levels in all treatments, but neither adding context to the setting nor allowing the buyer to interrogate the seller affects the level of accuracy. Another interesting study, by Chen and Houser (2017), focuses on informal written communication in a variant of the trust game introduced by Charness and Duwfenberg (2006), and finds that longer messages, and messages that are considered to be promises, are more likely to facilitate trust. An important sub–literature within this strand focuses on politicians' campaign promises and their effects on election outcomes. Callander and Wilkie (2007) analyse a theoretical model in which candidates for office face incentives to lie, but voters prefer honest candidates, leading to moderate levels of lying overall. Experimental studies by Corazzini et al. (2014) and Feltovich and Giovannoni (2015) have found that campaign promises are typically believed by voters, and are initially truthful (for inexperienced subjects in the role of candidates) but lying emerges quickly as subjects gain experience.

A bridge between the lying and lie–detection literatures could be called "normative" lie detection: analysis of statements in order to determine characteristics that could be used to predict whether someone is lying, separate from the question of whether human observers are able to discern lying. Burgoon et al. (2016) examine corporate executives' statements about earnings in quarterly conference calls. Perhaps the most interesting of the results they report is that fraud–related utterances contained more specific details, and more complex wording, than nonfraudulant utterances. (Braun et al., 2015, observe a similar result in politicians' statements.) Larcker and Zakolyukina (2012) find that lying CEOs use more extreme positive emotion and fewer anxiety words. Walcyk et al. (2005) find that lying about personal information and experiences is associated with longer response times.

Theoretical study of lying in business organisations dates back at least to Schein (1979).[5] In economics, theoretical study of sender–receiver games like ours, where the sender is allowed to send unverifiable information, dates back to the seminal work of Crawford and Sobel (1982), who show that a better–informed sender will send only noisy information in equilibrium. This canonical setup is modified by Kartik (2009), who posits a cost of lying for the sender. He finds that incomplete separation emerges, with a subset of senders all choosing the highest (most self–flattering) possible message, with the size of this subset determined by the lying costs. Experimental studies of sender–receiver games with a single sender have generally found that the sender transfers more information than predicted by equilibrium (Dickhaut et al., 1995; Cai and Wang, 2006; Sánchez-Pagés and Vorsatz, 2007; de Haan et al., 2015).

There have been numerous theoretical extensions of the Crawford and Sobel (1982) model to the case of two senders and one receiver, where senders could be informed (Gilligan and Krehbiel, 1989; Krishna and Morgan, 2001; Gick, 2008; Li, 2008), or uninformed (Austen-Smith, 1990a, 1990b, 1993b). Alternative communication modes (Austen-Smith, 1990b, 1993a) and a multidimensional policy space (Milgrom and Roberts, 1986; Austen-Smith, 1993b; Battaglini, 2002, 2004; Ambrus and Takashi, 2008; Lai et al., 2015) have also been studied theoretically. Experiments involving two senders are quite recent, with most concentrating on the effects of increasing the number of senders and extending the policy space to multiple dimensions (Lai et al., 2015; Vespa and Wilson, 2016). A notable exception is a recent study by Brosig-Koch and Heinrich (2018), who examine factors (including the characteristics of messages) affecting which sender is chosen by receivers. Their setting is one of moral hazard – with senders' qualities chosen by them, instead of being assigned by the experimenter as in the current paper. Finally, Minozzi and Woon (2016) provide a theoretical and experimental study of a multi–sender environment in which senders' and receivers' preferences are positively correlated; their experimental results show persistent exaggeration by senders in the direction of their "bias", contrary to the predictions of their theoretical model (see also Minozzi and Woon, 2018).

The current paper contributes to this literature. First, we add to the small literature on signalling games with multiple senders, and we do so using a setting with severe conflicts of interest. Together, these make the theoretical predictions (under standard preferences) straightforward and transparent: receivers should ignore messages in all three of our treatments, and should fare no better than chance in selecting senders. This is in marked contrast to the sender–receiver settings typically analysed theoretically and experimentally; these tend to have complex equilibria that many subjects would find difficult to compute (to say nothing of the likelihood that such equilibria are common knowledge amongst the subjects). Second, we compare not only structured versus unstructured communication, as many previous studies have, but also two kinds of unstructured commu-

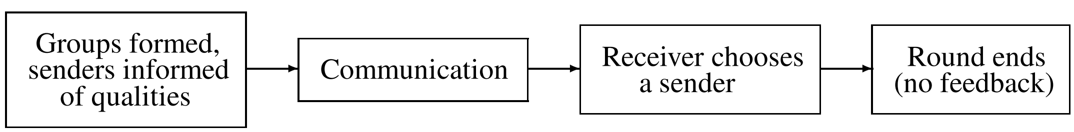
[5]As she notes, however, the study of lying in political settings dates back much further, to Machiavalli's *The Prince* (1513).

nication: one–way (akin to text messages) in our Rich treatment versus two–way (like back–and–forth chats) in our Chat treatment. This design feature complements our use of multiple senders: the ability of the receiver to interrogate a sender ought to be more valuable when she receives information from another sender.

3 Experimental design and procedures

The structure of a play of the game is shown in Figure 1. Senders' qualities are multiples of 5 between 0 and 50 inclusive. They are randomly drawn; the qualities of both senders in a group could not be the same, but otherwise, all quality pairs were equally likely. A play of the game begins with senders being informed of both their own and their rival's quality; the receiver is not informed of either quality. Then there is a round of communication that depends on the treatment (see below); once it is finished, the receiver chooses exactly one of the senders to match with, at which point the game ends. The receiver's payoff is the quality of the matched sender, while a sender earns 50 if matched and 0 if not.

Figure 1: Sequence of decisions in a play of the game



Our treatment variable is the form that communication takes between senders and receiver. In the *Byte* treatment, communication consists of a single message from each sender (chosen simultaneously), chosen via radio buttons from the set of possible qualities (i.e., multiples of 5 from 0 up to 50). The *Rich* treatment is similar except for the message space: each sender can write a single free–form message. In the *Chat* treatment, communication is via two two–way chat windows: one between the first sender and the receiver, the other between the second sender and the receiver.

3.1 Experimental procedures

The experiment took place between 5 December and 8 December, 2016, at the Economics Laboratory of Boğaziçi University in Istanbul, Turkey. Subjects were primarily undergraduate students, invited via ORSEE (Greiner, 2015). No–one took part more than once; otherwise, there was no exclusion of subjects. The experiment was run on networked personal computers, and programmed using z–Tree (Fischbacher, 2007). Subjects sat in individual carrels in a single room, were visually isolated from each other, and were asked to turn off their mobile phones and not to communicate with each other except via the computer program.

At the beginning of each session, subjects were given separate written instructions, which were also read aloud in an attempt to make them common knowledge.[6] Subjects were randomly assigned to roles: there were

[6] English translations of a sample set of instructions, the list of questionnaire questions, and sample screen–shots can be found in Appendices A, B and C respectively.

10 senders and 5 receivers per session. Roles were fixed for all rounds, but groups (each with 2 senders and 1 receiver) were formed so that no two subjects were grouped together more than once during the session. Each session comprised a main treatment (Byte, Rich or Chat) lasting for 5 rounds, followed by a questionnaire. (The small number of rounds was necessary to ensure subjects did not interaction a second time.) There was no instructions quiz, but subjects could ask questions privately at any time during the session. There were no practice rounds.

Each round started with a reminder to subjects of their roles, and senders were informed about their own quality and the quality of the other sender in the group. In the Byte treatment, senders' messages were then simultaneously chosen via a set of radio buttons, with "blank" messages impossible. The receiver observed both senders' messages on the same screen. The Rich treatment was identical except that senders wrote free–form messages of up to 500 characters including spaces. Senders were given 90 seconds to write a message, and blank messages were possible. Subjects were asked in the instructions not to send messages divulging identifying personal information, using abusive language, or making threats or promises regarding behaviour outside the game.

In the Chat treatment, the communication stage involved two–way chat messaging between the receiver and (separately) each sender. During the 210 seconds available for chat, the receiver's screen displayed two chat windows (one for each sender), while each sender's screen displayed one window. (Thus direct communication between the senders was impossible.) The same rules regarding message content held as in the Rich treatment, but there were no restrictions on the timing of messages other than the 210–second time limit (e.g., subjects did not have to wait for a reply before sending another message).

Once communication ended, receivers were prompted to choose between the senders; once all had done so, the round ended and the next round began. There was no end–of–round feedback. After the fifth round ended, subjects completed a questionnaire, which included a set of demographic questions (e.g., gender, age, number of siblings, living arrangement, university major, number of economics units taken) followed by some attitudinal questions (general willingness to take risks, trust, adherence to norms of civic cooperation). Then, subjects were paid, privately and in cash. Earnings were the sum of a show–up fee of 10 Turkish Liras (TRY) and the average of the five rounds' payoffs (in TRY).[7] Sessions lasted about 35 minutes, and average earnings were 35 TRY for senders and 36.64 TRY for receivers (35.55 TRY overall).

3.2 Research questions

Our main research questions are as follows:

Question 1 *Are there systematic differences across treatments in how likely receivers are to choose the higher–quality sender?*

Question 2 *How truthful, and how informative, are senders' messages? Do these vary across treatments?*

[7] At the time of the experiment, 1 TRY corresponded to about USD 0.29 at market exchange rates. However, the lower cost of living in Turkey compared to many developed countries made the stakes correspondingly higher in real terms. For comparison, at the time of the experiment, the minimum monthly wage in Turkey was 1647 TRY, and a lunch at the school cafeteria cost about 6–10 TRY.

The reasoning behind these questions is straightforward. Question 1 speaks to whether receivers *do* benefit from any particular form of communication, while Question 2 speaks to whether receivers *could* benefit from it.[8] Standard theory makes stark predictions here, and in particular does not predict systematic differences across our three treatments. There are no equilibria under standard (own–money–payoff–maximising) preferences in which senders' messages are informative and understood as so by receivers. There are only "babbling" equilibria where receivers ignore the messages they are sent, and are indifferent between choosing either of the senders. Even behavioural theories that incorporate a "cost of lying" may not imply any treatment effects. For example, if the cost of lying is a lump sum or dependent only on the size of the lie (the difference between true and stated quality) or on the harm done to the receiver (in money terms), there will be no differences across our treatments.

On the other hand, if lying has a cost that depends on the relationship between the sender and receiver, non–zero treatment effects are possible. Some previous experimental work (e.g., Bohnet and Frey, 1999; Rankin, 2006) suggests that lowering social distance can lead to more pro–social behaviour.[9] If our communication treatments have differing effects on social distance, it stands to reason that social distance would be reduced the most by communication in our Chat treatment, less in our Rich treatment, and less still in our Byte treatment. We would then expect less lying by senders – and more successful choices by receivers – as we move from Byte to Rich and thence to Chat.

A second avenue for treatment effects is via senders' beliefs about the receiver's ability to detect lies. If these beliefs vary across treatments, and if senders believe receivers may be reluctant to choose to match with someone suspected of lying, senders may lie to different extents across the treatments, separately from any effect on social distance or on preferences more generally. Intuitively, it is plausible that as communication becomes richer, receivers' lie–detection ability would improve, so that senders would be more truthful. Both of these effects would work in the same direction, so we would again expect receivers to be more successful as we move from Byte to Rich and thence to Chat.

4 Results

A total of 225 subjects participated in the 15 sessions. Some session information is shown in Table 1.

[8] Note that in Question 2, we distinguish between truthfulness and informativeness. In principle, these are indeed distinct notions (consider a broken clock, which tells the truth twice a day, versus a clock that runs 5 minutes fast, which never tells the truth but is arguably more informative), and given our Question 1, our focus should be on whether messages are informative rather than their truthfulness. However, we expect that in practice, truthfulness and informativeness will be highly correlated, and moreover, a truthful sender arguably has the *intention* to be informative with her messages.

[9] These results are also in the spirit of theoretical work where individuals' pro-sociality is influenced by their beliefs about how pro-social their current opponent is (Levine, 1998; Ellingsen and Johannessen, 2008) – which may be viewed as another measure of social distance.

Table 1: Session information

Treatment	Type of communication	Session numbers	Total number of subjects
Byte	Quality only, from each sender to receiver	2, 5, 8, 11, 15	75
Rich	One 500–character message, from each sender to receiver	1, 6, 9, 10, 14	75
Chat	Multiple messages between receiver and each sender	3, 4, 7, 12, 13	75

4.1 Receiver choices

We consider two reasonable measures of receiver success. One is simply the fraction of times the receiver chose the correct (higher–quality) sender in a given round; we call these "correct choices". The second is the receiver's average payoff, put through an affine transformation in which the minimum and maximum possible payoff (from always choosing the wrong sender and always choosing the right sender, respectively) are 0 and 1 respectively; we call this "normalised payoff". Normalised payoff can be thought of as a weighted average of correct choices, where the weightings are based on the quality difference between the two senders (i.e., the amount at stake for the receiver). In particular, random choice should make both of these measures equal to one–half on average.

Table 2 shows treatment averages of these two statistics. Overall, receivers in the Byte and Rich treatments performed no better than chance in picking the higher–quality sender, whether or not choices are weighted by the stakes. There is no significant difference in higher–quality choices between these two treatments (robust rank–order test, session–level data, $p > 0.20$ for both all rounds together and the last round on its own).[10] By contrast, receivers have substantial success in the Chat treatment, choosing the higher–quality sender almost two–thirds of the time and with a corresponding increase in normalised payoffs. The former is significantly

Table 2: Receiver treatment–level outcomes (all rounds)

	Treatment		
	Byte	Rich	Chat
Correct choices by receiver	0.496	0.496	0.640
Normalised receiver payoff	0.533	0.488	0.647

[10]Unless stated otherwise, our non–parametric tests use session–level data (the smallest independent unit) and are two–sided. We note that even if we use only the first–round data (so that each receiver represents an independent observation, allowing greater statistical power in principle), differences between Byte and Rich are not significant (indeed, first–round success rates for receivers are 11/25 in both treatments). See Siegel and Castellan (1988) for descriptions of the particular tests used in this paper, and Feltovich (2005) for critical values of the robust rank–order test used here and later.

higher than the other two treatments (robust rank–order test, $p \approx 0.039$ versus Byte and $p \approx 0.014$ versus Rich), and the latter is at least weakly significantly higher than the other treatments (robust rank–order test, $p \approx 0.096$ versus Byte and $p \approx 0.044$ versus Rich).

Table 3 shows how receivers' correct choices vary by the stakes. There appears to be a small improvement from low to moderate stakes – with success rates increasing slightly in all three treatments – but these differences are not significant (a chi–square test on the pooled data, which over–states significance by ignoring dependence within sessions, returns a p–value above 0.3).[11] Also, there is no apparent improvement when moving from moderate to high stakes, though small numbers of observations make drawing conclusions difficult in that latter case. Thus we do not find conclusive evidence that receivers make better choices when the stakes are higher.

Table 3: Receiver frequency of choice of higher–quality sender (all rounds)

Treatment	Difference in sender qualities		
	5–15	20–35	40–50
Byte	0.47 (30/64)	0.52 (26/50)	0.55 (6/11)
Rich	0.49 (33/67)	0.53 (27/51)	0.29 (2/7)
Chat	0.61 (39/64)	0.71 (32/45)	0.56 (9/16)

These results are further confirmed by probit regressions. We estimate two pairs of models, each with a correct receiver choice as the dependent variable. In Models 1a and 1b, right–hand–side variables are indicators for the Byte and Chat treatments (so Rich is the baseline), the absolute difference between the senders' qualities, the products of these variables, and a constant term. Models 2a and 2b include all of these variables along with the round number and its products with the other variables (including the interaction terms from Models 1a and 1b), to control for time dependence. Models 1a and 1b differ only in whether the set of demographic and attitudinal variables collected at the end of each session are included (and similarly for 2a and 2b).[12] We use regular probit models instead of panel probits so that we can cluster standard errors at the session level.[13] These estimations, along with those elsewhere in this paper, were performed using Stata (version 12).

Table 4 shows the main results, including average marginal effects (computed using Stata's "margins" command) and standard errors for each variable. Correct choices are more likely in the Chat treatment than in either the Byte or Rich treatment (evidenced by the p–value displayed below the Chat indicator's marginal effect, and by the significance of the marginal effect itself, respectively), while the insignificant marginal for the Byte treatment indicates no significant differences between the Byte and Rich treatments. Also, the marginal

[11] An additional chi–square test, using the three denominators for each treatment in Table 3, also returns a p–value above 0.3. This suggests that the distribution of quality differences (between lower– and higher–quality senders) is similar across treatments, even though these were drawn i.i.d. in each group–round of our experiment, rather than matched across treatments.

[12] These variables do turn out to be jointly significant, though none of the individual variables is significant once a correction for multiple comparisons is made (Benjamini and Hochberg, 1990). We do not display the variables individually in the table, to save space and since their effects are unconnected to our research questions.

[13] Panel linear models, which also allow clustering, yield similar results to those presented here.

Table 4: Receiver regression results (marginal effects at means, std. errors in parentheses)

Dependent variable: correct choice	[1a]	[1b]	[2a]	[2b]
Byte treatment	–0.002	0.033	–0.011	0.027
	(0.056)	(0.067)	(0.052)	(0.063)
Chat treatment	0.144^{***}	0.183^{***}	0.142^{***}	0.181^{***}
	(0.042)	(0.048)	(0.041)	(0.042)
p–value for differences between Byte and Chat	*0.010*	*0.044*	*0.004*	*0.029*
Round number			–0.002	–0.004
			(0.019)	(0.019)
Difference in qualities	0.002	0.002	0.001	0.002
	(0.002)	(0.002)	(0.002)	(0.002)
Demographic/attitudinal vars?	No	Yes	No	Yes
N	375	375	375	375
$\|ln(L)\|$	253.95	248.60	251.12	245.76

* (**,***): *Marg. effect significantly different from zero at the 10% (5%, 1%) level.*

effect of the difference in qualities is insignificant, confirming that receivers do not become more (or less) proficient when the stakes are higher. Finally, the round number has no effect: this may not seem surprising at first glance since no end–of–round feedback was given, though it also rules out receivers' learning though other routes (introspection, content of messages, etc.).

Result 1 *Receivers choose the higher–quality sender more in the Chat treatment than in the other two treatments.*

Result 2 *We do not find evidence that receivers are more (or less) likely to choose the higher–quality sender as the quality difference between senders increases.*

Because of the lack of a stake–size effect, here and in Tables 2 and 3, we henceforth focus on correct choices as our measure of receivers' success.

4.2 Sender message types

Messages were mandatory in the Byte treatment and optional in the Rich and Chat treatments; however the option to send messages in these latter two treatments was always taken (there were no "blank" messages). In the Rich treatment, senders' messages varied in length from 2 to 409 characters (including spaces), with a median of 129. (The 2–character message was "50", which was the sender's actual quality.) In the Chat treatment, senders sent between 2 and 38 messages, containing between 11 and 740 characters in total; medians were 12 messages and 312.5 characters.

We now turn to an analysis of the content of senders' messages. Such an analysis is straightforward in the Byte treatment, where messages were likely to have a literal meaning understood by both sender and receiver: the sender's quality. In the Rich and Chat treatments, however, the space of potential messages is much larger than the set of qualities, and there is no restriction of messages to be about quality. So, messages in these two treatments can contain statements about one's own quality, they can contain additional relevant information such as a statement about the rival sender's quality, they can contain less precise but still relevant information such as a claim to have the higher quality (without stating what that quality is), and they can contain seemingly irrelevant information. They can even contain statements that contradict other statements made by that same sender. So, even aside from the question of whether a statement is true or false (which applies in the Byte treatment as well as the Rich and Chat treatments), it is a non–trivial task for the researcher (and perhaps even the receiver) simply to determine what statements are being made.

In order to classify messages in the Rich and Chat treatments, we start by selecting four categories: "messageSelf", "messageRival", "messageHigher", and "messageLower". In the Rich treatment, messageSelf is defined if and only if the message contains a precise claim to having a particular quality, in which case its value is equal to that quality. MessageRival is similar but pertains to claims about the rival's quality. MessageHigher and messageLower are indicator variables that take on a value of one in the Rich treatment if the message contains a claim to have a higher (resp., lower) quality than the rival sender. So, the statement "My quality is 40, and this is higher than my rival's quality" is coded as messageSelf = 40, messageRival undefined, messageHigher = 1 and messageLower = 0, while "My quality is 25 and my rival's quality is 30" is coded as messageSelf = 25, messageRival = 30, messageHigher = 0 and messageLower = 1. (Note in this latter case that the claim of a lower quality was implicit.)

For the Chat treatment, the classification system was similar, except that it was the entire conversation between the receiver and a sender that receives a single classification, rather than individual messages. So if a sender sends the message "My quality is higher than my rival's", and then later in the stage sends another message "My quality is 35", then the conversation is coded as messageSelf = 35, messageRival undefined, messageHigher = 1 and messageLower = 0. Also, if a sender sends the message "My quality is 40" and then later in the same conversation sent "My quality is 25", the conversation is coded as messageSelf = 25 (the later message taking precedence), messageRival undefined and messageHigher = messageLower = 0.

The classifications themselves were performed by five research assistants. A bilingual research assistant translated all of the messages from Turkish into English, and also classified them. Two Turkish–speaking research assistants classified the original messages without seeing the translations, and two English–speaking research assistants classified the translations without seeing the originals. All classifications were performed independently, and research assistants had access to the experimental instructions but were not told anything about our research questions, nor could they see the senders' true qualities. After all of the classifications were done, a message (in Rich) or conversation (in Chat) was judged to have a non–negative value of messageSelf or messageRival, or a positive value of messageHigher or messageLower, if at least four of the five research assistants had coded it the same way.[14]

Table 5 shows how common the various categories of message are in each treatment. We see a substantial

Table 5: Sender message content, by treatment and sender quality (all rounds)

	Byte	Rich			Chat		
		Low	High	Total	Low	High	Total
Precise claim of own quality	1.000	0.592	0.624	0.608	0.896	0.968	0.932
Precise claim of rival quality	—	0.480	0.536	0.508	0.856	0.936	0.896
Precise claim of both qualities	—	0.472	0.496	0.484	0.848	0.936	0.892
Imprecise claim of higher quality	—	0.336	0.352	0.344	0.056	0.032	0.044
Imprecise claim of lower quality	—	0.000	0.000	0.000	0.008	0.000	0.004
Claim of higher quality (precise or not)	—	0.808	0.848	0.828	0.808	0.952	0.880
Claim of lower quality (precise or not)	—	0.000	0.000	0.000	0.104	0.016	0.060

(Column group heading: Treatment and sender quality)

difference between the Rich and Chat treatments here: while overall claims of higher quality are roughly equally common between the two treatments, the precision of these claims is much higher in the Chat treatment.[15] About 90 percent of senders make precise quality claims in the Chat treatment, compared to just over half in the Rich treatment. These differences are statistically significant (robust rank–order test, $p \approx 0.008$ for both own–quality and rival–quality claims). On the other hand, *imprecise* claims of higher quality are significantly more common in the Rich treatment than in the Chat treatment ($p \approx 0.008$).

Result 3 *Senders are more likely to make precise claims in the Chat treatment than in the Rich treatment, while imprecise claims are more common in the Rich treatment than in the Chat treatment.*

Perhaps unsurprisingly, we also see that there are almost no imprecise claims of lower quality, and very few precise ones, though it might be noteworthy that such claims are non–existent in the Rich treatment and merely rare in the Chat treatment.

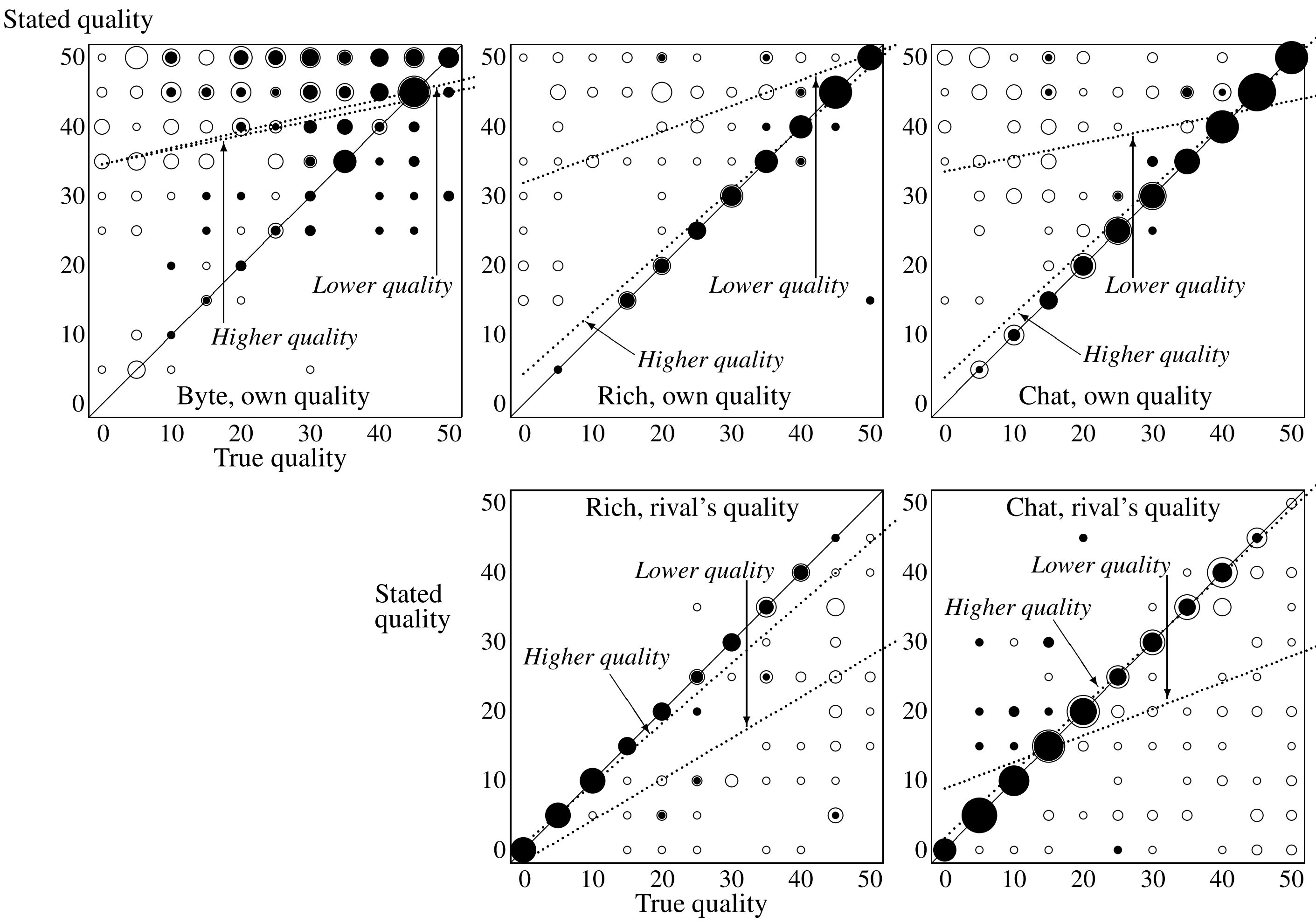
4.3 Sender message truthfulness

The top three panels of Figure 2 show – for the Byte, Rich and Chat treatments respectively – each individual pair of own true quality (horizontal axis) and stated own quality (vertical axis), for all observations in which a precise claim is made. The bottom two panels do the same for rival quality. Within each panel, the area of a dark circle at some particular coordinates is proportional to the number of observations at those coordinates

[14] As an indication of inter–coder reliability, we computed the correlation between responses for each pair of coders (ten possible pairings) for each of the four variables and both Rich and Chat treatments. These 80 correlations ranged from +0.496 to +1, with a median of +0.876.

[15] As already noted, messages in the Byte treatment were constrained to be precise claims about one's own quality.

Figure 2: Scatter–plots, senders' own and rival true–versus–stated quality, individual–level data (dark circles=higher quality sender, light circles/rings=lower quality sender) and OLS trends for higher/lower quality senders



in which the sender had the *higher* true quality. The area of a light circle or light ring is similarly proportional to the number of observations where the sender had the *lower* quality.[16] Finally, each panel also contains the 45–degree line corresponding to truthful messages, and two ordinary–least–squares trend lines: one for higher–quality senders and one for lower–quality senders.

Table 6 provides corresponding summary statistics. For all senders and also separately for lower– and higher–quality senders, and separately for messages about own and rival quality, the table shows the distribution of precise claims between over– and under–statements (stated quality more than and less than true quality, respectively) and truthful statements (stated quality equal to true quality). Also shown is the average magnitude

[16]Thus, a dark circle with a light ring around it means that a particular quality–message combination had come from both higher– and lower–quality senders. A dark circle by itself, or a light circle by itself, depict combinations only belonging to higher– or to lower–quality senders respectively. In all cases, the area of the outer (or only) circle is proportional to the total number of observations of that quality–message combination from all senders.

of over–statements of precise claims of own quality (defined as the difference between stated and true quality if the former is larger, or zero if the latter is at least as large), and the average magnitude of under–statements of precise claims of rival quality (defined similarly). Finally, the bottom of the table shows the fraction of times an imprecise claim of higher quality is true.

Table 6: Truthfulness of sender claims

	All senders			Lower–quality senders			Higher–quality senders		
	Byte	Rich	Chat	Byte	Rich	Chat	Byte	Rich	Chat
Precise claims of own quality									
Fraction truthful	0.224	0.500	0.549	0.112	0.068	0.152	0.336	0.910	0.917
Fraction of under–statements	0.072	0.026	0.004	0.024	0.013	0.000	0.120	0.038	0.008
Fraction of over–statements	0.704	0.471	0.440	0.864	0.919	0.847	0.544	0.051	0.066
Average size of over–statement (overall)	15.34	10.10	10.60	23.32	20.61	21.12	7.36	0.13	0.87
(given over–statement)	22.87	22.01	23.79	27.31	22.50	24.89	15.81	13.75	12.22
Precise claims of rival's quality									
Fraction truthful		0.551	0.629		0.117	0.327		0.940	0.906
Fraction of over–statements		0.008	0.063		0.017	0.038		0.000	0.085
Fraction of under–statements		0.445	0.303		0.867	0.635		0.074	0.009
Average size of under–statement (overall)		8.50	5.78		16.67	13.13		1.19	0.94
(given under–statement)		19.46	21.30		19.42	21.25		20.00	25.00
Imprecise claims of higher quality									
Fraction truthful		0.512	0.364						

The figure and table show several results. Exaggeration abounds: many statements about own quality have coordinates above the 45–degree line (over–statements) while many about rival quality are below it (under–statements). However, truth telling is also common – making up almost half of all precise statements – and maximal lying (claiming a quality of 50 for oneself or 0 for the rival) is quite rare. Rather, over– and under–statements are spread over the interval between truthfulness and maximal lying, indicating both "small" and "large" lies. This reluctance to lie maximally may be due to preferences (e.g., a distaste for lying that increases in the "size" of the lie), or to a belief that smaller lies are more likely to be believed by the receiver.

Overall, exaggeration is much more pronounced amongst lower–quality senders, who tell the truth with a frequency ranging from about one–third of the time down to less than 10 percent of the time across the five cases. By contrast, higher–quality senders are largely truthful, with the notable exception of the Byte treatment, where they lie about as much as their rivals. (We conjecture that being unable to send a message about the rival's quality leads higher–quality senders in the Byte treatment to over–state their own quality to guard against likely over–statements by the rival.) Non–parametric tests largely confirm these apparent differences. Higher–quality senders are significantly more likely to tell the truth, and have significantly lower average over–statements, in all treatments and for both own– and rival–quality messages (two–sided Wilcoxon signed–ranks test, session–level data, $p = 0.0625$ for both comparisons).[17]

Moving on to treatment effects, the fraction of truthful messages about own quality is significantly higher overall in both Rich and Chat treatments than in the Byte treatment (two–sided robust rank–order test, session–level data, $p \approx 0.008$ for both comparisons), and it is weakly significantly higher in the Chat treatment than in the Rich treatment ($p \approx 0.053$). The average over–statement shows similar patterns – with less over–statement in either Rich or Chat than in Byte ($p \approx 0.081$ for both comparisons), while there is no significant difference between Chat and Rich treatments ($p > 0.20$).

For messages about the rival's quality, the fraction that is truthful is again higher in Chat than Rich, and this difference is weakly significant ($p \approx 0.056$), while the average over–statement is significantly lower in Chat than Rich ($p \approx 0.008$). Direct comparisons within the Chat treatment between messages about own quality and messages about rival quality show that the latter are more likely to be truthful and have less average mis–statement – with the latter defined as over–statement for messages about own quality, and under–statement for rival quality (Wilcoxon signed–ranks test, $p \approx 0.063$ for both comparisons). Similar comparisons within the Rich treatment yield no significant results ($p > 0.20$ for both comparisons).

By contrast, there are no significant differences – either between treatments or between own– and rival–quality messages within a treatment – in the average mis–statement given that there was a mis–statement (over–statement about own quality or under–statement about rival quality), suggesting that differences in mis–statements are driven by differences in the frequency, but not the severity, of false statements.

Before moving on, we note some significance results for higher– and lower–quality senders separately; these should be interpreted with a degree of caution since our original research questions did not distinguish between these two groups. For higher–quality senders' messages about their own quality, lying is more common and over–statement higher on average in the Byte treatment than in the Chat and Rich treatments (robust rank–order test, $p \approx 0.008$ for all four comparisons), while there are no significant differences between these latter two treatments ($p > 0.20$ for both comparisons). Their messages about the rival's quality do not differ in truthfulness between the Chat and Rich treatments ($p > 0.20$), but the average extent of under–statement is significantly lower in the Chat treatment ($p \approx 0.014$), due to a (counter–intuitive) small but non–zero frequency of *over*–reporting rival quality. For lower–quality senders, there are no significant differences in either truthfulness or average mis–statement between the Byte treatment and either of the other two treatments ($p > 0.20$ for all comparisons). There is significantly less lying in the Chat treatment than in the Rich treatment about both own and rival quality ($p \approx 0.044$ for both comparisons), though the corresponding differences in mis–statement are insignificant ($p > 0.20$ for over–statement of own quality, $p \approx 0.130$ for under–statement of rival quality).

Additionally, there are patterns in the kinds of lies senders tell which, while tangential to our main research

[17] The p–value of 0.0625 is not below the 5–percent threshold normally used to distinguish significance from "weak significance". However, it should be noted that this is the lowest possible p–value in a two–sided within–subjects comparison using session–level data (the more commonly used one–sided test would yield a p–value of 0.03125). Given that the maximum significance is found in all five of the possible comparisons, we are comfortable in calling this result significant rather than only weakly significant.

questions, are interesting. For example, "switching" (reporting the rival sender's quality as one's own, and one's own quality as the rival's) by the lower–quality sender is common. When a lower–quality sender sends precise signals about own and rival quality, he switches about 20 percent of the time in the Rich treatment and 30 percent in the Chat treatment. This is a substantial fraction: more frequent than truth telling by lower–quality senders (see Table 5).) We find no evidence that switching is more successful than other kinds of lying (in the sense of being chosen more often by the receiver), so its prevalence might be attributable to placing a lower cognitive load on the sender, relative to other lies.

In Table 7, we report some associations between message truthfulness and effort spent on messages. While we have treated messages as cheap talk thus far, we should point out that this is only approximately true: typing and entering messages requires some non–zero effort, and it is reasonable to suppose that writing more or longer messages requires more effort. If writing lies is more effort–costly than writing the truth (due to an aversion to lying, or perhaps the cognitive load from trying to keep one's story straight), then there should be a positive correlation between length of messages and truthfulness. Table 7 shows that for both Rich and

Table 7: Pairwise correlations between message effort and message truthfulness

	Total message length (characters)		Number of messages (Chat)
	Rich	Chat	
Truthful message about self	$+0.165^{**}$	$+0.182^{***}$	$+0.132^{**}$
Truthful message about rival	+0.085	+0.086	$+0.119^{*}$

* (**,***): *Correlation significantly different from zero at the 10% (5%, 1%) level.*

Chat treatments, for messages about oneself and the rival sender, and for two measures of message effort (total number of characters in Rich and Chat, total number of messages in Chat), the correlation with the truthful–message indicator is positive, though statistical significance varies.

4.4 Sender message regressions

We continue the analysis of sender messages with regressions, reported in Table 8. Since messages about own and rival quality are sent concurrently, we use simultaneous–equation specifications. Our first specification (model 3) has indicators for truthful statements about own and rival quality as the two dependent variables; the second (model 4) uses the amount of over–statement of own quality and amount of under–statement of rival quality. We use the seemingly–unrelated–variable technique for both specifications. The two "truthful statement" variables are indicators, while the over–/under–statement variables are approximately continuous, so we estimate probit and linear models respectively, and as before, we cluster standard errors at the session level. To arrive at our sample, we begin with the sender data from all treatments, and then drop those observations where no precise quality claim is made, leaving 635 observations.

Our main explanatory variables are indicators for the Byte and Chat treatments for statements about own quality, and for the Chat treatment for statements about rival quality (so Rich is the baseline). As controls, we include an indicator for being the higher–quality sender and its products with the treatment indicators. Additional controls are the magnitude of the difference in qualities, the total message length (in characters) and an indicator for female (Capraro, 2018, reports sex differences in lying based on a meta–analysis of lab experiments), as well as subject random effects and a constant term. As before, we include the remaining demographic and attitudinal variables in the regressions, though to save space we do not include them in Table 8.[18]

Table 8: Simultaneous–equation regression results for sender precise quality claims (average marginal effects, std. errors in parentheses)

	[3]	[4]
Dependent variable:	Truthful, own quality	Size of over–statement, own quality
Byte treatment	–0.128	6.501^{***}
	(0.081)	(1.431)
Chat treatment	0.066	–0.939
	(0.075)	(1.599)
p–value for differences between Byte and Chat	$p \approx 0.073$	$p \approx 0.002$
High–quality sender	0.580^{***}	-18.739^{***}
	(0.036)	(0.920)
Difference in qualities	–0.0004	0.237^{***}
	(0.0008)	(0.038)
Total message length	0.0002	0.003
	(0.0003)	(0.005)
Female	-0.073^{*}	1.867^{*}
	(0.039)	(1.134)
Dependent variable:	Truthful, rival quality	Size of under–statement, rival quality
Chat treatment	0.231^{***}	-11.775^{***}
	(0.101)	(1.721)
High–quality sender	0.329^{***}	-16.065^{***}
	(0.052)	(0.993)
Difference in qualities	–0.000	0.162^{***}
	(0.001)	(0.041)
Total message length	0.0005^{**}	-0.012^{***}
	(0.0002)	(0.005)
Female	–0.043	5.625^{***}
	(0.048)	(1.230)
Demographic/attitudinal vars?	Yes	Yes
N	635	635

* (**,***): *Marg. effect significantly different from zero at the 10% (5%, 1%) level.*

[18] These variables turn out to be jointly significant, but nearly all are not significant individually after correction for multiple comparisons.

The marginal effects of the Byte–treatment indicator are negative in Model 3 and positive in Model 4, indicating both more lying and "larger lies" about own quality in the Byte treatment compared to the Rich treatment, though only the latter effect is significant. The four marginal effects of the Chat–treatment indicator imply less lying and "smaller lies" in the Chat treatment than in the Rich treatment, though the differences for own–quality messages are not significant. As transitivity would suggest, the Chat treatment has less lying and smaller lies about own quality compared to the Byte treatment.

A few other noteworthy results are visible in Table 8. The marginal effects of the "high–quality sender" indicator are consistent with results shown earlier that senders lie less when they have the higher quality. The effect of the difference in qualities suggests that as the stakes for the receiver rise, senders are not more or less likely to lie, but their lies become bigger. This is the opposite of what would be implied by a cost of lying that increases in the harm done to others, but consistent with a desire to be chosen by the receiver.[19] There is some evidence from the "female" indicator that females lie more (and more often) than males. Finally, consistent with the correlations presented in Table 7, we find that longer messages are more likely to be truthful regarding the rival's quality, though there is no significant association with the sender's own quality.

Result 4 *Senders' messages about their own quality are more truthful in the Chat and Rich treatments than in the Byte treatment, though there are no significant differences between the Chat and Rich treatments.*

Result 5 *Senders' messages about the rival's quality are more truthful in the Chat treatment than in the Rich treatment.*

4.5 A classification of message corpora

The results in the previous section suggest that differences in outcomes across treatments may be attributable to differences in message truthfulness. Of course, receivers typically do not know whether a given message is true or false. In this section, we look at a distinct but related notion. We take advantage of the fact that receivers receive messages from two senders to create a proxy for truthfulness: whether the messages are consistent with one another. This necessitates examination of the entire corpus of messages, from both senders to a given receiver in a round.[20] We call the corpus *jointly consistent* if any of these three conditions holds:

1. One of the senders makes a claim (precise or imprecise) to have either a higher or a lower quality than the other, and the other sender does not contradict that claim.

2. Both senders make precise claims of own quality, these claims are distinct, and their ordering is not contradicted by anything else in the corpus of messages.

[19] Additional regressions, not reported here, show that the effect of the difference–in–qualities variable are robust to adding the own and rival qualities (and their interactions) to the model, so the result seen here is not an artefact of any presumed high correlation between these variables and the absolute quality difference.

[20] We borrow the term "corpus" from linguistics, where it can pertain to a set of messages on which some analysis is performed. "Corpora" is the plural of corpus.

3. Both senders make precise claims of the rival's quality, these claims are distinct, and their ordering is not contradicted by anything else in the corpus.

An example of the first condition is if the first sender says "my quality is 40, the other sender's is 30" and the second sender says nothing or only irrelevances.[21] By contrast, if the second sender had said "my quality is 50, the other sender's is 40" or "I have the higher quality", the messages would not have been jointly consistent (in either of these cases, they would have been categorised as "jointly inconsistent", as described below). An example of the second condition is if the first sender says "my quality is 40" and the second sender says "my quality is 30", with nothing else said. (Note that based on this condition, messages in the Byte treatment are always jointly consistent unless both senders claim the same quality.) The third condition was never observed in the data, but we include it for completeness.

Joint consistency – which we will often abbreviate as "consistency" – simply means that the senders' messages have a common implication for the receiver's choice (i.e., which sender has the higher quality). It is clearly a weaker condition than truthfulness; a corpus of messages can have a common implication even if one or both senders is lying. Indeed, some plausible patterns of lying (e.g., in the Byte treatment, one sender truthfully reporting 40 and the other reporting 50 when his true quality is 20) would still yield consistency. Conversely, in the Rich and Chat treatments, it is possible for both senders to be truthful without their messages being consistent (e.g., if one sender says "my quality is 30" and the other says "my rival's quality is 30"). If the messages are consistent, the receiver has two options: choose the sender indicated by the messages as having the higher quality, or the other sender. We call these "following" and "flouting" the consistent messages, respectively.

There are three ways that the senders' messages can fail to be consistent. The most important of these is when senders make claims that are inconsistent with each other. We call these *(jointly) inconsistent* messages. One pattern of jointly inconsistent messages – the lower–quality sender lying enough to "beat" the higher–quality sender's truthful message – was mentioned above. Other examples exist, such as both senders making imprecise claims to have the higher (or lower) quality, or both claiming to have an own quality of 40. A second way messages can fail to be consistent is if neither sender makes even an imprecise claim about quality; we call this the case of *(jointly) irrelevant* messages. A third way, mentioned above, is when messages are neither jointly irrelevant nor jointly inconsistent, but still provide insufficient information for the receiver to choose a sender (e.g., both senders make a precise statement about one sender but say nothing about the other). We call these *(jointly) uninformative* messages. As it happens, the cases of uninformative and irrelevant messages do not occur in our data; we therefore do not discuss them further.

[21] Of course, the messages would also have been jointly consistent if the second sender had said something like "my quality is lower than the other sender's" or "my quality is 30 and the other sender's is 40". As observed in Table 5, however, admissions of the lower quality are rare. So jointly consistent messages in our setting are likely to involve a claim by one sender that it not denied by the other, rather than a claim by one that is explicitly verified by the other.

4.6 Receiver responses to sender messages

Table 9 displays receivers' success frequency for each treatment according to two factors: the consistency of the corpus of messages as described in the section above, and the precision of these messages as discussed earlier. Recall from Section 4.2 that each sender can send a precise message (or not) about both his own quality and his rival's quality; so for a given receiver, the number of precise signals received ranges from 0 to 4. In the case of consistent messages, we additionally break down the success rate by whether the messages' implication was followed or flouted.

After *inconsistent* messages, receivers are roughly as successful as chance would predict, with the possible exception of the Chat treatment where they appear to do better.[22] In this case, the number of precise signals does not appear to have an effect.

When messages are *consistent*, the frequency with which they are followed varies across treatments, from about half the time (53 percent) in the Byte treatment to 66 percent in the Rich treatment to 85 percent in the Chat treatment. In the Rich and Chat treatments, receivers chose correctly exactly half of the time when flouting consistent messages, and hence would not have done better by following them – even though they were substantially more successful in those cases where they actually did follow them. In the Byte treatment, by contrast, receivers would have been successful about 60 percent of the time after consistent messages if they had simply followed them, compared to an observed success rate just under 50 percent. The number of precise signals appears to be positively correlated with receiver success in the Rich treatment, but not in the other two treatments.

Further evidence of the effects of treatment, message content, and message precision is presented in Table 10, based on a probit model similar to those in Table 4. Our sample is all of the receiver data, and the dependent variable is an indicator for choosing the higher–quality sender. Our main right–hand–side variables are the indicators for Byte and Chat treatments (so Rich is the baseline) and for consistent messages (so that the baseline is inconsistent message), and the number of precise signals. We also include the product of this last variable with the Chat indicator, and the product of the consistent–message indicator with all of these other variables (in particular, including the three–way interaction beween Chat treatment, consistent messages, and number of precise signals), as well as a constant term.[23] Table 10 shows marginal effects and standard errors. Rather than the usual average marginals, we present marginals conditional on particular values for the treatment dummies, the consistent–messages dummy, and/or the number of precise signals.

The average marginal effect of the Byte treatment (compared to the baseline case of the Rich treatment) on correct choices by receivers is negative but insignificant (point estimate –0.049, standard error 0.047), but as the table shows, its effect is positive but insignificant after inconsistent messages, and significant and negative when messages are consistent. The latter suggests that consistent messages are less useful in the Byte treatment,

[22] As a rough measure of performance versus chance, consider chi–square tests with the null hypothesis that success for a receiver in a given round is just an i.i.d. 50–50 coin toss. Such a null hypothesis can be rejected for inconsistent messages in the Chat treatment, but only at the 10–percent level ($p \approx 0.069$), while it cannot be rejected in the Byte or Rich treatments ($p > 0.20$ in both cases).

[23] Additional probits, not reported here, indicate that demographics and attitudes have no systematic association with receiver success, nor do objective characteristics of messages such as their lengths.

Table 9: Receiver frequency of correct choices, conditional on sender messages

	Number of precise signals					
	0	1	2	3	4	Total
Byte treatment						
Senders send inconsistent messages			0.545			0.545
			(12/22)			(12/22)
Senders send consistent messages...						
...and followed			0.582			0.582
			(32/55)			(32/55)
...and flouted			0.375			0.375
			(18/48)			(18/48)
...overall success rate			0.485			0.485
			(50/103)			(50/103)
Rich treatment						
Senders send inconsistent messages	0.667	0.500	0.300	0.438	0.486	0.469
	(10/15)	(5/10)	(6/20)	(7/16)	(17/35)	(45/96)
Senders send consistent messages...						
...and followed	0.250	0.667	0.750			0.632
	(1/4)	(2/3)	(9/12)			(12/19)
...and flouted	0.333		0.571			0.500
	(1/3)		(4/7)			(5/10)
...overall success rate	0.286	0.667	0.684			0.586
	(2/7)	(2/3)	(13/19)			(17/29)
Chat treatment						
Senders send inconsistent messages	1.000		0.500	0.833	0.567	0.576
	(1/1)		(1/2)	(5/6)	(51/90)	(58/99)
Senders send consistent messages...						
...and followed		1.000	0.875	1.000	0.917	0.909
		(1/1)	(7/8)	(1/1)	(11/12)	(20/22)
...and flouted		1.000	0.333			0.500
		(1/1)	(1/3)			(2/4)
...overall success rate		1.000	0.727	1.000	0.917	0.846
		(2/2)	(8/11)	(1/1)	(11/12)	(22/26)

where as mentioned earlier, nearly all corpora of messages are consistent. The difference in the two effects is significant ($p \approx 0.011$). By contrast, the Chat treatment has no significant effect after consistent messages compared to the baseline Rich treatment, while its effect after inconsistent messages is positive and at least weakly significant, suggesting that even inconsistent messages in this treatment may carry some information for the receiver. However, differences between these last two effects, holding constant the number of precise signals, are typically insignificant (p–values of 0.82, 0.099, and 0.25 for 0, 2, and 4 precise signals respectively).

Table 10: Receiver regression results (Dependent variable: correct choice) – marginal effects (MEs) and std. errors

ME of....conditional on...	0 precise signals	2 precise signals	4 precise signals
Byte treatment...			
...consistent messages		-0.211^{***}	
		(0.073)	
...inconsistent messages		+0.064	
		(0.071)	
Chat treatment...			
...consistent messages	+0.339	+0.107	–0.030
	(0.349)	(0.190)	(0.102)
...inconsistent messages	$+0.389^{**}$	$+0.310^{*}$	$+0.143^{*}$
	(0.181)	(0.162)	(0.076)
Consistent messages...			
...Byte treatment		–0.060	
		(0.077)	
...Rich treatment	-0.218^{**}	$+0.214^{***}$	$+0.506^{***}$
	(0.109)	(0.076)	(0.097)
...Chat treatment	–0.268	+0.012	$+0.333^{***}$
	(0.186)	(0.097)	(0.114)

	Consistent messages	Inconsistent messages
Number of precise signals...		
...Rich treatment	$+0.172^{**}$	–0.027
	(0.073)	(0.029)
...Chat treatment	+0.062	-0.091^{*}
	(0.062)	(0.050)

*Notes: N=375, $|ln(L)| \approx 248.03$. * (**,***) = ME significantly different from zero at the 10% (5%, 1%) level.*

The effect of consistent messages on correct choices varies by treatment. It is insignificant in the Byte treatment, while in the other two treatments it depends on the number of precise signals: negative (and significant in the Rich treatment) when no precise signals are received, positive (and again significant in the Rich treatment) when 2 precise signals are received, and significantly positive in both Rich and Chat treatments when the maximum of 4 precise signals are received. All six of the associated pairwise within–treatment comparisons are significant ($p \approx 0.023$ for 0 versus 2 precise signals in the Chat treatment, $p < 0.005$ for the other five comparisons), as are the two corresponding three–way within–treatment comparisons ($p < 0.001$ for joint comparison of 0 versus 2 and 2 versus 4 in both Rich and Chat). The interaction between consistent messages and precise signals is also visible in the positive marginal effect of the number of precise signals in both Rich and Chat treatments when messages are consistent, and the negative effects when messages are inconsistent, though only two of these four marginals are significant. The corresponding within–treatment differences are significant ($p \approx 0.002$ and $p < 0.001$ for consistent versus inconsistent messages in the Rich and Chat treatments respectively), while the differences between Rich and Chat treatments are not ($p \approx 0.27$ and $p \approx 0.25$ for consistent and inconsistent messages respectively).

Result 6 *Receivers are more likely to choose higher–quality senders when (i) messages are consistent and (ii) when signals are precise. Either of these two criteria individually is not sufficient.*

Result 6, summarising Tables 9 and 10, suggests that the differences we have seen across our Byte, Rich, and Chat treatments arise from differences in how likely these treatments are to yield the combination of (i) consistent messages between the two senders and (ii) precise claims of the senders' qualities.[24] In particular, the better performance of receivers in the Chat treatment compared to the other two treatments is due to the receiver's ability in that treatment to extract more precise statements from the senders (as also shown in Table 5), making conversations with both consistency and precise statements more common.

5 Summary and discussion

The setting we consider – a choice by an uninformed receiver between two well–informed senders – captures many real–life situations. Given that the receiver has minimal ex–ante information about the senders (only the range of potential sender qualities), and given that all talk by the senders is cheap in the game–theoretic sense, we should have low expectations regarding the receiver's ability to pick the better sender. Indeed, receivers on average perform no better than coin tosses in two of our treatments: *Byte*, where each sender sends a single message in a restricted domain (whole numbers) meant to indicate his quality; and *Rich*, where each sender sends a single message in natural language, with nearly unrestricted content. By contrast, receivers choose correctly about two–thirds of the time on average in our third treatment, *Chat*, where communication allows for multiple back–and–forth natural–language messages between the receiver and each individual sender. This treatment effect is significant, and at around 15 percentage points, rather striking.

Some insight into *why* receivers fare better in the Chat treatment comes from examining the communication between senders and receivers. In the Chat treatment, we observe that communications are more likely to have the combination of (i) consistency in their implication for the receiver's decision (i.e., which sender should be chosen) and (ii) precise claims about both senders' qualities. Within treatments, conversations with both of these features result in a high degree of receiver success, while when only one (or neither) is present, receivers fare poorly. Indeed, once we control for these factors, the treatment itself has little explanatory power for receivers' success, suggesting that these factors are precisely the mechanism through which our Chat treatment works.

[24] The only apparent exception – the poorer performance following consistent messages in the Byte treatment – is attributable not to any presumed low quality of consistent messages in that treatment, but rather to the high frequency with which these messages' implication is flouted by receivers (47 percent of the time) resulting in a lower success rate of receivers in that case (38 percent, compared to 58 percent success when following consistent messages in the Byte treatment).

This pair of criteria for success makes intuitive sense. If messages are precise but not consistent, the receiver would probably be unable to decipher their combined implication for her decision. If messages are consistent but not precise, the signal they impart is likely to be weak, and possibly containing "lies of omission". But together, precise and consistent messages are likely to be useful to the receiver. Importantly, both of these criteria are discernible by the receiver: verifiable based entirely on information the receiver has at the time. By contrast, while (e.g.) truthfulness would also be a useful characteristic of messages, the receiver never knows whether messages are truthful until it is too late.

Our results have implications for cheap–talk studies broadly. The literature contains many examples where unstructured communication performs better than highly restricted communication, and a smaller number of experiments comparing face–to–face with other kinds of unstructured communication. But there have been few controlled comparisons between distinct types of anonymous unstructured communication, and those that we know of have tended not to find systematic differences. Our finding that back–and–forth chat outperforms one–time unstructured messages suggests that not all anonymous unstructured communication should be considered equivalent. Along with our other main finding – that one–time unstructured messages perform no better than highly structured messages – this implies that the important dichotomy in settings like ours is not structured–versus–unstructured communication, but between one–way and two–way communication.

This begs the question of what is different about our setting compared to those where any kind of communication, or at least any communication in natural language, improves outcomes. We can only conjecture, but one potential explanation concerns the payoff structure of our setting compared to many others in this literature. The latter often involves social dilemmas – most notably the prisoners' dilemma – where by definition cooperation can make all players better off, meaning that the payoffs of different players are partly aligned. By contrast, our setting has very little alignment of payoffs, with the two senders having completely opposed preferences to each other and (on average) orthogonal preferences to the receiver. When there is little or no scope for mutual gains, the stigma associated with lying is likely to be lower than when such gains are possible, and it may be more robust to reductions in social distance on their own.[25] When lying is viewed less as anti–social behaviour and more on a par with bluffing in poker, the kinds of communication that matter will instead be those providing more interaction between players.

This last point should not be surprising, given the role these more interactive kinds of communication have in many real–world settings. Anthropologists have long recognised the importance of explicit requests in soliciting generous behaviour from others, even in foraging societies where such generosity is common and repeated–game considerations should be strong (see, e.g., Peterson, 1993). Examples also abound in western societies: court trials in most countries allow for cross–examination of witnesses, press conferences are typi-

[25] Dictator–game experiments would seem to be an exception to this, as the players' preferences are strictly opposed, but even highly structured communication has been found to increase offers (Rankin, 2006). However, Dana et al. (2006) find that many subjects will choose an outcome where they receive $9 and a would–be recipient earns nothing over being the dictator in a standard dictator game for $10, as long as in the former, the would–be recipient is not told that a game has been played (see also Dana et al., 2007 and Larsen and Capra, 2009). This pattern of behaviour suggests that giving in the dictator game has less to do with generosity and more to do with other situational factors (e.g., aversion to disappointing another person, even behind the veil of anonymity).

cally viewed as shams unless journalists can ask follow–up questions, and much more about political candidates is learned from debates than campaign speeches. In these real settings, it is unlikely that the additional communication works by lowering social distance or by making lying more repugnant; more likely, it works by making lying either more cognitively taxing or more transparent to others.

We close with some suggestions for future research. Even though both the Byte and Rich treatments allow only one message from sender to receiver, the use of natural language in the latter gives rise to the possibility that the message contains multiple pieces of information, while messages in the former necessarily comprise only one piece of information. Byte therefore differs from Rich and Chat in not allowing a message about the *rival's* quality. An alternative "Byte2" treatment could have senders sending *two* numerical messages: one for their own quality and one for the rival's, and would therefore lie in between the Byte and Rich treatments. We conjecture that receivers would not fare better than in our actual Byte treatment, given the lack of observed differences between the Byte and Rich treatments, but we would expect changes to the pattern of lying. Specifically, high–quality senders would likely be largely truthful, while low–quality senders would typically over–state their own quality and under–state the rivals' – as we observed in our Rich and Chat treatments.

Another extension might involve a treatment that lies between our Rich and Chat treatments. A "Rich–plus" treatment would have messages from senders to receivers as in our Rich treatment, but preceded by the option for the *receiver* to send a message to each sender. This would allow for a limited version of the two–way communication present in our Chat treatment, but with no opportunity for the receiver to confront senders with new information. Such a treatment would help to disentangle the effect of back–and–forth communication from that of allowing receivers to request truthfulness or informativeness from the senders.

One might also examine other implementations of back–and–forth communication. In our Chat treatment, the receiver could converse with each sender, but senders could not communicate directly with each other. Alternatively, there could be a "public chat room" (with messages viewable by the receiver and both senders), either with or without the private chat rooms between the receiver and individual senders, or indeed a private chat room for only the two senders. We conjecture that allowing more public communication would weakly increase truthfulness and receivers' success, though the extent of this improvement may be small since receivers already perform well in our current Chat treatment.

Finally, it might be worthwhile to examine settings with one receiver and more than two senders. We expect that adding a third, or a fourth, sender will change the environment less than going from one to two senders does, but since the theoretical prediction for correct choices varies inversely with the number of senders, increasing that number will allow more room to detect treatment effects. We encourage future research on these extensions and others.

References

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

Ambrus, A. and S. Takahashi (2008), "Multi–sender cheap talk with restricted state spaces", *Theoretical Economics* 3, pp. 1–27.

Ariely, D. (2012), *The Honest Truth About Dishonesty: How We Lie to Everyone – Especially Ourselves*, New York: HarperCollins.

Austen-Smith, D. (1990a), "Information transmission in debate", *American Journal of Political Science* 34, pp. 124–152.

Austen-Smith, D. (1990b), "Credible debate equilibria", *Social Choice and Welfare* 7, pp. 75–93.

Austen-Smith, D. (1993a), "Interested experts and policy advice: multiple referrals under open rule", *Games and Economic Behavior* 5, pp. 3–43.

Austen-Smith, D. (1993b), "Information acquisition and orthogonal argument", in *Political Economy: Institutions Competition and Representation (Proceedings of the Seventh International Symposium in Economic Theory and Econometrics)*, eds. Barnet, W.A., Melvin. H.J. and N. Schofield. Cambridge; New York and Melbourne: Cambridge University Press, pp. 407–436.

Azfar, O. and W.R. Nelson, Jr. (2007), "Transparency, wages, and the separation of powers: an experimental analysis of corruption", *Public Choice* 130, pp. 471–493.

Balliet, D. (2010), "Communication and cooperation in social dilemmas: a meta–analytic review", *Journal of Conflict Resolution* 54, pp. 39–57.

Barnes, J.A. (1994), *A Pack of Lies: Towards a Sociology of Lying*, Cambridge: Cambridge University Press.

Batson, C.D., D. Kobrynowicz, J.L. Dinnerstein, H.C. Kampf, and A.D. Wilson (1997), "In a very different voice: unmasking moral hypocrisy", *Journal of Personality and Social Psychology* 72, pp. 1335–1348.

Batson, C.D., E.R. Thompson, G. Seuferling, H. Whitney and J.A. Strongman (1999), "Moral hypocrisy: appearing moral to oneself without being moral", *Journal of Personality and Social Psychology* 77, pp. 525–537.

Battaglini, M. (2002). "Multiple referrals and multidimensional cheap talk", *Econometrica* 70, pp. 1379–1401.

Battaglini, M. (2004), "Policy advice with imperfectly informed experts", *Advances in Theoretical Economics* 4, issue 1, article 1.

Belot, M., V. Bhaskar and J. van de Ven (2012), "Can observers predict trustworthiness?" *Review of Economics and Statistics* 94, pp. 246–259.

Belot, M. and J. van de Ven (2017), "How private is private information? The ability to spot deception in an economic game", *Experimental Economics* 20, pp. 19–43.

Benjamini, Y. and Y. Hochberg (1995), "Controlling the false discovery rate: a practical and powerful approach to multiple testing", *Journal of the Royal Statistical Society, Series B (Methodological)* 57, pp. 289–300.

Berry, L.M. (2003), *Employee Selection*. Belmont, CA: Wadsworth.

Bochet, O., T. Page and L. Putterman (2006), "Communication and punishment in voluntary contribution experiments", *Journal of Economic Behavior & Organization* 60, pp. 11–26.

Bohnet, I. and B. Frey (1999), "Social distance and other–regarding behavior in dictator games: comment", *American Economic Review* 89, pp. 335–339.

Bond, C.F., Jr. and B.M. DePaulo (2006), "Accuracy of deception judgments", *Personality and Social Psychology Review* 10, pp. 214–234.

Bond, C.F., Jr. and B.M. DePaulo (2008), "Individual differences in judging deception: accuracy and bias", *Psychological Bulletin* 134, pp. 477–492.

Boush, D.M., M. Friestad and P. Wright (2015), *Deception in the Marketplace: The Psychology of Deceptive Persuasion and Consumer Self–Protection*. Routledge, New York.

Braun, M., L.M. Van Swol and L. Vang (2015), "His lips are moving: Pinocchio effect and other lexical indicators of political deceptions", *Discourse Processes* 52, pp. 1–20.

Brosig, J., A. Ockenfels and J. Weimann (2003), "The effect of communication media on cooperation", *German Economic Review* 4, pp. 217–242.

Brosig-Koch, J. and T. Heinrich (2018), "The role of communication content and reputation in the choice of transaction partners: a study based on field and laboratory data", *Games and Economic Behavior* 112, pp. 49–66.

Buller, D.B. and J.K. Burgoon (1996), "Interpersonal deception theory", *Communication Theory* 3, pp. 203–242.

Burgoon, J.K., W.J. Mayew, J.S. Giboney, A.C. Elkins, K. Moffitt, B. Dorn, M. Byrd and L. Spitzley (2016), "Which spoken language markers identify deception in high–stakes settings? Evidence from earnings conference calls", *Journal of Language and Social Psychology* 35, pp. 123–157.

Cai, H. and J.T.Y. Wang (2006), “Overcommunication in strategic information transmission games”, *Games and Economic Behavior* 56, pp. 7–36.

Callander, S. and S. Wilkie (2007), “Lies, damned lies, and political campaigns”, *Games and Economic Behavior* 60, pp. 262–286.

Capraro, V. (2018), “Gender differences in lying in sender–receiver games: a meta–analysis”, *Judgment and Decision Making* 13, pp. 345–355.

Charness, G. and M. Dufwenberg (2006), “Promises and partnership”, *Econometrica* 74, pp. 1579–1601.

Charness, G., F. Feri, M.A. Meléndez-Jiménez and M. Sutter (2019), “An experimental study on the effects of communication, credibility, and clustering in network games”, CESifo working paper no. 7659.

Chen, J. and D. Houser (2017), “Promises and lies: can observers detect deception in written messages”, *Experimental Economics* 20, pp. 396–419.

Cliffe, L., M. Ramsay and D. Bartlett (2000), *The Politics of Lying: Implications for Democracy*, London: MacMillan Press.

Conrads, J. and S. Lotz (2015), “The effect of communication channels on dishonest behavior”, *Journal of Behavioral and Experimental Economics* 58, pp. 88–93.

Conrads, J. and T. Reggiani (2017), “The effect of communication channels on promise–making and promise–keeping: experimental evidence”, *Journal of Economic Interaction and Coordination* 12, pp. 595–611.

Corazzini, L., S. Kube, M.A. Maréchal and A. Nicolò (2014), “Elections and deceptions: an experimental study on the behavioral effects of democracy”, *American Journal of Political Science* 58, pp. 579–592.

Cosmides, L. and J. Tooby (1992), “Cognitive adaptations for social exchange”, in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds. J. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press, pp. 163–228.

Cowley, E. and C.I. Anthony (2019), “Deception memory: when will consumers remember their lies?” *Journal of Consumer Research* 46, pp. 180–199.

Crawford, V. and J. Sobel (1982), “Strategic information transmission”, *Econometrica* 50, pp. 1431–1451.

Dana, J., D.M. Cain and R.M. Dawes (2006),“What you don’t know won’t hurt me: costly (but quiet) exit in dictator games”, *Organizational Behavior and Human Decision Processes* 100, pp. 193–201.

Dana, J., R. Weber and J. Kuang (2007), “Exploiting moral wriggle room: experiments demonstrating an illusory preference for fairness”, *Economic Theory* 33, pp. 67–80.

Darke, P.R. and R.J.B. Ritchie (2007), “The defensive consumer: advertising deception, defensive processing, and distrust”, *Journal of Marketing Research* 44, pp. 114–127.

Dawkins, R. (1978), *The Selfish Gene*. New York: Oxford University Press.

Dawkins, R. and J.R. Krebs (1978), "Animal signals: information or manipulation?", in *Behavioral Ecology*, eds. J.R. Krebs and N.B. Davies. Oxford: Blackwell Publishers, pp. 282–309.

DePaulo, B.M., D.A. Kashy, S.E. Kirkendol, M.M. Wyer and J.A. Epstein (1996), "Lying in everyday life", *Journal of Personality and Social Psychology* 70, pp. 979–995.

DePaulo, B.M., J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton and H. Cooper (2003), "Cues to deception", *Pschological Bulletin* 129, pp. 74–118.

DePaulo, B.M., J.I. Stone and G.D. Lassiter (1985), "Deceiving and detecting deceit", in *The Self and Social Life*, ed. B.R. Schlenker. New York: McGraw-Hill, pp. 323–370.

Dickhaut, J., K. McCabe and A. Mukherji (1995), "An experimental study on strategic information transmission", *Economic Theory* 6, pp. 389–403.

Duffy, J. and N. Feltovich (2002), "Do actions speak louder than words? An experimental comparison of observation and cheap talk", *Games and Economic Behavior* 39, pp. 1–27.

Duffy, J. and N. Feltovich (2006), "Words, deeds, and lies: strategic behaviour in games with multiple signals", *Review of Economic Studies* 73, pp. 669–688.

Dwenger, N. and T. Lohse (2019), "Do individuals successfully cover up their lies? Evidence from a compliance experiment", *Journal of Economic Psychology* 71, pp. 74–87.

Ekman, P. and M. O'Sullivan (1991), "Who can catch a liar?" *American Psychologist* 46, pp. 913–920.

Ellingsen, T. and M. Johannesson (2008), "Pride and prejudice: the human side of incentive theory", *American Economic Review* 98, pp. 990–1008.

Feltovich, N. (2005), "Critical values for the robust rank–order test", *Communications in Statistics—Simulation and Computation* 34, pp. 525–547.

Feltovich, N. and F. Giovannoni (2015), "Selection vs. accountability: an experimental investigation of campaign promises in a moral–hazard environment", *Journal of Public Economics* 126, pp. 39–51.

Fischbacher, U. (2007), "z–Tree: Zurich toolbox for ready–made economic experiments", *Experimental Economics* 10, pp. 171–178.

Fischbacher, U. and F. Föllmi-Heusi (2013), "Lies in disguise: an experimental study on cheating", *Journal of the European Economic Association* 11, pp. 525–547.

Frohlich, N. and J. Oppenheimer (1998), "Some consequences of e–mail versus face–to–face communication in experiment", *Journal of Economic Behavior & Organization* 35, pp. 389–403.

Gick, W. (2008), "Cheap talk equilibria – a note on two senders", working paper, Harvard University.

Gilligan, T. and K. Krehbiel (1989), "Asymmetric information and legislative rules with a heterogeneous committee", *American Journal of Political Science* 33, pp. 459–490.

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

Gneezy, U. (2005), "Deception: the role of consequences", *American Economic Review* 95, pp. 384–394.

Gunia, A. and E. Levine (2019), "Deception as competence: the effect of occupational stereotypes on the perception and proliferation of deception", *Organizational Behavior and Human Decision Processes* 152, pp. 122–137.

de Haan, T., T. Offerman and R. Sloof (2015), "Money talks? An experimental investigation of cheap talk and burned money", *International Economic Review* 56, pp. 1385–1426.

Harada, T., S. Itakura, F. Xu, K. Lee, S. Nakashita, D.N. Saito and N. Sadato (2009), "Neural correlates of the judgment of lying: a functional magnetic resonance imaging study", *Neuroscience Research* 63, pp. 24–34.

Hartwig, M., P.A. Granhag, and L.A. Stromwell (2004), "Police officers' lie detection accuracy: interrogating freely versus observing video", *Police Quarterly* 7, pp. 429–456.

Iyer, R., A.I. Khwaja, E.F. Luttmer and K. Shue (2015), "Screening peers softly: inferring the quality of small borrowers", *Management Science* 62, pp. 1554–1577.

Kocher, M.G., S. Schudy and L. Spantig (2018), "I Lie? We Lie! Why? Experimental evidence on a dishonesty shift in groups", *Management Science* 64, pp. 3995–4008.

Kartik, N., 2009, "Strategic communication with lying costs", *Review of Economic Studies* 76, pp. 1359–1395.

Konrad, K.A., T. Lohse and S. Qari (2014), "Deception choice and self–selection: the importance of being earnest", *Journal of Economic Behavior & Organization* 107, pp. 25–39.

Krishna, V., and J. Morgan (2001), "A model of expertise", *Quarterly Journal of Economics* 116, pp. 747–775.

Lai, E.K., W. Lim, and J.T.-Y. Wang (2015), "An experimental analysis of multidimensional cheap talk", *Games and Economic Behavior* 91, pp. 114–144.

Larson, T. and C.M. Capra (2009), "Exploiting moral wiggle room: illusory preference for fairness? A comment", *Judgment and Decision Making* 4, pp. 467–474.

Levine, D.K. (1998), "Modeling altruism and spitefulness in experiments", *Review of Economic Dynamics* 1, pp. 593–622.

Li, M. (2008), "Two (talking) heads are not better than one", *Economics Bulletin* 3, pp. 1–8.

Lipman, M. and W.R. McGraw (1988), *A Tremor in the Blood: Uses and Abuses of the Lie Detector*. New York: McGraw–Hill.

Machiavelli, N. (1513), *The Prince*, Appleton–Century–Crofts, 1964.

Meibauer, J. (2011), "On lying: intentionality, implicature, and imprecision", *Intercultural Pragmatics* 8, pp. 277–292.

Milgrom P. and J. Roberts (1986), "Relying on information of interested parties", *RAND Journal of Economics* 17, pp. 18–32.

Minozzi, W. and J. Woon (2016) "Competition, preference uncertainty, and jamming: a strategic communication experiment", *Games and Economic Behavior* 96, pp. 97–114.

Minozzi, W. and J. Woon (2018) "Competition and exaggeration in experimental cheap talk games", working paper, University of Pittsburgh.

Minson, J.A., E.M. Van Epps, J.A. Yip and M.E. Schweitzer (2018), "Eliciting the truth, the whole truth, and nothing but the truth: the effect ofquestion phrasing on deception", *Organizational Behavior and Human Decision Processes* 147, pp. 76–93.

Newman, M.L., J.W. Pennebaker, D.S. Berry and J.M. Richards (2003), "Lying words: predicting deception from linguistic styles", *Personality and Social Psychology Bulletin* 29, pp. 665–675.

O'Connor, A.M., T.D. Lyon and A.D. Evans (2019), "Younger and older adults' lie–detection and credibility judgments of children's coached reports", forthcoming, *Psychology, Crime, & Law*.

Peterson, N. (1993), "Demand sharing: reciprocity and the pressure for generosity among foragers", *American Anthropologist (new series)* 95, pp. 860–874.

Rankin, F.W. (2006), "Requests and social distance in dictator games", *Journal of Economic Behavior & Organization* 60, pp. 27–36.

Sally, D. (1995), "Conversation and cooperation in social dilemmas: a meta–analysis of experiments from 1958 to 1992", *Rationality and Society* 7, pp. 58–92.

Sánchez-Pagés, S. and M. Vorsatz (2007), "An experimental study of truth–telling in a sender-receiver game", *Games and Economic Behavior* 61, pp. 86–112.

Schein, V.E. (1979), "Examining an illusion: the role of deceptive behaviors in organizations", *Human Relations* 32, pp. 287–295.

Siegel, S. and N.J. Castellan, Jr. (1988), *Nonparametric Statistics for the Behavioral Sciences*, McGraw–Hill, New York.

Trivers, R. (1985), *Social Evolution*. Menlo Park, CA: Benjamin/Cummings.

Vartanian et al., O., P. Kwantes and D.R. Mandel (2012), "Lying in the scanner: localized inhibition predicts lying skill", *Neuroscience Letters* 529, pp. 18–22.

Vespa, E. and A.J. Wilson (2016), "Communication with multiple senders: an experiment", *Quantitative Economics* 7, pp. 1–36.

von Hippel, W., E. Baker, R. Wilson, L. Brin and L. Page (2016), "Detecting deceptive behaviour after the fact", *British Journal of Social Psychology* 55, pp. 195–205.

Vrij, A. (2008), *Detecting Lies and Deceit: Pitfalls and Opportunities*, 2nd Edition. West Sussex (England): Wiley.

Vrij, A., P.A. Granhag and S. Porter (2010), "Pitfalls and opportunities in nonverbal and verbal lie detection", *Psychological Science in the Public Interest* 11, pp. 89–121.

Walczyk, J.J., J.P. Schwartz, R. Clifton, B. Adams, M. Wei and P. Zha (2005), "Lying person–to–person about life events: a cognitive framework for lie detection", *Personnel Psychology* 58, pp. 170.

Wright, R. (1995), *The Moral Animal: Why We Are The Way We Are: The New Science of Evolutionary Psychology*. Oxford: Oxford University Press.

Zuckerman, M., B.M. DePaulo, and R. Rosenthal (1981), "Verbal and nonverbal communication of deception", in *Advances in Experimental Social Psychology* (vol. 14), ed. L. Berkowitz. New York: Academic Press, pp. 1–59.

A English translation of experimental instructions

The instructions below are translated from those used in the Rich treatment. The instructions from the other treatments are similar and available upon request from the corresponding author.

Welcome,

Today's experiment is about the economic choices of people in certain situations. The experiment will last for 5 rounds and you will earn a monetary payment as a result of your decisions. Your earnings will depend on your and other players' decisions and chance. These earnings will be paid in cash at the end of the experiment.

No other participant will learn about the choices you make and the amount you earn. The decisions you make during the experiment will be recorded with a participant number randomly assigned to you by the computer and never be matched with your first and last name.

If you have a question, please raise your hand. An experimenter will come and answer your question. From now on, participants are not allowed to talk to each other, in which case we will have to terminate the experiment.

Roles:

At the beginning of the experiment you will be assigned to one of the Buyer and Sender roles and you will play in this role throughout the experiment.

Groups:

In each round, groups of three persons are formed, one of which is the Buyer and the other two are the Sender(s). Groups are constantly changing. None of the participants will be in the same group more than once. So there will always be different people in your group.

Game:

Each round in the game will proceed as follows:

A) A number is assigned to each person in the Sender role separately. These numbers are randomly selected among 0–5–10–15–20–25–30–35–40–45–50. The Senders can observe both the number assigned to them and the number assigned to the other Sender. Receivers can not observe these numbers. The numbers assigned to Senders are set to be different from each other.

B) The Sender sends a text message to the Receiver. The size of this message can range from 0 to 500 characters. The senders send the messages independently and without seeing each other's messages. The duration for sending a message is 90 seconds.

C) The Receiver observes the messages from the Senders on the same screen. In this screen, Senders are defined as first sender and second sender. This naming is done separately and completely randomly in each round.

D) Earnings are calculated. A new round begins.

Message Rules:

(1) Messages must be in Turkish.

(2) You cannot share information about your actual identity in the messages (name, department, address, telephone, your location in the lab, physical appearance etc.).

(3) You cannot send messages that contain threats, insults, or offensive words.

Every message that fits these rules can be sent. Messages may or may not include statements about the number assigned to you or the number assigned to other sender.

Recipients should remember that incoming messages may contain incorrect information.

Earnings:

For each round,

The buyer earns as much money as the number that has been assigned to the Sender he/she chooses. So,

- If the first Sender is chosen, Receiver earns as much as the number assigned to the first Sender.

- If the second Sender is chosen, Receiver earns as much as the number assigned to the second Sender.

The Receiver chosen by the Receiver earns 50TL.

The Receiver not chosen by the Receiver earns 0TL.

Payments

The average of your earnings for the 5 rounds will be used for actual payouts. Therefore, it will be appropriate to make a careful decision every round.

B English translation of post–experiment questionnaire

Age of the subject (in years).

Sex of the subject (1=male, 0=female).

Living: living arrangement for the subject (0=student housing, 1=with family, 2= with friends, 3=alone).

Siblings: number of siblings of subject.

Older siblings: number of siblings who are older than the subject.

Trust: "Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?" (Be careful 0 ... 10 can be trusted)

Risk: "How willing are you to take risks in general?" (0 lowest – 10 highest)

Please tell me for each of the following statements whether you think it can always be justified, never be justified, or something in between (Never justified 0 ... 10 Always justified):

Civic1: claiming government benefits to which you are not entitled

Civic2: avoiding a fare on public transport

Civic3: cheating on taxes if you have the chance

Major: subject's major (2=economics, 1=business, political science or international trade, 0=other).

Econ: number of economics classes (censored at 4).

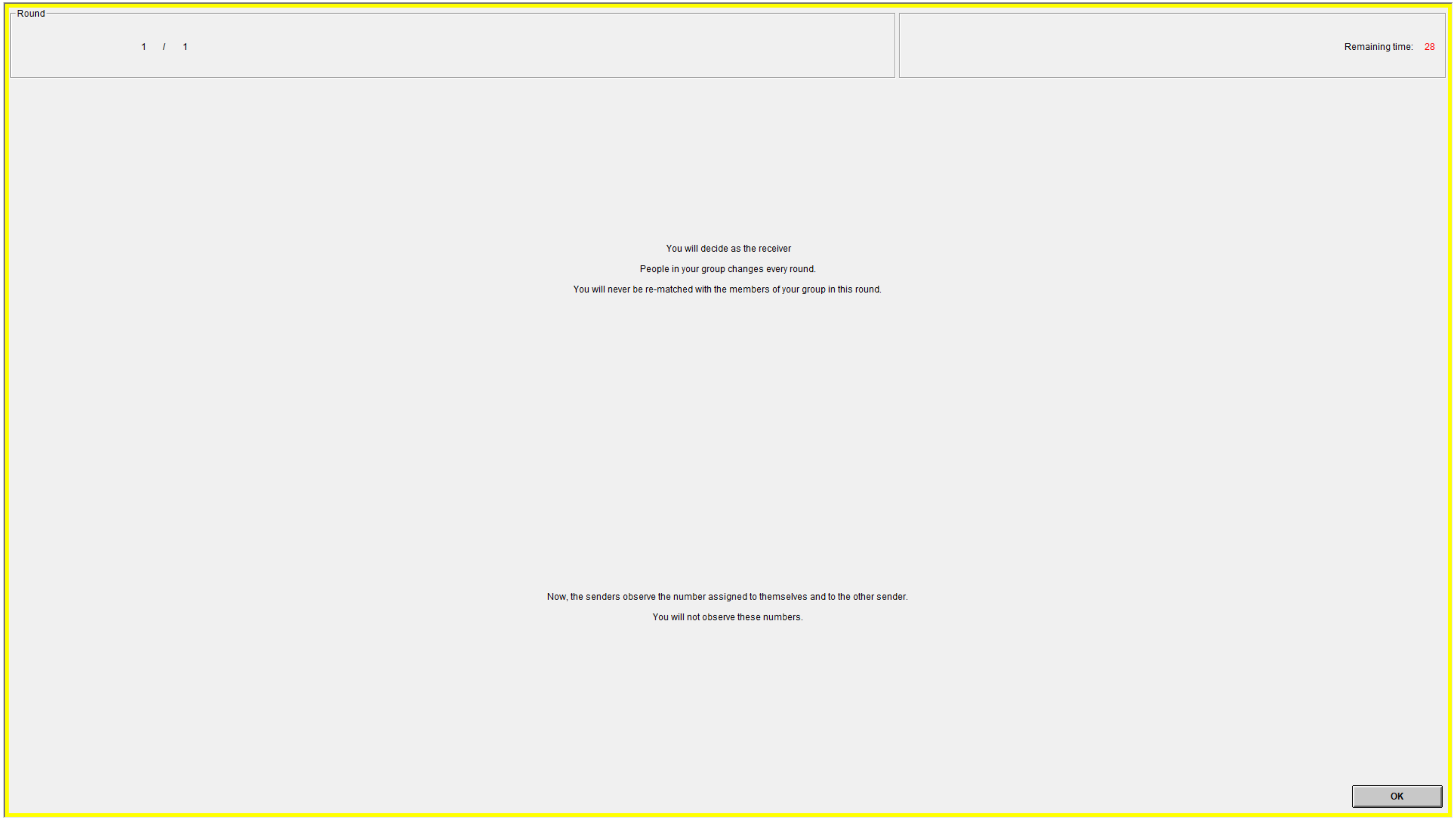
Friends: number of people known in the session

Member: Membership in organisations (student clubs, political parties, NGOs) (0:no, 1: yes)

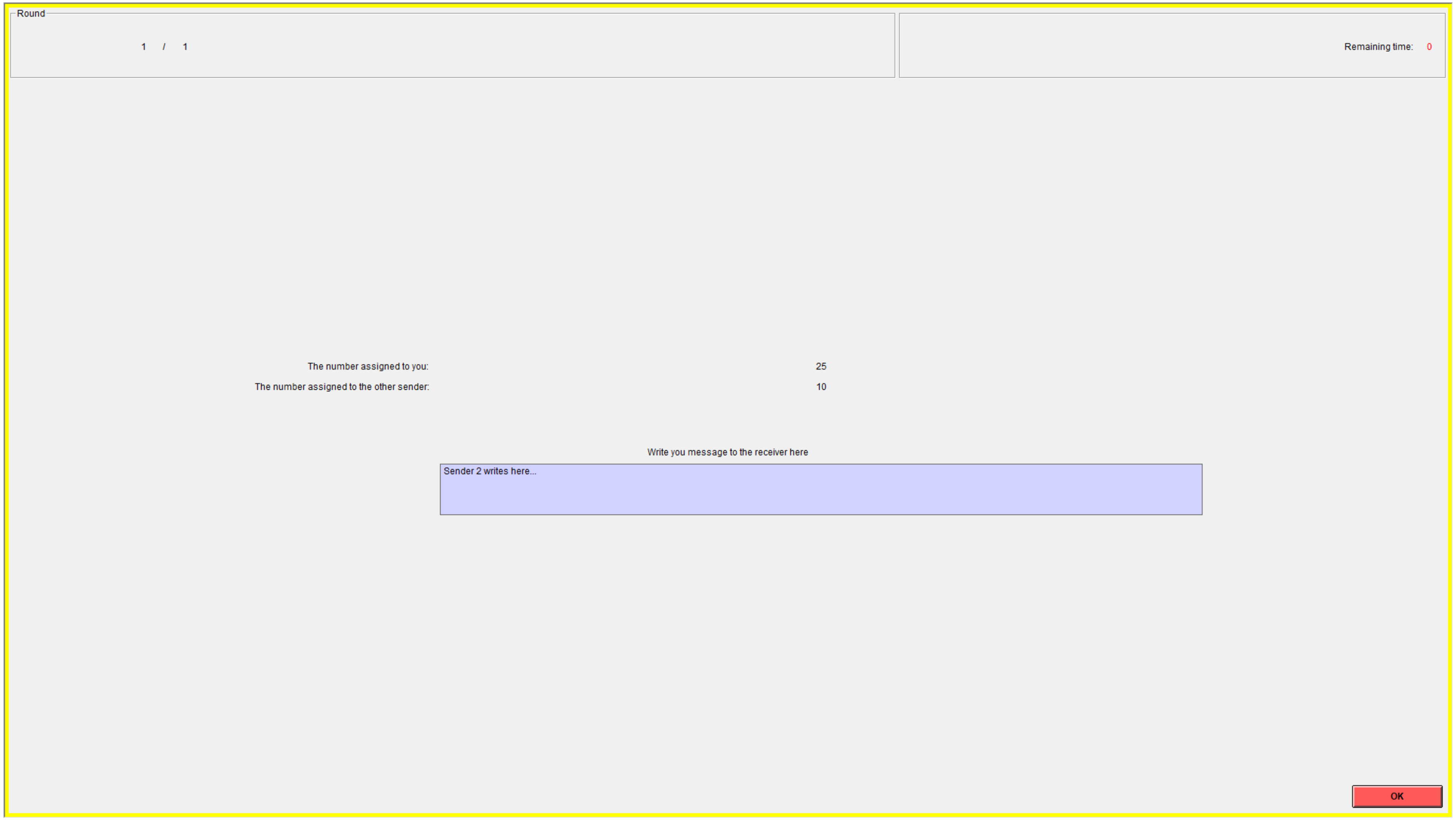
Rely: "How much can we trust the data coming from you in this experiment?" (0 lowest – 10 highest)

C Sample screenshots (translated from original Turkish)

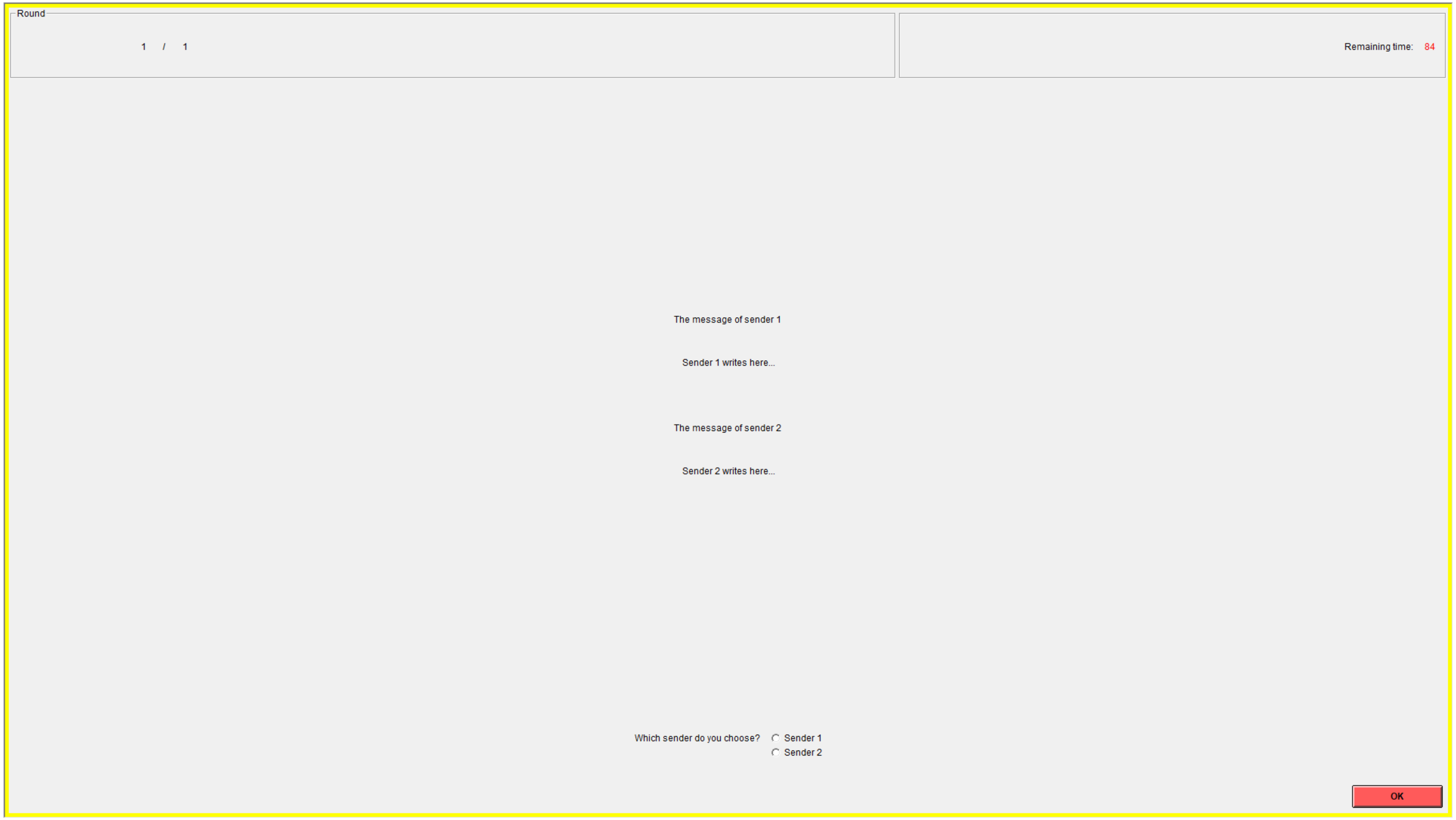
Receiver introductory screen:



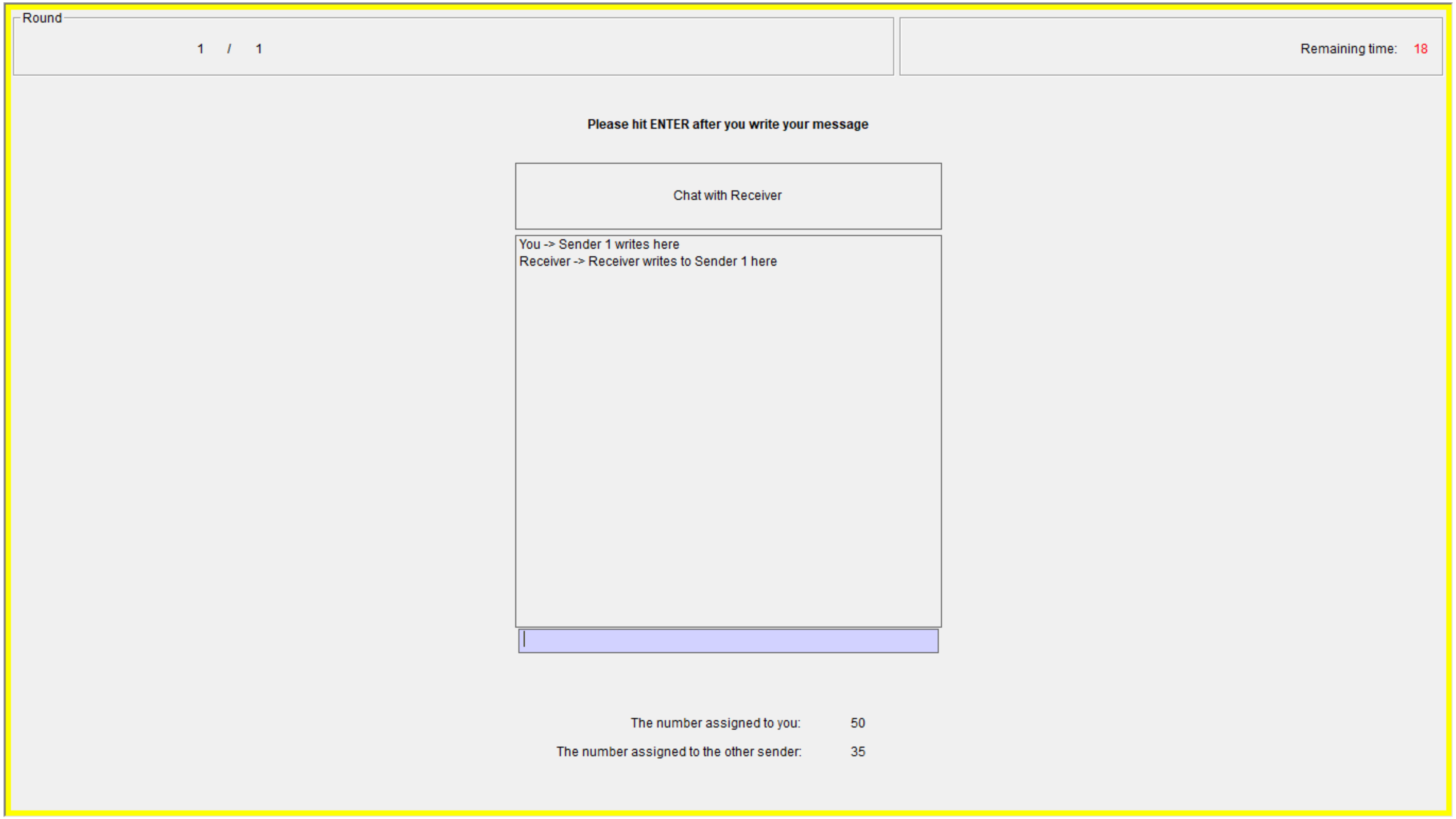
Sender message choice (Rich treatment):



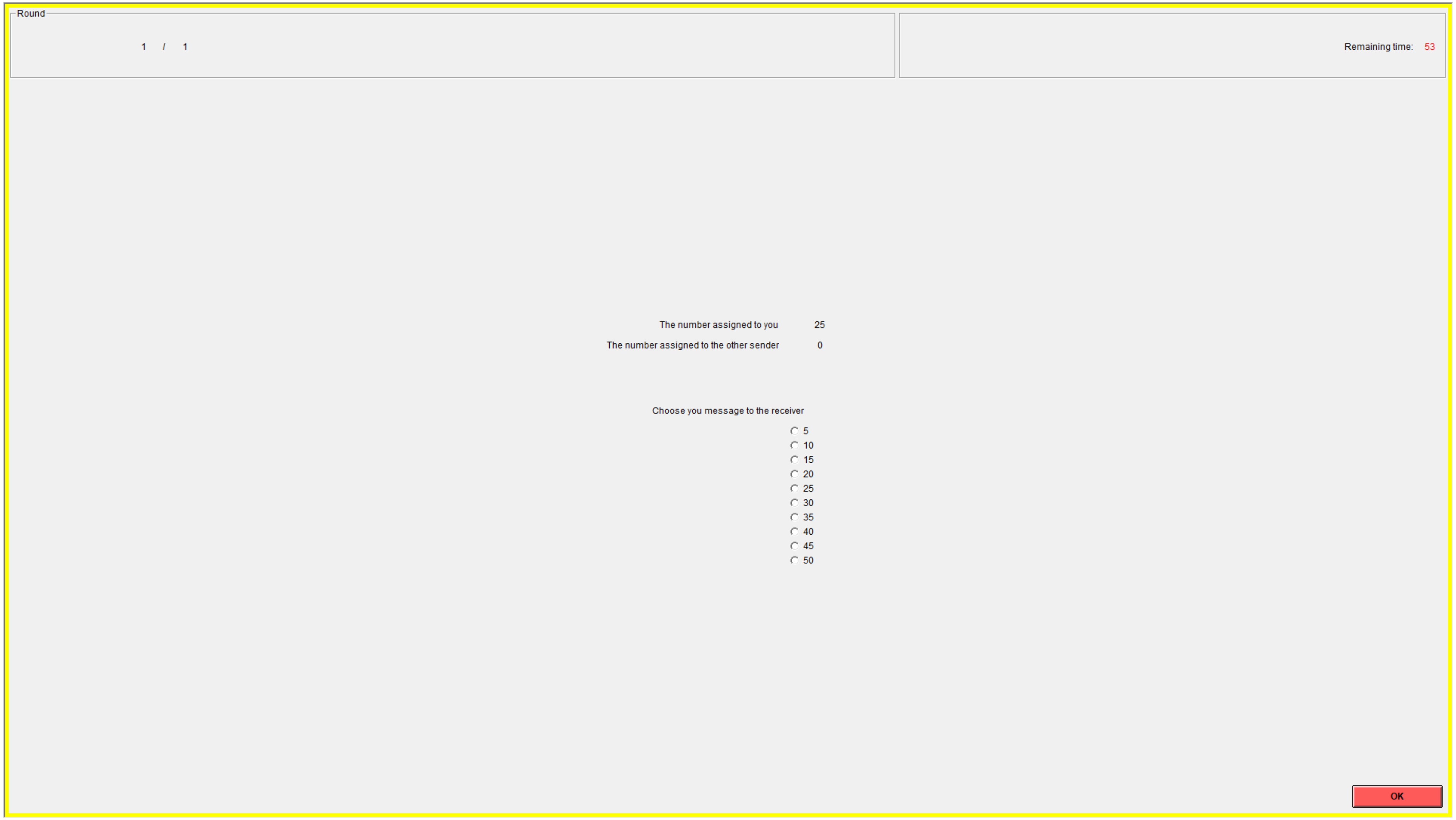
Receiver observes messages, chooses sender (Rich treatment):



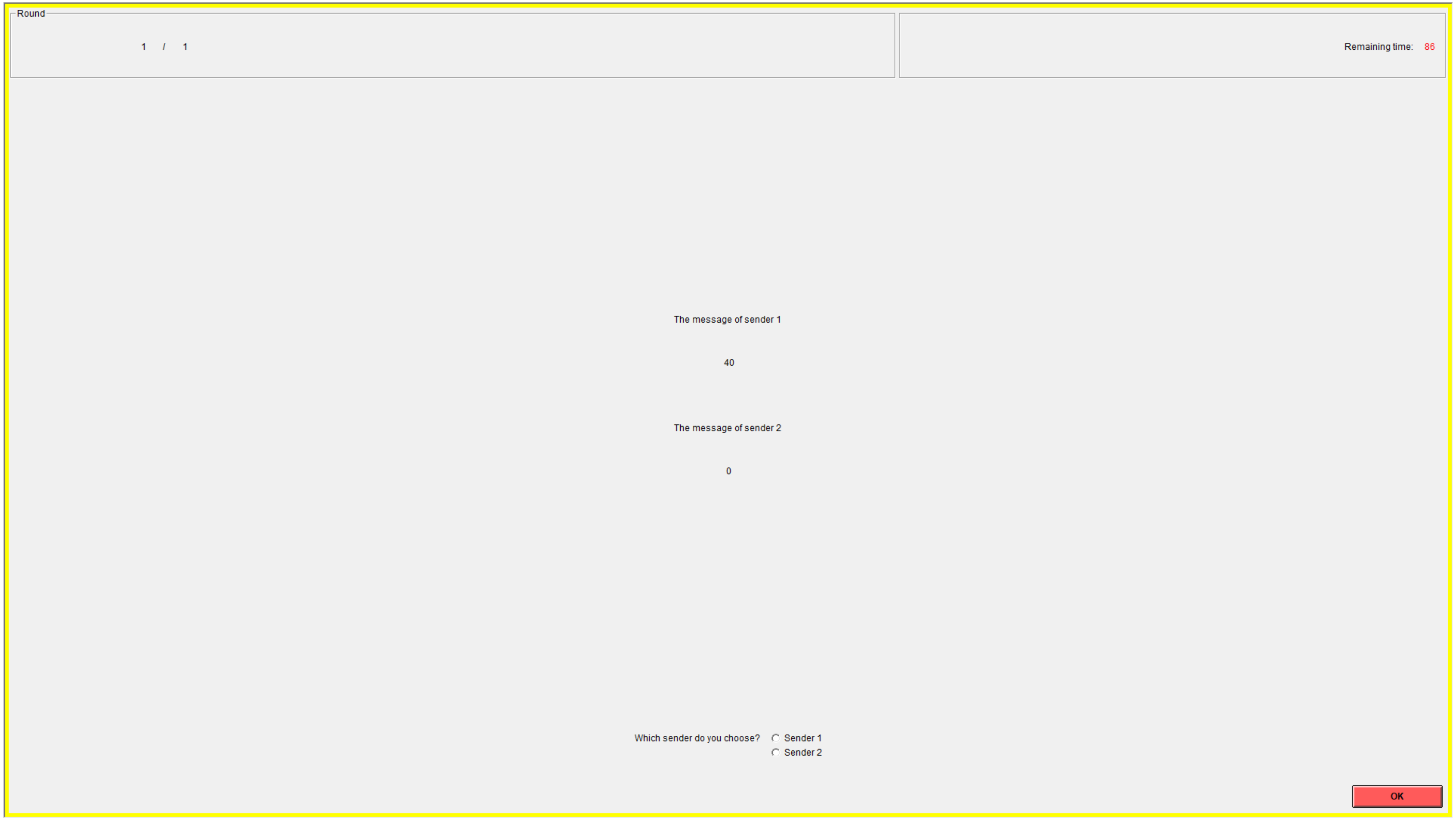
Sender chat screen (Chat treatment):



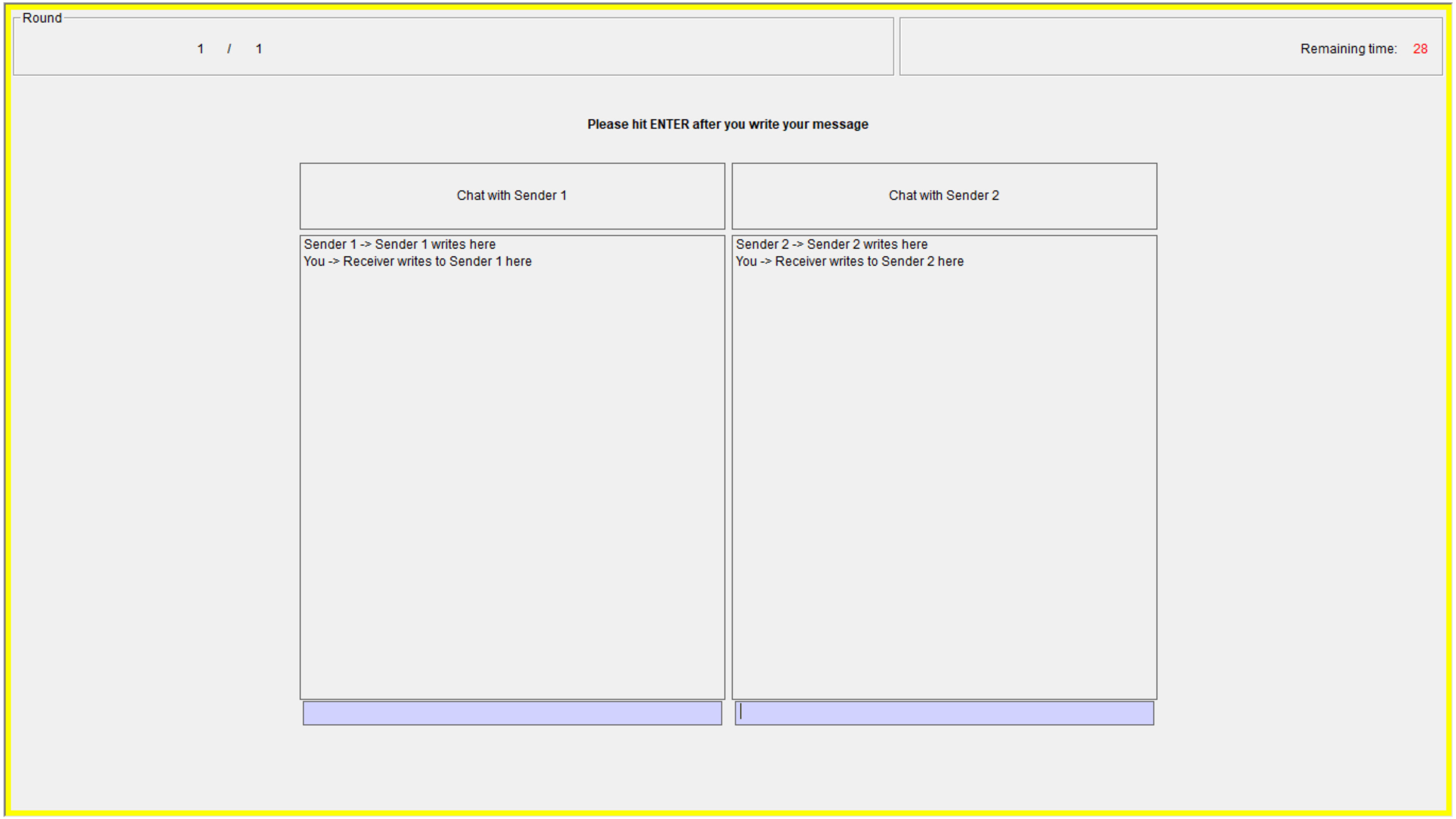
Sender message choice (Byte treatment):



Receiver observes messages (Byte treatment):



Receiver chat screen (Chat treatment):



Receiver observes chat, chooses sender (Chat treatment):