American Economic Journal: Microeconomics 2021, 13(2): 141–173
https://doi.org/10.1257/mic.20180217

Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure†

By Ginger Zhe Jin, Michael Luca, and Daniel Martin*

This paper uses laboratory experiments to directly test a central prediction of disclosure theory: that strategic forces can lead those who possess private information to voluntarily provide it. In a simple sender-receiver game, we find that senders disclose favorable information, but withhold unfavorable information. The degree to which senders withhold information is strongly related to their stated beliefs about receiver actions, and their stated beliefs are accurate on average. Receiver actions are also strongly related to their stated beliefs, but their actions and beliefs suggest that many are insufficiently skeptical about nondisclosed information in the absence of repeated feedback. (*JEL* C70, D82, D83)

From the number of calories in a croissant to the fuel efficiency of a car, businesses routinely have private information about the quality of their products that potential customers would like to know. Businesses then face a decision—should they reveal or withhold this information?

A central result in information economics is that market forces can lead firms to voluntarily and completely disclose such information, as long as the information is verifiable and the costs of disclosure are small (Viscusi 1978, Grossman and Hart 1980, Grossman 1981, Milgrom 1981). The mechanism behind this result is simple: consumers will treat all nondisclosing companies the same, so the best businesses among those that do not disclose have an incentive to separate themselves through disclosure. Applied iteratively, this logic produces unraveling in the quality of nonreporting firms, so that in equilibrium consumers correctly infer the very worst from nondisclosure. In other words, no news is bad news.

The policy relevance of the unraveling result is clear. Information can be important for markets to function properly, and this result suggests that voluntary disclosure

* Jin: Department of Economics, University of Maryland College Park, 3115 F Tydings Hall, College Park, MD 20742 (email: jin@econ.umd.edu); Luca: Harvard Business School, Soldiers Field Road, Boston, MA 02163 (email: mluca@hbs.edu); Martin: Kellogg School of Management, Northwestern University, 2211 Campus Drive, Evanston, IL 60208 (email: d-martin@kellogg.northwestern.edu). Michael Ostrovsky was coeditor for this article. Patrick Rooney provided excellent research assistance. All errors are ours. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Funding for this project came exclusively from our academic research budgets. Harvard University IRB approval was obtained for all experiments performed at the CLER laboratory at the Harvard Business School (IRB 14-1027), and New York University IRB approval was obtained for all experiments performed at the CESS laboratory at New York University (IRB 10-8117).

† Go to https://doi.org/10.1257/mic.20180217 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

could solve asymmetric information problems across a variety of domains. Moreover, this result points to policies that should increase the extent of disclosure. For example, some cities send hygiene scorecards that restaurants can voluntarily post on their doors, which is an attempt to make disclosure both low cost and verifiable.

Yet, the unraveling logic rests on strong assumptions around the ability of consumers to make inferences about a business's decision to withhold information. In practice, voluntary disclosure is far from complete (Mathios 2000, Luca and Smith 2015, Bederson et al. 2018). Many restaurants did not post their hygiene grades unless required to, many universities only publicized rankings in which they did well, and many grocery store food items did not include nutritional information until it was mandated. However, because multiple factors can lead to failures of voluntary disclosure, it is difficult to cleanly test the unraveling prediction and isolate the role of consumer inference about nondisclosure in the field.

The goal of our paper is to investigate the unraveling predictions using lab experiments that are complex enough to capture the main strategic tensions of the theory yet simple enough for subjects to easily understand these tensions. In our experiments, there are two players: an information sender (e.g., the firm) and an information receiver (e.g., the consumer). The sender receives private information that perfectly identifies the true state (e.g., the firm's true quality level). The sender then makes a single decision: whether or not to disclose this information to the receiver. As a result, the sender cannot misrepresent the state.[1] In many markets, such as those with truth-in-advertising laws, firms choose whether or not to reveal information, but can only disclose verifiable information. By prohibiting dishonest reporting, we mirror this feature of many disclosure settings, and also reproduce the assumptions underlying the unraveling prediction.

After the sender decides whether or not to disclose their private information, the receiver must guess the state. If the sender has revealed the state (a whole number between 1 and 5), the receiver knows it with certainty. Otherwise, the receiver must infer the true state based on the sender's decision to withhold information and on the distribution from which states are drawn, which is common knowledge. Reflecting many market transactions, the sender and receiver do not have aligned interests. The sender earns more when the receiver guesses that the state is higher, and the receiver earns more when their guess is closer to the true state.

With these payoffs, the logic of unraveling leads to a unique sequential equilibrium: senders should always reveal their information (unless the state takes the lowest possible value, in which case they are indifferent between revealing and not), and receivers should correctly guess that the state takes the lowest possible value when senders do not reveal their information.

Consistent with the theoretical predictions, we find high overall rates of disclosure—and almost full disclosure at the highest states. However, in contrast with the full unraveling prediction, some senders do not disclose intermediate states. Importantly, and for the first time in the literature, we elicit sender beliefs about receivers and find evidence that nondisclosure is driven by the belief that not

[1] This is in contrast with existing experiments on strategic information transmission where senders can engage in "cheap talk" (Cai and Wang 2006; Wang, Spezio, and Camerer 2010).

disclosing intermediate states is optimal given receiver actions. Moreover, we find that sender beliefs are correct on average, which means that the nondisclosure we observe is optimal.

These departures from equilibrium leave senders better off because on average receivers guess that nondisclosed states are higher than they actually are. What drives this "overguessing" by receivers? Even with the simplicity of our experimental design, one possibility is just that some receivers are confused by the actions of senders or other aspects of the game.[2] Alternatively, receivers may think senders are confused, which given boundary effects, could make the overguessing of low numbers a best response. We address both possibilities in our structural model of receiver guesses, and we find that role switching is likely to reduce confusion about the actions and payoffs of their opponents, although many of our results are robust to having senders and receivers switch roles.

Another possible reason for overguessing, which has been suggested often in literature, is that receivers are naïve about the strategic use of nondisclosure. To assess this possibility, we elicit receiver beliefs about how often senders fail to report for each secret number, and based on a regression analysis, we find that the extent of receiver overguessing is strongly related to these elicited beliefs about sender strategies. This suggests that receivers overguess nondisclosed states because they are insufficiently skeptical about undisclosed information—the extent to which no news is bad news.

However, because a receiver's guess might be impacted by other factors, there are limits to how clearly our regressions can establish a relationship between beliefs and actions. To enhance our understanding of this relationship, we use a partial-equilibrium structural model to predict receiver guesses based on their elicited beliefs. In this model, we assume that receivers use their beliefs about sender strategies to form beliefs about the likelihood of each state when the secret number is not reported and then to choose the action that maximizes their expected utility given these beliefs.[3] If we assume that receivers have beliefs in line with their elicited beliefs, this model predicts choices better than an alternative model that assumes all subjects have accurate beliefs about sender strategies.[4]

We also find that receiver mistakes decrease with immediate, direct, and repeated feedback, which eventually leads senders, in this case, toward disclosing all but the worst states, a direction more in line with theory.[5] However, we find that when there is no immediate and repeated feedback or when feedback is at the aggregate level, the rate of convergence is much slower. Moreover, if we also fix subjects in

[2] For example, receivers may not realize that senders cannot misrepresent the state. However, if this was the case, we might expect that receivers would underguess instead of overguess, given the possibility that senders inflate their reports.

[3] We allow their decisions to be impacted by boundary effects, social preferences, and confusion, and we estimate the probability of confusion and social preferences out-of-sample using receiver guesses when the state is actually disclosed.

[4] We did not incentivize subjects to provide accurate beliefs, which could introduce noise in elicited beliefs. Thus, the fit of our model could potentially be improved by incentivizing the process of belief elicitation.

[5] Our structural model does not capture this feature of receiver mistakes because it assumes that beliefs remain fixed over rounds. Thus, the fit of our model could be improved by allowing for dynamic trends.

the same role throughout the experiment, then we find that the decrease in receiver mistakes and sender nondisclosure over rounds is no longer statistically significant.

These results help to shed light on the economics of voluntary disclosure. In situations where immediate and direct feedback about nondisclosed information is limited, our findings suggest there is reason to be skeptical that full unraveling will occur. One reason for limited feedback could be infrequent transactions, as in the markets for house or car purchases. Even in markets with frequent transactions, feedback can be limited when consumers are inattentive or unable to process feedback about nondisclosed information. For example, when restaurants choose not to disclose their hygiene ratings or calorie counts, it can be difficult or time-consuming for consumers to assess this information immediately after having completed their meal. The impact of feedback is especially policy-relevant, as policymakers have discussed various informational interventions related to disclosure.

This paper provides three main contributions, which are discussed in more detail in the next section. First, we complement existing empirical studies from the lab and field by providing evidence of strategic naïveté in verifiable disclosure using a novel and clean laboratory experiment. Second, we elicit beliefs from both senders and receivers, which provides direct evidence on the underlying mechanisms behind nondisclosure. Third, we vary the feedback provided to subjects, and explore its role in producing convergence to equilibrium behavior.

The rest of the paper is organized as follows. Section I describes three related literatures and our contributions to each. Section II lays out the disclosure game and its equilibrium features. Section III describes our experimental design, and Section IV reports the results of our experiments. A brief discussion of the results is offered in Section V.

I. Related Literature

Our paper draws on and contributes to three literatures: the literature on voluntary disclosure, the literature on communication experiments, and the literature on beliefs and play in games.

A. *Voluntary Disclosure*

Voluntary disclosure is appealing from a policy perspective because it can improve consumer welfare even without mandatory disclosure policies, which are often opposed by industry groups and challenging to implement and enforce. The classic unraveling result suggests that the same benefits as mandatory disclosure can be achieved simply by ensuring that disclosed information is verifiable and the related costs are low. This has inspired a number of measures, including standardized information displays, certification agencies, and truth-in-advertising laws.

In practice, voluntary disclosure is observed in many industries, but is far from complete.[6] As summarized in Dranove and Jin (2010), this incompleteness has

[6] See Mathios (2000); Jin (2005); Fung, Graham, and Weil (2007); and Luca and Smith (2015) for specific examples.

motivated two strands of theories to account for why unraveling does not occur. One strand emphasizes external factors such as disclosure cost and consumer knowledge before disclosure, while the other strand focuses on a seller's strategic incentives. As an example of the latter, sellers may choose not to obtain data on product quality in order to avoid future demand for disclosure (Matthews and Postlewaite 1985).

Other examples of strategic incentives include product differentiation (Board 2009) and countersignaling (Feltovich, Harbaugh, and To 2002; Harbaugh and To 2020). The seller's strategic incentive can also be dynamic: one may refrain from disclosure even if he has favorable information at hand, as he fears that today's disclosure may make it harder to explain nondisclosure in the future when the information turns out nonfavorable (Grubb 2011). In another example of dynamic incentives, a pharmaceutical firm may prefer to be silent about the potential health risks of its products because of litigation risk, but this may crowd out positive disclosures (Marinovic and Varas 2014).

In this paper, we use lab experiments to exclude these additional reasons for limited disclosure, and therefore create a test environment closer to the original, classical theory. The strategic incentives for disclosure that we study are also present in the subsequent literature, so our results can potentially inform the wider literature as well. For instance, persistent naïveté about nondisclosure could be combined with any of the additional forces given above to produce new predictions for verifiable disclosure.

Our work also draws on the literature in behavioral economics, which has posited that if buyers are naïve about the quality of nondisclosed information, sellers may not disclose all of their private information. Eyster and Rabin (2005) consider this possibility in the context of their "cursed equilibrium" concept, and Gabaix and Laibson (2006) and Heidhues, Kőszegi, and Murooka (2017) consider it in the context of shrouded attributes. Mullainathan, Schwartzstein, and Shleifer (2008) present a model of coarse thinking that highlights how informational spillovers from one environment to another can make nondisclosure more persuasive than it would otherwise be. Likewise, in the accounting literature, Hirshleifer and Teoh (2003) consider the impact of naïveté on financial disclosures. In their model, receivers can be naïve about nondisclosed information, but they can also be inattentive to disclosed information.

Our findings are also consistent with growing field evidence on attention and inference in disclosure contexts. Brown, Camerer, and Lovallo (2012) find that firms with lower quality movies choose to engage in "cold openings" (i.e., they withhold movies from critics until the movie is released). Their data suggest that customers do not fully infer that movies with cold openings tend to be worse. Brown, Camerer, and Lovallo (2013) demonstrate how data on movie openings can be used to differentiate between equilibrium and non-equilibrium behavior (specifically related to the extent that naïveté limits unraveling in settings of verifiable disclosure).

We add to these literatures in several ways. Our primary contribution is to provide evidence of receiver naïveté in voluntary disclosure through a controlled laboratory experiment where beliefs are elicited. In addition, we show this naïveté is not easily eliminated, even if receivers are provided information about aggregate disclosure behavior and have played as senders for many rounds. Moreover, we show that

subjects drawn from the same population appear strategically sophisticated in the role of sender, but strategically naïve in the role of receiver.[7] This suggests that receiver choices may be more "strategically complex" than sender choices. For instance, receivers need to undertake hypothetical thinking, which has been shown to be difficult in voting games (Esponda and Vespa 2014).

B. *Communication Experiments*

Our design borrows many features from the cheap talk experiments of Cai and Wang (2006) and Wang, Spezio, and Camerer (2010). For instance, we follow both of these experiments in describing the sender's type using "secret" numbers and in starting messages to the receiver with "The number I received is." In addition, our type space and payoffs are similar to those found in Wang, Spezio, and Camerer (2010). The key difference in our experimental design is that the sender's messages must be truthful. Hence, our experiment tests models of verifiable disclosure, rather than cheap talk.[8]

In their investigation of cheap talk, Cai and Wang (2006) find that senders give more informative messages and receivers rely more on senders' messages compared to equilibrium. Instead, we find that senders give less informative messages and that receivers rely less on senders' messages compared to equilibrium. However, when there is a strong conflict of interest, the equilibrium predictions for sender behavior in cheap talk and voluntary disclosure games are opposite: no information transmission in cheap talk and full information transmission with voluntary disclosure.

While the equilibrium predictions are different, the departures from equilibrium are similar in size. In their maximum bias treatment, Cai and Wang (2006) find a correlation between states and actions of 0.21 instead of the theoretical prediction of 0. Across treatment, we find a correlation between states and actions of 0.77 instead of the theoretical prediction of 1, which is a difference of 0.23.

Even though our participants in the experiment are not able to lie, they are able to deceive by strategically withholding information. In this regard, we contribute to a growing literature around the difference between lying and deception (Sobel 2020). We find that very few senders disclose secret numbers of 1, but not reporting 1 is not deceptive if senders think that receivers are certain the nondisclosed number is 1. However, we find evidence using elicited beliefs that most senders believe receivers will guess above 1 for nondisclosed secret numbers, which suggests that many senders are being "deceptive" when they do not disclose a secret number of 1 (based on the definition of deception given in Sobel 2020). One central question in this literature is whether there is a difference between the decision to lie and the decision to deceive. One way we can get a sense of this is to look at the fraction of people who withhold the worst state and the fraction of people who lie in the worst

[7] In fact, asymmetry at the aggregate level suggests asymmetry at the individual level because subjects play as both senders and receivers in some treatments. A similar asymmetry appears in the cheap talk treatment of Forsythe, Lundholm, and Rietz (1999).

[8] Montero and Sheth (2019) extend our experimental design to allow for consultation among receivers and restaurant hygiene framing, and Sheth (2019) extends it to allow for competition among senders. Our findings on receiver naïveté are largely robust to these design changes.

state. While we do not have a treatment with lying, we can compare with results from similar experiments in the literature. Specifically, the fraction of people who withhold 1 in our experiment ranges across treatments between 88.8 percent and 89.3 percent, and the fraction of people who give a message above 3 when the state is 1 in the maximum bias treatment of Cai and Wang (2006) is 70.4 percent.[9]

There are only a limited number of experiments that explicitly study verifiable disclosure. Forsythe, Isaac, and Palfrey (1989) study disclosure in asset markets, as in Milgrom and Roberts (1986). Their experiment features a sender (the asset seller) who decides whether to disclose the asset's quality to receivers (the potential asset buyers) who compete with each other through an auction mechanism. They find that behavior converges to the predictions of unraveling after repeated experience: most sellers move to disclosing the asset's value and most buyers appear to have adjusted their beliefs appropriately (although one market did not unravel, which appears to be due to a single buyer who did not have skeptical beliefs). Subjects in their experiment receive feedback after each round, so our finding that behavior converges to the predictions of unraveling with repeated feedback replicates their result in a simpler setting.

King and Wallin (1991) and Dickhaut et al. (2003) show what happens in asset markets when there is a possibility that sellers may not be informed about an asset's quality (and thus have nothing to report). By looking across three probabilities of being informed (70 percent, 90 percent, and 100 percent), King and Wallin (1991) provide evidence that more information is disclosed (the maximum nondisclosed value evidence decreases) when the sellers are more likely to be informed and that buyers react to recent seller disclosures (though not necessarily in an optimal way). Dickhaut et al. (2003) also allow sellers to be partially informed (and partially disclose), and they find that most sellers do not fully disclose (they follow a sanitation strategy instead) and that buyers appear to be more naïve when the seller is less likely to be informed.

Forsythe, Lundholm, and Rietz (1999) compare voluntary disclosure and cheap talk in reducing adverse selection when sellers do not have to sell the asset. They find that although voluntary disclosure reduces adverse selection and there is general consistency with equilibrium predictions, sellers do not always disclose the assets value and buyers do not always appear to be fully skeptical of nondisclosures.

Concurrent to but separate from our study are three new papers that use experiments to study verifiable disclosure. Benndorf, Kübler, and Normann (2015) study a disclosure game in a labor market setting where multiple senders compete through the use of disclosure. Unlike our experiments, the receiver in their experiment is a computer that uses an automated strategy, so there is no room for receiver naïveté to impact disclosure. Hagenbach and Perez-Richet (2018) investigate a verifiable disclosure game where sender payoffs are not necessarily monotonic in the state space. They find that receivers correctly account for disclosure in their decisions, and that senders who do not have an incentive to masquerade as another type ("satisfied"

[9] See Table 6 of Cai and Wang (2006). The fraction of people giving a message that does not include 1 when the state is 1 (the lying rate for a state of 1) is not provided.

senders) disclose fully, while senders who have an incentive to masquerade as another type ("envious" senders) do not disclose fully. In addition, they find that 96 percent of receiver decisions and 85 percent of sender decisions are consistent with iterated elimination of obviously dominated strategies based on obvious dominance. Li and Schipper (2018) implement a disclosure game and find high levels of reasoning based on iterated admissibility. They show theoretically that the level of mutual cautious belief in rationality required for full disclosure is higher when there are more states, and they test this prediction experimentally by varying the number of states both between and within-subject. Consistent with this prediction, they find more unraveling when there are fewer states.

In two experiments that study lying aversion, senders have three options: tell the truth, lie, or not disclose. Nondisclosure takes the form of vague messages in the case of Serra-Garcia, van Damme, and Potters (2011) and silence in the case of Sánchez-Pagés and Vorsatz (2009), so the latter is closer to our experiment. However, unlike in our experiments, in Sánchez-Pagés and Vorsatz (2009) nondisclosure carries a cost. Even with this cost, some senders choose not to disclose. Serra-Garcia, van Damme, and Potters (2011) find that intermediate senders sometimes use vague messages, which receivers do not make correct inferences about. Agranov and Schotter (2012) study the use of both vague and ambiguous messages, and they find that an announcer in coordination games might want to use such messages.

Relative to this literature, we believe that our experiment is the first to elicit beliefs about receiver guesses and sender strategies, to vary the feedback provided to subjects, to contrast fixed roles with role switching, or to provide information about aggregate sender behavior. All four design elements, separately and in combination, are used to generate new insights on voluntary disclosure.

C. *Beliefs and Play in Games*

Central to any strategic interaction is the set of beliefs that people hold about each other. This has given rise to work in the experimental economics literature based on the following question: Do people hold correct beliefs about how other people play, and do they best respond to these beliefs?

While economists typically infer beliefs from actions, stated beliefs can provide further evidence on both the belief formation process and the ways in which people react to their own stated subjective beliefs. For example, Costa-Gomes and Weizsacker (2008) find that subjects often do not best respond to stated beliefs about sender strategies, while Rey Biel (2009) finds much higher rates of best responding to these beliefs in simpler games.

We find a strong, positive, and statistically significant correlation between elicited beliefs about receiver guesses and the probability of the sender reporting the underlying state. Likewise, we find a strong, positive, and statistically significant correlation between implied beliefs about the underlying state (based on elicited beliefs about sender strategies) and receiver guesses of the underlying state, which suggests that their actions incorporate beliefs about sender strategies. Furthermore, we demonstrate asymmetry between the correctness of sender beliefs and the correctness of receiver beliefs. While sender beliefs are correct on average, average receivers are

insufficiently skeptical about undisclosed information. This asymmetry rationalizes why senders withhold nonfavorable information more than theory predicts.

II. The Disclosure Game

The one-shot disclosure game we study involves two agents: an information sender and an information receiver. At the beginning of the game, nature determines the state b (which can be interpreted as the sender's type) by taking a draw from a probability distribution F with full support over a finite state space B, which is a subset of the real numbers. The sender knows the realized state, but ex ante, the receiver knows only the distribution of possible states.

The sender has two possible actions, and the receiver is aware that these are the only two actions available to the sender. The sender can either report the state to the receiver or make no report. This report must be truthful and cannot be vague. Thus, the set of actions M available to a sender of type b is just $M(b) = \{b, null\}$.[10]

Regardless of whether or not they receive a report from the sender, the receiver takes an action a from a finite space A, which is also a subset of the real numbers and contains B. We interpret this action as guessing the type of the sender.

The true state and the receiver's action determine the payoffs for the two parties. The sender's utility is given by a function $U_s(a)$, which is strictly increasing in the receiver's action, and independent of the state. The receiver's utility is given by a function $U_R(a,b)$, which is concave in the receiver's action a and reaches its maximum when a is equal to b. In other words, the receiver benefits more from selecting an action that is closer to the true state, while the sender benefits the most when the receiver's action is as high as possible. These utility functions produce a strong conflict of interest when the state is low.

When the set of receiver actions A is sufficiently rich, the techniques found in Milgrom (1981) can easily be adapted to show that in every sequential equilibrium of this disclosure game, the sender always reports the state (unless it is the minimum element in B), and if there is no report, the receiver takes the action that is the minimum element in B. In other words, the sender always reports their type (unless it is the worst possible type), and the receiver always guesses the sender is the worst possible type if they do not report. When the realized state is the minimum element in B, the sender is indifferent between reporting or not, so any mixture over these actions is consistent with equilibrium.

III. Experimental Design

In our experiments, subjects completed 45 rounds and then, depending on the session, specific additional tasks. Subjects were told at the beginning of the experiment that they would complete additional tasks but were given no details about the tasks they would face later. The online Appendix contains the full set of instructions given

[10] In the model of Milgrom (1981), senders are allowed to report a range of states, but we consider a simpler message space in order to reduce the strategic complexity of the game, which could add cofounding factors.

before the start of the experiment. Instructions for an additional task were presented to subjects on the computer screen just before the start of a task.[11]

At the end of each session, subjects were paid, privately and in cash, their show-up fee plus any additional earnings from the experiment. Over the course of the experiment, subjects had the opportunity to accumulate or lose "Experimental Currency Units" (or ECUs). At the end of the experiment, each subject's ECU balance was converted into US dollars at a treatment and role-specific rate, and their final payment was rounded up to the nearest nonnegative whole dollar amount.

A. *Main Sessions*

Our main sessions were conducted at the Computer Lab for Experimental Research (CLER) facility at the Harvard Business School (HBS). In this laboratory, subjects are separated with dividers, and each subject was provided with a personal computer terminal.

In Each Round.—In each round, subjects were randomly matched into pairs. To reduce reputational effects, subjects were matched anonymously and were told that it was very unlikely they would be paired with the same subject in consecutive rounds. For a session size of 14, the actual likelihood of being paired with the same subject in consecutive rounds is 7.7 percent.

In each round and for each pairing, one subject was the sender and the other subject was the receiver. To reduce framing effects, the sender was referred to as the "*S* Player," and the receiver was referred to as the "*R* Player."

For each pair, the computer drew a whole number from 1 to 5, called the "secret" number. Thus, the state space was $B = \{1, 2, 3, 4, 5\}$. Each of these numbers was equally likely to be drawn, and both senders and receivers were made aware of this probability distribution over the state space.

Each sender was shown the secret number for their pairing and then made their decision while the receivers waited. Senders were given the option to either "report" or "skip," with no time limit on their decision.

After all senders made their decisions, the receivers' screens became active. If a sender decided to report their secret number, the receiver they were paired with was shown this message: "The number I received is," followed by the actual secret number. If a sender decided instead to skip any reporting, the area for messages on the receiver's screen was left blank. Subjects were told that these were the only two actions available to senders, so that if the area for messages on the receiver's screen was left blank, the instructions were clear that it was because the sender chose not to report the secret number.

Below the area for messages, receivers were asked to guess the secret number, and these guesses could be any half unit between 1 and 5. Thus, the set of actions

[11] Our experiment was programmed and run using the z-Tree software package (Fischbacher 2007).

was $A = \{1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5\}$.[12] There was also no time limit for receiver decisions.

Receiver payoffs in each round were $ECU_R = 110 - 20\left|b - a\right|^{1.4}$, where b is the secret number and a is the receiver's guess.[13] These payoffs decrease strictly as the guess moves further from the secret number in either direction. The sender payoffs in each round were $ECU_S = 110 - 20|5 - a|^{1.4}$.[14] These payoffs are independent of the secret number and increase strictly with receiver guesses (because guesses cannot be higher than 5). Because we use just a small number of states and actions, the payoffs could be shown in a table, so that subjects did not need to know or interpret these functional forms.

These payoff functions satisfy the assumptions made in the disclosure game outlined previously. They also satisfy Milgrom's assumptions that receiver payoffs are concave, singled-peaked, and weakly increasing in the state and that sender payoffs are strictly monotonic in the action for a given state. While Crawford and Sobel (1982) make different assumptions,[15] the quadratic payoff example in their paper satisfies Milgrom's assumptions when the bias parameter is large enough. One meaningful difference between our payoffs and the quadratic payoffs example in Crawford and Sobel (1982) is that the sender's payoff function is state-independent in our experiment, which we chose to simplify the payoffs in our experiment.

With these payoff functions, there was a clear misalignment of interests between senders and receivers. Receiver payoffs were higher when their guesses were closer to the secret number, and sender payoffs were higher when the receiver made higher guesses. Subjects were told in the instructions about these two features of sender and receiver payoffs.

Treatment Variation.—Our primary treatment variations occurred along two dimensions. First, we varied the information provided as feedback to subjects after each round (no feedback versus feedback), and second, we varied the way that roles were assigned (fixed role versus random role). Both sources of variation were used to study the channels through which subjects learn in this setting. We ran three combinations: "no feedback and fixed role," "no feedback and random role," and "feedback and random role."

In our "no feedback" treatments, subjects were given no information after completing each round. After all receivers had made their decisions, subjects proceeded to a screen that required them to click "OK" to start the next round. After all

[12] The action space of receivers was made sufficiently rich that the unique sequential equilibrium involves full unraveling.

[13] We allowed subjects to accrue ECU in all rounds because payoffs could vary substantially between roles and realizations of the state, and we wanted performance to play a larger role than luck in final payments. Cai and Wang (2006) use similar payoff functions and also paid subjects every round. However, this approach introduces the possibility of wealth and portfolio effects. To ameliorate such effects, subjects were not told the cumulative payoffs they had earned so far in the experiment.

[14] With these payoff functions it is possible to have negative payoffs in a round, which could generate distortions related to avoiding actions that generate negative payoffs. A small change to the payoff function would eliminate negative payoffs and allow for testing of robustness along this dimension.

[15] See Seidmann and Winter (1997) for a comparison of the payoffs assumptions made in Milgrom's voluntary disclosure paper and Crawford and Sobel's cheap talk paper, and the application of Crawford and Sobel's payoff assumptions to Milgrom's voluntary disclosure game.

subjects had pressed this button, the next round began. In contrast, in our "feedback" treatment, subjects were told four pieces of information after each round: the actual secret number, whether the sender reported the secret number, the receiver's guess of the secret number, and their own payoff. After all subjects pressed the "OK" button on the screen containing this feedback, the next round began.

In our "fixed role" treatment, subjects were randomly assigned a role at the beginning of the session, and they stayed in that role throughout the entire experiment. Instead, in our "random role" treatments, subjects were randomly assigned roles before each round, so that roles might change after each round. In both cases, a subject was equally likely to be assigned either role. As a result, the likelihood of a subject experiencing both roles by round 5 in the "random role" treatments is 93.75 percent. In the random role treatment, ECU were converted into US dollars at a rate of 200 to 1, but to equalize expected payments across subjects in the fixed role treatment, ECU were converted at a rate of 150 to 1 for senders and 250 to 1 for receivers.

To reduce social considerations, subjects in the feedback treatment were not told the payoff for the other player in their pairing, though it could be deduced using the payoff table. In addition, between rounds subjects only received feedback about their pairing, not all pairings in the session.

Belief Elicitation and Additional Tasks.—After completing 45 rounds, subjects were asked to guess both the rate at which senders reported each secret number and the average receiver guess of nonreported secret numbers during the preceding 45 rounds. The purpose of these questions was to assess whether subject beliefs about sender strategies influenced their decisions as receivers and whether subject beliefs about receiver guesses influenced their decisions as senders. These guesses were not incentivized, which could potentially add measurement error. In a recent paper, Trautmann and van de Kuilen (2015) argue that incentivizing beliefs might reduce "noise" in elicitation or the possibility of demand effects to impact elicitation.[16]

We waited until all 45 rounds were complete to elicit beliefs in order to avoid distorting choices during those rounds. While we asked subjects to assess behavior over all 45 rounds, because beliefs may change over the course of the experiment, especially in the face of feedback, stated beliefs may reflect beliefs of opponent actions in just the final few rounds.[17]

After beliefs were elicited, subjects completed an additional task. In our main sessions, subjects either completed a "high incentives" task or an "aggregate feedback" task. The "high incentives" task was designed to better understand the impact of incentives on mistakes. This task mirrored an earlier choice, but with much higher incentives. In this task, subjects played once more in the role of sender or once more in the role of receiver, but in both cases they played against a computer instead of a human, and the computer played a strategy that was designed to mimic the

[16] They study belief elicitation in an ultimatum game and find, given three possible utility functions, that just 30–40 percent of choices can be explained by elicited beliefs when they are incentivized, and that the rate drops to 20 percent when elicited beliefs are not incentivized.

[17] We will examine this possibility during our analysis of the experimental data. For example, we compare stated beliefs also to opponent actions in the last block of rounds (see Figure 1).

unobserved decision of a previous opponent. Subjects were not reminded of the choices they had made previously.

When in the role of receiver, subjects were told that the computer sender would not report the secret number and that the secret number would be from a random past round in which the secret number was not reported. When in the role of sender, subjects were told that if they reported the secret number, the computer receiver would guess that number, and if they did not report, the computer receiver would repeat the guess of a receiver from a random past round where the secret number was not reported. To get as much information as possible from the sender decisions, we used the "strategy" method in which senders made a decision for each possible secret number before seeing the actual secret number. The payoffs from this task were added to the ECU earned in the first 45 rounds.

Importantly, for this decision the payoffs of the subject were ten times the rate in the initial 45 rounds. This design allows us to hold fixed the strategy of the opponent, so that the impact of incentives can be isolated to just one side of the pairing. Niederle and Vesterlund (2007) use a similar approach to hold fixed opponent strategies.[18]

In addition, to better understand the interaction between beliefs and information, we used a second additional task in which subjects were shown information about the play of all subjects in the first 45 rounds, guessed again a nondisclosed secret number from a previous round, and then played 5 more rounds just as in the first 45 rounds. The payoffs from this task were added to the ECU earned in the first 45 rounds.

We call this task the "aggregate feedback" task because subjects were shown the number of times that each secret number was reported and not reported for all subjects during the first 45 rounds of their session. This provided enough information to determine both the average reported secret number and the average nonreported secret number.

B. *10-State Robustness Sessions*

To get a sense for how the size of the state space might impact our findings, we ran additional robustness sessions with 10 secret numbers, where $B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, which is twice as large as the state space in the main sessions. Here again we allow receivers to guess half-unit intervals, so the action space is $A = \{1, 1.5, 2, 2.5, \ldots, 9, 9.5, 10\}$.

To keep payoffs in a comparable range to the main sessions, the distance from the ideal action is divided in half in the payoff functions, so that receiver payoffs are $ECU_R = 110 - 20\left|(b-a)/2\right|^{1.4}$ and sender payoffs are $ECU_S = 110 - 20\left|(10-a)/2\right|^{1.4}$. As a result, in these robustness sessions the payoffs for senders and receivers when the receiver guesses a and the secret number

[18] There are two potential confounds for this task. First, in this choice, there are no payoff consequences for their opponent, so social preferences related to the opponent's payoffs are no longer in play. Second, the random round could be drawn from any part of the experiment, so if there is a large time trend in behavior, the subject may choose differently because of additional uncertainty over actions.

is b is the same as when the receiver guesses $a/2$ and the secret number is $b/2$ in the main sessions.

Aside from increasing the set of secret numbers and changing the payoff table, the experimental design and instructions are the same as in the main sessions. In order to ensure sufficient statistical power, we did not vary treatments in the robustness sessions—all subjects completed the "no feedback and random role" treatment. We also conducted these robustness sessions in the CLER facility at HBS.

IV. Results

This section examines sender and receiver behavior in both our main sessions and robustness sessions. To provide a complete picture of behavior, we look both at choices pooled across rounds and how choices evolve from round to round.

A. *Subjects in the Main Sessions*

In our main sessions, we observed 324 subjects complete a total of 14,580 decisions, which corresponds to 7,290 pairings. Over 23 sessions, the mean session size was approximately 14. We used a show-up fee of $5, and on average subjects earned $26.60. The minimum payment was $14, and the maximum payment was $37.

We assigned 114 subjects to "no feedback and fixed role" sessions, 120 subjects to "no feedback and random role" sessions, and 90 subjects to "feedback and random role" sessions. All subjects in the no feedback and fixed role sessions completed the "high incentives" additional task, and all subjects in the no feedback and random role and feedback and random role sessions completed the "aggregate feedback" additional task.

In terms of demographics, we had roughly an equal number of men and women, and a large majority of subjects were undergraduates and native English speakers. Around 15 percent of subjects reported having a friend present in the room during the session. In the regressions presented in this paper, we either control for these demographic factors or use subject fixed effects.

B. *Sender Disclosures and Receiver Guesses*

When pooling across rounds, our primary qualitative findings are that senders disclose favorable states more often than less favorable states and receivers tend to overguess nondisclosed states, and these features are summarized by Table 1.

Looking first at senders, the average reporting rate is above 80 percent when the draw is equal to the average state (a secret number of 3) and above 90 percent when the draw is 4 or 5. For lower draws, however, the average reporting rate drops to 44.5 percent for draws of 2 and 11.0 percent for draws of 1. The theoretical predictions are a reporting rate of 100 percent for draws of 2 and anywhere between 0 percent and 100 percent for draws of 1.

Between treatments, the sender reporting rate differs the most for draws of 2. When the draw is 2, the reporting rate is not significantly different (for a two-sided test of proportions) between the no feedback and fixed role treatment and the no

Table 1—Summary of Player Actions in Main Sessions

	No feedback fixed role		No feedback random role		Feedback random role	
Variables	*N*	Mean	*N*	Mean	*N*	Mean
Report (secret number = 1)	490	0.110	568	0.107	421	0.112
Report (secret number = 2)	529	0.420	552	0.426	406	0.502
Report (secret number = 3)	533	0.818	507	0.779	400	0.848
Report (secret number = 4)	508	0.945	540	0.926	383	0.956
Report (secret number = 5)	505	0.968	533	0.929	415	0.949
Secret number (no report)	884	1.734	1,014	1.802	675	1.680
Guess (report = 1)	54	1.250	61	1.533	47	1.298
Guess (report = 2)	222	2.203	235	2.243	204	2.157
Guess (report = 3)	436	3.002	395	3.061	339	3.071
Guess (report = 4)	480	3.897	500	4.009	366	4.016
Guess (report = 5)	489	4.707	495	4.825	394	4.968
Guess (report = blank)	884	2.243	1,014	2.283	675	1.897
Guess − secret number (no report)	884	0.508	1,014	0.481	675	0.217

feedback and random role treatment (p-value = 0.840), but is significantly different between either of those treatments and the feedback treatment (p-value = 0.0117 for the fixed role treatment and p-value = 0.0185 for random role treatment).[19] This difference is reflected in the average secret number when the secret number is not reported. When there is no disclosure, the average secret number is smaller in the feedback treatment than in the no feedback treatments, though it is only statistically significant (for a two-sided t-test) between the no feedback and random role and feedback and random role treatments (p-value = 0.0148).[20]

Table 1 also presents average receiver guesses by treatment, conditional on whether the sender reports 1, 2, 3, 4, 5, or nothing. Because senders are not allowed to misreport, one may expect receivers to guess exactly the reported number if the sender discloses it. This expectation is largely confirmed when the reported number is 3 or 4, but with some deviation when the reported number is close to either extreme. In particular, receivers tend to overguess at the low extreme (1, 2) and underguess at the high extreme (5). Overguessing and underguessing of disclosed secret numbers at the extremes is analyzed in detail using a structural estimation in a subsequent section.

When senders choose not to disclose, receivers guess 2.243 on average in the no feedback and fixed role treatment and 2.283 on average in the no feedback and random role treatment, and this is not significantly different (p-value = 0.3472 for a two-sided t-test). In the feedback and random role treatment, the average guess is lower (1.897), and this is significantly different from the other treatments at the 1 percent level for a two-sided t-test.[21] There is a similar pattern for the average amount of overguessing (how far the guess is above the actual secret number). The

[19] The p-values for a logit regression of reporting rate onto treatment dummies with session-level clustering for the respective comparisons are 0.900, 0.087, and 0.209, given cluster sizes of 14, 18, and 14.

[20] The p-value for a linear regression of nonreported secret number onto treatment dummies with session-level clustering for this comparison is 0.295.

[21] The p-values for a linear regression of guess of nonreported secret number onto treatment dummies with session-level clustering are also significantly different for these comparisons at the 1 percent level.

Table 2—Summary of Player Actions in No Feedback and Random Role Treatment for Main Sessions (5 Secret Numbers) and Robustness Sessions (10 Secret Numbers)

	Main sessions no feedback random role		Robustness sessions no feedback random role	
Variables	*N*	Mean	*N*	Mean
Report (1-baseline or 1-robustness)	568	0.107	216	0.130
Report (2-robustness)			189	0.222
Report (2-baseline or 3-robustness)	552	0.426	199	0.412
Report (4-robustness)			185	0.714
Report (3-baseline or 5-robustness)	507	0.779	201	0.836
Report (6-robustness)			187	0.898
Report (4-baseline or 7-robustness)	540	0.926	193	0.959
Report (8-robustness)			173	0.960
Report (5-baseline or 9-robustness)	533	0.929	170	0.965
Report (10-robustness)			177	0.983
Secret number (no report)	1,014	1.802	581	2.616
Guess (report = 1)	61	1.533	28	1.429
Guess (report = 2)	235	2.243	42	2.190
Guess (report = 3)	395	3.061	82	3.244
Guess (report = 4)	500	4.009	132	4.201
Guess (report = 5)	495	4.825	168	5.244
Guess (report = 6)			168	6.202
Guess (report = 7)			185	7.122
Guess (report = 8)			166	8.018
Guess (report = 9)			164	8.963
Guess (report = 10)			174	9.920
Guess (no report)	1,014	2.283	581	3.419
Guess − secret number (no report)	1,014	0.481	581	0.803

amount of overguessing is similar and not significantly different between the two treatments without feedback (p-value = 0.6557), and the treatment with feedback is significantly different from the other treatments at the 1 percent level for a two-sided t-test.[22] In all three treatments, overguessing is significantly different from zero at the 1 percent level.

Robustness Check: 10-State Sessions.—There are 84 subjects in our 10-state robustness sessions, and as mentioned previously, all subjects in those sessions were assigned to the no feedback and random role treatment. Table 2 provides the summary of player actions in the no feedback and random role treatment for the main sessions and for the 10-state robustness sessions. As when there are five secret numbers, the reporting rate increases monotonically with the secret number in the sessions with ten secret numbers. The reporting rate for a secret number of 3 in the robustness sessions is 41.2 percent, which is comparable to and not statistically different from the reporting rate for a secret number of 2 in the primary study

[22] The p-values for a linear regression of guess of nonreported secret number onto treatment dummies with session-level clustering are significantly different between the feedback treatment and either of the other treatments (p-value = 0.002 for the fixed role treatment and p-value = 0.020 for random role treatment).

of 42.6 percent ($p = 0.7379$ for a two-sided test of proportions). In addition, the reporting rate in the robustness sessions for secret numbers of 5, 7, and 9 are similar and not statistically different from the reporting rates for 3, 4, and 5 in the main sessions.

A secret number of 3 in robustness sessions and a secret number of 2 in main sessions are also comparable in the sense that a risk-neutral sender would not want to report secret numbers of 3 in the robustness sessions (when pooling choices across rounds). The average guess for a nonreported secret number is 3.419 with a 95 percent confidence interval of 3.254 to 3.584. As in the main sessions, the average guess is above the average actual nonreported secret number in the robustness sessions. In the robustness sessions, the average nonreport secret number is 2.616, which is 0.803 below the average guess.

We also find a similar pattern in overguessing and underguessing when secret numbers are reported. Once again, guesses are higher than reported secret numbers for low secret numbers and lower than reported secret numbers for high secret numbers, though the effect is smaller for high secret numbers.

In short, our primary qualitative findings for sender and receiver behavior in the main sessions—incomplete disclosure by senders and overguessing of nonreported states by receivers—are robust to enlargement of the state space.

C. *Departures from the Highest Expected Payoff*

To quantify the impact of incomplete disclosure and receiver overguessing on payoffs, we measure how far a subject is from taking the payoff-maximizing action in each decision problem, which provides a rough sense for the size and consequences of the "mistakes" they are making. To do this, we construct the average opponent strategy from our data, determine the expected payoffs for each possible action, and then calculate how far the expected payoffs for the taken action are from the highest expected payoff for taking any action.[23] For senders, the possible actions are reporting or not reporting the secret number. For receivers, we just consider the guesses that are available to them, which is important because they are limited to guessing half units.[24]

Table 3 reports the monetary losses that result from the actions taken in our main sessions, our robustness sessions, and in our high incentives task. Across these settings, we find that receivers are between 9.3 percent and 13.3 percent away on average from the highest expected payoff and that senders are between 3.4 percent and 7.3 percent away on average from their highest expected payoff. In all settings and roles, the losses are significantly different from 0 at a 1 percent level (using a two-sided t-test). We also show the losses of receivers relative to the payoffs they would make in the unraveling equilibrium. These losses range from 19.9 percent to 26.4 percent.

[23] Because the minimum possible payoff can be negative, we normalized payoffs by the distance from the minimum possible payoff for the realized state.

[24] We thank an anonymous referee for suggesting this method.

Table 3—Payoff Losses for Senders and Receivers (Main Sessions, Robustness Sessions, and High Incentives Task)

Variables	*N*	Mean	SD
Panel A. Fraction of highest expected payoffs not earned			
Receiver loss (no feedback and fixed role)	884	0.105	0.146
Receiver loss (no feedback and random role)	1,014	0.132	0.215
Receiver loss (feedback and random role)	675	0.0972	0.177
Receiver loss (robustness sessions)	581	0.106	0.188
Receiver loss (high incentives task)	57	0.0927	0.128
Sender loss (no feedback and fixed role)	2,565	0.0546	0.148
Sender loss (no feedback and random role)	2,700	0.0477	0.137
Sender loss (feedback and random role)	2,025	0.0684	0.169
Sender loss (robustness sessions)	1,890	0.0733	0.154
Sender loss (high incentives task)	285	0.0344	0.110
Panel B. Fraction of equilibrium payoffs not earned			
Receiver loss (no feedback and fixed role)	884	0.222	0.128
Receiver loss (no feedback and random role)	1,014	0.264	0.187
Receiver loss (feedback and random role)	675	0.213	0.163
Receiver loss (robustness sessions)	581	0.199	0.170
Receiver loss (high incentives task)	57	0.211	0.112
Panel C. Fraction of distance from action with highest expected payoffs			
Receiver mistake (no feedback and fixed role)	884	0.157	0.133
Receiver mistake (no feedback and random role)	1,014	0.171	0.170
Receiver mistake (feedback and random role)	675	0.141	0.148
Receiver mistake (robustness sessions)	581	0.157	0.162
Receiver mistake (high incentives task)	57	0.144	0.131

Notes: In panel A, the payoff loss is the fraction of the highest expected payoffs that was not achieved for a given decision (relative to the expected payoffs for that decision). In panel B, the payoff loss is the fraction of the equilibrium payoffs that was not achieved for a given decision (relative to the expected payoffs for that decision). All payoffs have been normalized by the minimum payment that could be achieved with a decision. Expectations are formed treating all secret numbers as equally likely and assuming there is an equal chance of facing players in the other role from any round of the same session. Receiver losses are just for rounds where the secret number is not reported. In panel C, we calculate the distance from the guess with highest expected payoff divided by the width of the state space (5 or 10).

In addition, Table 3 illustrates two more findings. First, senders have smaller losses than receivers in all settings. Second, when incentives are increased tenfold, receivers continue to have similar percentage losses. Receiver losses in the first 45 rounds of the no feedback and fixed role treatment of the main sessions are not significantly different from receiver losses in the high incentives task ($p = 0.5510$). This does not appear to be driven by low power, as sender losses are significantly different in the same comparison ($p = 0.0260$).

These losses are reflected in the size of receiver mistakes, as measured by the distance from the guess with highest expected payoff (normalized by the size of the state space). Across treatments, receivers are 14.1 percent to 17.1 percent away from the guess with the highest expected payoff. Even when payoffs are increased tenfold, receivers are still 14.4 percent away, and as the state space is stretched in the robustness treatment, receivers are still 15.7 percent away.

Our calculations take an ex ante perspective, so when determining the highest expected payoff for receivers, we assume that all states are equally likely to happen and determine the average sender behavior separately for each state. In addition, we pool together all rounds when determining average sender and receiver behavior,

TABLE 4—SUMMARY OF PLAYER MISTAKES RELATIVE TO A PLAYER'S SELF-REPORTED BELIEF AND DYNAMIC RESPONSE TO LAST BLOCK OF ROUNDS (DEFINE EVERY 5 ROUNDS AS ONE BLOCK)

	No feedback and fixed role	No feedback and random role	Feedback and random role
Panel A. Senders			
Number of unique subjects that have acted as sender	57	120	90
Belief of receiver guess if no report (subject-specific, self-reported)	2.273	2.143	1.617
Action consistent with self-reported belief (report if belief $\leq$ draw, not report if belief $\geq$ draw)	0.875	0.872	0.858
Rounds 1–15	0.856	0.816	0.793
Rounds 16–30	0.876	0.887	0.861
Rounds 31–45	0.891	0.913	0.921
Action consistent with average receiver behavior in last block of rounds (report if average guess given reporting $\geq$ average guess given nonreporting)	0.824	0.805	0.863
Rounds 6–15	0.814	0.773	0.798
Rounds 16–30	0.813	0.814	0.864
Rounds 31–45	0.843	0.817	0.907
Sender loss (fraction of highest expected payoffs not earned given average receiver behavior in last block of rounds)	0.051	0.068	0.065
Rounds 6–15	0.052	0.088	0.085
Rounds 16–30	0.062	0.061	0.074
Rounds 31–45	0.039	0.062	0.043
If draw = 1 or 2:			
Action consistent with self-reported belief	0.815	0.831	0.778
Rounds 1–15	0.805	0.762	0.696
Rounds 16–30	0.825	0.853	0.777
Rounds 31–45	0.814	0.885	0.858
Action consistent with average receiver behavior in last block of rounds	0.665	0.650	0.706
Rounds 6–15	0.679	0.630	0.591
Rounds 16–30	0.663	0.707	0.718
Rounds 31–45	0.659	0.594	0.777
Sender loss	0.085	0.109	0.131
Rounds 6–15	0.083	0.150	0.173
Rounds 16–30	0.098	0.083	0.146
Rounds 31–45	0.072	0.104	0.085

(*continued*)

which is equivalent to assuming that a subject is equally likely to face an opponent from any round. While this is exactly the way that opponents are determined in the high incentives task, such an assumption may not be suitable for treatments where we observe learning.

To adjust for dynamic considerations, in Table 4 we calculate sender and receiver losses with the assumption that a subject is equally likely to face any opponent from the previous block of 5 rounds. However, the magnitude of the losses does not change dramatically with the change in reference group. Across treatments in the main sessions, average sender losses range from 5.1 percent to 6.8 percent, and the average receiver losses range from 9.4 percent to 12.9 percent.

Table 4—Summary of Player Mistakes Relative to a Player's Self-Reported Belief and Dynamic Response to Last Block of Rounds (define every 5 rounds as one block) (*continued*)

	No feedback and fixed role	No feedback and random role	Feedback and random role
Panel B. Receivers			
Number of unique subjects that have acted as receiver	57	120	90
Conditional on sender not reporting			
Guess of secret number	2.243	2.283	1.897
Rounds 1–15	2.325	2.464	2.190
Rounds 16–30	2.211	2.251	1.904
Rounds 31–45	2.180	2.088	1.495
Actual secret number	1.734	1.802	1.680
Rounds 1–15	1.811	1.884	1.895
Rounds 16–30	1.781	1.826	1.618
Rounds 31–45	1.590	1.670	1.464
Guess − actual secret number	0.102	0.096	0.043
Rounds 1–15	0.103	0.116	0.059
Rounds 16–30	0.086	0.085	0.057
Rounds 31–45	0.118	0.084	0.006
Mistake: (guess − guess with highest expected payoff given average sender disclosure in last block of rounds)/5	0.159	0.167	0.138
Rounds 6–15	0.167	0.169	0.152
Rounds 16–30	0.146	0.164	0.156
Rounds 31–45	0.169	0.175	0.105
Receiver loss (fraction of highest expected payoffs not earned given average sender disclosure in last block of rounds)	0.107	0.129	0.094
Rounds 6–15	0.113	0.143	0.106
Rounds 16–30	0.093	0.117	0.112
Rounds 31–45	0.118	0.129	0.062

D. *Stated Beliefs: Senders*

As described in Section IV, after all 45 rounds were completed and before any additional tasks were undertaken, we asked subjects to guess the average receiver guess when the secret number was not reported over all 45 rounds. The responses of those who played the role of sender at least once are given in panel A of Table 4.

Subjects in the fixed role treatment have the highest average guess (2.273), followed by those in the no feedback and random role treatment (2.143), and the feedback and random role treatment (1.617). The first two are similar in size and are not significantly different ($p = 0.4330$ for a two-sided t-test). However, stated beliefs are significantly different between both of the no feedback treatments and the feedback treatment at a 1 percent level. As shown in panel B of Table 4, these guesses are close to the actual averages (2.243, 2.283, and 1.897), particularly for the no feedback treatments.

We find that sender disclosure decisions are largely consistent with their beliefs about receiver guesses. Regardless of treatment, over 85 percent of decisions are consistent with reporting if and only if the secret number is below the belief of the average receiver guess of the nonreported secret number. Over the last 15 rounds, this rate rises to 89.1 percent in the fixed role treatments, to 91.3 percent in the no

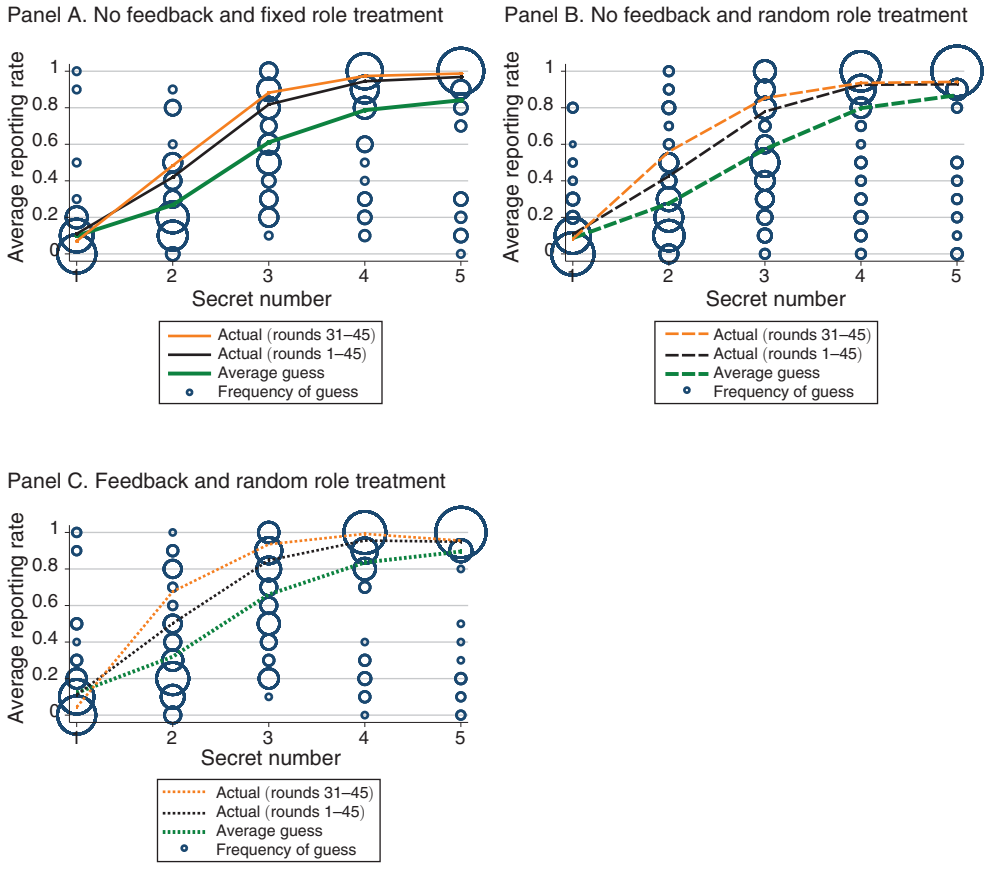

FIGURE 1. SENDER DISCLOSURE RATES AND GUESSES OF SENDER DISCLOSURE RATES (MAIN SESSIONS)

feedback and random role treatment, and to 92.1 percent in the feedback and random role treatment. It may not be surprising that the rate of consistency increases given that beliefs are collected after all 45 rounds are complete.

E. *Stated Beliefs: Receivers*

After all 45 rounds were complete, we also asked subjects who played the role of receiver at least once to guess the percentage of senders who reported each secret number over all 45 rounds. The frequency of their responses is given in the three panels of Figure 1, along with the average guesses and the actual rates each secret number was reported.

Larger bubbles correspond to a larger number of guesses in an interval of 5 percentage points, and the thick line corresponds to the average guess. While there is clear heterogeneity in the stated beliefs of subjects, the bulk of guesses follow the average rate. One exception is that there are a large number of guesses above the average guess for a secret number of 3. However, the two upper lines, which represent the actual reporting rate, pass through this region, so these higher guesses represent

accurate beliefs. The thin line represents the reporting rate for the last 15 rounds, and the medium line represents the average reporting rate for all 45 rounds.

Regardless of the treatment, average guesses for reporting rates of draws of 2 and 3 are well below the actual averages, and they are significantly different at a 1 percent level for all treatments (for a two-sided *t*-test). Although the differences in this gap between treatments are not significant, the feedback treatment has the largest gap between average guess and actual average for draws of 2 (18.3 percent). Though subjects were asked to guess the reporting rate for all 45 rounds, it seems plausible that their guesses would be closer to the actual reporting rate for the more recent rounds. However, the gap between average guesses and actual rates increases if we just consider actual rates in the last 15 rounds.

Using Bayes' Rule, the likelihood that senders report each secret number can be mapped into the likelihood that a nonreported secret number is either 1, 2, 3, 4, or 5. Thus, if we assume that receivers update beliefs correctly, elicited beliefs about sender strategies imply certain beliefs about how likely a nonreported secret number is 1, 2, 3, 4, or 5 and beliefs about the average nonreported secret number. The former will be useful in the following structural analysis, and the latter will be useful in a subsequent regression analysis.

F. *Structural Model*

To study the relationship between elicited beliefs and receiver guesses, we estimate a structural model of receiver decision-making that accounts for boundary effects and allows for confusion and social preferences. In this model, a receiver who faces nondisclosure first forms beliefs γ of the probability of each secret number based on their beliefs F of the probability a sender of each type will report, and then takes an action a (makes a guess) that maximizes their expected utility. Because secret numbers are equally likely, this decision rule is given by the following optimization problem:

$$\max_{a \in A} \sum_{b \in B} \gamma(b \,|\, null)\, U_R(a, b) \tag{1}$$

$$\text{where} \quad \gamma(b \,|\, null) \;=\; \frac{F\big(null \,|\, b\big)}{\sum_{b' \in B} F\big(null \,|\, b'\big)}.$$

We allow that with some probability the receiver makes a random guess. Specifically, in our structural model there is a parameter η that captures the frequency with which the receiver makes a uniform random guess. In other words, η percent of the time a receiver randomly guesses—and each action has a $1/9$ chance of being selected when this occurs. There are multiple ways to interpret this parameter, including receiver confusion.[25] By this, we mean receivers who overlook the strategic element

[25] While random guessing is a frequent assumption for confused play, such as "Level-0" play in the Level-*k* model, it does not capture all forms of confusion. However, it allows us to avoid specifying the exact form of confusion, which is required for other approaches (see Martin and Munoz-Rodriguez 2019).

of their decision or some aspect of the game. We include this parameter because we do not want to confound confusion with naïveté in our estimates of naïveté.

We also allow for some probability that receivers adopt social preferences as in Fehr and Schmidt (1999). In our structural model, there is a parameter δ that captures the frequency with which receivers have the utility function $U_R(a,b) - \beta\big(U_R(a,b) - U_S(a,b)\big)$, where the parameter β captures the receiver's aversion to advantageous inequality.[26] We include these parameters because we do not want to confound overguessing due to social preferences and overguessing due to naïveté.

Together, these two forces give us three parameters to estimate: the probability of uniform random guessing η, the probability of using social preferences δ, and the parameter of the Fehr-Schmidt model of social preferences β. In estimating these parameters, we take the likelihood of action a as the sum of three values: $(1 - \eta - \delta)$ if a maximizes the utility function without social preferences, $(1/9)\eta$ for the chance it is randomly selected, and δ if a maximizes the social preference utility function given parameter β. We estimate these parameters using guesses when the secret number is reported (out-of-sample) and then use them to predict guesses when the secret number is not reported (within-sample).

Guesses of Disclosed Secret Numbers.—As discussed above, some subjects do not guess the secret number correctly, even when it has been disclosed. These mistakes suggest that some subjects may be confused about the game or may understand the game but choose to guess differently from the true disclosed number because of social preferences. Measuring the extent of these behavioral factors when secret numbers are disclosed can be helpful, as the same behavioral factors may affect receiver behavior when they face the more complicated situation of nondisclosure.

We estimate confusion and the social preferences of the subjects jointly. The parameters of this model were estimated to maximize likelihood of the observed receiver guesses (of disclosed secret numbers), using the Nelder–Mead method with 1,000 random started values, and the standard errors were computed using 1,000 bootstrapping samples. These results, along with the likelihoods and predicted levels of overguessing, are provided in Table 5, panel A.

We find that receivers are confused 18.8 percent of the time in the no feedback and fixed role treatment, though this estimate includes times that the receiver was confused but guessed correctly anyway. This fraction declines to 11.9 percent if we add role switching and, down further, to 6.3 percent if we add both role switching and round-by-found feedback. This pattern suggests that role switching and full feedback help to reduce subject confusion.

In contrast, the estimates of social preferences are comparable across treatments. These results suggest that only 3.1 percent to 4.6 percent of receiver decisions are impacted by social preferences. For the advantageous inequality parameter, the

[26] Because sender payoffs are concave, this functional form makes the consequences of advantageous inequality particularly strong for low secret numbers. An alternative reason why social preferences might be particularly strong for low secret numbers is that guesses of 1 and 1.5 generate negative payoffs for the sender. We thank an anonymous referee for raising this possibility.

Table 5—Summary of Structural Estimation of Receiver Decisions (main sessions)

Panel A. Reported secret numbers

	No feedback and fixed role		No feedback and random role		Feedback and random role	
Variable	Actual	Model	Actual	Model	Actual	Model
Average log likelihood		−0.868		−0.661		−0.436
Total log likelihood		−1,459		−1,114		−588
Parameter (confusion)		0.188		0.119		0.063
Parameter (social preferences)		0.031		0.046		0.041
Parameter (Fehr-Schmidt)		0.419		0.536		0.409
Secret number	Guess – secret number (mean values)					
1	0.250	0.423	0.533	0.329	0.298	0.167
2	0.203	0.219	0.243	0.188	0.157	0.104
3	0.002	0.016	0.061	0.046	0.071	0.021
4	−0.103	−0.172	0.009	−0.095	0.016	−0.043
5	−0.294	−0.376	−0.175	−0.237	−0.032	−0.126
Average distance (unweighted)		0.071		0.088		0.078

Panel B. Nonreported secret numbers

	No feedback and fixed role			No feedback and random role			Feedback and random role		
Variable	Actual	No naïveté	Elicited beliefs	Actual	No naïveté	Elicited beliefs	Actual	No naïveté	Elicited beliefs
Average log likelihood		−2.950	−2.242		−3.189	−2.818		−3.555	−3.462
Total log likelihood		−2,608	−1,981		−3,234	−2,857		−2,400	−2,337
Secret number	Guess – secret number (mean values)								
1	1.196	0.829	1.237	1.236	0.747	1.145	1.145	0.656	0.938
2	0.279	−0.171	0.242	0.316	−0.253	0.159	0.159	−0.344	0.066
3	−0.655	−1.171	−0.694	−0.571	−1.253	−0.831	−0.831	−1.344	−0.933
4	−1.679	−2.171	−1.794	−1.800	−2.253	−1.893	−1.893	−2.344	−1.869
5	−2.938	−3.171	−2.803	−2.697	−3.253	−2.809	−2.809	−3.344	−2.830
Average distance (unweighted)		0.412	0.073		0.494	0.142		0.585	0.144

estimate is 0.419 for the treatment of no feedback and fixed roles, which implies that it is optimal for a receiver with such social preferences to overguess secret numbers of 1 by 1.5, secret numbers of 2 by 1, and secret numbers of 3 and 4 by 0.5. The value of 0.409 has similar implications, except that secret numbers of 1 are only overguessed by 1 instead of 1.5.

Model of Naïveté for Nondisclosed Secret Numbers.—To accommodate strategic uncertainty when secret numbers are not disclosed, we initially assume that receivers are risk-neutral, expected-utility maximizers over ex post payoffs and hold correct beliefs about the frequency that senders of each type report the secret number. This corresponds to the columns labeled "No naïveté" in Table 5, panel B.

To see if elicited beliefs help in explaining receiver mistakes, we further assume that receivers vary in their strategic beliefs, and that they hold beliefs in line with their elicited beliefs. The predictions of our model with these assumptions is presented in the columns labeled "Elicited beliefs" in Table 5, panel B. This assumption substantially increases the model's likelihood in the no feedback treatments: from −2,608 to −1,981 when there are fixed roles and −3,234 to −2,857 when roles are random. It also improves the model's ability to explain overguessing. The average distance between actual and predicted guesses is reduced from a range of 0.412 to 0.585 to a range of 0.073 to 0.144. As discussed previously, this fit could potentially be improved further by incentivizing beliefs and by accounting for dynamic effects in the model.

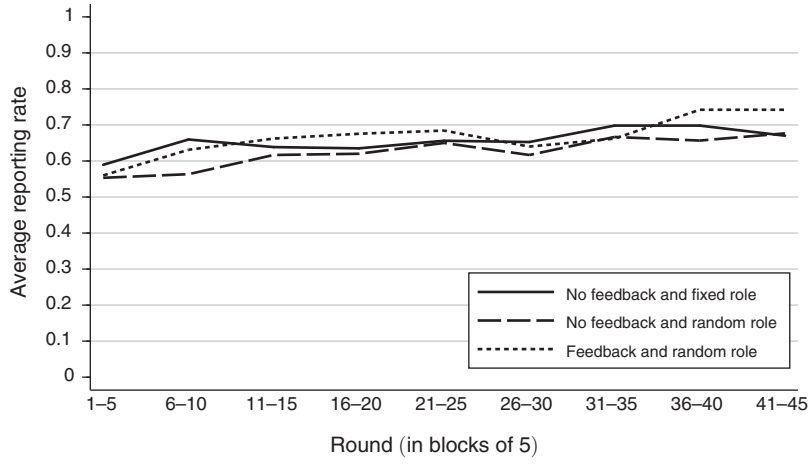

FIGURE 2. SENDER DISCLOSURE RATES BY ROUND (MAIN SESSIONS)

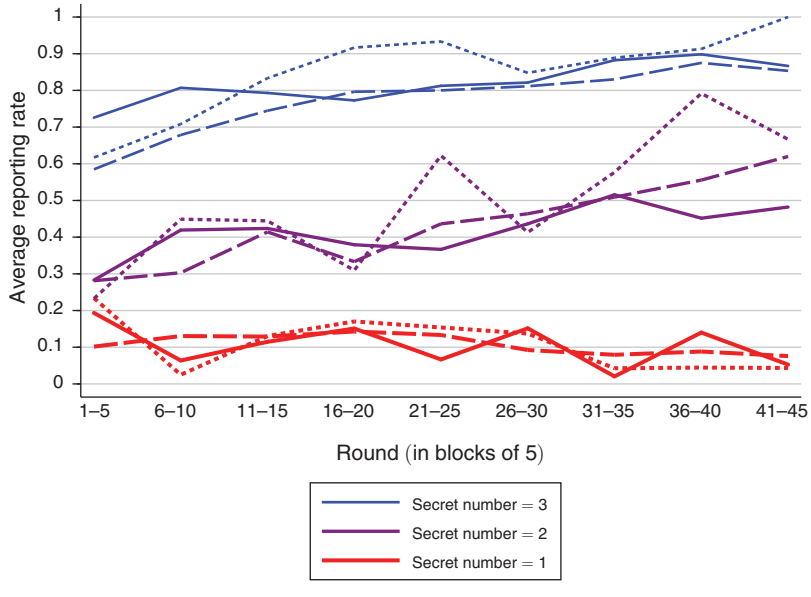

FIGURE 3. SENDER DISCLOSURE RATES BY ROUND FOR SECRET NUMBERS OF 3, 2, OR 1 (MAIN SESSIONS)

G. *Round-by-Round Dynamics: Sender Disclosures*

Next, we analyze the dynamics of behavior across 45 rounds and with varying levels of feedback, starting with sender behavior. Figure 2 shows the overall sender reporting rate for each treatment by block of 5 rounds, and Figure 3 shows the sender reporting rate for each treatment for each of the first three secret numbers separately. As in Figure 1, the dotted line represents the feedback treatment and the

Table 6—Regressions on Sender Disclosures (main sessions)

Sample	All draws		Only draws of 2		Rounds 6–45	
Dependent variable	Report or not		Report or not		Distance from highest expected payoff (fraction)	
	(1)	(2)	(3)	(4)	(5)	(6)
Dummy = 1 if in the first 5 rounds	−0.0442	−0.0464	−0.0652	−0.0636		
	(0.0209)	(0.0209)	(0.0604)	(0.0648)		
Round number (1 to 45)	0.00113	0.00118	0.00277	0.00207	−0.000413	−0.000417
	(0.000352)	(0.000341)	(0.00119)	(0.00126)	(0.000241)	(0.000226)
Round number × random role × no feedback	0.000686	0.000329	0.00378	0.00225	1.95e-05	7.20e-05
	(0.000591)	(0.000574)	(0.00208)	(0.00237)	(0.000430)	(0.000508)
Round number × random role × feedback	0.00180	0.00167	0.00706	0.00651	−0.000878	−0.000737
	(0.00106)	(0.00108)	(0.00307)	(0.00356)	(0.000460)	(0.000473)
Dummy = 1 if sender belief of receiver guess upon non-report is below the actual draw	0.241		0.321		−0.0402	
	(0.0423)		(0.0473)		(0.0177)	
Dummy = 1 if draw = 2	0.219	0.335			0.00586	−0.0138
	(0.0341)	(0.0262)			(0.0191)	(0.0224)
Dummy = 1 if draw = 3	0.523	0.700			−0.0194	−0.0468
	(0.0473)	(0.0238)			(0.0184)	(0.0190)
Dummy = 1 if draw = 4	0.607	0.837			−0.0442	−0.0808
	(0.0496)	(0.0173)			(0.0175)	(0.0185)
Dummy = 1 if draw = 5	0.613	0.839			−0.0493	−0.0847
	(0.0519)	(0.0193)			(0.0207)	(0.0195)
Individual demographics	X	Absorbed	X	Absorbed	X	Absorbed
Session fixed effects	X	Absorbed	X	Absorbed	X	Absorbed
Subject fixed effects		X		X		X
Observations	7,224	7,224	1,477	1,477	5,742	5,742
R^2	0.512	0.580	0.180	0.629	0.075	0.224

Notes: In parentheses are robust standard errors clustered by session. We define every five rounds as one block. Highest expected payoff is based on the distribution of receiver behavior he/she has observed in the last block of the same session. Columns 5 and 6 exclude the first five rounds because we need to construct the initial condition from the first five rounds. In all regressions, the default is the no feedback & fixed role treatment, and draw = 1.

solid line represents the fixed role treatment. Without controlling for any other factors, it appears that there is an increasing trend for draws of 2 and 3 for the feedback treatment and possibly for the other treatments as well. However, these effects could be confounded with demographic differences between subjects or differences in the composition of each session.

Table 6 investigates these trends using regression analysis with demographic controls, session fixed effects or subject fixed effects, and robust standard errors clustered by session (the same regression specifications estimated with standard errors clustered by subject appear in the online Appendix). When not using subject fixed effects, we also include a dummy variable for whether the draw is above the sender's guess of the average receiver guess of nondisclosed numbers.

In specifications 1 and 2, all draws are included in the analysis, but separate dummy variables are provided for each of the draws. In keeping with our previous findings, the probability of reporting increases monotonically with the draw and the difference between a draw of 1 and any other draw is statistically significant at the 1 percent level holding round, draw, and individual characteristics fixed. This is also true if standard errors are clustered by subject instead of by session.

In terms of dynamics, we find that the reporting rate is much lower in the first five rounds, and this effect is statistically significant at the 5 percent level for both

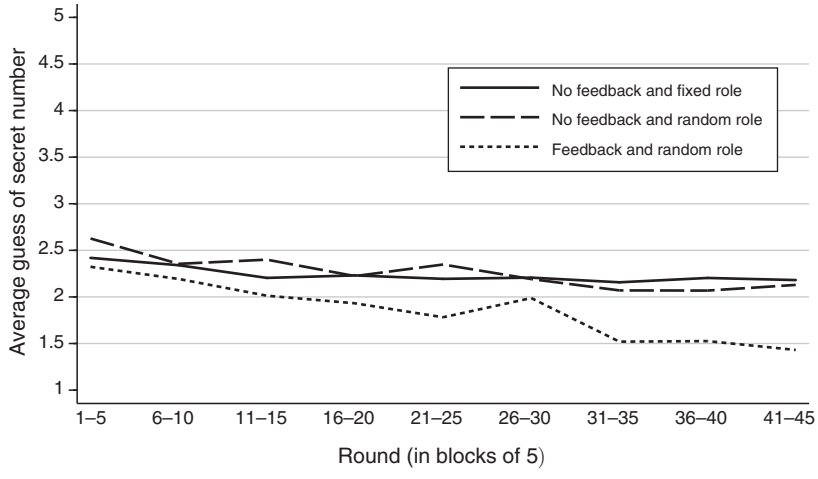

Figure 4. Receiver Guesses of Nondisclosed Secret Number by Round (main sessions)

specifications (regardless of the level of clustering). There appears to be a small positive time trend on the reporting rate even for the no feedback treatment, which is significant at the 1 percent level.

If we look just at draws of 2 (specifications 3 and 4), there is only a statistically significant time trend for the no feedback and fixed role treatment without subject fixed effects and when clustering at the sessions level, but there is a statistically significant difference between this treatment and the feedback treatment for both specifications and clustering levels. The effect size is also relatively large for the difference in time trends between these two treatments.

In these regressions, we also see that the impact of beliefs is largely and highly significant, regardless of which fixed effects or clustering level are used. Controlling for the draw itself, along with time trends and individual characteristics, the probability of reporting increases almost 25 percent if the sender's stated beliefs are that the receiver's average guess is lower than this draw.

H. *Round-by-Round Dynamics: Receiver Guesses*

Figure 4 plots the average receiver guess for each block of 5 rounds, conditional on senders not reporting. While all three treatments start out at a similar place, the feedback treatment appears to diverge from the others and has a clearer time trend.

Panel B of Table 4 provides more detail on these trends, but again without any controls. The average guess of nonreported secret numbers appears to be decreasing for all treatments, but this is a bit misleading, as the actual secret number is also decreasing for all treatments. If we look instead at the amount of overguessing, it appears that the feedback treatment has the biggest drop, with almost no overguessing in the final block of 15 rounds. While overguessing is significantly different from zero in the final block of rounds for the no-feedback treatments, it is not significantly difference different from zero based on a two-sided t-test (p-value $= 0.7160$).

Table 7—Regressions on Receiver Guesses of Nonreported Secret Numbers (main sessions)

Dependent variable	Receiver guess		Distance from highest expected payoff (fraction)	
	(1)	(2)	(3)	(4)
Dummy = 1 if in the first five rounds	0.159	0.111		
	(0.0547)	(0.0514)		
Round number (1–45)	−0.00358	−0.00382	0.000247	0.000133
	(0.00281)	(0.00287)	(0.000385)	(0.000403)
Round number × random role × no feedback	−0.00466	−0.00564	−0.000162	−0.000369
	(0.00361)	(0.00315)	(0.000662)	(0.000669)
Round number × random role × feedback	−0.0173	−0.0182	−0.00197	−0.00208
	(0.00421)	(0.00387)	(0.000503)	(0.000489)
Implied average nonreported number given receiver stated beliefs	0.695		0.120	
	(0.104)		(0.0217)	
Individual demographics	X	Absorbed	X	Absorbed
Session fixed effects	X	Absorbed	X	Absorbed
Subject fixed effects		X		X
Observations	2,551	2,551	2,204	2,204
R^2	0.315	0.680	0.204	0.636

Notes: In parentheses are robust standard errors clustered by session. We define every five rounds as one block. Highest expected payoff is based on the distribution of sender behavior in the last block of the same session. Columns 3 and 4 exclude the first five rounds because we need to construct the initial condition from the first five rounds.

Table 7 shows the output of a number of regressions on the dynamics of receiver guesses. All specifications include both a version with individual demographic controls and a version with subject fixed effects, and these regressions are estimated with clustering by session (once again, versions estimated with standard errors clustered by subject appear in the online Appendix). Looking at specifications 1 and 2, there is a large and statistically significant impact on guessing in the first five rounds: guesses are much higher during those rounds. In the no feedback and fixed role treatment, there is not a statistically significant time trend otherwise. However, there is a large and statistically significant difference in time trends between this treatment and the feedback treatment. These results are robust to the level of clustering used.

As with senders, beliefs appear to have a large and highly significant impact on guesses. Controlling for time trends and individual factors, we find in specification 1 that there is a very high correlation between the implied belief of the average secret number and what subjects guess for both levels of clustering.

To investigate why receiver guesses decrease over time in the feedback treatment, we examine whether changes from one guess of a nondisclosed secret number to the next guess of a nondisclosed secret number are related to the types of mistakes made and whether feedback was received. Table 8 shows the results of several regression specifications based on this objective. For specifications 1 and 2, we find strong evidence that subjects who were informed in the feedback treatment that they guessed too high decreased their guesses the next time they had an opportunity to do so. This

TABLE 8—REGRESSIONS ON RECEIVER GUESSES OF NONREPORTED SECRET NUMBERS (MAIN SESSIONS)

Sample	All last guesses		Last guess = 2, 2.5, or 3	
Dependent variable	Guess with no report − last guess with no report			
	(1)	(2)	(3)	(4)
Dummy = 1 if in the first 5 rounds	−0.027	−0.014	0.065	0.048
	(0.038)	(0.046)	(0.076)	(0.080)
Round number (1–45)	−0.000	−0.000	0.000	−0.002
	(0.001)	(0.001)	(0.001)	(0.002)
Round number × random role × no feedback	0.002	0.003	0.003	0.002
	(0.002)	(0.002)	(0.002)	(0.003)
Round number × random role × feedback	0.001	−0.000	−0.003	−0.010
	(0.002)	(0.002)	(0.006)	(0.007)
Overguessed last time	−0.191	−0.249	−0.072	−0.094
	(0.051)	(0.061)	(0.068)	(0.081)
Overguessed last time × random role × no feedback	−0.129	−0.228	0.028	−0.020
	(0.085)	(0.112)	(0.097)	(0.112)
Overguessed last time × random role × feedback	−0.386	−0.519	−0.384	−0.360
	(0.134)	(0.204)	(0.134)	(0.161)
Underguessed last time	−0.005	0.005	−0.073	−0.095
	(0.056)	(0.062)	(0.074)	(0.088)
Underguessed last time × random role × no feedback	0.103	0.131	0.209	0.200
	(0.092)	(0.118)	(0.115)	(0.133)
Underguessed last time × random role × feedback	0.037	0.015	0.124	0.120
	(0.071)	(0.091)	(0.146)	(0.197)
Individual demographics	X	Absorbed	X	Absorbed
Session fixed effects	X	Absorbed	X	Absorbed
Subject fixed effects		X		X
Observations	2,287	2,306	1,151	1,154
R^2	0.078	0.124	0.087	0.386

Note: Robust standard errors clustered by session.

effect is statistically significant at the 1 percent level for both specifications and both levels of clustering.[27]

There are two potential concerns with these results. First, if subjects have access to other sources of learning, we would expect this effect to exist for all treatments, and it does. However, the effect is stronger for subjects receiving feedback. Second, mean reversion from extreme guesses could produce similar results because very high guesses are almost surely overguesses and very low guesses are almost surely not overguesses. In specifications 3 and 4, we examine whether the effects we observe in specifications 1 and 2 hold also for intermediate guesses. We find that the effect in the no-feedback treatment diminishes, but the effect in the feedback treatment stays large. It is statistically significant in both specifications and both levels of clustering.

[27] Unlike overguessing, there is not a statistically significant relationship between receiving feedback and underguessing for any of the specifications. However, this asymmetry could be driven by the fact that underguessing occurs less often, so may be underpowered.

Table 9—Summary of Player Actions before and after Information on Aggregate Reporting in the "Random Role" Treatments (main sessions)

	No feedback rounds 41–45		No feedback after information		Feedback rounds 41–45		Feedback after information	
Variables	*N*	mean	*N*	mean	*N*	mean	*N*	mean
Report (secret number = 1)	66	0.0758	72	0.0278	46	0.0435	44	0.0227
Report (secret number = 2)	50	0.620	57	0.474	36	0.667	48	0.792
Report (secret number = 3)	75	0.853	47	0.809	48	1	45	0.956
Report (secret number = 4)	61	0.967	54	0.963	47	0.979	46	0.957
Report (secret number = 5)	48	0.917	70	0.943	48	0.979	42	1
Secret number (no report)	97	1.649	115	1.609	58	1.328	57	1.351
Guess (no report)	97	2.129	115	2.170	58	1.431	57	1.614

Note: Just five rounds before and after information intervention.

I. *Learning from Aggregate Reporting*

The final question we address is whether providing information about aggregate sender strategies impacts behavior. As shown previously, we find evidence that reported beliefs about sender strategies are not skeptical enough and are strongly correlated with actual guesses, so information about sender strategies could potentially improve guesses and disclosure rates.

To test this, we examine the choices of subjects who completed the aggregate feedback additional tasks, which are all subjects in the two random role treatments. As mentioned previously, after 45 rounds the subjects are shown the corresponding aggregate information. We then have subjects play five more rounds, and we compare the reporting rates and guesses in these rounds to the reporting rates and guesses made in the last five rounds of the first 45 rounds. These results are provided in Table 9.

This table compares the reporting rates for the five rounds just before and just after the information is provided. For senders in the feedback treatment, the reporting rates do not significantly differ for a two-sided *t*-test, nor does the average secret number when senders make no disclosure. The same is true if we look at subjects in the no feedback treatment. The table also compares the guesses made when senders did not report their secret number. If anything, the information intervention causes guesses to rise on average. However, there is not a significant difference in guesses before and after the informational intervention for either of the treatments.

V. Discussion

In this paper, we implemented a simple disclosure game in the lab, measured beliefs, and varied the feedback that subjects received. We found that behavior was largely in line with the theoretical prediction of unraveling, but that systematic departures from these predictions were strongly related to receiver beliefs that appear insufficiently skeptical of nondisclosure. We also found that repeated and direct feedback was required to eliminate these departures and reach full disclosure.

However, we acknowledge that there are other possible reasons for these departures, including the possibility that participants were simply confused about the instructions. We partially account for this possibility in our structural model.

It is also possible that receiver reactions to nondisclosure depend on the message provided along with nondisclosure. For instance, even though the instructions clearly state that senders must choose whether to disclose (and receivers are also put into the sender role to help with learning), it is possible that receivers do not fully appreciate the strategic nature of nondisclosure. Further experiments could help to tease out the underlying mechanisms. For example, the interface could be redesigned to explicitly state that the sender chose not to reveal any news (rather than just being given a blank screen). This has the potential to have both informational effects (if receivers do not fully understand that senders have the option to disclose) and attention effects (by making the strategic nature of the decision more top of mind). Along these lines, if the sender was restricted to send a first-person message such as "I choose not to disclose," then it could make the sender uncomfortable with nondisclosure. We leave the exploration of these forces to future research.

Our findings shed light on a fundamental inference problem that prevents full unraveling in voluntary disclosure, and the conditions under which full unraveling is most likely to occur. In contexts with little or no feedback, receivers appear not to be sufficiently skeptical about undisclosed information. In our experiments, this can persist for the full 45 rounds of the experiment, and as a result, information senders can profit by limiting disclosure. However, round-by-round feedback about mistakes can result in behavior that converges towards the predictions based on full unraveling.

Our results also shed light on the factors that may limit voluntary disclosure in the field, and the situations in which we might expect voluntary disclosure to be an effective policy. These findings suggest that unless buyers receive fast and precise feedback about mistakes after each transaction, market forces can be insufficient to close the information gap between sellers and buyers.

For the products that naturally offer such feedback—say cereals that taste crunchy and t-shirts that hold color fast—voluntary disclosure may converge to the unraveling predictions after a buyer purchases the product many times. However, for product attributes with less immediate feedback—such as the fat content of salad dressing and the cleanliness of a restaurant kitchen—voluntary disclosure may not converge to the unraveling results. In these situations, mandatory disclosure may be necessary if the policy goal is complete disclosure.

However, there is growing evidence that mandating disclosure may not always be sufficient for achieving the desired outcomes. For example, Loewenstein, Sunstein, and Golman (2014) provide a review of papers that show "limited attention, motivated attention, and biased assessments of probability can undermine the goal of promoting informed consumer choice, potentially rendering disclosure ineffective." Moreover, Jin, Luca, and Martin (forthcoming) show that when mandated disclosures can be made complex, senders can successfully hide bad news in complexity, even when the returns to doing so are transparent. Like this paper, they find evidence that one reason why hiding bad news is successful is the strategic naïveté of some receivers.

REFERENCES

Agranov, Marina, and Andrew Schotter. 2012. "Ignorance Is Bliss: An Experimental Study of the Use of Ambiguity and Vagueness in the Coordination Games with Asymmetric Payoffs." *American Economic Journal: Microeconomics* 4 (2): 77–103.

Bederson, Benjamin B., Ginger Zhe Jin, Phillip Leslie, Alexander J. Quinn, and Ben Zou. 2018. "Incomplete Disclosure: Evidence of Signaling and Countersignaling." *American Economic Journal: Microeconomics* 10 (1): 41–66.

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann. 2015. "Privacy Concerns, Voluntary Disclosure of Information, and Unraveling: An Experiment." *European Economic Review* 75: 43–59.

Board, Oliver. 2009. "Competition and Disclosure." *Journal of Industrial Economics* 57 (1): 197–213.

Brown, Alexander L., Colin F. Camerer, and Dan Lovallo. 2012. "To Review or Not to Review? Limited Strategic Thinking at the Movie Box Office." *American Economic Journal: Microeconomics* 4 (2): 1–26.

Brown, Alexander L., Colin F. Camerer, and Dan Lovallo. 2013. "Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews." *Management Science* 59 (3): 733–47.

Cai, Hongbin, and Joseph Tao-Yi Wang. 2006. "Overcommunication in Strategic Information Transmission Games." *Games and Economic Behavior* 56 (1): 7–36.

Costa-Gomes, Miguel A., and Georg Weizsäcker. 2008. "Stated Beliefs and Play in Normal-Form Games." *Review of Economic Studies* 75 (3): 729–62.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.

Dickhaut, John, Margaret Ledyard, Arijit Mukherji, and Haresh Sapra. 2003. "Information Management and Valuation: An Experimental Investigation." *Games and Economic Behavior* 44 (1): 26–53.

Dranove, David, and Ginger Zhe Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." Journal of Economic Literature 48 (4): 935–63.

Esponda, Ignacio, and Emanuel Vespa. 2014. "Hypothetical Thinking and Information Extraction in the Laboratory." *American Economic Journal: Microeconomics* 6 (4): 180–202.

Eyster, Erik, and Matthew Rabin. 2005. "Cursed Equilibrium." *Econometrica* 73 (5): 1623–72.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114 (3): 817–68.

Feltovich, Nick, Richmond Harbaugh, and Ted To. 2002. "Too Cool for School? Signalling and Countersignalling." *RAND Journal of Economics* 33 (4): 630–49.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics* 10 (2): 171–78.

Forsythe, Robert, R. Mark Isaac, and Thomas R. Palfrey. 1989. "Theories and Tests of 'Blind Bidding' in Sealed-Bid Auctions." *RAND Journal of Economics* 20 (2): 214–38.

Forsythe, Robert, Russell Lundholm, and Thomas Rietz. 1999. "Cheap Talk, Fraud, and Adverse Selection in Financial Markets: Some Experimental Evidence." *Review of Financial Studies* 12 (3): 481–518.

Fung, Archon, Mary Graham, and David Weil. 2007. Full Disclosure: The Perils and Promise of Transparency. New York: Cambridge University Press.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121 (2): 505–40.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* 24 (3): 461–83.

Grossman, S.J., and O.D. Hart. 1980. "Disclosure Laws and Takeover Bids." *Journal of Finance* 35 (2): 323–34.

Grubb, Michael D. 2011. "Developing a Reputation for Reticence." *Journal of Economics and Management Strategy* 20 (1): 225–68.

Hagenbach, Jeanne, and Eduardo Perez-Richet. 2018. "Communication with Evidence in the Lab." *Games and Economic Behavior* 112: 139–65.

Harbaugh, Richmond, and Theodore To. 2020. "False Modesty: When Disclosing Good News Looks Bad." *Journal of Mathematical Economics* 87: 43–55.

Heidhues, Paul, Botond Kőszegi, and Takeshi Murooka. 2017. "Inferior Products and Profitable Deception." *Review of Economic Studies* 84 (1): 323–56.

Hirshleifer, David, and Siew Hong Teoh. 2003. "Limited Attention, Information Disclosure, and Financial Reporting." *Journal of Accounting and Economics* 36 (1–3): 337–86.

Jin, Ginger Zhe. 2005. "Competition and Disclosure Incentives: An Empirical Study of HMOs." *RAND Journal of Economics* 36 (1): 93–112.

Jin, Ginger Zhe, Michael Luca, and Daniel J. Martin. Forthcoming. "Complex Disclosure." *Management Science*.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin. 2021. "Replication data for: Is No News (Perceived As) Bad News? An Experimental Investigation of Information Disclosure." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E118203V1.

King, Ronald R., and David E. Wallin. 1991. "Voluntary Disclosures When Seller's Level of Information Is Unknown." *Journal of Accounting Research* 29 (1): 96–108.

Li, Ying Xue, and Burkhard C. Schipper. 2018. "Strategic Reasoning in Persuasion Games: An Experiment." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3127357.

Loewenstein, George, Cass R. Sunstein, and Russell Golman. 2014. "Disclosure: Psychology Changes Everything." *Annual Review of Economics* 6: 391–419.

Luca, Michael, and Jonathan Smith. 2015. "Strategic Disclosure: The Case of Business School Rankings." *Journal of Economic Behavior and Organization* 112: 17–25.

Marinovic, Ivan, and Felipe Varas. 2014. "No News Is Good News: Voluntary Disclosure in the Face of Litigation." Stanford University Graduate School of Business Research Paper 13-19.

Martin, Daniel, and Edwin Muñoz-Rodriguez. 2019. "Misperceiving Mechanisms: Imperfect Perception and the Failure to Recognize Dominant Strategies." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3316346.

Mathios, Alan D. 2000. "The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market." *Journal of Law and Economics* 43 (2): 651–78.

Matthews, Steven, and Andrew Postlewaite. 1985. "Quality Testing and Disclosure." *RAND Journal of Economics* 16 (3): 328–40.

McKelvey, Richard D., and Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games." *Games and Economic Behavior* 10 (1): 6–38.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *Bell Journal of Economics* 12 (2): 380–91.

Milgrom, Paul, and John Roberts. 1986. "Relying on the Information of Interested Parties." *RAND Journal of Economics* 17 (1): 18–32.

Montero, M., & Sheth, J. 2019. Naivety about hidden information: An experimental investigation. CeDEx Discussion Paper Series ISSN 1749–3293.

Mullainathan, Sendhil, Joshua Schwartzstein, and Andrei Shleifer. 2008. "Coarse Thinking and Persuasion." *Quarterly Journal of Economics* 123 (2): 577–619.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122 (3): 1067–1101.

Rey-Biel, Pedro. 2009. "Equilibrium Play and Best Response to (Stated) Beliefs in Normal Form Games." *Games and Economic Behavior* 65 (2): 572–85.

Sánchez-Pagés, Santiago, and Marc Vorsatz. 2009. "Enjoy the Silence: An Experiment on Truth-Telling." *Experimental Economics* 12 (2): 220–41.

Seidmann, Daniel J., and Eyal Winter. 1997. "Strategic Information Transmission with Verifiable Messages." *Econometrica* 65 (1): 163–69.

Serra-Garcia, Marta, Eric van Damme, and Jan Potters. 2011. "Hiding an Inconvenient Truth: Lies and Vagueness." *Games and Economic Behavior* 73 (1): 244–61.

Sheth, Jesal. 2019. "Disclosure of information under competition: An experimental study." CeDEx Discussion Paper Series ISSN 1749–3293.

Sobel, Joel. 2020. "Lying and Deception in Games." *Journal of Political Economy*, 128 (3): 907–47.

Trautmann, Stefan T., and Gijs van de Kuilen. 2015. "Belief Elicitation: A Horse Race among Truth Serums." *Economic Journal* 125 (589): 2116–35.

Viscusi, W. Kip. 1978. "A Note on 'Lemons' Markets with Quality Certification." *Bell Journal of Economics* 9 (1): 277–79.

Wang, Joseph Tao-yi, Micahel Spezio, and Colin F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games." *American Economic Review* 100 (3): 984–1007.