Complex Disclosure

Ginger Zhe Jin
University of Maryland and NBER

Michael Luca
Harvard Business School

Daniel Martin
Northwestern Kellogg School of Management

July 27, 2019

Abstract

We present evidence that complex disclosure can result from strategic incentives to shroud information. In our lab experiment, senders are required to report their private information truthfully, but can choose how complex to make their reports. We find that senders use complex disclosure over half the time. This obfuscation is profitable because receivers make systematic mistakes in assessing complex reports. Regression and structural analysis suggest that these mistakes could be driven by receivers who are naive about the strategic use of complexity or overconfident about their ability to process complex information. We further evaluate the importance of these two explanations in our experiment.

1 Part of the research was conducted when Jin took leave at the Federal Trade Commission. The views expressed are those of the authors and do not necessarily represent those of the U.S. Federal Trade Commission or any individual Commissioner. Martin would like to thank both the Paris School of Economics and the Camargo Foundation for their hospitality during the writing of this paper. Early stages of this project were supported by the French National Research Agency, through the program Investissements d'Avenir, ANR-10--LABX_93-01. We would like to thank Patrick Rooney, Byron Perpetua, and Philip Marx for excellent assistance. All rights reserved. All errors are ours.
1. Introduction

Firms are often required to disclose contract terms and other relevant information to consumers. For example, credit card companies are required to disclose interest rates. Tech companies are required to disclose privacy policies. And public firms are required to disclose financial performance. Ben-Shahar and Schneider (2014) and Lowenstein et al. (2014) provide a long list of hard-to-understand disclosures and have argued that mandated disclosure has failed as public policy as a consequence. Financial experts even blame the complexity of financial products for the 2008 financial crisis, although the risks embedded in these products were supposedly disclosed to a ratings agency.²

Some of these disclosures are complex by necessity – simply because firms need to provide very detailed information. However, because firms have the ability manipulate the complexity of their reports, some of these disclosures may be far more complex than they need to be. For example, credit card companies can present payment schedules, penalties, and fees clearly or bury potentially important details in the fine print.³ Privacy policies can be written in easy to understand language or shrouded in pages of complex legalese. When public firms make financial disclosures, they can summarize them in several paragraphs or run them as long as 257 pages.⁴

The extent to which firms can exploit consumers by increasing complexity depends crucially on how consumers respond when they observe complex information. If consumers are sufficiently skeptical about firms that use complex disclosures and account for this in their decision-making process, then firms that offer better terms or higher-quality products will want to present this information as clearly and simply as possible to prevent themselves from being mistaken for worse firms. As a result, we would expect only the worst firms to use complex disclosures, which is similar to the “unraveling” results in voluntary disclosure (Viscusi 1978; ² https://www.ft.com/content/24f73610-c91e-11dc-9807-000077b07658 accessed on September 26, 2017.
³ The Truth in Lending Act of 1968 (TILA) requires lenders to disclose consumer credit terms and cost in a standardized way. The Real Estate Settlement Procedures Act of 1974 (RESPA) requires lenders and others involved in mortgage lending to provide borrowers with pertinent and timely disclosures regarding the nature and costs of a real estate settlement process. In 2015, the US Consumer Finance Protection Bureau consolidated the disclosure requirements under TILA and RESPA, resulting in the Loan Estimate Form and the Closing Statement Form, which standardize the content and format of disclosure in mortgage lending. ⁴ The SEC does not impose a limit on the length of a filing, and the average 10-K has grown from roughly 30,000 words in 2000 to 42,000 words in 2013, with GE’s 2014 10-K stretching to 103,484 words and 257 pages. Source: https://www.wsj.com/articles/the-109-894-word-annual-report-1433203762.
Grossman and Hart 1980; Grossman 1981; Milgrom 1981). However, systematic mistakes by consumers when facing complex reports can give rise to strategic incentives to complexify, motivating companies to choose complexity over simplicity in their disclosures.

In reality, it is difficult to determine if consumers are sufficiently skeptical about firms that use complex disclosures and account for this in their decision-making process, because it is hard to discern whether disclosures are complex by necessity or the outcome of firms strategically choosing to make information unnecessarily complex. To overcome this difficulty, we design a laboratory experiment to study the strategic use of complexity in a controlled setting. In our experiment, complexity arises only from its strategic use, the conflicting interests of senders and receivers are clear, the amount and over-use of complexity are quantifiable, and the beliefs of agents are easily elicited.

There are two roles in our experiment: a sender (e.g., the firm) and a receiver (e.g., the consumer). Subjects are randomly paired in each round, one randomly assigned to be the sender and the other to be the receiver. In each round, the sender observes a new state (which is an integer drawn uniformly from 1 to 10) and chooses how complex to make their report of the state. Based on the sender’s report, the receiver then makes a guess of the state.

We impose a strong and clear conflict of interest: the sender would like receivers to guess that the true state is as high as possible, and the receiver would like to guess as accurately as possible. In our main sessions, we debrief both players at the end of each round about the true state, the sender’s choice of complexity, and the receiver’s guess in that round. This way, subjects have many opportunities to learn the strategic forces in the game and the consequences of their actions.

When a sender’s report is simple, the state is presented as a single integer. When their report is complex, the state is presented as several computer-generated numbers (up to 20) that add up to the state. While this is just one of many ways to operationalize complexity, it has the advantage that individuals have experience with the task and may hold well-formed beliefs about their ability to internalize complexity of this form. In

---

5 Roles were randomly assigned so that subjects could experience both roles, which allowed subjects to be well informed about the actions and payoffs available in both roles.
6 As a robustness check, we also ran sessions without feedback. In addition, we ran a robustness check where we limit the number of complexity levels available to senders. See the appendix for more details.
addition, this task allows us to easily measure the extent of complexity and the size of the sender and receiver mistakes.

Although senders can make their reports complex, they cannot misrepresent the underlying state, as the computer-generated numbers will always add up to the state. Thus, the theoretical predictions for behavior in our experiment are stark. If complexity generates sufficient uncertainty about the state, it is easy to show using the unraveling arguments mentioned previously that the unique sequential equilibrium is full use of simple reports for all states above the worst, even though there is a conflict of interest between senders and receivers in our experimental game.

Instead, defining “low complexity” as messages with 5 or fewer numbers to sum, senders use low complexity less than 50% of the time, even if we look just at the second half of rounds. When not using low complexity, senders mostly use high complexity (defined as messages with 15 or more numbers to sum), and they do so in a systematic way. Over rounds, senders gravitate towards two extremes: using low complexity for high states (secret numbers of 8 and up) and using high complexity for low states (secret numbers of 3 or less). When the state is neither high nor low, senders use high complexity approximately 33% of the time, even in the second half of rounds.

Why is complex disclosure so prevalent? One possibility is that senders use complex disclosure more than they should. However, we find that using high complexity to hide both low and middle states is profitable for senders because receivers guess higher than the actual secret number in both low and middle states, and this persists into the second half of rounds. Because sender behavior is largely consistent with these strategic incentives, the average losses of senders are small and decrease over rounds.

So why do receivers systematically guess higher than the actual secret number when disclosures are complex? The driving force of complexity in disclosure is that it adds noise to messages because receivers are unable to fully internalize complex information. Because of this noise, receivers face an involved decision-making process:

---

7 Senders use middle complexity at a similar rate across states, so the only variation across states is the frequency that low and high complexity are used.
8 We define the “optimal” action for senders as the one that has the highest expected payoff, and we measure losses in terms of expected payoff.
9 The primary sources of losses are choosing high complexity at high states and choosing low complexity at low states. These mistakes, along with their possible sources, are examined in Section 4.3.
they must infer why complexity was used and how much noise complexity added to the message, and then adjust their guesses accordingly.

There are several ways that this process can be distorted. For example, a receiver who is naive about the strategic use of complexity will not adequately adjust their guesses for the bad news communicated by complexity, which will lead them to systematically guess above the actual secret number. In fact, naivete is the leading assumption that justifies the strategic use of complexity in theoretical models (for example, see Gabaix and Laibson 2006; Spiegler 2006; Carlin 2009; and Armstrong and Vickers 2012), and it has been found to impact other forms of disclosure (Cai and Wang 2006; Jin, Luca, and Martin 2018).

However, naivete is not the only way this process can be distorted. A receiver who is overconfident about their ability to understand complex reports will think there is less noise in the message than there is, so they will put too much weight on their reading of the message. Because complexity is more likely to be used for low secret numbers, if receivers overweight their signal (their reading of the message) relative to their prior (their skepticism about complex messages), it will push their guesses up on average.

Overconfidence about ability has been found in a number of other domains and can take a wide variety of forms (Moore and Healy 2008; Grubb 2015). In addition, there is mounting evidence that ego-utility is an important driver of overconfidence about ability and can lead to asymmetric updating in beliefs after receiving feedback about performance (Eil and Rao 2011; Mobius, Niederle, Niehaus, and Rosenblat 2011).

To study the impact of naivete and overconfidence on receiver guesses, we elicit beliefs from subjects about the strategic implications of complexity and about their math ability. Based on these beliefs, 12.6% of subjects appear naive about complexity, and 33.8% appear overconfident about their math ability. We find strong evidence based on regressions that when receivers are either more naive about complexity or more overconfident about their math ability, they over-guess by a larger amount.

We expand on these results with a structural analysis. Because the impact of naivete and overconfidence on receiver guesses is indirect and distorted by boundary effects, it is not obvious whether the extent of receiver mistakes we observe can be attributed to this degree of naivete, this degree of overconfidence, or a combination of
both biases. Thus, we use a partial-equilibrium structural model to explore the impact of different mechanisms for receiver mistakes. In our structural model, receivers observe a level of complexity, apply their strategic beliefs about the link between complexity and states, receive a noisy signal of the state based on the complexity level (due to math errors), update their beliefs about the state, and then make a guess. Importantly, we estimate math errors out-of-sample using a math test in which subjects face high complexity without any strategic considerations.

As a baseline, we close the model by assuming that strategic beliefs are correct and belief updating is perfect, and this version of the model fails to predict the over-guessing we observe at middle states and the extent of under-guessing we observe at high states. If we assume instead that 12.6% of receivers are naive, we are better able to predict receiver guesses and the overall levels of over-guessing. However, adding overconfidence about math ability to the baseline model at the levels we observe (instead of naivete) leads to even more accurate predictions of receiver guesses and levels of over-guessing. The combination of overconfidence and naivete also explains the data well, but does little to improve predictions over overconfidence alone. This structural analysis demonstrates that a simple behavioral model with naivete and overconfidence can explain the choices we observe in our experiment.

Evidence that strategic disclosure can be impacted by naivete is consistent with analogous voluntary disclosure experiments in which senders can either disclose or not disclose their information. However, we find two main differences between non-disclosure and complex disclosure. First, complexity persists and is effective even with repeated feedback, unlike non-disclosure. Second, we find evidence that complex disclosure is impacted by an additional bias: overconfidence about ability. This factor is unique to complex disclosure, where noise in messages arises from limits in ability, not external randomization devices.

To our knowledge, this is the first lab experiment to examine the strategic use of complexity for a simple sender-receiver framework in which preferences are known. It

---

10 See DellaVigna (2018) for a survey of the growing use of structural models to separate behavioral forces.
11 Overconfidence about ability is less relevant when noise is added mechanically, such as when messages are randomly selecting from an interval (as in Cai and Wang 2006).
provides evidence from a controlled environment that complex disclosure can result from strategic incentives to shroud information. In addition, our results suggest that strategic naivete and overconfidence about the ability to internalize complexity might be important drivers of these incentives.

Naivete has been proposed in the theoretical literature as reason for the strategic use of complexity, but to the best of our knowledge, overconfidence about ability has not been previously proposed as an explanation for complex disclosure.\textsuperscript{12} Of course, caution should be taken when extrapolating the results of a simple, stylized lab experiment to the field. However, our results suggest that this previously over-looked force might be worth exploring in field studies on complex disclosure and adding to theoretical models of complex disclosure. If this form of overconfidence is found to matter in those settings also, it might be worth addressing in regulatory policy on complex disclosure, as it has a distinct set of policy implications. For example, asymmetry in belief updating could make feedback less effective at reducing receiver mistakes, as we observe in our experiment.

The rest of the paper is organized as follows. Section 2 reviews three related literatures and articulates our contribution to each. Section 3 presents our experimental design. Section 3 discusses our experimental results, looking both at overall behavior and dynamic patterns using regression analysis. Section 4 concludes with potential policy implications.

1. Experimental Design

In this section, we present the game of complex disclosure that we implement in the lab. In order to isolate the forces of interest, the game we use takes a simple form. It is based on the sender-receiver framework used to study cheap talk by Cai and Wang (2006) and voluntary disclosure by Jin et al. (2018). We extend this framework to require senders to truthfully disclose the state and allow them to choose the complexity of their

\textsuperscript{12} Grubb (2015) presents evidence of how other forms of overconfidence interact with complex disclosures, such as overconfidence about the precision of estimates, overconfidence about self-control, and overconfidence about attention to fulfilling contract terms.
messages. After presenting our game, then we give other details about the experimental design.

2.1 The Complex Disclosure Game

In each round, subjects were paired together, and in each pairing, one subject was randomly assigned to be the sender and the other to be the receiver (with equal likelihood). To reduce framing effects, the sender was referred to as the “A Player”, and the receiver was referred to as the “B Player”.

In each round and for each pair, the computer drew a whole number from 1 to 10, called the “secret” number. Each of these numbers was equally likely to be drawn, and both senders and receivers were made aware of this probability distribution.

Each sender was shown the secret number for their pairing and then made their decision about report complexity while the receivers waited. In our main sessions, the sender chose a “report length”, which was a whole number $c$ between 1 and 20. The computer program then randomly selected $c$ integers between -10 and 10 until those numbers added up to the true state $b$. Both senders and receivers were told how these numbers were generated.

After all senders made their decisions, the receivers’ screens became active. If a sender decided on a report of length $c$, the receiver they were paired with was shown this message: “The number I received is”, followed by a table of $c$ integers ranging from -10 to 10 that added up to the secret number. The instructions explicitly stated that the sender only chooses the report length $c$ and that the table of random numbers is generated by the computer. In the appendix, we present the full instructions and an example of a report with maximum length ($c=20$).

Below the area for the sender’s message, receivers were asked to make a guess $a$ of the secret number $b$, and this guess could be any integer between 1 and 10. The receiver had 60 seconds to view the sender’s report and make a guess. If nothing was guessed after that time, a random guess is entered for the receiver. In our main sessions, less than 4% of receivers hit this time limit.
Receiver payoffs, denominated in “Experimental Currency Units” (ECU), were $ECU_R = 110 - 20|b - a)/2|^{1.4}$, where $b$ is the secret number and $a$ is the receiver’s guess. These payoffs decrease monotonically as the guess moves further from the secret number. The sender payoffs in each round were $ECU_s = 110 - 20|(10 - a)/2|^{1.4}$. These payoffs are independent of the secret number and increase monotonically with receiver guesses because guesses cannot be higher than 10. These payoffs are similar to the quadratic specification found in Crawford and Sobel (1982) when there is a large bias towards higher actions. 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.

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 about these two features of sender and receiver payoffs in the instructions.

2.2 Experimental Sessions

Our 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. Subjects do not have to be Harvard University students, but we restricted subject to be no older than 25 years old. The software used to run the experiments was the z-Tree software package (Fischbacher 2007).

Each session consisted of 30 rounds of the disclosure game. 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%. The purpose of switching roles is to ensure that both sides have a good sense for the incentives and
actions available to the other side. In a related experiment, Kalayci and Potters (2011) implement a laboratory experiment where sellers have control over the complexity of product quality, but in their experiment, buyers are given no information about the objectives and incentives of sellers, so it is very difficult to know what buyers believe about why sellers make quality complex.

At the end of each session, subjects were privately paid in cash a show up fee of $5 plus all additional earnings they accumulate over the course of the session. ECU were converted to U.S. dollars at a rate of 150 to 1 (rounded up to the nearest dollar). While it is possible for subjects to end up with a negative balance of ECU, because subjects are paid for every round, this outcome is extremely unlikely and never came close to occurring in the sessions we ran. However, because subjects are paid for every round, there is the potential for intentional variation in play (a “portfolio” strategy), but we find little evidence of such behavior.

2.3 Feedback, Beliefs, Math Test, and Demographics

Our main sessions provide round-by-round feedback. Subjects were told four pieces of information after each round: 1) the actual secret number; 2) the report length chosen by the sender; 3) the receiver’s guess of the secret number, and; 4) their own payoff. After all subjects pressed the “OK” button on the screen containing this feedback, the next round began. 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.

Once all rounds are completed, subjects were asked questions about their beliefs of how other subjects played in their session. First, subjects were asked to guess the average report length that senders chose for each secret number. Second, subjects were asked to guess the average secret number when the sender chose complexity levels between 1 and 5, between 6 and 10, between 11 and 15, and between 16 and 20. The purpose of these questions was to assess whether subject beliefs about sender strategies influenced their decisions as receivers. These guesses were not incentivized, which
introduces the possibility of additional noise. However, we find that even with this additional noise, elicited beliefs are fairly accurate for most subjects.

In some sessions, subjects were asked to complete a four-question math test after answering the two belief questions. For each question in this test, subjects were asked to add up 20 numbers, and were paid $4 if a randomly selected question was correctly answered. Subjects were told that the numbers would sum up to an integer between 1 and 10, that all integers were equally likely, and that the 20 numbers would be generated in the same fashion as in the disclosure game. After completing the math test, subjects answered two additional belief questions. First, they were asked to guess the number of questions on the math test (from 0 to 4) that they thought they answered correctly. Second, they were asked to guess the average number of questions they thought others answered correctly. These belief questions were also not incentivized.

At the end of the experiment, subjects were asked to complete a questionnaire that includes questions about demographic details. Specifically, subjects are asked for their gender, if they are a native English speaker, their year in school, and if they have a friend participating in that session.

2. Experimental Results

In our 29 main sessions, we have 294 subjects, all of whom experience both roles (sender and receiver) and receive round-by-round feedback for 30 rounds. Roughly 41% of the subjects are male, 72% are undergraduate students, 85% are native English speakers, and 14% report that they have a friend in the same session. These demographic distributions are similar to the ones reported by Jin, Luca, and Martin (2018), who also conducted experiments in the CLER lab.

3.1.1 Summary of Behavior and Mistakes
In contrast to the unraveling prediction, the average choice of complexity is 9.728 and increases almost monotonically as the secret number gets smaller.\textsuperscript{19} For the two smallest secret numbers (1 and 2), a majority of senders choose the maximum complexity (report length $c=20$) and over 72% choose high complexity ($c\geq 15$). For the two highest secret numbers (9 and 10), a majority of senders choose the simplest report ($c=1$) and over 72% choose low complexity ($c\leq 5$).

Figure 1 depicts the distribution of complexity choices for each secret number, where the size of the bubble represents the number of senders choosing a specific complexity level conditional on a specific secret number. Most senders concentrate on high complexity when the secret number is below 5, and switch to low complexity when the secret number is above 5. If the secret number is exactly 5, sender choices are dispersed across all levels of complexity.

Turning to receivers, the median receiver guess is correct for every secret number, but the average guess is significantly different from the secret number for every secret number except for 6 and 7 (using a two-sided t-test and a significance level of 5%). On average, the bias in receiver mistakes reveals much greater over-guessing for low secret numbers (1.183 for secret numbers of 1 and 0.936 for 2) than under-guessing for high secret numbers (-0.367 for 9 and -0.403 for 10).

To further explore receiver behavior, we define the size of receiver mistakes as the absolute distance between the receiver guess and the secret number. The average receiver mistake size is the highest for the lowest secret number ($c=1$) and decreases almost monotonically with secret number.\textsuperscript{20} This is consistent with the fact that senders present simpler reports for higher secret numbers, which reduces the potential for math errors, shortens the response time for receivers, and lowers the probability of receivers not making a decision within the 60 second time limit. For the less than 4% of receivers that are over the time limit, the computer generates a random guess, which can lead to large mistakes. Excluding these observations, receiver mistakes remain large for the smallest secret numbers (0.946 for secret numbers of 1 and 0.777 for 2) as compared to

\textsuperscript{19} In a regression of complexity choice onto secret number with individual fixed effects and robust standard errors, the coefficient is negative (-1.496) and statistically significant (p<0.001).

\textsuperscript{20} In a regression of mistake size onto secret number with individual fixed effects and robust standard errors, the coefficient is negative (-0.085) and statistically significant (p<0.001).
the mistake sizes for large secret numbers (between 0.370 and 0.379 for secret numbers between 6 and 10). In fact, mistake sizes are significantly different between secret numbers of 1 and 10 using a two-sided Wilcoxon rank-sum test (p-value<0.001).

On average, we observe a small amount of under-guessing for complexity up to a length of 4. For complexity between 5 and 12, the number of observations is smaller, and the average guess fluctuates between over-guess and under-guess. Once complexity is over 12, we observe consistent over-guessing that peaks at the highest level of complexity (0.655 for length 20). Interestingly, the size of receiver mistakes is less monotonic, but is clearly much higher for high complexity than for low and medium complexity.21 These results are robust to excluding rounds where receivers did not make their decision within the time limit. Without those rounds, the magnitude of receiver mistakes is slightly lower for the two highest complexity levels (0.783 versus 0.748 for length 19, and 1.284 versus 1.008 for length 20).

Absent behavioral factors, one would imagine that receiver mistakes should be zero for the simplest reports because such report reveals the secret number exactly. In contrast, the data shows an average mistake of 0.243 for length 1 and 0.257 for length 2, which are significantly different from 0 using a two-sided t-test. When the secret number is presented simply, receivers tend to have largest mistakes (0.6 on average) for the lowest secret number (1). This is consistent with social preferences because a simple report of a low states is helpful for receivers but harmful for senders, so some receivers may be willing to reciprocate to “honest” senders by sacrificing their own payoff to reward this behavior. However, it may also reflect confusion of about the game. These possibilities are discussed and analyzed further in the appendix.

3.1.2 Dynamics

Table 1 presents the results of our regressions based on sender behavior, and Table 2 presents those based on receiver behavior. For senders, the dependent variables are sender choice of complexity and the payoff departure from the highest expected

21 In a regression of mistake size onto complexity with individual fixed effects and robust standard errors, the coefficient is positive (0.054) and statistically significant (p<0.001).
payoff. In the first and third columns of Table 1, we include subject demographics and session fixed effects. Taking a secret number of 1 as the default, Table 1 shows that senders choose significantly less complexity and depart less from the highest payoff when their secret number increases. This is consistent with our results without subject, session, or round controls.

To capture sender learning, we include the round number (1-30) and the interaction with whether the secret number is in the medium (4-6) or high range (7-10). These coefficients suggest that senders learn to *increase* complexity for low states (1-3) but *decrease* complexity for medium and high states. We also include a dummy for the first five rounds, in case the initial learning about the game creates a level effect in choice of complexity. There is little evidence for a difference when controlling for other factors.

Columns (2) and (4) include sender fixed effects, which absorb individual demographics. Results for most coefficients are similar to what we have without individual fixed effects, suggesting that sender choice and learning are not driven by unobserved individual characteristics.

Turning to receivers, Table 2 attempts to understand the absolute size of receiver mistakes and receiver’s payoff losses when controlling for time trends. Because we want to study the mistakes that receivers actively made, we focus our analysis on the 96% of receiver guesses that are made before the time limit. Since receivers observe the sender’s choice of complexity, we include a separate dummy for each complexity level. We control for the same subject-level variables as in the sender regression, but we also include the receiver’s response time.

Compared with the default complexity (1), Table 2 shows that receiver mistakes drop significantly for some low complexity levels but increase significantly for some levels of high complexity. This pattern is similar with and without subject fixed effects. Results on payoff losses are less consistent, but once we control for subject fixed effects, payoff losses increase significantly between the default complexity and some high complexity levels. Receivers do appear to lower their guess for high complexity over time (after we control for subject fixed effects). As a result, they depart less from the highest expected payoff with high complexity.
3.2 Reasons Behind Receiver Mistakes

In this section, we study the reasons for the mistakes that receivers make when the secret number is presented in a complex way. Along the way, we also explore the reasons behind the mistakes made with simple reports, but our primary focus is on complex reports because the vast majority of receiver mistakes occur when the secret number is disclosed with high complexity, and it is these mistakes that justify the complexity that is observed in our experiment. As a consequence, in the subsequent analyses we only use receiver guesses from rounds where senders chose high complexity and where receivers made a guess before the time limit.

The two primary forces we consider for receiver over-guessing of complexity reports are naivete and overconfidence about ability, but we also consider several other possibilities, such as pure boundary effects, social preferences, confusion, and risk preferences in the structural analysis that appears in the appendix. To help identify naivete and overconfidence, we elicited subject beliefs about the strategic implications of complex disclosure and their performance in a short math test.

3.2.1 Beliefs about Senders and Math Ability

As mentioned previously, after all 30 rounds of the game were completed, we asked subjects to report what they think the secret number was on average in their session when the report complexity was 1-5, 6-10, 11-15, and 16-20. We refer to a subject’s guess of the average secret number when the report complexity was 16-20 as their “complex guess”, and we classify subjects as being “naive" if their complex guess is higher than the actual average secret number when complexity was 16-20 in their session. Across all 294 subjects, 12.6% are classified as naive. When naive, the average amount of naivete is 3.491, which is 98.9% above the actual average secret number in their session.28 Out of the 160 subjects who also completed the math test, 9.4% are classified as naive.

---

28 This is calculated as the difference between complex guess and the actual average divided by the actual average (when the complex guess is higher than the actual average).
The average answers are presented, along with the actual average secret number for each complexity level in Figure 2A. Given that just 12.6% of subjects are classified as naive, it is unsurprising that the average complex guess (2.510) is lower than both the actual average secret number for such reports (3.626) and the average guess in the game for such reports (4.191).

We also asked subjects to guess the average sender choice of complexity for each secret number, and the average responses are provided in Figure 2B. Assuming receivers use these stated beliefs as their prior beliefs at the beginning of each round and only use the observed complexity level (not the content of each report) to determine the value of the secret number, we can infer what they should have guessed via Bayes’ Rule. This value (referred to as the “inferred guess”) is on average 2.546 for high complexity (c≥15), which is also lower than the average actual guess in the game for such reports (4.222).

After these strategic beliefs were elicited, 160 subjects were asked to complete a math test that consisted of four questions. Each question required them to sum 20 numbers in a table similar to the most complex table in our game, and these questions were incentivized for correct answers. 54 subjects got all four questions correct (33.75%), 48 get one wrong (30%), 27 get two wrong (16.88%), and the remaining get either three wrong (10.62%) or all wrong (8.75%).

For the 160 subjects that completed the math test, we classify 33.8% as overconfident because they think they answered more questions correctly than they actually did. When overconfident, the average amount of overconfidence is 1.53. For the overconfident subjects who answered at least one correctly, this is 76.43% over their actual math test performance on average.29

Out of all subjects who completed the math test, 41.88% believe they got all four questions correct and 72.5% believe they got three or four correct. Both of these rates are higher than the actual fraction of subjects who got this many correct (33.75% and 63.75% respectively). When asked to predict the average number of questions that other subjects

29 This is calculated as the difference between their guess of number answered correctly and the actual number they answered correctly divided by the actual number they answered correctly (when their guess of number answered correctly is larger than the actual number answered correctly).
answered correctly, the average prediction was 2.694, which is close to actual average (also 2.694).

3.2.2 Regressions of Receiver Mistakes on Beliefs

Based on regressions of receiver mistakes onto the extent of naivete and overconfidence (presented in Table 3), we find evidence that when receivers are more naive about complexity and when they are more overconfident about their math performance, receivers over-guess by a larger amount. The relationship between these measures and the size of mistakes is also statistically significant, negative, and of a similar magnitude.

These regressions strongly suggest that both overconfidence and naivete are related to receiver mistakes. However, they do not tell us how much of receiver mistakes are explained by these forces, in part because it does not account for boundary effects. To get a more precise answer, we develop and estimate a structural model which appears in the appendix.

4. Conclusion and Policy Implications

Our results highlight the incentives for firms to strategically complexify information disclosed to consumers, potentially harming consumers and undermining the effectiveness of disclosure. The patterns we observe have policy implications as well. For example, many obfuscation theories assume naivete in (a fraction of) consumers, hence consumer education that reduces naivete should alleviate the seller’s incentives to obfuscate. But sophistication does not save them from obfuscation if they are overconfident about their ability to comprehend complex reports. Policy tools that target such overconfidence can be different from education efforts that target consumer naivete. Our results also suggest that a mandate on simplicity can be as important as a mandate on truthful disclosure. More generally, this highlights the potential for regulation aimed at encouraging disclosure to be both simple and salient.
Another policy implication is seen in sender behavior. Surprisingly, round-by-round feedback does not reduce obfuscation. If anything, learning encourages senders to understand receiver mistakes in low states and exploit it via obfuscation.

A final policy implication is related to disclosure in general. Our results suggest that the unraveling prediction is fragile. Although immediate and repeated feedback can steer voluntary disclosure towards the predictions of unraveling, it fails once we change the setting a little away from simple, voluntary disclosure. How to harvest the benefits of the incentives produced by unraveling remains a challenge in the real world.
References


