Bayesian Persuasion in the Lab*

OSub Kwon†

August 30, 2020

Abstract

This experiment focuses on testing Bayesian persuasion (Kamenica and Gentzkow, 2011) through minimal design. We adopt an experimental design in which the Sender chooses a partition of the state space instead of an information structure. This experimental design makes Bayesian persuasion highly interpretable and, more importantly, eliminates the burden of Bayesian updating for the subjects. We find that 1) the Senders overall behavior is qualitatively optimal in the sense that they set the posterior probability of the weaker signal near zero, but 2) they quantitatively do not best respond to the Receivers in the sense that the stronger signals are systematically lower than what the Receivers require, resulting in a persistently high rejection rate of the stronger signal. Moreover, the uncertainty about the requirement of the Receivers is the key impeding factor for the Senders to persuade in that 3) once we replace the Receivers with a robot that plays a known strategy, most Senders learn to play the optimal strategy. This suggests that Bayesian Persuasion, a supposedly difficult problem, is easy to learn once all other sources of difficulties that are not essential to the key strategic element are lifted, although strategic uncertainty from the Receiver side can impede learning.

1 Introduction

In both the political and economic realms, there are numerous cases in which one party attempts to sway another party into taking an action by providing information. Among these is information design, or Bayesian Persuasion (Kamenica and Gentzkow, 2011). In this setting, the Sender does not start with private information and, therefore, does not choose

* I thank Yaron Azrieli, James Peck, John Rehbeck, and Huanxing Yang for their invaluable advice. I am particularly grateful to John Kagel and Dan Levin for their support in all stages of this research. All errors are my own.

†Department of Economics, The Ohio State University. Email: kwon.406@osu.edu

how to communicate it as in Cheap Talk (Crawford and Sobel, 1982). Instead, he chooses what information is available to the Receiver, and the Receiver sees the realization without any distortion[1].

It is now widely recognized in theory that, given this freedom of choice, the Sender can choose a particular correlation between the signals and the states to increase the chance of inducing the desired action. However, in the field, the choice of correlation is obfuscated by too many other sources of difficulties that the theory (at least in its basic version) does not discuss. Thus, it is difficult for one to tell whether, as the theory predicts, a Sender can manipulate the correlation outside a lab environment.

For example, Kamenica and Gentzkow (2011) argue that, when a prosecutor initiates an investigation to persuade a judge to convict a suspect, "we can think of the choice of the investigation as consisting of the decisions on whom to subpoena, what forensic tests to conduct, what questions to ask an expert witness, etc." This example raises several issues. Often, as the number of tests grows, the signal space grows exponentially, which raises the issue of whether the prosecutors can carry out the contingent reasoning[2] needed to identify all possible combinations of the evidence. Even if the prosecutor can do this, he will worry how the judge, who may not properly do Bayesian updating[3] or have a different prior[4], will interpret the evidence. Even if he knows how the judge will interpret each combination of evidence, the judge's payoff can still be unknown to the prosecutor. In other words, it is not certain how the judge's beliefs will map onto her actions. Thus, we can see that the prosecutor has substantial freedom to design the investigation, but we do not know, cognitively speaking, how he could leverage this freedom to choose the right correlation between the signals and the states.

In this experimental study, we take a minimalist approach and test the bare bones Bayesian Persuasion to see, absent all other sources of difficulties, to what extent the key strategic element of the theory is alive. To minimize all other sources of difficulties that are not essential to the key strategic element (including those mentioned above), we design an experiment in which each Sender chooses a partition of the augmented state space, which Green and Stokey show is (1978) an equivalent way to represent information to information

[1]Such phenomena are more prevalent than one might think. For example, suppose a Sender and a Receiver sit in front of the same television, and the Sender holds the remote control. This situation can be thought of as Bayesian Persuasion because the Sender can choose which channel to watch and thereby choose what information is generated, but the Receiver sees exactly what the Sender sees. Thus, many situations have some element of Bayesian Persuasion, although they may not satisfy all the assumptions of the basic version of the theory.

[2]See Esponda and Vespa (2019) and Martínez-Marquina A, Niederle M, Vespa (2019).

[3]See Charness and Levin (2005) and the reference therein.

[4]Alonso and Câmara (2016) consider the case of different priors between the Sender and the Receiver. Kosterina (2019) considers the case in which the Sender does not know the belief of the Receiver.

structure[5]. The key feature of this design is that neither Sender nor Receiver needs to use the Bayes formula to update their belief given a signal. This isolates the key strategic element of Bayesian Persuasion and massively reduces the complexity of the problem without distorting the meaning of Bayesian Persuasion or restricting the Sender's choice of information.

To understand how choosing a partition works, consider a college (Sender) that wants to induce an employer (Receiver) to hire one of its graduates (as in, e.g., Boleslavsky and Cotton (2015)). This graduate comes from a college pool of 400 good students and 600 bad students in the college[6]. The employer is willing to hire a graduate if and only if the graduate has a grade that indicates that she is more likely to be a good student than a bad student. To maximize the probability that a particular student is hired by the employer, the college designs a grading policy that essentially partitions the student pool into different subgroups characterized by grades (e.g., those who get Grade A and those who get Grade B).

As it turns out, the college can get the employer to hire more students than the actual number of good students. The trick is twofold. First, by categorizing many good students as Grade A, the college can make Grade A a stronger signal (that the student is a good one) than Grade B. Because the weaker signal (Grade B), which does not induce employment, has no benefit to the college, it must give all the good students Grade A to ensure that all the good students are employed. This makes Grade B an unambiguous signal that the student is a bad one. Second, because the strength of the stronger signal (Grade A) has no use beyond inducing employment, and because the objective is to maximize the probability of employment, the college must also give some bad students Grade A. This increases the frequency of Grade A and depresses the strength of it as a signal, which does not matter as long as it still induces employment. Thus, in this example, the optimal thing for the college to do is to give Grade A to all 400 good students and 399 bad students (and Grade B to 201 bad students). This way, Grade A induces a posterior slightly higher than 50% (400/799) and the college can get the employer to hire 79.9% of its students.

Our experimental design directly reflects this example, except that good and bad students are replaced by red and blue balls, and the two grades are replaced by two urns. As in the example, the optimal strategy of the Sender always entail two features: 1) he should set the weaker signal to be fully revealing and 2) he should set the stronger signal just above the cutoff at which the Receiver is willing to take the Sender's desired action. This logic applies

[5]Let Θ be the payoff relevant state space with a prior distribution defined on it. We call $(S, \{\pi_\theta\}_{\theta\in\Theta})$ an information structure if each π_θ is a probability distribution of signals given the state θ. We call ($\Theta \times [0, 1]$, μ) an augmented state space if μ is the probability measure on $\Theta \times [0, 1]$ whose marginal distribution agrees with the prior distribution on Θ. Green and Stokey (1978) show that any information structure can be replicated by a partition on this augmented state space and vice versa.

[6]We switch to from a Prosecutor-Judge example to a College-Employer example because it is more difficult (although not impossible) to conceptualize a pool of 400 guilty suspects and 600 innocent suspects.

regardless of where this cutoff is, which we will influence by changing the Receiver's payoff.

We find that, after several rounds, the strategies of Senders are generally qualitatively optimal in the sense that they set the weaker signal to be nearly fully revealing. However, their strategies are quantitatively suboptimal in the sense that the stronger signals are systematically and persistently weaker than what Receivers require, resulting in a consistently high rejection rate of the stronger signal. This comes as a surprise because undershooting (setting the posterior weaker than what is required) is a more costly mistake than overshooting. This reflects a lack of understanding by Senders of the strength of the signals Receivers require.

This lack of understanding is the only factor that impedes learning the optimal information design. When Receivers are replaced with a robot that plays a known cutoff strategy, most Senders learn to choose the optimal strategy. In fact, in terms of setting the weaker posterior to zero, the Senders who face a robot Receiver learn faster than the Senders who face human Receivers. Thus, contrary to what one may initially think, the key strategic element of the Bayesian persuasion problem is only moderately difficult and it can be learned by most of the human subjects in our setting, although the strategic uncertainty on the Receiver side can impede learning.

Our study is a part of a conversation with the fast growing theoretical literature that relates Bayesian Persuasion (also known as information design) to real-world applications. Examples in the literature include price discrimination ((Bergemann et al. 2013), financing (Szydlowski 2016), and voting (Alonso and Câmara 2016), etc.

Although Bayesian persuasion is ubiquitous and important, and the theoretical investigation has been very successful, it is, with only few exceptions, understudied in the lab.

Nguyen (2017) shows that when they are given a limited choice of information structures and much repetition, Senders mostly choose the information structure predicted by the theory. However, because the choices are restricted, it is difficult to determine the extent to which the Senders grasp the strategic element fully and are not forced by the restriction of choices.

Frechette et al. (2019) experimentally examine an ambitious model that bridges a wide range of communication models, including Bayesian Persuasion, Cheap Talk (Crawford and Sobel, 1982), and the disclosure of verifiable message (e.g., Grossman, 1981; Milgrom 1981;). Because of how Frechette et al. bridge different types of models, they must literally let the Sender choose an information structure (i.e., our design cannot replicate their treatments). They find that the behavior is generally in line with the comparative statics predicted by the theory. However, in their baseline treatment (Bayesian Persuasion), they find a very large and persistent heterogeneity of Sender behavior: some set the posterior of the stronger signal higher than 0.5, near 0.5, and below 0.5. Although our experiment is less ambitious in scope, we obtain a much sharper result. For example, our Receivers rarely take the Sender's desired action when the posterior is lower than 0.5; our Senders rarely set the stronger posterior below 0.5 consistently; and almost all Senders choose the optimal strategy against a robot Receiver. Thus, at least some of the heterogeneity that Frechette et al. find can be attributed to the complexity of choosing an information structure and using the Bayes formula to calculate the posterior probabilities; neither of these exist in our experiment.

Au and Li (2018) use an experiment similar to ours to examine a variation of the Bayesian Persuasion model that incorporates the reciprocity concern; it focuses mainly on the Receiver. Their main treatment variable is the prior probability. They find that, when the prior is high, the Receiver is more likely to go against the Sender's preferred action given the same posterior. They explain on the basis of negative reciprocity: the same just-over-the-threshold posterior that starts from a high prior is regarded by Receivers as ill-intentioned unlike a posterior that starts from a low prior. In contrast, our experiment focuses on examining the Sender's behavior through a minimal experiment, and so we use different treatment variables. Due to some subtle differences, our experiment also is more in line with the partition interpretation of information, which we discuss in Section 2.3.

We believe Bayesian Persuasion is understudied in the lab precisely because it is a such a difficult problem; indeed, it is difficult to implement in a way that is interpretable to the human subjects. Consequently, we are unsure whether we can obtain meaningful results from a lab experiment. We make two contributions to the literature. First, we show that, by adopting the partition interpretation of information, which is well understood in the theoretical literature, Bayesian Persuasion can be massively simplified to a point where most people can understand it. Second, we provide a simple and portable design that can be readily extended by using our knowledge about this partition interpretation.

Because it is a laboratory test of a communication model, our experiment is related to a large experimental literature on cheap talk[7] and (veriable) information disclosure[8]. What distinguishes our experiment from these experiments is that the Receiver should take the "message" from the Sender at face value because the Sender can neither distort (as in cheap talk) nor hide (as in information disclosure) the realization of the signal, even if this is against the interest of the Sender.

In our experiment, the Sender essentially chooses a menu of lotteries for the Receivers. To do this optimally, the Sender needs to correctly predict the Receiver's risk attitude. Thus,

[7]See, for example, Dickhaut et al. (1995) , Forsythe et al. (1999); see Blume et al. (2017) for a survey on this topic.

[8]See, for example, Jin et al. (2017), Hagenbach (2018).

our experiment is related to studies that examine how one predicts other people's choices under risk[9]. Some contributions to this literature find that people predict that others are more risk seeking than they actually are. These studies are consistent with our findings that the Sender set the stronger posterior as if he thinks the Receiver is more risk seeking than she actually is. However, because in that line of literature the evidence is mixed (with the exception of the common finding that people are not good at predicting other peoples risk attitudes), and because our setting is different, it is not clear how the findings of these studies can be applied here.

The rest of the paper is organized as follows. In Section 2, we outline the basic structure of the experimental design, its interpretation, and our theoretical predictions. In Section 3, we report our main results. In Section 4, we briefly discuss the results and how the design can be extended.

2 Experimental Design

2.1 The Basic Structure

The basic structure of the experiment is summarized as follows. In each round of the game, a Sender (Role 1) is randomly matched with a Receiver (Role 2). The roles are fixed throughout the experiment. The task that the two parties face concerns the color of a randomly chosen ball. The Sender's goal is to induce the Receiver to guess Red while the Receiver's goal is to guess correctly.

The experiment consists of three steps. Figure 1 is an illustration of how the experiment may unfold.

Step 1: A ball is randomly chosen among 10 red balls and 15 blue balls. Neither the Sender nor the Receiver knows the color of this chosen ball.

Step 2: The Sender chooses how to divide the 25 balls into two urns, Urn A and Urn B[10]. Note that, at this point, either Urn A or Urn B must contain the chosen ball in Step 1.

Step 3: After the Sender's decision, Urn A or Urn B, whichever contains the chosen ball, is assigned to the Receiver. The Receiver knows how many red balls and blue balls are in the urn she is assigned. Then, she guesses the color of the chosen ball.

Payoff: The Sender gets a payoff of \$10 if the Receiver guesses Red and a payoff of \$0 otherwise[11]. The Receiver's payoff depends on a treatment variable q, and it is summarized

[9]See, for example, Hsee and Weber (1997, 1999) , Siegrist et al. (2002), Eckel and Grossman (2008)

[10]In practice, he chooses how many red balls and blue balls to be put into Urn A. Then, all the rest of the balls are automatically put into Urn B.

[11]In practice, we change these numbers across treatments to balance the Senders' payoff in different

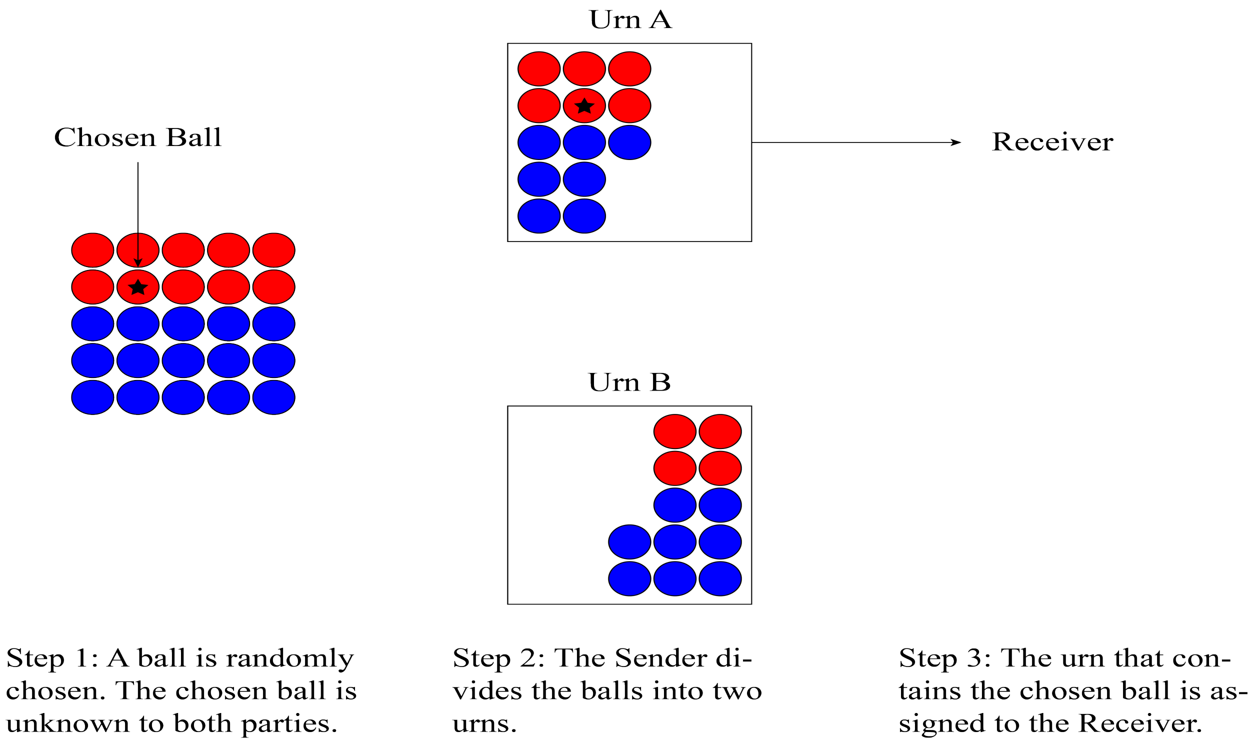


Figure 1: Basic Structure of the Experiment

as follows:

$$\begin{cases} \$10 & \text{if the guess is correct} \\ \$10(1-q) & \text{if the guess is Red but the color is Blue} \\ \$10q & \text{if the guess is Blue but the color is Red} \end{cases}$$

That is, whenever the Receiver guesses correctly, the Receiver gets \$10. But when the Receiver guesses incorrectly, the cost of an incorrect Red guess can be different from the cost of an incorrect Blue guess. The higher the q, the costlier it is to guess Red when the color of the chosen ball is blue. It is easy to verify that, under this payoff scheme, q is the the threshold belief at which a risk-neutral Receiver is indifferent between guessing Red and guessing Blue.

2.2 Theory

2.2.1 Theoretical Interpretation

In the Kamenica and Gentzkow (2011) setting, the Sender chooses an information structure defined on the payoff relevant state space. Here in our experiment, the Sender chooses a partition of the state space that has both the payoff relevant aspect (i.e., the color of the

treatments. However, the difference between the high payoff and the low payoff is always \$10.

drawn ball) and the payoff irrelevant aspect (i.e., the identity of the drawn ball). Green and Stockey (1978) show that these are two equivalent notions of information. Thus, the Sender in our experiment indeed designs the information for the Receiver, making our experiment a Bayesian Persuasion experiment.

We use our experiment design for two reasons. First, this partition representation of information makes it particularly obvious that the Sender designs how the Receiver narrows down the possibilities. This enables the subjects to interpret the problem more easily.

Second, the partition interpretation of information allows subjects to directly choose a distribution of posteriors[12]. To see this, simply note that the number of balls put in each new urn determines the probability that it contains the initially drawn ball and, thus, that it is given to the Receiver, while the fraction of the red balls in an urn determines the posterior belief that the drawn ball is red. Thus, a partition is directly related to a distribution of posteriors. Because the whole procedure directly reflects Bayesian updating, the Bayes plausibility constraint (i.e., that the expected posterior should be equal to the prior) is automatically satisfied.

2.2.2 Theoretical Predictions

In this part we describe what the Sender should do to maximize payoff. For ease of exposition, we introduce a definition. Because the overall fraction of the red balls is equal to 40%, if one urn has a fraction higher than 40%, the other must have a fraction lower than 40%. Thus, only one urn can reasonably induce the Receiver to guess Red. Throughout, we will call that urn *the stronger urn* and the corresponding posterior *the stronger posterior*.

Table 1: Theoretical Point Predictions

q	Component of Stronger Urn	Prob. of Red Guess
0.5	(Red: 10, Blue: 9)	76%
0.6	(Red: 10, Blue: 6)	64%
0.8	(Red: 10, Blue: 2)	48%

Suppose the Receiver plays a cutoff strategy, that is, she chooses to guess Red if and only if the posterior belief is larger than a certain cutoff. Then the Sender's optimal strategy has

[12] Au and Li (2018) also let Senders choose a partition (with five elements) of some balls. However, in their experiment, an urn is randomly selected and then a ball is drawn from the selected urn in a way mathematically equivalent to ours. We adopt our design for two reasons. Conceptually, our design is in line with the "deterministic view" (the state comes first and then the signal follows) of information often used in the literature. Behaviorally, our design makes it easy to interpret what the Sender does and is easy for us to convince the subjects that the Sender has no control over the overall probability that the drawn ball is red (because the ball realizes before the Sender moves).

two features.

First, he should put all the red balls in the stronger urn. To see this, suppose that he does not; then he can move a red ball from the weaker urn to the stronger urn. This makes the stronger urn more frequent and even stronger, which weakly increases the probability of a red guess.

Prediction 1: The Sender should always put all the red balls in the stronger urn. That is, the weaker posterior should be equal to 0.

Second, the Sender should put some blue balls in the stronger urn so that the stronger urn becomes more frequent, but still induces a posterior higher than the cutoff. The optimal persuasion is achieved when the stronger posterior is just above the cutoff.

Prediction 2: The Sender should put some blue balls in the stronger urn in addition to all the red balls so that the stronger urn just induces the red guess.

Note that there is a large jump in the payoff around the cutoff: when the stronger posterior is slightly above the cutoff, the Sender induces the red guess with the highest probability; but when the stronger posterior is slightly below the cutoff, the Sender induces the red guess with 0 probability. Thus, the Sender should always ensure that the stronger urn induces a red guess. To the extreme, the Sender can always choose to completely separate the red balls and the blue balls to ensure a 40% probability of inducing the red guess.

Prediction 3: The stronger urn should always induce a red guess. The probability of the red guess should always be higher than 40%.

Table 1 gives the point predictions of the theory if the Receiver is risky neutral. Recall that in our experiment, the treatment variable q coincides with the risky neutral cutoff belief of the Receiver. Note that, if the Receiver is risk averse, the cutoff is generally not equal to q. Specifically, when $q > 0.5$, a risk-averse Receiver is more reluctant to make the costlier mistake and, thus, a higher fraction of the red balls is required to guess Red. Therefore, when $q > 0.5$, the threshold belief of a risk-averse Receiver is higher than q. However, $q = 0.5$ is a special case where the requirement that the Receiver is risk neutral is not needed: even when she is risk averse, when $q = 0.5$, the costs of the two different types of mistakes are the same. Thus, she must always guess the color that has the higher fraction.

2.3 Treatments

Table 2 summarizes the treatments of the experiment. In the main treatments (Treatment 1, 2a, and 3), the number of balls is fixed at 25 and q is the treatment variable. In each treatment, the subjects face two q in a session. For example, in the treatment with q = (0.5, 0.8), we let the subjects play the above game with $q = 0.5$ for 12 rounds and then switch to

$q = 0.8$ for another 12 rounds.

Table 2: Treatments

	q	No. of Balls	Receiver
Treatment 1	(0.5, 0.8)	(Red: 10, Blue: 15)	Human
Treatment 2a	(0.6, 0.8)	(Red: 10, Blue: 15)	Human
Treatment 2b	(0.6, 0.8)	(Red: 400, Blue: 600)	Human
Treatment 3	(0.8, 0.6)	(Red: 10, Blue: 15)	Human
Treatment 4	(0.6, 0.8)	(Red: 400, Blue: 600)	Robot

We recognize that the case of $q = 0.5$ is important and special in that the best response of the Receiver is independent of her risk preference. Thus, we include a similar case of $q = 0.6$ to see whether there is any qualitative difference compared to the case of $q = 0.5$. Adopting a within-subjects design (changing q halfway) allows us to determine whether the Sender has a robust understanding of the problem and how responsive Senders are to the change in the incentive of the Receivers. To determine whether the main results are driven by the order effect we include Treatment 3 as a robustness check.

A small number of balls arguably keeps things simple, but it leads to heavy discretization which may affect the results[13]. Thus, as another robustness check, we replicate Treatment 2a with 1000 balls. It turns out that enlarging the strategy space has little effect on behavior. Consequently, in the main analysis we pool the data from Treatment 2a and Treatment 2b and report the comparison between the two treatments in Appendix A.

Finally, in Treatment 4, we replace human Receivers with a robot Receiver. Unlike a human Receiver, a robot Receiver plays a cutoff strategy known to the Sender: the robot guesses Red if and only if the fraction of the red balls in the urn it receives is higher than q. This robot treatment serves as an important benchmark for understanding, with no strategic uncertainty from the Receiver side, the extent to which the Senders can design information optimally.

2.4 Procedure

Subjects were recruited through emails and were predominantly undergraduate students at the Ohio State University. Each session included 10-20 subjects. At the beginning of the session, each subject was given a written copy of the instructions, which were also read

[13]Note that even when there are 25 balls, the Sender still has 88 ($11 \times 16 \div 2$) different partitions to choose from; thus the strategy space is not so small. A potential problem, however, is that the number of reasonable strategies can be small when $q = 0.8$, which may force the Sender to completely separate the red balls and the blue balls.

aloud by the experimenter. After a round of clarification, each subject is randomly assigned a role (the Sender or the Receiver), which is fixed throughout the session. Then, the subjects played two practice rounds. In each practice round, they experienced each of the two q's within a session. After the practice rounds, the subjects played 12 rounds for each q. Each session lasted approximately 90 minutes. For every 12 rounds, we randomly picked one round for the payment in addition to a $5 participation fee. This resulted in an average payment of about $20 for each subject.

It is worth mentioning that we took three extra steps to simplify the tasks of the subjects. First, we calculated the relevant probabilities for both Senders and Receivers[14]. Especially for the Senders, we updated all the probabilities in real time as they decide how to divide the balls. Second, as Senders choose how to divide the balls, we updated in real time for each urn the choice of lotteries the Receivers would face given their ball divisions (the human treatment) or the Receiver's response (the robot treatment). Third, we elicited the Receiver's responses for both urns through the strategy method and reported them at the end of a round so that Senders can learn.

3 Results

3.1 Receiver's Behavior

We start with a discussion of the behavior of the Receivers because their choices are simple in the sense that they only have to choose between two lotteries and because the regularity of the Receiver behavior is important for us to evaluate the performance of the Senders.

Figure 1 depicts the aggregate response of the Receivers to different posteriors. We see that, for each q, the aggregate probability of the red guess increases continuously in the posterior belief[15], often with a sharp increase around q. We also see that when the posterior is lower than 0.5, almost no Receiver guesses Red regardless of q —a reasonable response. Thus, the only interesting region of posteriors is $[0.5, 1]$.

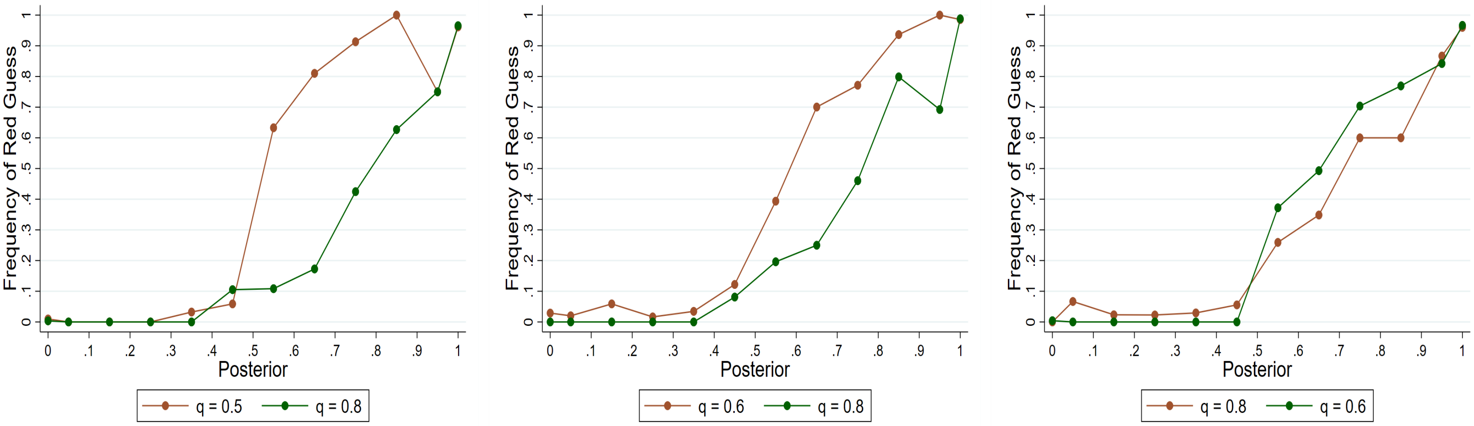
In general, Receivers are responsive to the change in q. To show this, we partition the posterior region $[0.5, 1]$ into three subregions ($[0.5, 0.6)$, $[0.6, 0.8)$, and $[0.8, 1]$) and determine whether the average frequency of red guesses in each subregion differs across q[16]. Table 2

[14]This can be done naturally because all the probabilities can be calculated by simple divisions without the Bayes formula.

[15]In Figure 1, there is some non-monotonicity within the interval $[0.9, 1]$. We attribute this non-monotonicity to the lack of data because, as is noted in Section 3.3, the posteriors within this region are surprisingly unpopular choices among the Senders.

[16]Across treatments, overall we see almost no statistically significant difference (the Wilcoxon ranksum test, clustered at the Receiver level) in the Receiver response in each subregion under each q. The only

Figure 2: Receiver Aggregate Response



(a) Treatment 1 (b) Treatment 2(a and b) (c) Treatment 3

shows the frequency of the red guess in each subregion under each q. We see that q generally is a good predictor of how the Receivers will react: when the posterior is lower than q, most of the time Receivers guess Blue, while when the posterior is higher than q, most of the time Receivers guess Red. Moreover, as we move to the higher subregion, the difference in the Receiver's response becomes smaller. In particular, we see no statistically significant difference between $q = 0.5$ and $q = 0.6$ in the subregion $[0.8, 1]$.

Table 3: Frequency of Red Guess in Each Subregion

	$[0.5, 0.6)$	$[0.6, 0.8)$	$[0.8, 1]$
$q = 0.5$	65.43%	85.35%	96.67%
$q = 0.6$	36.12% ($-29.31\%^{***}$)	70.24% ($-15.11\%^{*}$)	91.60% (-5.07%)
$q = 0.8$	16.89% ($-19.23\%^{***}$)	34.59% ($-35.65\%^{***}$)	82.85% ($-8.75\%^{**}$)

Difference with the previous row reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (the Wilcoxon ranksum test, clustered at the Receiver level).

Note also that in the case of $q = 0.5$, not all Receivers guess Red under a posterior slightly higher than 50%, although doing so is optimal regardless of risk attitude. This is clearly not due to random error: when the posterior is lower than 50% they almost always guess Blue. This irregular behavior of the Receivers is not particular to our experiment. Frechette et al. (2019) find that some Receivers guess Blue unconditionally even when (in our terms) $q = 0.5$; Au and Li (2018) find that Receivers do not always guess Red under a just-over-50% posterior. In the end-of-experiment survey, some subjects reported that they

exception is between Treatment 2 and 3, in $[0.8, 1]$ under $q == 0.6$.

are more inclined to guess Blue simply because the prior probability that the drawn ball is Red is lower than 50%.

In summary, although the aggregate behavior of the Receivers is not completely in line with a cutoff strategy, it is monotone and responsive to the change in q.

At the individual level, does a Receiver play a cutoff strategy? To answer this question, for each Receiver and q, we look for the highest posterior under which she guessed Blue[17], and then we count how many times she guessed Red under a lower posterior. This gives us a count of the violations of a cutoff strategy from above.

Similarly, we look for the lowest posterior under which she guessed Red and count how many times she guessed Blue from below. In this way, we obtain two measures of violations that will coincide with each other if the Receiver has no violations.

The two measures are very demanding. For example, if once by error a Receiver guesses Red under the weaker signal, she is likely to have many violations from below.

Table 4: No. of Receivers Compatible with Cutoff Strategies

	q	0 Violation	$\leq$ 2 Violations (Above)	$\leq$ 2 Violations (Below)
Treatment 1	0.5	40%	71.43%	68.57%
Treatment 1	0.8	48.57%	80%	80%
Treatment 2(a and b)	0.6	38.80%	74.63%	64.18%
Treatment 2(a and b)	0.8	67.16%	86.57%	86.57%
Treatment 3	0.8	52%	72%	76%
Treatment 3	0.6	56%	72%	76%

Table 3 shows that for each q, most Receivers have less than 2 violations both from above and below. In fact, 26.77% (41.73%) of the Receivers pass the even more demanding requirement to have no (less than 2) violation(s) throughout the whole session. Overall, each individual Receiver's behavior is consistent with a cutoff strategy.

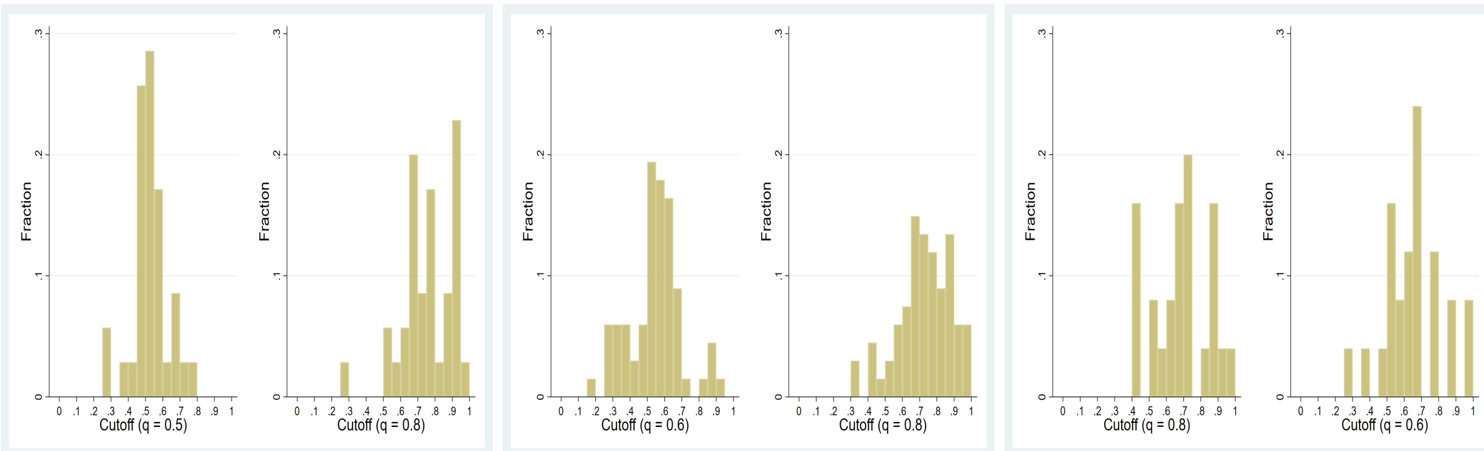
Figure 2 shows the distribution of the estimated cutoffs, which is defined as the average of the the lowest posterior with a Red guess and the highest posterior with a Blue guess[18]. Although Receivers are mostly compatible with cutoff strategies, they do not use the same cutoff; thus explaining why the aggregate response is a continuously increasing function rather than a step function. In particular, when $q = 0.5$, most Receivers set their cutoffs around 0.5, but 34.29% of them still have estimated cutoffs higher than 0.55. This explains why the aggregate response is very steep around the region (0.4, 0.5) but does not go all the way to 100% guessing red when the posterior is slightly higher than 0.5. However, we also

[17]If a Receiver empirically only guessed Blue, we take this posterior as 1.

[18]This measure can be imprecise if the two posteriors are very different. However, for 74.02% of the Receivers, the difference between the two posteriors is less than 0.2.

see that the higher the q, the higher the Receiver cutoff on average (the Wilcoxon ranksum test, $p < 0.001$ for all pairs of q's). Thus, although there is some heterogeneity in the cutoff, the distribution of the cutoffs is consistent with the basic comparative statics.

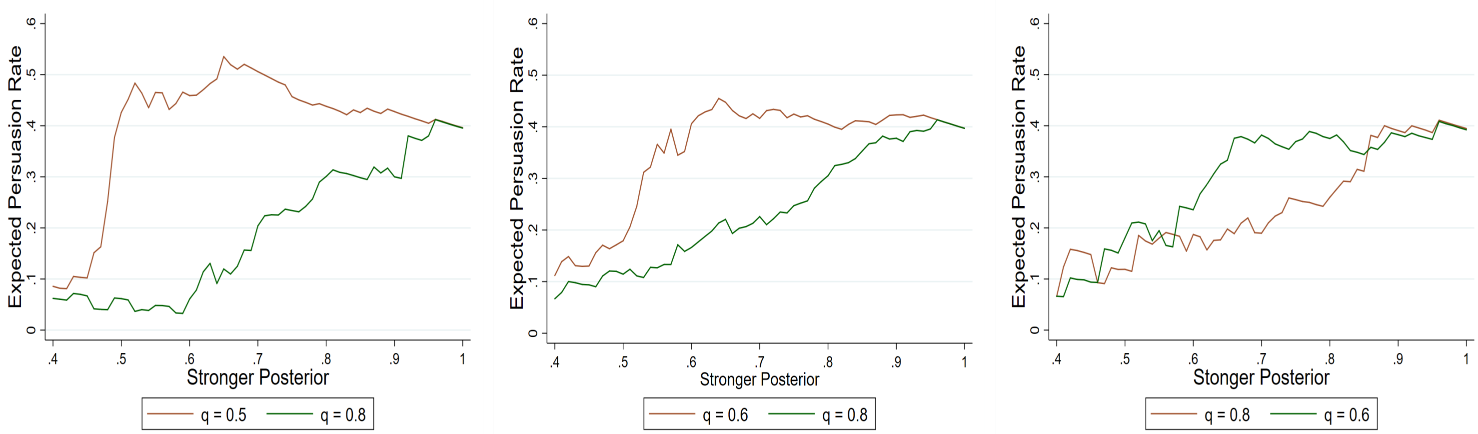
Figure 3: Estimated Cutoff Distribution by Treatment



(a) Treatment 1 (b) Treatment 2(a and b) (c) Treatment 3

The fact that overall Receivers play cutoff strategies and that they are responsive to q even at the individual level means that, as is argued in Section 2.4, the optimal strategy of the Senders is to put in one urn all the red balls plus, depending on how the Receivers react, some additional blue balls.

Figure 4: Empirical Best Response Analysis by Treatment



(a) Treatment 1 (b) Treatment 2(a and b) (c) Treatment 3

On the basis of this knowledge, we now turn to estimating the empirical best response of the Senders to the Receivers. For each q of each treatment, we estimate the expected persuasion rate[19]. that a Sender can achieve if 1) he puts all the red balls in the stronger urn and 2) he sets the posterior of the stronger urn to be p, where $p \in [0.4, 1]$. Note that the number of red balls put in the stronger urn and the stronger posterior together define the Sender's strategy.

Figure 3 shows the estimated expected persuasion rate as a function of the stronger posterior. We find that when $q = 0.5$ or 0.6, the persuasion rate function is steeper on the left than on the right. This means that a small amount of undershooting (of the stronger posterior) relative to q is more costly than the equal amount of overshooting (of the stronger posterior). This is reasonable because setting the stronger posterior lower than what a Receiver requires will never lead to persuasion. But setting the stronger posterior higher than what a Receiver requires will only lead to a lower frequency of the stronger signal and may even increase the overall persuasion rate, depending on how many more Receivers the signal is able to persuade. This asymmetry of response leads to a sharp kink in the expected persuasion rate function, which is yet another piece of evidence that overall the Receivers behave regularly.

Importantly, in all treatments, when $q = 0.8$, the Sender's best response is to set the stronger posterior to 1. Two factors contribute to this response. First, when q is high, it is very costly for a Receiver to make a wrong Red guess; consequently, many Receivers behave in a risk-averse manner. Second, the maximum frequency $0.4/p$ that can induce a given stronger posterior p is a convex function. That is, by putting one less blue ball in the stronger urn, the Sender can increase the stronger posterior by a larger amount as the posterior approaches 1. This makes completely separating the red balls and the blue balls an attractive choice. Any posterior lower than 100% is likely to result in a persuasion rate that is lower than 40% when $q = 0.8$.

3.2 Putting All Red Balls in One Urn

We now turn to one of the key theoretical predictions that the Sender will set the weaker posterior to be 0. In this experiment, this corresponds to putting all the red balls in the stronger urn.

From Figure 4 we see that, regardless of the treatment, the average proportion of red balls put in the stronger urn is high at the beginning (around 75%) and increases over time (ending up at around 90%). Surprisingly, there is no significant change in the average proportion in round 13, when q changes.

At the individual level, there is some heterogeneity in whether Senders learn to put all

[19] We assume that, if a Receiver gets a posterior of 1, she will guess Red. This assumption is approximately true. In our data, when the posterior is 1, the Receiver guesses Blue only 8/317 (2.52%) of the time.

Figure 5: Prop. of Red Balls in the Stronger Urn Over Time

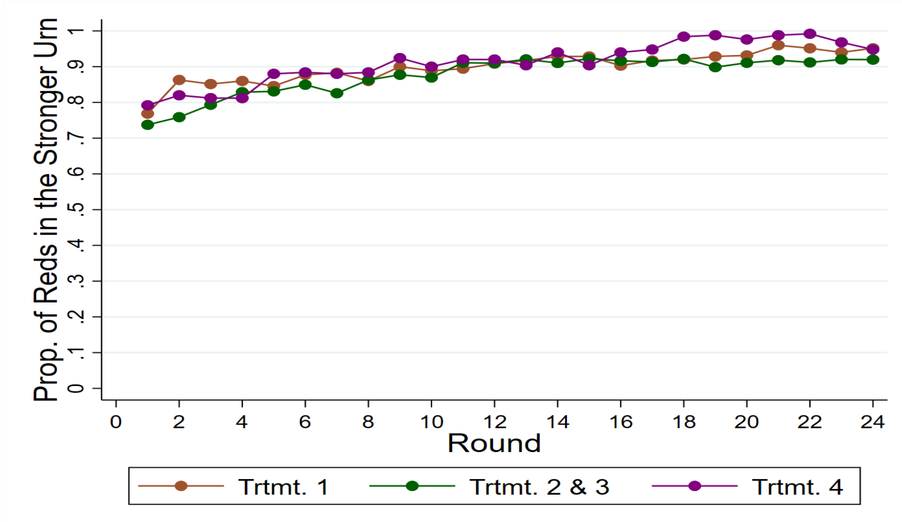


Table 5: No. of Senders Constantly Putting 90% (100%) of the Red Balls in the Stronger Urn

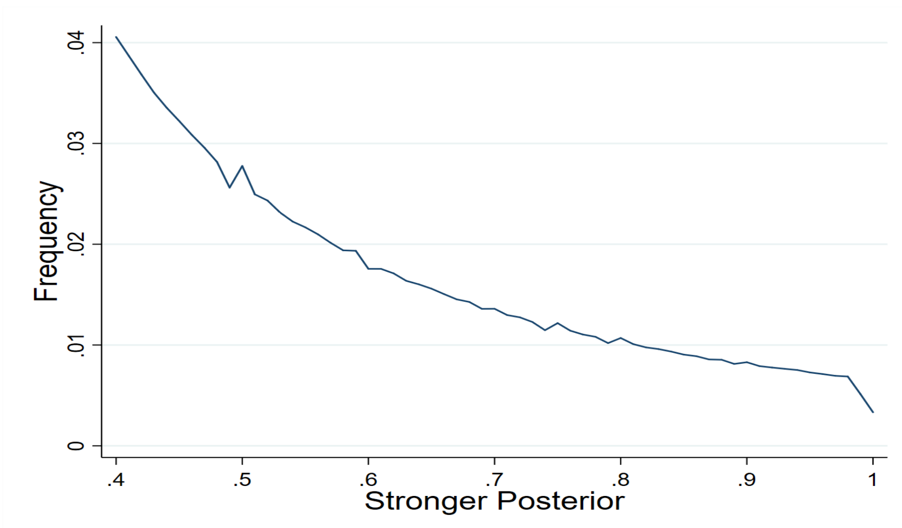
	From Round 1	From Round 7	From Round 13	From Round 19
Treatment 1	20% (20%)	45.71% (31.43%)	54.29% (37.14%)	74.29% (51.43%)
Treatment 2(a and b)	19.4% (13.43%)	40.3% (32.84%)	56.72% (46.27%)	70.15% (58.21%)
Treatment 3	20% (12%)	40% (36%)	60% (60%)	84% (76%)

the red balls in the stronger urn. However, once they learn to do so, many Senders choose to stick to this concentration until the end of the experiment. Table 3 shows the proportion of Senders who, starting from rounds 1, 7, 13, and 19 (i.e., the beginning of each quarter), persistently put at least 90% (or 100%) of the red balls in the stronger urn until the end of the experiment. About 20% of the Senders put at least 90% of the red balls in the stronger urn in the first round and continue to do so throughout the experiment. More than 70% (50%) of the Senders learn by round 19 to put 90% (100%) of the red balls in the stronger urn, and they stick with this strategy until the end. This suggests that for most Senders, concentrating the probability mass on one signal is a very normatively appealing rule, which is not difficult for them to learn within a session.

3.3 Setting the Right Stronger Posterior

We now turn to how the Senders set the strength of each signal. We focus on the posterior induced by the stronger urn (i.e., the stronger posterior) because the weaker urn almost always induces a Blue guess[20].

Figure 6: The Distribution of the Stronger Posterior from Uniform Randomization



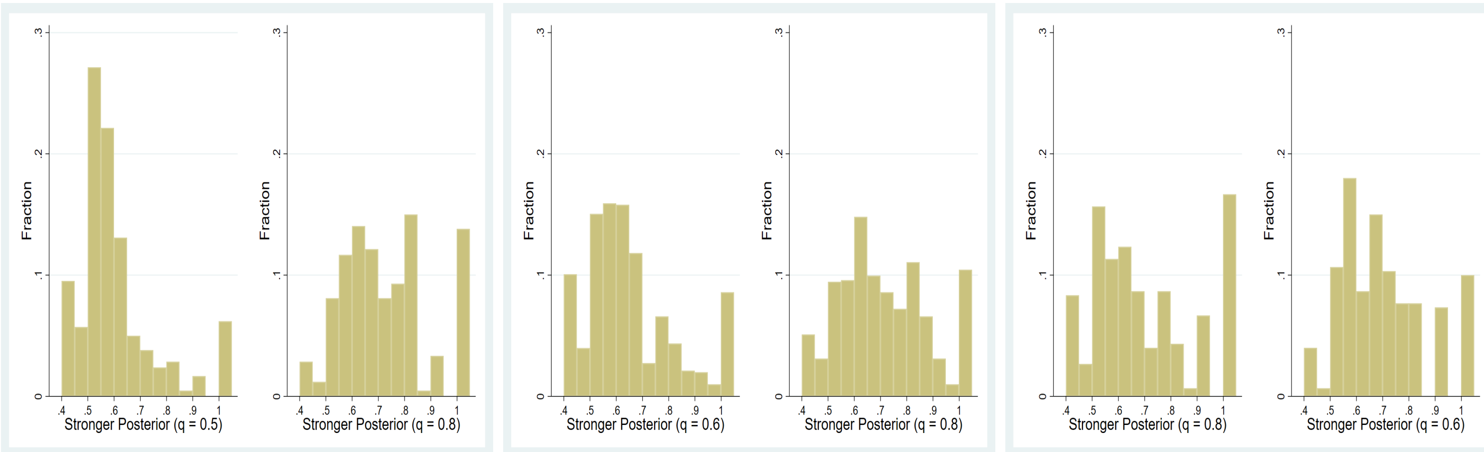
Since the strategy space is large, we set a benchmark by asking how a Sender who simply uniformly randomizes the number of red and blue balls put in one urn would set the stronger posterior. Figure 5 shows the (simulated) distribution of the stronger posterior generated by a uniformly randomizing Sender. As we can see, posteriors closer to the prior (40%) are more likely to be chosen by a uniformly randomizing Sender. Interestingly, with about a 33.12% probability, a uniformly randomizing Sender will set the stronger posterior to be lower than 50% —a choice that is very hard to reconcile with the regularity of the Receiver's behavior. Therefore, whether the stronger posterior is higher than 50% can serve as a basic but reasonably strong rationality test.

Figure 6 reports the empirical distribution of the stronger posterior by treatment and q. First, we find that, in contrast to Figure 5, Senders do not consistently set the stronger posterior to be lower than 50%. In our data, in the first half, the stronger posterior is set lower than 50% in only 13.78% of the games, and in the second half, this percentage falls to 6.36%. This means that the Senders in our experiment generally pass the basic rationality test in all treatments.

Second, we find that Senders are directionally responsive to the changes in q. On average, Senders set the stronger posterior higher under higher q than under lower q[21]. We find no significant difference in the stronger posteriors set under the same q in different treatments. Thus, in terms of direction, the Senders react rationally to the change in q.

[20]In our data, the weaker signal induces a Red guess in only 1.33% (41/3087) of the times.
[21]The Wilcoxon Ranksum test, $p < 0.01$ for all pairs of q, clustered at the Sender level.

Figure 7: Empirical Distribution of Stronger Posterior by Treatment and q



(a) Treatment 1 (b) Treatment 2(a and b) (c) Treatment 3

Third, we find that Senders generally set the stronger posterior larger than q for $q = 0.5$ but not for other q. When $q = 0.5$, the stronger posteriors are set lower than q for only 15.24% of the time, whereas when $q = 0.6$ and $q = 0.8$, this percentage is 41.85% and 68.44%, respectively. Because we see in Figure 1 that most of the Receivers guess Blue under a posterior lower than q and in Figure 3 that undershooting relative to q is a more costly mistake than overshooting, we can conclude that Senders are generally undershooting relative to the best response. In standard terms, this behavior can only be rationalized if the Senders believe that the Receivers are risk loving, which generally is not true.

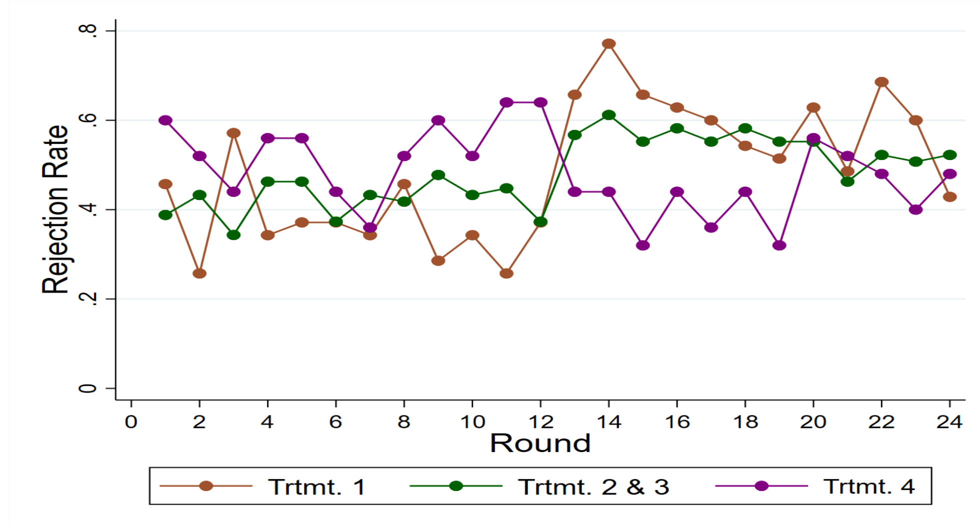
The overall undershooting leads to frequent rejection of the stronger signal (i.e., the Blue guess under the stronger signal). Figure 7 shows the rejection rate over time. Although there is a weak trend that the rejection rate falls, the rejection rate over time is persistently high. Moreover, in all treatments, the higher the q, the more frequently the stronger signals are rejected: when q is low ($q = 0.5$ or 0.6), the average rejection rate is between 35% and 45% (in all treatments), and when q is high ($q = 0.8$), the average rejection rate is between 50% and 60%. This is in sharp contrast to the equilibrium prediction that the stronger signal always induces a Red guess.

Given that undershooting is a more costly mistake than overshooting, why is there a high rejection rate of the stronger signal? We examine several conjectures that can be tested by our data[22].

On the Receiver side, the high rejection rate of the stronger posterior is not likely to be caused by the non-monotonicity of some Receiver's response. To show this, we compare the

[22]Note that any explanation related to the Bayesian updating is automatically ruled out because the subjects do not bear the burden of calculating the posteriors.

Figure 8: Rejection Rate Over Time



rejection rate of Receivers who have less than 2 violations of cutoff strategies throughout the whole session to all the other Receivers. We find that the rejection rate is high in both groups and that when q = 0.8, the group with fewer violations rejects more (61.56% vs. 52.51%).

Given that the non-monotonicity of the Receivers does not drive the result and that Senders often set the stronger posterior lower than q, we turn to the Sender side for explanations.

First, we see that the rejection of the stronger signal is not restricted to a particular subset of Senders. When $q = 0.5$ or $q = 0.6$, more than 60% of the Senders have a stronger posterior rejected at least 4 times (i.e., 1/3 of the time). When $q = 0.8$, the percentage is over 80%. This means that the high rejection rate of the stronger signal is a systematic pattern, and it is not driven solely by some Senders.

Second, we see that the rejection is still prevalent in those games in which the Sender chooses reasonable strategies, although slightly less frequently than overall. For example, when $q = 0.8$, and in those games where the Senders put 90% of the red balls in the stronger urn and set the stronger posterior to be larger than 0.5, the stronger urn is still rejected 50.43% of the time. This means that the undershooting and the high rejection rate of the stronger signals is not likely to be simply an artifact of Sender confusion about the task.

Third, it is unlikely that the undershooting is caused by the discreteness of the strategy space. Recall that in our main treatments, the number of balls is set to be 25 (10 red balls and 15 blue balls). Because the partitions of the 25 balls that generate a stronger posterior higher than 0.8 are rare, the discreteness of the strategy space may —from an ex ante point of view —be a cause of undershooting when $q = 0.8$. However, in Treatment 2b, in which the total number of balls is 1000, when $q = 0.8$, the rejection rate is still as high as 52.14%, which is not far from 57.55% in Treatment 2a.

Fourth, the high rejection rate in the second half of Treatment 2(a and b) (where $q = 0.8$) is not solely driven by behavioral inertia from the first half (where $q = 0.6$). We know this because in Treatment 3, where the order of q's is reversed, we still find a very high rejection rate when $q = 0.8$.

Fifth, the social preference of the Senders does not plausibly explain the undershooting of the Senders because, by overshooting, the Senders can generate more information for the Receivers in the sense of Blackwell (1953) (holding the number of red balls constant). Because overshooting also yields more payoff for the Senders than undershooting, it leads to a Pareto improvement compared to undershooting .

We conjecture that the rejection rate is high because the Senders suffer from a systematic prediction error. On one hand, the Senders understand that the posterior has no use beyond inducing the Receivers to guess Red (as is evidenced by the results in the next section), so they try to match the cutoffs of the Receivers. On the other hand, the Senders do not predict well how Receivers choose between the lotteries and they overestimate the degree to which Receivers are risk loving[23]. As a result, the Senders constantly set the stronger posteriors lower than what the Receivers require expecting that the Receivers will guess Red.

3.4 Robot Treatment

At this point, we see that most Senders play a qualitatively optimal strategy in the respect that they put most of the red balls in the stronger urn but do not set the stronger posterior strong enough to persuade Receivers. This raises a natural question: Suppose there is no uncertainty about what the Receivers require; can Senders arrive at the optimal persuasion strategy? In Treatment 4, we let Senders persuade a robot that plays a known cutoff strategy.

For this Robot treatment, we have a remarkably simple result: most people can learn to pin down the exact optimal strategy. On one hand, Figure 8 shows that Senders in the Robot Treatment learn faster than Senders in the Human treatment how to put all the red balls in one urn. Table 5 shows that Senders in the Robot treatment put all the red balls in one urn faster than do Senders in the Human Treatments.

On the other hand, Senders in the Robot Treatment perform very well at setting the stronger posterior just above what the Receiver requires. Figure 9 shows that when $q = 0.6$, the Senders set the stronger posterior within the interval $(0.6, 0.65]$ 66.67% of the times. When $q = 0.8$, they set the stronger posterior within the interval $(0.8, 0.85]$ 83.59% of the

[23]Hsee and Weber (1997) provides evidence that people overestimate how risk loving others are

Figure 9: Prop. of Red Balls in the Stronger Urn Over Time (with Robot Trtmt.)

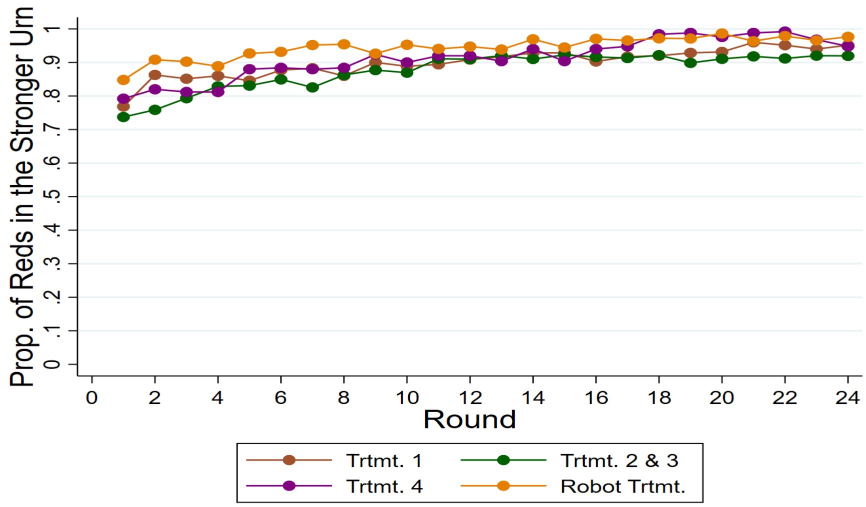


Table 6: No. of Senders Constantly Putting 90% (100%) of the Red Balls in the Stronger Urn

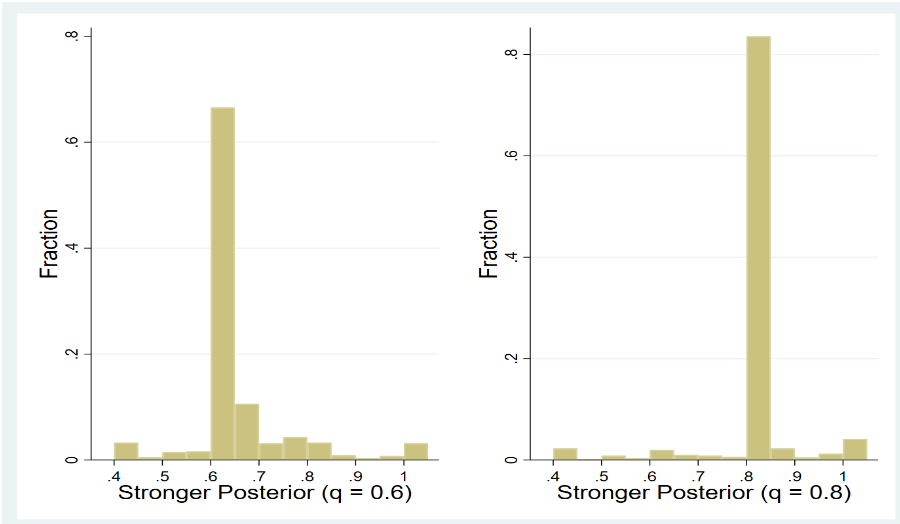
	From Round 1	From Round 7	From Round 13	From Round 19
Treatment 1	20% (20%)	45.71% (31.43%)	54.29% (37.14%)	74.29% (51.43%)
Treatment 2(a and b)	19.4% (13.43%)	40.3% (32.84%)	56.72% (46.27%)	70.15% (58.21%)
Treatment 3	20% (12%)	40% (36%)	60% (60%)	84% (76%)
Robot Treatment	45.45% (40.91%)	66.67% (56.06%)	72.73% (63.64%)	86.36% (78.79%)

times. Interestingly, Senders perform better when $q = 0.8$ in the Robot treatment, which is the opposite of what we find in the Human treatments. This is clearly due to learning.

We can also ask how many times Senders are able to combine the two aspects. We find that when $q = 0.6$, in 55.80% of the games Senders put 90% of the red balls in the stronger urn and set the stronger posterior so that it is 5% or less above q. When $q = 0.8$, this percentage is 80.81%. Thus, most of the Senders who know how to set the right posterior also know to put most of the red balls in the stronger urn.

In conclusion, when Senders are faced with a Robot Receiver that plays a known cutoff strategy, the optimal persuasion simply requires the Senders to put as many balls in one urn as possible under the constraint that the fraction of the red balls in that urn exceeds a certain cutoff. Most Senders easily realize how to do this.

Figure 10: The Empirical Distribution of Stronger Posterior in the Robot Treatment



4 Discussion

In summary, we find that the Senders can overall understand the normative appeal of setting the weaker posterior to 0, which makes their strategies qualitatively optimal. However, they set the stronger posterior lower than what the Receivers require, which leads to the frequent rejection of the stronger signal. The frequent rejection of the stronger signal is puzzling because the Sender can always avoid it by providing full information and get higher payoff empirically. Yet, it is a very robust phenomenon across treatments and subjects. But when the Senders know what posterior can persuade the Receivers, they often choose to exactly match that posterior and therefore arrive at the exact optimal strategy.

Our results suggest that, although Bayesian persuasion is only a recent discovery, its basic strategic element, when cleanly isolated, can be easily understood by our human subjects who presumably have no relevant knowledge. Our experimental design makes such clean isolation possible while keeping the problem interpretable. The key step is to reinterpret Bayesian persuasion in terms of information partition.

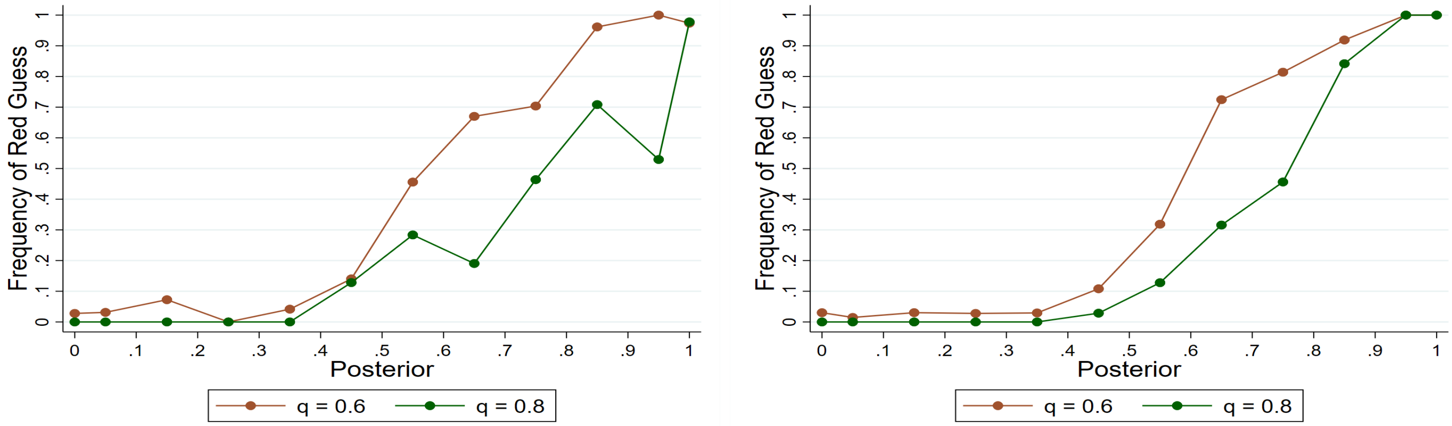
Our experimental design allows for a variety of extensions based on our knowledge about the information partition. For example, we know that we can combinetwo pieces of information by taking the meet[24] of their respective information partition. Consequently, one way to implement competitive Bayesian Persuasion (e.g., Gentzkow and Kamenica (2016)) between multiple Senders is to let each Sender choose a partition and then give the Receiver the meet

[24]Recall that the meet $\mathcal{C}$ of two partitions $\mathcal{A}$ and $\mathcal{B}$ is the coarsest partition that is finer than the two partitions. Formally $\mathcal{C} = \{A \cap B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\}$.

of the partitions. One caveat is that our design always give the Receiver the correct way to interpret the signals. Thus, we cannot replicate the treatments of Frechette et al. (2019) where the interpretation of the signals are subject to the Sender's manipulation.

Although our design is minimal and may exclude some interesting behavioral aspects, we believe that as long as the strategic element under study is orthogonal to those aspects, a minimal design like ours provides a useful starting point to understand the extent to which the strategic element is alive. Here, we find that the strategic element of Bayesian Persuasion is not only alive but very strong, and we expect it will also be strong in many other extensions.

Figure A1: Receiver Aggregate Response



(a) Treatment 2a

(b) Treatment 2b

Appendix A: Comparison of Treatment 2a and 2b

This section of the Appendix reports the comparison of Treatments 2a and 2b. Recall that in these two treatments, the only difference is the total number of balls. Here we argue that the qualitative findings are the same in Treatment 2a and Treatment 2b, and, therefore, that the total number of balls is unlikely to be an important factor in the experiment. To avoid redoing all the analysis, we focus on several key aspects: 1) how Receivers behave monotonically and are responsive to q; 2) whether Senders can learn to put all the red balls in one urn; 3) how Senders set the stronger posterior; and 4) whether there is a high and persistent rejection rate.

Figure A1 compares the aggregate response of the Receivers in Treatment 2a and Treatment 2b. In both treatments, the aggregate Receiver response is continuous and monotone. Moreover, we see that the two treatments are quantitatively similar in the aggregate response, except that $q = 0.8$ and the posterior is within $(0.9, 1)$ where we get few data. For each of the three posterior subregions ($[0.5, 0.6)$, $[0.6, 0.8)$, and $[0.8, 1]$), we find no statistically significant difference in the aggregate response between the two treatments (the Wilcoxon ranksum test clustered at the Receiver level).

Figure A2 reports the proportion of red balls in the stronger urn over time. We see that in both of the treatments, Senders learn at approximately the same speed to put all the red balls in one urn.

Figure A3 reports the empirical distribution of the stronger posterior. We see that in both treatments 1) the stronger posterior below 0.5 is not frequent, 2) the stronger posterior below q is very frequent when $q = 0.8$, and 3) the stronger posterior is generally set higher

Figure A2: Prop. of Red Balls in the Stronger Urn Over Time

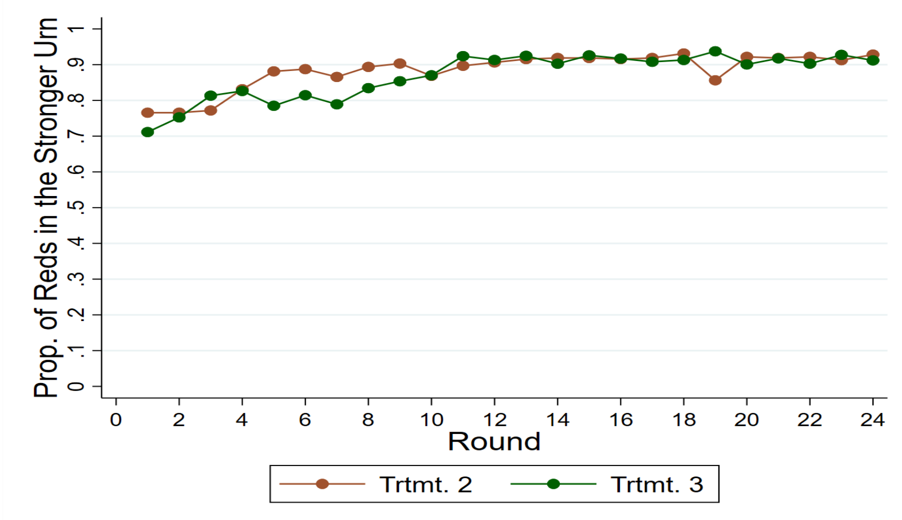
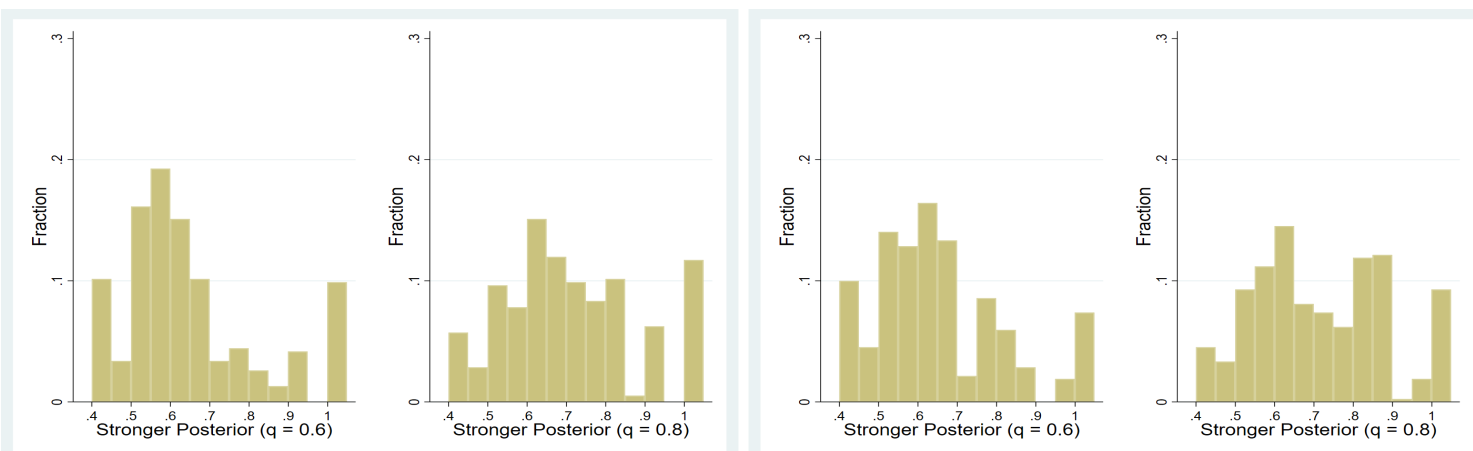


Figure A3: Empirical Distribution of Stronger Posterior by Treatment and q



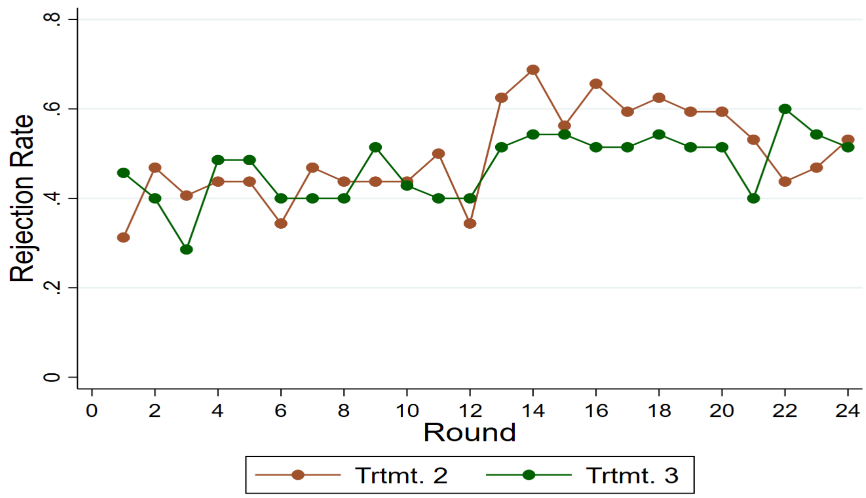
(a) Treatment 2a

(b) Treatment 2b

under a higher q.

Figure A4 reports the rejection rate over time. We see that in both of the two treatments, the rejection is frequent and robust.

Figure A4: Prop. of Red Balls in the Stronger Urn Over Time



References

Alonso, R. and O. Camara (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory 165*, 672–706.

Alonso, R. and O. Câmara (2016). Persuading voters. *American Economic Review 106*(11), 3590–3605.

Au, P. H. and K. K. Li (2018). Bayesian persuasion and reciprocity: Theory and experiment.

Bergemann, D., B. Brooks, and S. Morris (2015). The limits of price discrimination. *American Economic Review 105*(3), 921–57.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The annals of mathematical statistics*, 265–272.

Blume, A., E. K. Lai, and W. Lim (2017). Strategic information transmission: A survey of experiments and theoretical foundations. *Report.[1457]*.

Boleslavsky, R. and C. Cotton (2015). Grading standards and education quality. *American Economic Journal: Microeconomics 7*(2), 248–79.

Charness, G. and D. Levin (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *American Economic Review 95*(4), 1300–1309.

Crawford, V. P. and J. Sobel (1982). Strategic information transmission. *Econometrica: Journal of the Econometric Society*, 1431–1451.

Dickhaut, J. W., K. A. McCabe, and A. Mukherji (1995). An experimental study of strategic information transmission. *Economic Theory 6*(3), 389–403.

Eckel, C. C. and P. J. Grossman (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization 68*(1), 1–17.

Esponda, I. and E. Vespa (2019). Contingent thinking and the sure-thing principle: Revisiting classic anomalies in the laboratory. Technical report, mimeo.

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

Fréchette, G. R., A. Lizzeri, and J. Perego (2019). Rules and commitment in communication: an experimental analysis. Technical report, National Bureau of Economic Research.

Green, J. R. and N. Stokey (1978). *Two representations of information structures and their comparisons*. Institute for Mathematical Studies in the Social Sciences.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *The Journal of Law and Economics 24*(3), 461–483.

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

Hsee, C. K. and E. U. Weber (1997). A fundamental prediction error: Self–others discrepancies in risk preference. *Journal of experimental psychology: general 126*(1), 45.

Hsee, C. K. and E. U. Weber (1999). Cross-national differences in risk preference and lay predictions. *Journal of Behavioral Decision Making 12*(2), 165–179.

Jin, G. Z., M. Luca, and D. Martin (2017). Is no news (perceived as) bad news? an experimental investigation of information disclosure. Technical report, Harvard Business School.

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

Kosterina, S. (2018). Persuasion with unknown beliefs. *Work. Pap., Princeton Univ., Princeton, NJ*.

Martínez-Marquina, A., M. Niederle, and E. Vespa (2019). Failures in contingent reasoning: The role of uncertainty. *American Economic Review 109*(10), 3437–74.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 380–391.

Nguyen, Q. (2017). Bayesian persuasion: Evidence from the laboratory. Technical report, Working Paper.

Siegrist, M., G. Cvetkovich, and H. Gutscher (2002). Risk preference predictions and gender stereotypes. *Organizational Behavior and Human Decision Processes 87*(1), 91–102.

Szydlowski, M. (2016). Optimal financing and disclosure. *Available at SSRN 2735981*.