Communication, Perception and Strategic Obfuscation

Geoffroy de Clippel* Kareen Rozen*

Revised July 2020

Abstract

We study the empirical content of simple Sender-Receiver games in which disclosures are mandatory but may be obfuscated. We focus on the fungibility between strategic inference and costly perception, developing a stylized theoretical framework that highlights this channel. Our framework yields crisp testable implications for equilibrium play, and naturally lends itself to an experimental design. Our laboratory results show that a large majority of Senders strategically obfuscate; and an aggregate analysis of Receiver's stochastic-choice data suggests Receivers adjust their perception in response to strategic inference.

*Department of Economics, Brown University. This material is based upon work supported by the National Science Foundation under Grant No. 1559398 (*Communication, Perception and Strategic Obfuscation*, from June 2016 to September 2019). We are grateful to Pedro dal Bó, Mark Dean, Brian Knight, Henrique Roscoe De Oliveira, and Ran Spiegler for valuable comments and suggestions, and to Tommaso Coen and Zeky Murra for research assistance.

1 Introduction

Communication is a key component of many interactions. Pharmaceutical companies advertise medications to consumers; job candidates recount their qualifications to potential employers; attorneys provide evidence to defend clients against litigants; researchers describe their studies to potential participants; food manufacturers report ingredients to shoppers. These examples share three common features. First, the favorite action of the informed party only sometimes aligns with that of the uninformed decision-maker. Second, regulations (or the potential for serious repercussions) force the informed party to truthfully disclose their information. Finally, without resorting to dishonesty, the informed party can try to mitigate detrimental information by making it harder to understand (e.g., job candidates dress up resumes, attorneys submit entire hard drives into evidence instead of just the incriminating files). Concerns about strategic obfuscation have been raised in some of these settings, with the challenge that obfuscation is in the eye of the beholder. The 'Belmont report',[1] which lays out the ethical principles governing human-subjects research, warns that "presenting information in a disorganized and rapid fashion, allowing too little time for consideration or curtailing opportunities for questioning, all may adversely affect a subject's ability to make an informed choice." Similarly, the FDA considers presentational choices such as typesetting when deciding whether an advertisement violates legal requirements to provide a "fair balance" of risks and benefits.[2] And the federal law mandating the disclosure of genetically-modified ingredients itself spurred controversy with its requirement that food packaging *either* label GMOs in clear language, *or* include a QR code linking to a site where the information can be found (the food industry backed the latter approach).[3]

But if obfuscation occurs only when information is detrimental, then witnessing obfuscation reveals information by itself, and impacts the effort a rational agent should exert to sift through the message's content. In other words, strategic sophisti-

[1] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979), downloadable from the Office for Human Research Protections.

[2] See the FAQ Does the law say anything about the design of ads for prescription drugs?

[3] See Public Law 114-216. While proponents touted the law for mandating disclosure, some advocacy groups derided it as the "DARK Act," for Deny Americans the Right to Know. The controversy is described, for instance, in the WSJ article Consumer Advocates Wary of Digitally Coded Food Labels, and the Huffington Post entry Obama Expands Monsanto Doctrine By Signing DARK Act And Invalidating Vermont GMO Labeling Law.

cation can potentially substitute for costly perception. We are particularly interested in this question in view of the recent literature on the empirical content of optimal attention in individual decision-making problems (Caplin and Dean, 2015; Caplin and Martin, 2015; de Oliveira et al., 2017; Ellis, 2018; Dean and Nelighz, 2019). A goal of our paper is to illustrate how these ideas might play out in interactive decision-making, highlighting the interplay between strategic inferences and inattention.

For this, we focus on a highly stylized, easily interpretable model of communication. There are two possible actions and two equally-likely states of the world, corresponding to whether or not the Sender's and Receiver's preferred actions agree. The Receiver receives a fixed benefit for taking the action that matches the state; while the Sender benefits whenever one particular action is taken. There are two possible types of messages for each state, *transparent* and *obscure*, both of which fully reveal the state. The Receiver immediately understands the state from a transparent message. By contrast, the only thing the Receiver can immediately tell from an obscure message is that it is not transparent; costly effort is required to understand it. The Sender can condition his preferred message type on the state, but his communication goal may be imperfectly realized. Consider, for instance, an article under journal review. The authors may believe their writing is transparent, but the referees may disagree; and conversely, the referees may easily recognize some issue despite the authors' attempts at obfuscation. We encapsulate the potential for such disagreement in the *precision level of communication* (p), which is the probability with which a Sender who aims to send an obfuscated message (or aims to send a transparent one) in a given state achieves his goal.

Following Caplin and Dean (2015), we model the process that allows the Receiver to decipher information as a costly task whose cost is unknown to the modeler. Embedded in a game-theoretic framework, a novel feature of our problem is that the Receiver's beliefs about the state after receiving an obscure message, *but before exerting any effort to decipher it,* depend on what the Sender is expected to do. In an undominated Bayesian-Nash equilibrium, a strategically-sophisticated Sender should aim to obfuscate when his favored action is worst for the Receiver. Indeed, this is the only weakly-dominant strategy, and is the unique best reply if there is any chance at all that the Receiver fails to match the state following an obfuscated message. In equilibrium, as the precision of communication increases, the Receiver becomes more convinced that any observed obfuscation is attributable to intention, and thus more convinced that the Sender's favored action is the wrong one to take. These beliefs inform how much effort to exert in deciphering the message.

The Receiver's strategic inference and ensuing effort exertion both manifest themselves in the probabilities, conditional on each state, that he takes the right action despite obfuscated information. For instance, if he eschews effort entirely and simply chooses the Sender's worst action, then the Receiver takes the right action for sure in the opposing-interest state, but always fails to do so in the common-interest state. If he instead makes a decision based on some costly perception strategy, his success in the common-interest state may increase at the expense of his success in the opposing-interest state. Elaborating on Caplin and Dean (2015)'s approach, we derive the testable implications on the Receiver's state-contingent stochastic choice data for it to be consistent with a rational Receiver's decisions in equilibrium. Collecting such data at the individual level is hard in general, but even harder in our setting when p is large (as obfuscated messages are then received only rarely in the common-interests state). It is important then to note that the result on the empirical content of equilibrium naturally extends into a test on aggregate data in situations where Senders and Receivers are randomly matched and have potentially heterogenous perceptual costs, as expected in a laboratory experiment.

Our framework translates directly into an experimental design, which we explore using three treatments: low, medium and high precision levels, corresponding to values of $p \in \{51\%, 70\%, 90\%\}$. To bring obscure and transparent messages to life, we import the novel 'colored balls'-design of Dean and Nelighz (2019)[4], who experimentally test the rational inattention model through how success rates vary with incentives. They represent the true state (either Red or Blue) by a square matrix of 100 red and blue balls: though the balls are randomly placed in the matrix, exactly 51 of them color-match the state. We use precisely this construction for obfuscated messages, and use the same balls ordered neatly by color for transparent messages. Thus messages do not differ in substance beyond their clarity.

The data analysis in Section 3.2 substantiates the overall strategic sophistication and rationality of Senders and Receivers. We find that 77% of all Senders *strategically* obfuscate: they aim for clarity in the common-interests state, but aim to obfuscate in the opposing-interests state. Receivers' success probabilities vary with p, showing that strategic inferences can impact attention in games. We find evidence that perception

[4]Also appearing in an earlier working paper which Dean and Nelighz (2019) subsumes.

is adjusted optimally, as the testable implications on aggregate stochastic choice data are satisfied (or nearly satisfied, in one instance). Some interesting demographic differences are discussed.

Given this evidence that average Receivers do adjust their perception with strategic inference, Section 4 concludes by highlighting some (potentially counterintuitive) implications for welfare. First, we point out that a naive regulator (one who does not recognize that obfuscated messages carry information beyond their immediate content) can grossly overestimate or underestimate the Receiver's welfare gain from mandating information disclosure. Second, strategic sophistication can be detrimental for the Receiver. Finally, greater alignment of preferences between the Sender and Receiver does not guarantee greater success at guessing the state.

Related work on communication

Communication games have traditionally been studied in one of two extreme settings. In the cheap-talk setting of Crawford and Sobel (1982), information is soft: messages need not bear any verifiable relation to the truth, but could have meaning in equilibrium. At the other extreme is hard information, see Grossman (1981); Milgrom and Roberts (1986): messages are immediately verifiable, and the absence of a message could be revealing in equilibrium (leading to information unraveling).

Our work is in the spirit of Dewatripont and Tirole (2005), which introduced information whose softness is intermediate and endogenous into communication games.[5] They derive equilibria in a stylized model with communication subject to moral hazard in teams: the Sender and Receiver have increasing and convex costs of effort, and the Receiver assimilates the Sender's information with probability xy when the Sender exerts an effort $x \in [0, 1]$ and the Receiver exerts an effort $y \in [0, 1]$. Most of their work presumes simultaneous effort choices, so there is no role for strategic inference. Inferences do matter though in two variations they discuss: playing two successive rounds of the above game, and a one-shot interaction where the Receiver directly observes x before choosing y. Our works share the broad features that there is a probabilistic nature to whether the Receiver understands her message, and that

[5] See also Persson (2018) for an extension of Dewatripont and Tirole (2005) to decisions with multi-dimensional considerations: a capacity-constrained decision maker gets information from an expert on each of these dimensions, leading to information overload as a manipulation device. Like in Dewatripont and Tirole (2005), the expert prefers a particular action but does not know the state.

the Sender's choice impacts this distribution (whatever it is). However, our work is interested in testable implications that hold for any cost function, and a framework for experimentally highlighting the role of strategic sophistication. To focus on that channel, we abstract from Sender costs and study the extreme setting where the Receiver can freely distinguish between transparent and non-transparent information. This is the only means through which our Receiver can draw inferences about the Sender's unobservable communication goal.

This paper contributes more generally to the emerging theoretical literature on attention in games. Some works consider the behavior of firms facing consumers whose perception is exogenously given, as in Gabaix and Laibson (2006) and Bordalo et al. (2015); or where consumers can optimally allocate a fixed total effort among different dimensions, as in Spiegler (2006) and de Clippel et al. (2014). Another strand of the literature derives equilibria of games with players who can endogenously choose their perceptual efforts at a cost, often modeled using the Shannon mutual-information function applied in Sims (2003); see, for instance, Matějka (2015), among others. Most of these works study settings where there can be no strategic inference about private information, and therefore no fungibility with optimal attention. An exception is Martin (2016b), who considers a firm's strategic pricing when facing a consumer who is rationally inattentive about its product's quality. He shows there is a mixed-strategy equilibrium in which the high-quality seller always sets a high price, while the low-quality seller randomizes between low and high prices. The buyer's attention responds to the seller's mixed strategy in equilibrium, using Sims' linear parametrization of attention costs.

In a companion paper, Martin (2016a) provides experimental data to illustrate and calibrate the above model. A seller owns a hypothetical product that has low or high value to the buyer with whom he is randomly matched. Knowing the buyer's value, he chooses between a low-price and a high-price offer. Accepting a low-price offer is always profitable for the buyer, but accepting a high-price offer is profitable only if the product is of high value to the buyer. Not knowing his value, the buyer can examine a string of twenty randomly generated numbers (between -100 and 100) whose sum is the true value. Using time responses and the frequency of purchasing mistakes, Martin provides supporting evidence that the attention buyers pay to learn the product's value is impacted by the seller's price. Focusing on a rationally-inattentive representative buyer, the best approximation of buyers' average behavior is obtained with a marginal attention cost of 11.9. Interestingly, explaining sellers' average behavior using the equilibrium described in the previous paragraph leads to a comparable estimate for the sellers' belief regarding buyers' marginal attention cost.

Our paper differs from Martin (2016a,b) on multiple dimensions. First, the Sender in our game decides whether or not he tries to obfuscate the information, while in Martin's framework the seller chooses a price and information is always obfuscated. Second, we characterize testable implications of our model: that is, properties on observables which remain valid whatever the participants' utilities and attention costs. We do not place restrictions on the form of attention costs. The Receiver is not required, for instance, to process information as in Sims' model of rational inattention (see Dean and Nelighz (2019) for an experimental comparison of the relative performance of different cost functions in individual decision-making). Third, the testable implications are derived while allowing the precision parameter to vary, which leads to cross-observation tests of consistency with equilibrium play. In our framework, the Receiver cannot perfectly observe the Sender's choice, which is the key to our treatments. In addition to the more stringent testable implications of equilibrium play, simply witnessing that success rates at guessing the true state vary with the precision level provides evidence in a clean treatment-control design that Receivers do adjust their attention level based on the strategic inference they can draw in the games they play.

The experimental literature has examined many different aspects of communication. Blume et al. (2020) surveys a large literature on experimental studies of cheap talk. A smaller literature studies information unraveling when information is hard; see Jin et al. (2016) and references therein. Fréchette et al. (2019) explores an umbrella framework nesting cheap talk, hard information, and Bayesian persuasion. They relax the commitment assumption in Bayesian persuasion through a probability the Sender can revise his choice. They do not, however, consider obfuscated information. Jin et al. (2019) studies obfuscation, albeit with a different goal. Their Senders know the state $s \in \{1, 2, \dots, 10\}$ and send a string with $c \in \{1, 2, \dots, 20\}$ numbers whose sum equals the state. The Receiver has 60 seconds to guess the sum (else a random guess is made), and is paid for accuracy. The Sender's payoff increases in the guess. In theory, information should unravel the same way it does in the case of voluntary disclosure, as one can draw comparable inferences from witnessing obfuscation or information withholding. But Jin et al. (2019) shows that unraveling does not occur the way it does under voluntary disclosure, and advances possible explanations. By contrast, unraveling does not occur at equilibrium in our model: quite realistically, perceiving a message as obfuscated is informative about the Sender's intention, but does not reveal for sure that the Sender aimed to obfuscate. This allows us to focus on our main question of interest, the testable implications for the substitutability between optimal perception and strategic inference.

2 Theoretical benchmark

We consider a Sender-Receiver game with two possible states of the world, $\Omega = \{\omega_1, \omega_2\}$, and a prior $\pi(\omega_i)$ over the states. The Receiver has two possible actions, $A = \{a_1, a_2\}$, with action a_i being the Receiver's preferred action in state ω_i. The Sender, on the other hand, strictly prefers the Receiver to choose action a_2 no matter the state, and would like to persuade the Receiver to pick it; hence we refer to their interaction as a persuasion game. Both Sender and Receiver are expected utility maximizers, and have strictly increasing utility functions over money. For simplicity, we assume the Receiver receives $\$m_R$ whenever he chooses the action that matches the state, and nothing otherwise; while the Sender receives $\$m_S$ whenever the Receiver chooses action a_2, and nothing otherwise. Thus ω_1 is a state of *opposing interests,* because the action the Receiver prefers in that state is worst for the Sender. By contrast, ω_2 is a state of *common interests.*

The Sender is informed of the true state and must communicate it to the Receiver, but need not make this information easily understood. Formally, for each state ω, the Sender can aim to *communicate clearly* or aim to *obfuscate*. The *precision level* $p \in (1/2, 1)$ calibrates how likely the Sender's communication goal is achieved. Specifically, if the Sender aims to communicate clearly in state ω, then with probability p the Receiver's message will be *transparent* (denoted $T(\omega)$) and with probability $1 - p$ the message will be *obscure* (denoted $O(\omega)$). Oppositely, if the Sender aims to obfuscate in state ω, then the Receiver's message will be obscure with probability p and transparent with probability $1 - p$. The precision level p will be a parameter we vary across treatments in the experiment.

As the Receiver does not know the true state, there are four possible message types she may receive prior to deciding on an action. With a transparent message $T(\omega)$, the Receiver understands at once that the state is ω, and will take the action that matches the state. With an obscure message, however, the Receiver's only way to distinguish $O(\omega_1)$ from $O(\omega_2)$ is to exert effort to decipher the message. In line with recent models of optimal attention in decision theory, he chooses a perception strategy $(\mathcal{S}, \mu)$, where $\mathcal{S}$ is any finite set whose elements $s \in \mathcal{S}$ are called *signals* and $\mu(s|\omega)$ is the probability of signal s when state is ω. He also chooses a *decision rule* which fixes the action to pick following each signal. A decision rule can be described as a partition of $\mathcal{S}$ into two subsets, $\mathcal{S}_1$ and $\mathcal{S}_2 = \mathcal{S} \setminus \mathcal{S}_1$, such that the action a_i is taken following signals in the subset $\mathcal{S}_i$. The choice of a perception strategy and decision rule is associated with a cost $c_R(\mathcal{S}_1, \mathcal{S}, \mu)$, on which we impose no functional form assumption besides modeling it as being subtracted from the expected utility of earnings.

2.1 Equilibrium Conditions

Our equilibrium notion is that of *undominated Bayesian Nash equilibrium*.[6] We now further detail the mutual best-response conditions. Consider the Sender's optimization problem first. The Receiver's equilibrium perception strategy and decision rule define a *success probability* $\ell(\omega_i)$ for $i = 1, 2$, the probability of guessing the state ω_i correctly when receiving an obfuscated message. With $\tau(\omega)$ denoting the probability that the Sender aims to communicate clearly in state ω, the Sender's equilibrium conditions are

$$\begin{cases} \tau(\omega_1) \in \arg\max_{x \in [0,1]} \{u_S(m_S) - \nu(x, \ell(\omega_1))[u_S(m_S) - u_S(0)]\} \\ \tau(\omega_2) \in \arg\max_{x \in [0,1]} \{u_S(0) + \nu(x, \ell(\omega_2))[u_S(m_S) - u_S(0)]\}, \end{cases} \tag{1}$$

where $\nu(x, y) = [x(p+(1-p)y)+(1-x)(py+1-p)]$ is the unconditional probability that the Receiver guesses the state correctly, when the Sender aims to communicate clearly

[6]The game also admits BNEs with weakly dominated strategies when perception costs are high enough. In those equilibria, an obfuscated message prompts the Receiver to pick a_1, the Sender communicates clearly given ω_2, and is indifferent between obfuscating or communicating clearly given ω_1. But if there is any doubt in the Sender's mind that an obfuscated message results in a_1 for sure, then aiming to obfuscate given ω_1 is strictly preferred, which eliminates dominated equilibria. While incorporating them in the analysis is conceptually easy, we think it is preferable for expositional clarity to focus on undominated equilibria, a mild refinement. This is especially reasonable given that the Sender faces a random Receiver in our experiment (more on this in Section 2.4), which means that he'd play a dominated strategy only if he is sure that *all* subjects in the room overlook the content of an obfuscated message and pick a_1 for sure.

with probability x, and the Receiver guesses correctly with probability 1 following a transparent message but only with probability y following an obfuscated message. Clearly, the Sender's payoff is decreasing in ν in the opposing-interests state (top equation), and increasing in ν in the common-interests state (bottom equation).

As for the Receiver's best response, notice that her beliefs about the state after receiving an obscure message, *but before exerting any effort to decipher it*, are given by:

$$\hat{\pi}(\omega) = \frac{\pi(\omega)\,(\tau(\omega)(1-p) + (1-\tau(\omega))p)}{\sum_{i=1,2} \pi(\omega_i)\,(\tau(\omega_i)(1-p) + (1-\tau(\omega_i))p)}. \quad (2)$$

Then, her perception strategy and decision rule maximize

$$u_R(0) + \sum_{i=1,2} \hat{\pi}(\omega_i)\mu(\mathcal{S}_i|\omega_i)(u_R(m_R) - u_R(0)) - c_R(\mathcal{S}_1, \mathcal{S}, \mu) \quad (3)$$

under the constraint that, if signal σ belongs to $\mathcal{S}_i$, then following that signal the Receiver prefers a_i over a_j. This means that the Receiver's posterior probability

$$\hat{\mu}(\omega|\sigma) = \frac{\mu(\sigma|\omega)\hat{\pi}(\omega)}{\sum_{\omega' \in \Omega} \mu(\sigma|\omega')\hat{\pi}(\omega')}$$

for state ω, conditional on getting the signal σ from an obscure message, satisfies

$$\hat{\mu}(\omega_i|\sigma) \geq \hat{\mu}(\omega_{-i}|\sigma), \quad (4)$$

for all $\sigma \in \mathcal{S}_i$ and for all $i = 1, 2$.

2.2 Observables and Equilibrium Consistency

Consider repeated observations from several different persuasion games, which differ only in the precision level p. Of course, perception strategies and decision rules are not observable. Instead, we provide testable implications on a *dataset* $\{(p^j; \tau^j, \ell^j)|j = 1, \ldots, J\}$, where p^j is the precision level of the persuasion game j, $\tau^j = (\tau^j(\omega_1), \tau^j(\omega_2))$ specifies the probability the Sender aims to communicate clearly as a function of the state in game j, and $\ell^j = (\ell^j(\omega_1), \ell^j(\omega_2))$ specifies the probability the Receiver chooses the correct action as a function of the state in game j.

The dataset is *consistent with equilibrium play* if there exist utility functions u_S, u_R, a perception-cost function c_R, and for each j a perception strategy and decision rule $(\mathcal{S}_1^j, \mathcal{S}^j, \mu^j)$ that combine with τ^j to form an equilibrium of the corresponding Sender-Receiver game given by p^j, such that $\ell^j(\omega_i) = \mu^j(\mathcal{S}_i|\omega_i)$ for $i = 1, 2$. This amounts to a revealed-preference exercise in a game with state-dependent stochastic choice data. In other words, we are interested in finding all predictions of our Sender-Receiver games that remain valid whatever the utility and perception-cost functions. While we focused on an equilibrium notion, we will see that the testable implications we derive also remain valid under alternative assumptions on participants' expectations.

Collecting individual success rates is experimentally demanding in general, and this is especially true in our Sender-Receiver game when p is large (see discussion in Section 3). By contrast, collecting aggregate success rates is more easily done. In that case, of course, there may be heterogeneity in utilities and perception costs. In subsection 2.4, we explain how our testable implications can be adapted to such situations.

2.3 Testable Implications

We are now ready to state and prove the main theoretical result.

Proposition 1. *The dataset is consistent with equilibrium play if, and only if, all the following conditions hold:*

(i) For each persuasion game j, the Sender aims to obfuscate in the opposing-interests state and aims to communicate clearly in the common-interests state: $\tau^j(\omega_1) = 0$ and $\tau^j(\omega_2) = 1$;

(ii) The Receiver's belief upon the receipt of an obfuscated message in game j is:

$$\hat{\pi}^j(\omega_1) = \frac{p^j \pi(\omega_1)}{p^j \pi(\omega_1) + (1 - p^j)\pi(\omega_2)};$$

(iii) In each persuasion game j, the Receiver's expected success rate upon the receipt of an obfuscated message must be at least as high as the expected success rate from choosing action a_1:

$$\hat{\pi}^j(\omega_1)\ell^j(\omega_1) + (1 - \hat{\pi}^j(\omega_1))\ell^j(\omega_2) \geq \hat{\pi}^j(\omega_1);$$

(iv) *For any pair of persuasion games with* $\hat{\pi}^j(\omega_1) > \hat{\pi}^k(\omega_1)$ *(equivalently* $p^j > p^k$*, given (ii)) we have* $\ell^j(\omega_1) - \ell^j(\omega_2) \geq \ell^k(\omega_1) - \ell^k(\omega_2)$.

We now discuss and prove part of the result, while also pointing out that the testable implications we uncovered remain valid with other assumptions on beliefs. The Receiver's strategy pins down success probabilities $\ell(\omega_i)$ for $i = 1, 2$. Whatever these probabilities are, notice that the Sender's payoff in (1) is decreasing in x in the opposing-interests state, increasing in x in the common-interests state, and strictly so if the Receiver does not always succeed at guessing the state correctly. Thus aiming to obfuscate given ω_1 and to communicate clearly given ω_2, is not just a characteristic of equilibrium play: it is the unique weakly dominant strategy for the Sender, and the unique best response with success rates strictly below one. Hence property (*i*) is robust to alternative specifications of Sender's expectations about Receivers.

We already described in Equation (2) how Bayes' rule pins down the probability that the Receiver attaches to both states upon receipt of an obfuscated message, given his belief about the Sender's strategy. This belief is determined in (*i*) at equilibrium, which establishes (*ii*). However, one must recognize the possibility that, in practice, a Receiver's updated probability of the state may differ. For instance, he may fail to update probabilities accurately, fail to recognize that the strategy in (*i*) is dominant, or fear that the Sender might not recognize that strategy is dominant. Thus, while we perform our benchmark analysis of the data using the equilibrium beliefs in (*ii*), we will also discuss its robustness to alternative beliefs of the Receiver.

The empirical content of equilibrium play on the Receiver's side boils down to the testable implications of optimal perception, but given beliefs updated through strategic inference. This is reminiscent of the individual decision-making problems studied by Caplin and Dean (2015). A difference is that they compare behavior as payoffs change, keeping states' probabilities unchanged. The very nature of our analysis entails (endogenous) changes in probabilities instead, while keeping the payoffs constant across observations. This difference can be dealt with through a mathematical transformation: probabilities premultiply utilities under expected utility, so one can reinterpret probability changes as payoff changes (see details in the Appendix). The special structure of the resulting problems leads to a new feature in the testable implications: one can restrict attention to pairwise comparisons in (*iv*), instead of dealing with a condition analogous to Rockafeller (1970)'s cyclical monotonicity. We discuss this further below.

To better understand (*iii*), note that the Receiver's expected success rate upon the receipt of an obfuscated message is $\hat{\pi}^j(\omega_1)\ell^j(\omega_1) + (1 - \hat{\pi}^j(\omega_1))\ell^j(\omega_2)$. If simply picking a_1 led to a higher success rate, then the Receiver would be better off doing this and saving on perception costs. The same reasoning also applies to a_2, but the resulting inequality is redundant: upon receipt of an obfuscated message, but before attempting to decipher it, the opposing-interests state is more likely and so a_1 preferable to a_2. This remains true for a wide range of beliefs $\hat{\pi}^j(\omega_1)$ beyond those in equilibrium. Condition (*iii*) is the analogue in our framework of Caplin and Dean's NIAS condition. If we observe a single precision level ($J = 1$), then it is the only testable implication for the Receiver's choices.

By varying the precision of communication, and thus the Receiver's beliefs, more nuanced implications in (*iv*) obtain from the Receiver's choices *across* persuasion games. This relates to Caplin and Dean's NIAC condition, which ensures that total utility cannot increase by reassigning attention across any cycle of decision problems (of any length). Considering only pairwise comparisons (cycles of size two) is sufficient in our framework because payoffs are held constant across observations; see the proof in the Appendix. This leads to the simpler condition (*iv*) that *excess-success rates* weakly increase in the precision of communication. For instance, a Receiver who views ω_1 as more likely in game j than game k upon the receipt of an obfuscated message ($\hat{\pi}^j(\omega_1) > \hat{\pi}^k(\omega_1)$, which amounts to $p^j > p^k$ under (*ii*)), may exert less effort in deciphering it, while using a decision rule that further favors a_1. This will increase success in the opposing-interests state, and decrease success in the common-interests state, consistent with (*iv*). The proof in the Appendix shows that (*iv*) is indeed necessary, but also sufficient: no stronger condition on success rates can be found. In examples, one can construct perception costs for which success rates increase in both states, and others for which success rates decrease in both states. The only unequivocal prediction is that the spread between the common-interests success rate and that of the opposing-interests state increases when moving from k to j.

Conditions in (*iv*) are valid given any beliefs $(\hat{\pi}^j)_{j=1}^J$, however they arise, and turn out to be robust against an array of beliefs the Receiver may hold: that is, permitting violations of (*ii*). Indeed, when comparing the excess-success rates in two persuasion games j and k, the direction of inequality to check depends only on which of $\hat{\pi}^j(\omega_1)$ and $\hat{\pi}^k(\omega_1)$ is larger, not on their cardinal values. Of course, one would generally expect the belief $\hat{\pi}^j(\omega_1)$ to increase in the precision level p^j, since the Sender has no incentive to obfuscate in the common-interests state ω_2.

2.4 A population of Senders and Receivers

The analysis above presumes there is only one possible Sender and only one possible Receiver that can be matched. Suppose instead that Senders and Receivers are paired uniformly at random. All players strictly prefer more money to less, but may differ in their utility function and perception costs.

This is a setting of practical relevance for at least two reasons. First, a Sender may expect to face a population of Receivers with different perceptual costs, as would be the case with a population of consumers. Second, it can be quite time-consuming for subjects in experiments to generate individual-level, state-dependent stochastic choice data (in this case, the probability the Receiver chooses the right action following obfuscated information, conditional on each state of the world). This is especially so in our experiment, where obfuscated messages are generated endogenously, and we would expect to see few of them in the common-interests state under high precision levels p. Fortunately, Proposition 1 extends to characterize the testable implications in this setting. Regarding Senders, the argument for (i) immediately extends to show the same strategy remains weakly dominant, and strictly so if there is any chance of average success rates strictly below one (i.e., if there is a strictly positive probability that at least one Receiver imperfectly guess the state following an obfuscated message). Hence Receivers all share the same belief (ii) as before, upon seeing an obfuscated message but before attempting to decipher it.

As for the testable implications on success rates, notice that Proposition 1 holds for each Receiver i. This means each Receiver satisfies (iii) and (iv), which are linear in individual success rates. Let $\bar{\ell}^j(\omega)$ be the average success rate of Receivers in game j and state ω. Notice that, in our between-subject analysis, different sets of subjects play our communication game with different precisions levels. However, we can reasonably assume, as is standard, that the different treatments are drawn from the same population of characteristics. Summing (iii) and (iv) over all Receivers in each treatment implies:

$$\hat{\pi}^j(\omega_1)\bar{\ell}_i^j(\omega_1) + (1 - \hat{\pi}^j(\omega_1))\bar{\ell}_i^j(\omega_2) \geq \hat{\pi}^j(\omega_1), \text{ for all } j.$$
$$\bar{\ell}_i^j(\omega_1) - \bar{\ell}_i^j(\omega_2) \geq \bar{\ell}_i^k(\omega_1) - \bar{\ell}_i^k(\omega_2), \text{ when } \hat{\pi}^j(\omega_1) > \hat{\pi}^k(\omega_1).$$

These necessary conditions are the same as (*iii*) and (*iv*) in Proposition 1, but using *average* success rates. In the other direction, if average success rates satisfy these conditions, then Proposition 1 implies that the data can be explained by a hypothetical population of Receivers identical in their utility functions and perception costs, even if the members of the population are in fact heterogenous. Thus Proposition 1, applied to population averages, remains necessary and sufficient for consistency with equilibrium play even in this more general setting.

3 Experiment

In our experiment, subjects are matched to play Sender-Receiver games. Before describing how subjects are matched, we describe the game itself.

There are two possible keywords, Blue and Red, with a fifty percent chance of each. Given the keyword, the Receiver will receive one of two possible types of messages. In both types of messages, there are 100 balls of blue and red color arranged in a 10x10 matrix, with exactly 51 balls whose color matches the keyword and exactly 49 balls whose color mismatches it. However, the messages differ in how the balls are arranged. In one type of message, the balls are arranged by color (see the left panel of Figure 1). Such a message is transparent, in the sense that it immediately reveals the majority color, and hence the keyword. In the second type of message, the same balls are distributed uniformly at random into the 10x10 matrix (see the right panel of Figure 1). Such a message is obfuscated, in the sense that it takes some effort to garner information on the majority color, and hence the keyword. Both message types reveal the truth about the keyword (with differing levels of transparency), as the keyword is revealed by the majority color. Hence the Sender cannot lie; she can only obfuscate.

The Receiver's message will always be one of these two types, but the probability of the different types is impacted by the Sender. For each of the two possible keywords, the Sender is asked to choose whether they prefer it to be more likely (with probability $p > 1/2$) that the balls will be arranged by color or more likely (with the same probability p) that the balls will be in random order. Thus the Sender makes a contingent messaging plan, tailored to each keyword. We consider three treatments, corresponding to p being 51%, 70%, and 90%. The value of the probability p is held constant within each session, both to avoid pollution across precision levels and to

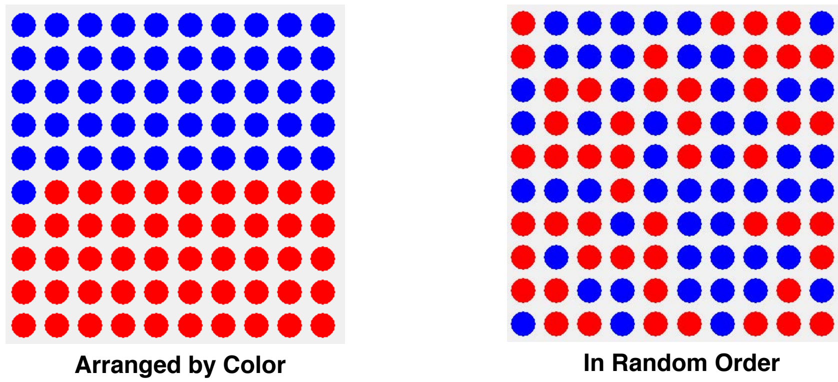


Figure 1: Examples of message types when the keyword is Blue.

ensure sessions are not unreasonably long.

In any Sender-Receiver matchup, the computer draws the keyword (Blue or Red) uniformly at random. The Sender's decision for that keyword is implemented according to the probability p to determine the Receiver's message. The Receiver is then asked to guess the keyword. The Receiver is permitted unlimited time to examine their message before making this choice. The Sender's payoff from the matchup is \$15 if the Receiver guesses that the keyword is Red, and \$0 otherwise. The Receiver's payoff from the matchup is \$15 for guessing the keyword correctly, and \$0 otherwise.

The timing of decisions and matching process are as follows. Each session is split into two phases. In the Sending phase, each subject acts as a Sender and selects a contingent messaging plan. Decisions made during the Sending phase will be used to determine messages in the Receiving phase of the experiment. In this second phase, each subject acts as a Receiver, and is matched forty successive times. Each match is with an independently and uniformly drawn Sender other than themselves. In each of these forty matches, the computer independently and uniformly draws the keyword and implements the matched Sender's decision for the drawn keyword. That is, the computer uses the relative likelihood the Sender picked for that keyword, to display a message with balls arranged either by color or randomly. The Receiver has an unlimited amount of time to examine this message before guessing the keyword.

The experiment is designed so that if Senders make their communication decision rationally, then the probability a Receiver gets an obfuscated screen in a match is the same across different precision levels p. Namely, it is:

$$\pi(\text{Red})(1-p) + \pi(\text{Blue})p = \frac{1}{2}(1-p+p) = \frac{1}{2}.$$

This means that on average, there is an equal burden for Receivers across treatments. What does change across treatments is the distribution of keywords *conditional* on seeing an obfuscated screen. If Senders act rationally, then conditional on an obfuscated screen, the probability of Blue is exactly the treatment's precision level p.

No feedback is provided in either the Sending or Receiving phases. The experiment concludes once the Receiving phase is complete. At that point, subjects may complete an optional exit survey. The computer determines each subject's payoff by randomly picking a role (Sender or Receiver) and a match in which the subject played that role. Each subject is paid, in cash, their payoff from that match in addition to the $10 show-up fee. Subjects are not told the roles other subjects played, the choices others made, or the payoffs others received.

3.1 Procedure

There were six experimental sessions, two for each value of $p \in \{51\%, 70\%, 90\%\}$. In aggregate, 131 subjects participated, with 42 in the 51% treatment, 44 in the 70% treatment, and 45 in the 90% treatment. All sessions were conducted at BUSSEL, the Brown University Social Sciences Experimental Laboratory, in April and May of 2018. The laboratory is equipped with sunken terminals and vertical privacy panels between desks. Subjects were allowed to participate in at most one session, and were recruited via the BUSSEL website.[7]

Sessions lasted approximately one hour. At the start of each session, the supervisor read aloud the experimental instructions, which were simultaneously available to each subject in a paper handout. The experimental interface was programmed using z-Tree (Fischbacher, 2007). Following the reading of instructions, subjects completed a short quiz via the z-Tree interface to confirm their understanding of how payoffs

[7]This site, available at bussel.brown.edu, offers an interface to register in the system and sign up for economic experiments. To do so, the information requested from subjects is their name and email address and, if applicable, their school and student ID number. The vast majority of subjects registered through the site are Brown University and RISD graduate and undergraduate students, but participation is open to all interested individuals of at least 18 years of age without discrimination regarding gender, race, religious beliefs, sexual orientation or any other personal characteristics.

from each match are generated. Before the start of the experiment, subjects also had an opportunity to familiarize themselves with the interface, using practice screens for the Receiving and Sending phases. Instructions and screenshots of the interface are available in the Online Appendix.

At the end of each session, subjects were presented with an optional exit survey via the z-Tree interface. This survey collected basic demographic information and allowed subjects to describe how they made their choices. Among those reporting gender, about 37% of subjects were male.[8] Subjects reported a wide variety of majors; among these about 32% reported an economics major or a joint major with economics. Subjects were paid their earnings in cash before leaving the laboratory.

3.2 Results

We begin by studying subjects' choices as Senders. There are four possible choice combinations in our experiment, capturing for which keywords (Blue and/or Red, where Blue is the opposing-interests state) they would like an obfuscated message to be more likely than a transparent one. For our purposes here, we may describe these combinations as follows: *Do Not Obscure in Any State* (None), *Obscure in All States* (All), *Obscure Only in the Common-Interests State* (Common), and *Obscure Only in the Opposing-Interests State* (Opposing). As seen from Proposition 1(i), a rational and self-interested Sender should choose *Obscure Only in the Opposing-Interests State*. The table below details the observed sending choices per treatment. The vast majority of Senders are in line with Proposition 1(i) for each treatment: 71.4%, 86.4%, and 73.3%, respectively. If Senders made their choices uniformly at random, we would expect only 25% to be in line with Proposition 1(i). For each treatment, a binomial test rejects, at all significance levels, the null hypothesis that the percent of Senders who follow the equilibrium is equal to that random-choice benchmark.

A minority of Senders aim to communicate clearly in all states, which would be consistent, for instance, with altruistically easing the Receivers' perceptual burden. This is the most common departure we find from equilibrium play. A couple of Senders per treatment aim to obfuscate in all states, which is patently inconsistent with altruism. Each treatment has one subject who obfuscates in the common-interest

[8] According to US News and World Report, 46% of Brown undergraduates are male.

Treatment	None	All	Common	Opposing	Total
51%	9	2	1	30	42
70%	4	1	1	38	44
90%	8	3	1	33	45
Total	21	6	3	101	131

Table 1: Senders' choices, by the states in which they aim to obfuscate.

state only, which could be consistent with intending optimal behavior but confusing the Blue and Red keywords' payoffs.

We find no statistically significant differences in the distribution of sending choices across the treatments (p-value 0.635), nor across sessions within the same treatment (p-values 0.218 for the 51% treatment, 0.363 for the 70% treatment, and 1.000 for the 90% treatment). Moreover, there are no statistically significant differences in sending choices across gender (p-value 0.807) or across economics- versus non-economics majors (p-value 0.284). We also do not find a difference, at any of the above levels of aggregation, in the proportion of Senders who follow the equilibrium prediction (p-values ranging from 0.195 to 0.627).

We now turn our attention to subjects' choices as Receivers. Do they play their part in the equilibrium? In our experimental framework with equally likely states, Receivers' beliefs in equilibrium should equal the precision level p in that treatment. To test conditions (*iii*) and (*iv*) of Proposition 1, we must estimate the Receivers' aggregate success probability $\ell^p(\omega)$ following obfuscated messages, for each state ω and each treatment p. To this end, we pool Receiver's choices for obfuscated messages across the sessions of a treatment p. This leads to 2,446 observations of Receivers' guesses for obfuscated messages, with 805 observations from the 51% treatment, 835 observations from the 70% treatment, and 806 observations from the 90% treatment.

We use logistic regression to estimate the success probabilities, along with their errors, for testing purposes.[9] Define the following dummy variables: $Correct_i$ indicates whether the Receiver in observation i guessed the keyword correctly, Red_i^p indicates whether the observation is from treatment p *and* the keyword was Red, and $Blue_i^p$ indicates whether the observation is from treatment p *and* the keyword was

[9] Logistic regression offers better properties for confidence intervals around probabilities than does linear regression. One could use a linear model for unbiased estimates of the probabilities themselves. For our data, with either method, the success rates are the same to at least six decimal places (and using the errors from linear regression, the qualitative conclusions would be unchanged as well).

Blue. Let $X = (Blue^{51}, Red^{70}, Blue^{70}, Red^{90}, Blue^{90})$ be the vector of explanatory variables. Thereby interacting the treatment with the keyword, we estimate

$$\ln\left(\frac{\mathrm{P}(Correct = 1|X)}{\mathrm{P}(Correct = 0|X)}\right) = \alpha_0 + \alpha_B^{51} Blue^{51} + \alpha_R^{70} Red^{70} + \alpha_B^{70} Blue^{70} + \alpha_R^{90} Red^{90} + \alpha_B^{90} Blue^{90},$$

using Huber-White errors clustered at the session level. Hence our estimates are robust to both heteroskedasticity and correlation in choices (both within and across Receivers in a session).

Remember that the opposing-interests state ω_1 corresponds to the Blue keyword and the common-interests state ω_2 corresponds to the Red keyword. We estimate:

Precision p	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$
51%	0.837	0.783
70%	0.878	0.669
90%	0.927	0.590

Table 2: Estimated state-dependent success probabilities of Receivers per treatment, rounded to three decimal places.

Of course, these success probabilities pertain only to Receiver's choices following obfuscated messages. Our framework assumes that Receivers choose the correct action following a transparent message. Out of the 2,794 transparent messages that our Receivers faced over the three treatments, there were only 12 instances of a Receiver clicking on the wrong keyword, an overall 0.0043 probability of failure.

As for obfuscated messages, are Receivers sophisticated about their meaning? Note that the only difference between treatments is the precision level associated with the Sender's choice (the keyword is always drawn uniformly at random). If Receivers were strategically unsophisticated, then receiving an obfuscated message would have no impact on beliefs following an obfuscated message, and success rates for ω_1 and ω_2 would each be independent of p. To the contrary, the null hypotheses that $\ell^{51}(\omega_1) = \ell^{70}(\omega_1) = \ell^{90}(\omega_1)$ and $\ell^{51}(\omega_2) = \ell^{70}(\omega_2) = \ell^{90}(\omega_2)$ are each rejected (p-values 0.0000 and 0.0002, respectively). This demonstrates Receivers exhibit some strategic sophistication, but does not yet imply consistency with equilibrium play.

Condition (*ii*) applied with equally-likely states means $\hat{\pi}^p(\omega_1) = p$. Conditions (*iii*) and (*iv*) in Proposition 1 are tested through linear inequalities on the estimated success probabilities. Condition (iv) simply requires the excess-success probabilities to increase in p. As can be computed from Table 2, the estimated excess-success probabilities are strictly increasing: 0.054 in the 51% treatment, 0.290 in the 70% treatment, and 0.337 in the 90% treatment. Moreover, these excess-success probabilities are statistically different from each other, with a p-value of 0.0050 when comparing the 70% and 90% treatments, and p-values of 0.0000 when comparing the other two pairs of treatments.

Condition (iii) requires $\ell^p(\omega_2) + \hat{\pi}^p(\omega_1)(\ell^p(\omega_1) - \ell^p(\omega_2)) - \hat{\pi}^p(\omega_1) \geq 0$, which represents how much extra probability of success is attained beyond guessing Blue after each obfuscated message, for equilibrium beliefs $\hat{\pi}^p(\omega_1) = p$. The point estimates strictly satisfy these inequalities for the 51% and 70% treatments, with slack of 0.300 for the 51% treatment and 0.115 for the 70% treatment (both p-values 0.0000). The inequality is violated for the 90% treatment by 0.007 probability points, which despite being very small is statistically significant (p-value 0.0000). There are some demographic differences here, discussed further below. Of course, for fixed success rates, the condition (iii) is more demanding as the belief on ω_1 increases: for a high enough belief, it may in fact be optimal to simply pick action a_1. The precision p is the equilibrium belief, but not all subjects obfuscated in the state ω_1. In fact, condition (iii) would hold strictly (and with statistical significance) under the rational-expectations beliefs, and condition (iv) would be the same as before, since the rational-expectations beliefs also increase in p.[10]

We also consider the testable implications for Receiver's choices when restricting to subpopulations for whom equilibrium beliefs may be most natural. First, subjects who themselves followed Proposition 1(i) may have realized that strategic obfuscation is the only undominated sending strategy, and expect others to realize the same. Second, our exit survey asked subjects to verbally describe how they thought other Senders chose. We categorized these responses as either expecting Senders to follow Proposition 1(i), expecting some other behavior, or not clearly describing any behavior.[11] Among the respondents to the survey question, 84.4% of those who followed

[10] Isolating p, condition (iv) is tantamount to requiring that the probability on ω_1 be less than or equal to $\ell^p(\omega_2)/(1 + \ell^p(\omega_2) - \ell^p(\omega_2))$. The 95% confidence interval for this ratio in the 90% treatment is $(0.884, 0.894)$. The point estimates strictly satisfy the condition for any belief below 0.889, and one moreover rejects the null that the condition holds with equality for beliefs below 0.884. For our data, $\pi^{90}_{RE}(\omega) = 0.8122$ (for the other treatments, the rational-expectations beliefs are $\pi^{51}_{RE}(\omega) = 0.5069$ and $\pi^{70}_{RE}(\omega) = 0.6729$).

[11] The survey was presented at the end of the experiment, to avoid contaminating behavior in the

Proposition 1(i) themselves expected others to follow it, while only 23.1% of those who violated Proposition 1(i) expected others to follow it. We then estimated the success rates and tested conditions (iii) and (iv) for subjects who followed Proposition 1(i) themselves, and for subjects whose survey response expresses that Senders followed Proposition 1(i). The conditions cannot be rejected for each of these populations of Receivers (including (iii) in the 90% treatment); see the Online Appendix.

3.2.1 Demographic differences

We summarize here some demographic differences across men, women, economics majors and non-economics majors (see the Online Appendix for details). As noted earlier, we do not find differences in sending choices across these groups. Conditions (iii) and (iv) all hold strictly (and with statistical significance) for non-economics majors. The same is true for women, with the exception of the inequality (iii) for the 90% treatment, which is violated by 0.019 probability points. Men, on the other hand, satisfy condition (iii). The point estimates of both economics majors and men violate one of the three required monotonicity conditions associated to (iv): in both cases, the excess-success probability is strictly higher in the 70% treatment than the 90% treatment, though the difference is not statistically significant.[12]

At the same time, economics and non-economics majors act similarly in the 51% and 90% treatments. The main statistically significant differences appear in the intermediate, 70% treatment. In this intermediate-precision treatment, economics majors have less success than non-economics majors when the keyword is Red but better success when the keyword is Blue, suggesting relatively less perceptual effort and a greater reliance on strategic inference.

For women, success in Red clearly decreases in the precision level. For men, this stays fairly constant. The success probabilities of men and women are remarkably sim-

experiment itself. It was also un-incentivized, and did not specifically ask for the subject's level of certainty regarding their answer. Responses were examined and categorized by hand. A response is categorized as a "1" if the subject clearly describes the behavior in Proposition 1(i), or they followed Proposition 1(i) themselves and say that others acted just as they did. A response is categorized as a "0" if the subject describes some other sending strategy, or the subject did not follow Proposition 1(i) themselves and says that others acted just as they did. To be conservative, we did not count towards the above two categories responses that don't describe a sending strategy, such as "try to confuse the Receiver" or "try to make the most money," even though such responses are suggestive of Proposition 1(i).

[12]Women and men are approximately equally represented among economics majors, though the majority of other majors are women.

ilar in the 70%, intermediate-precision treatment, but differ in the extreme-precision treatments. In the 51% treatment, women's success probabilities are uniformly higher than men's, and their excess-success probability is smaller. These features are reversed in the 90% treatment. This suggests that men and women might differ in their perception costs. This may be interesting to investigate further in view of gender differences noted in Croson and Gneezy (2009).

4 Discussion

We established in this paper that when Senders have the option to obfuscate, Receivers do use strategic inferences to adjust their perceptual choices. In this concluding section, we point to some implications this has.

Suppose the Sender has no incentive to disclose information without being required to do so. This would be the case, for instance, if $\pi(\omega_2) = 51\% > 1/2$ (in which case the Receiver chooses the Sender's preferred action in the absence of communication). What are the welfare implications of mandating information disclosure when the regulator cannot prevent obfuscation as in our framework? To answer this question, a naive regulator might only think about the immediate informational content of obfuscated messages. For instance, he may assess that presented with a complex product label, a consumer has a 90% chance of correctly guessing whether the product is worth buying, which entails a small perception cost c.[13] According to the naive regulator, mandating disclosure results in the following ex-ante gain for the Receiver (assuming $m_R = 1$ without loss):

$$[\pi(\omega_1)(0.9p + (1-p)) + \pi(\omega_2)(p + 0.9(1-p)) - c] - \pi(\omega_2), \tag{5}$$

where the last term, $\pi(\omega_2)$, represents the success probability in the absence of disclosure (i.e., from taking action a_2 given that ω_2 is more likely in this example). But, as our analysis highlights, obfuscated messages reveal information beyond their immediate content, and average Receivers factor this in. To fix ideas, say that $p = 80\%$.

[13] In the context of our experiment, this would amount to guessing or eliciting success probabilities and the perception cost in an experiment à la Caplin and Dean (2014), without a Sender.

Then the rational Receiver's updated belief is

$$\hat{\pi}(\omega_1) = \frac{p\pi(\omega_1)}{p\pi(\omega_1) + (1-p)\pi(\omega_2)} \approxeq 79.35\%.$$

For simplicity, suppose that the Receiver's only alternative perception strategy is to overlook obfuscated messages and pick the action that matches the most-likely state (that is, all other strategies are assumed to be prohibitively costly). As easily checked, the rational Receiver does opt to overlook an obfuscated message when her belief is $\hat{\pi}$, because

$$\hat{\pi}(\omega_1) + p\hat{\pi}(\omega_2) > \hat{\pi}(\omega_1)(0.9p + (1-p)) + \hat{\pi}(\omega_2)(p + 0.9(1-p)) - c,$$

for any $c \geq 0$. To be consistent with the above discussion about the naive regulator, it must be that $0.9 - c > \pi(\omega_2)$, or $c < 0.39$, for the Receiver to prefer the costly perception strategy should her belief be π instead. The Receiver's true ex-ante welfare gain from mandated disclosure is

$$[\pi(\omega_1) + \pi(\omega_2)p] - \pi(\omega_2). \tag{6}$$

Subtracting (6) from (5), we get:

$$0.9\pi(\omega_2)(1-p) - 0.1\pi(\omega_1)p - c = 0.0526 - c,$$

which can be positive or negative depending on the value of c, which is only required to be smaller than 0.39. This simple example illustrates how *a naive regulator may overestimate in some circumstances, and underestimate in others, the Receiver's welfare gain from mandating information disclosure.*[14]

Notice that (5) minus (6) also captures the difference of utility between that of a naive Receiver who does not use strategic inferences to update her beliefs, and that of a

[14]Mandating disclosure is always welfare improving for the Receiver. Indeed, she keeps the option of picking ω_2 without paying attention to obfuscated messages, and she'll pick a different perception strategy only if it is welfare improving. As can be checked easily, the naive regulator also views mandating disclosure as welfare improving for the Receiver. But she can grossly missestimate the magnitude of the benefit, leading to misguided policy decisions when weighing those benefits against the welfare implications for the Sender and the cost of mandating disclosure. As c gets close to 0.39, for instance, the naive regulator thinks that the Receiver's ex-ante benefit from mandating disclosure is close to zero, while in fact it is close to 0.388.

Payoff Structure	*State* ω_1	*State* ω_2
Opposed	$(1-p)+p\ell(\omega_1)$	$(1-p)+p\ell(\omega_2)$
Persuasion	$(1-p)+p\hat{\ell}(\omega_1)$	$p+(1-p)\hat{\ell}(\omega_2)$
Common	$p+(1-p)\ell(\omega_1)$	$p+(1-p)\ell(\omega_2)$

Table 3: Receiver's success probabilities as a function of state and payoff structure

rational Receiver. The above example thus also illustrate how *strategic sophistication may be detrimental for the Receiver in terms of her actual ex-ante payoff.*

Our analysis focused on a persuasion payoff structure. While keeping the Receiver's payoffs unchanged, we could vary how aligned the Sender's interests are with the Receiver's. Interests are opposed (common) if the Sender's payoff in state ω_2 (resp., ω_1) had instead been m_S if the Receiver chooses a_1 and 0 otherwise. It is a dominant strategy for the Sender to try to obfuscate (clarify) when interests are opposed (resp., common). Hence, under each of those payoff structures, the rational Receiver's updated belief conditional on receiving an obfuscated message coincides with her prior π. Because the Receiver's choice of perception strategy for an obfuscated message depends only on her beliefs, the Receiver will choose the same perception strategy, with success probabilities denoted ℓ, under the opposed- and common-interests payoff structures. The Receiver's success probabilities under the persuasion-payoff structure are denoted $\hat{\ell}$. With these notations in mind, Table 3 indicates the ex ante probability the Receiver chooses the correct action as a function of the state and payoff structure. Our paper highlights that $\hat{\ell}$ is typically different from ℓ: a rational Receiver can adjust her optimal persuasion strategy based on strategic inference. Not recognizing this, one might suspect that success probabilities increase as the Sender and Receiver's preferences get more aligned (from opposed interests to persuasion, and from persuasion to common interests). *That intuition turns out to be wrong: greater alignment of preferences does not guarantee greater success for the Receiver.*

To see this, remember that in the above example, the Receiver overlooks an obfuscated message under $\hat{\pi}$ but pays a small cost c upon receiving an obfuscated message to guess the state correctly 90% of the time under π. In that case, the success probability (before getting a message, as in Table 3) is 92% (98%) for both ω_1 and ω_2 when interests are opposed (resp. common). But for the persuasion-payoff structure, the success probability is 1 in ω_1 and 80% in ω_2. Hence actual success can decrease in

ω_2 (resp., ω_1) when moving from opposed interests to persuasion (resp., persuasion to common interests). In fact, even the expected success probability can decrease: $0.51*0.8+0.49$ for the persuasion case is strictly inferior to 0.92 in the case of opposed interests.

Appendix

Proof of Proposition 1. It remains to show that (*iii*) and (*iv*) capture the empirical content of the Receiver's problem given his belief defined in (*i*). As explained in the text, the argument builds on Caplin and Dean (2015), but accounts for the facts that (a) probabilities vary across problems in our framework, and (b) pairwise comparisons in (*iv*) in our framework while Caplin and Dean must consider consider cycles of any size in their NIAC condition. We note that it is straightforward to generalize the proof below to allow the Receiver's benefit in state ω_i from guessing correctly to be $m_{S,i}$, showing that our result on pairwise comparisons relies only on the payoff structure being identical across observations, but not necessarily across states.[15] We proceed in three steps.

Step 1: The Receiver's choices are consistent with costly information acquisition under rational-expectations beliefs if, and only if, they are consistent with costly information acquisition in a transformed *individual* decision-making problem with *uniform* beliefs about the state and state-dependent payoffs. To see this, define $\Delta u_i^j = 2\hat{\pi}^j(\omega_i)(u_R(m_S) - u_R(0))$. With this notation, the objective function (3) for a persuasion game j is:

$$u_R(0) + \frac{1}{2}\mu(\mathcal{S}_1|\omega_1)\Delta u_1^j + \frac{1}{2}\mu(\mathcal{S}_2|\omega_2)\Delta u_2^j - c_R(\mathcal{S}_1, \mathcal{S}, \mu),$$

and the constraint that the Receiver prefers action a_i following signal $\sigma \in \mathcal{S}_i$ can be written $\mu(\sigma|\omega_i)\Delta u_i^j \geq \mu(\sigma|\omega_{-i})\Delta u_{-i}^j$, $\forall \sigma \in \mathcal{S}_i$, $\forall i = 1, 2$. Thus, the problem in game j is equivalent to one with uniform beliefs over states, after rescaling the payoff gain in state i from choosing correctly to Δu_i^j.

Step 2: Using Caplin and Dean (2015, Theorem 1), consistency in the transformed

[15] The proof follows the same steps, but must instead define $\Delta u_i^j = 2\hat{\pi}^j(\omega_i)(u_R(m_{S,i}) - u_R(0))$ and $\Delta \ell^j = \ell^j(\omega_1)(u_R(m_{S,1}) - u_R(0)) - \ell^j(\omega_2)(u_R(m_{S,2}) - u_R(0))$.

problem is equivalent to the Receiver's data satisfying their NIAC and NIAS conditions. Translated to our setting and notation, and using the fact that $p > 1/2$, NIAS corresponds to condition (*iii*) in Proposition 1, while NIAC corresponds to the condition that for any integer $J \geq 2$ and any J-length cycle $(p_1, p_2, \ldots, p_J, p_1)$ of persuasion games,

$$\sum_{j=1}^{J} \left(\ell^j(\omega_1) - \ell^{j+1}(\omega_1)\right) \Delta u_1^j \geq \sum_{j=1}^{J} \left(\ell^j(\omega_2) - \ell^{j+1}(\omega_2)\right) \Delta u_2^j, \tag{7}$$

where $\ell^{J+1} = \ell^1$ by convention.

Step 3: We now show that in our setting, (7) reduces to the pairwise condition stated in Proposition 1(*iv*). We begin by developing (7) into the condition

$$\sum_{j=1}^{J} \left(\ell^j(\omega_1) - \ell^{j+1}(\omega_1) - \ell^j(\omega_2) + \ell^{j+1}(\omega_2)\right) \hat{\pi}^j(\omega_1) \geq \sum_{j=1}^{J} \left(\ell^j(\omega_2) - \ell^{j+1}(\omega_2)\right) = 0,$$

by using $\hat{\pi}^j(\omega_2) = 1 - \hat{\pi}^j(\omega_1)$ and cancelling the common factor $2(u_R(m_S) - u_R(0))$ in the Δu_i^j's. Letting $\Delta\ell^j = \ell^j(\omega_1) - \ell^j(\omega_2)$, (7) is then equivalent to:

$$\sum_{j=1}^{J} \left(\Delta\ell^j - \Delta\ell^{j+1}\right) \hat{\pi}^j(\omega_1) \geq 0. \tag{8}$$

For $J = 2$ (that is, a cycle (p^j, p^k, p^j)), condition (8) reduces to condition (*iv*):

$$(\hat{\pi}^j(\omega_1) - \hat{\pi}^k(\omega_1))(\Delta\ell^j - \Delta\ell^k) \geq 0. \tag{9}$$

We now prove by induction that if (9) holds for all pairs of persuasion games, then (8) must be satisfied for any cycle length $J > 2$. Suppose (8) holds for all cycles of length $J - 1$, and consider a cycle of length J. For notational convenience we may translate the elements of the cycle so that the J-th element corresponds to the highest level of $\hat{\pi}^j$ (the sum in (8) is invariant to where the cycle begins). Notice that

$\sum_{j=1}^{J} \left(\Delta\ell^j - \Delta\ell^{j+1}\right) \hat{\pi}^j(\omega_1)$ can be decomposed into:

$$\sum_{j=1}^{J-1} \left(\Delta\ell^j - \Delta\ell^{j(mod(J-1))+1}\right) \hat{\pi}^j(\omega_1) - \left(\Delta\ell^{J-1} - \Delta\ell^1\right) \hat{\pi}^{J-1}(\omega_1) \tag{10}$$
$$+ \left(\Delta\ell^{J-1} - \Delta\ell^J\right) \hat{\pi}^{J-1}(\omega_1) + \left(\Delta\ell^J - \Delta\ell^1\right) \hat{\pi}^J(\omega_1).$$

The first term in (10) corresponds to the sum in (8) for the $(J-1)$-length cycle $(p_1, p_2, \ldots, p_{J-1}, p_1)$ that omits p_J; this is nonnegative by the inductive hypothesis. To get back to the sum in (8) for the original cycle, the second term removes the link from p^{J-1} to p^J, and the next two terms add back the links from p^{J-1} to p^J, and from p^J to p^1. These final three terms in (10) sum to $\left(\hat{\pi}^J(\omega_1) - \hat{\pi}^{J-1}(\omega_1)\right)\left(\Delta\ell^J - \Delta\ell^1\right)$. By our choice of numbering scheme, $\hat{\pi}^J(\omega_1)$ is maximal among all $\hat{\pi}^j(\omega_1)$, so the first factor in this product is nonnegative. Similarly, the pairwise condition (9) applied to $(1, J)$ ensures $\Delta\ell^J \geq \Delta\ell^1$, so the second factor is nonnegative too. Thus (10) is nonnegative, implying (8) holds for the J-length cycle. □

References

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

Bordalo, P., N. Gennaioli, and A. Shleifer (2015). Competition for attention. *The Review of Economic Studies 83*(2), 481–513.

Caplin, A. and M. Dean (2015). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review 105*(7), 2183–2203.

Caplin, A. and D. Martin (2015). A testable theory of imperfect perception. *The Economic Journal 125*(582), 184–202.

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

Croson, R. and U. Gneezy (2009). Gender differences in preferences. *Journal of Economic Literature 47*(2), 448–74.

de Clippel, G., K. Eliaz, and K. Rozen (2014). Competing for consumer inattention. *Journal of Political Economy 122*(6), 1203–1234.

de Oliveira, H., T. Denti, M. Mihm, and K. Ozbek (2017). Rationally inattentive preferences and hidden information costs. *Theoretical Economics 12*, 621–654.

Dean, M. and N. Nelighz (2019). Experimental tests of rational inattention. *Working Paper*.

Dewatripont, M. and J. Tirole (2005). Modes of communication. *Journal of Political Economy 113*(6), 1217–1238.

Ellis, A. (2018). Foundations for optimal inattention. *Journal of Economic Theory 173*, 56–94.

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

Fréchette, G. R., A. Lizzeri, and J. Perego (2019). Rules and commitment in communication: An experimental analysis. *Working paper*.

Gabaix, X. and D. Laibson (2006). Shrouded attributes, consumer myopia, and information suppression in competitive markets. *The Quarterly Journal of Economics 121*(2), 505–540.

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

Jin, G. Z., M. Luca, and D. Martin (2016). Is no news (perceived as) bad news? an experimental investigation of information disclosure. *Working Paper*.

Jin, G. Z., M. Luca, and D. Martin (2019). Complex disclosure. *Working Paper*.

Martin, D. (2016a). Rational inattention in games: Experimental evidence. *Working paper*.

Martin, D. (2016b). Strategic pricing with rational inattention to quality. *Games and Economic Behavior 104*, 131–145.

Matějka, F. (2015). Pricing and rationally inattentive consumer. *Journal of Economic Theory 158*.

Milgrom, P. and J. Roberts (1986). Relying on the information of interested parties. *Rand Journal of Economics 17*(1), 18–32.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979). The belmont report: Ethical principles and guidelines for the protection of human subjects of research.

Persson, P. (2018). Attention manipulation and information overload. *Behavioural Public Policy 2*(1), 78–106.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*, 665–690.

Spiegler, R. (2006). Competition over agents with boundedly rational expectations. *Theoretical Economics 1*(2), 207–231.

Online Appendix

Communication, Perception and Strategic Communication

Geoffroy de Clippel and Kareen Rozen

We provide below some additional analysis mentioned in the main paper. We also provide the paper instructions that were handed out for the experiment, as well as screenshots of the z-Tree experimental interface. Both are for the 70% treatment. The other treatments only change the value 70% to either 51% or 90%, as applicable.

A Receivers who expect Proposition 1(i)

We now estimate success rates among only subjects who follow Proposition 1(i) when acting as Senders, as well as among only subjects whose survey response (describing what they thought other Senders did) was placed in the Proposition 1(i)-category. We use the same logistic regression model with robust standard errors, clustered by session. We estimate:

	Based on Sending choice			Based on survey response		
Precision p	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$
51%	0.889	0.775	0.114	0.849	0.759	0.090
70%	0.878	0.635	0.243	0.879	0.626	0.252
90%	0.942	0.510	0.432	0.971	0.583	0.387

Table 4: The estimated state-dependent success probabilities of Receivers per treatment, and excess-success probabilities, estimated only among the subpopulation of subjects who followed Proposition 1(i) as Senders, and among the subpopulation of subjects who anticipated Proposition 1(i) according to the exit survey.

Consider first subjects who followed Proposition 1(i) themselves. Among these subjects, we have 1,896 observations of choices in obfuscated messages, with 581 in the 51% treatment, 731 in the 70% treatment and 584 in the 90% treatment. For each pair of treatments, among this subset of subjects the null hypotheses that $\ell^{51}(\omega_1) = \ell^{70}(\omega_1) = \ell^{90}(\omega_1)$ and $\ell^{51}(\omega_2) = \ell^{70}(\omega_2) = \ell^{90}(\omega_2)$ are rejected (p-values 0.0148 and 0.0000, respectively). We use $\hat{\pi}^p(\omega_1) = p$, following condition (ii). As seen from Table 3, the excess-success probabilities indeed increase with the precision level

p, verifying condition (iv). For each of the three pairwise comparisons of treatments, the excess success probabilities are significantly different from each other (all p-values 0.0000). As for condition (iii), the left-hand side is strictly larger than the right-hand side by 0.323 for the 51% treatment ($p = 0.0000$) and 0.105 for the 70% treatment ($p = 0.0001$), and is smaller by 0.0015 for the 90% treatment, which is not statistically different from zero ($p = 0.8801$).

Now consider subjects whose survey response indicates they think other Senders followed Proposition 1(i). Among these subjects, we have 1,629 observations of choices in obfuscated messages, with 478 in the 51% treatment, 657 in the 70% treatment and 494 in the 90% treatment. For each pair of treatments, among this subset of subjects the null hypotheses that $\ell^{51}(\omega_1) = \ell^{70}(\omega_1) = \ell^{90}(\omega_1)$ and $\ell^{51}(\omega_2) = \ell^{70}(\omega_2) = \ell^{90}(\omega_2)$ are rejected (p-values 0.0001 and 0.0208, respectively). We use $\hat{\pi}^p(\omega_1) = p$, following condition (ii). As seen from Table 3, the excess-success probabilities indeed increase with the precision level p, verifying condition (iv). These excess-success probabilities are statistically different when comparing the 51% treatment with either the 70% or 90% treatments (p-values 0.0020 and 0.0003, respectively), and not statistically different when comparing the 70% and 90% treatments (p-value 0.1627). As for condition (iii), the left-hand side is strictly larger than the right-hand side by 0.295 for the 51% treatment (p-value 0.0000) and 0.104 for the 70% treatment (p-value 0.0004), and by 0.032 for the 90% treatment (p-value 0.0006).

B Demographic analysis of Receiver behavior

B.1 By major

We estimate success probabilities for subjects who report an economics major (E) and those who do not (N), using a logistic regression that interacts the keyword, the treatment, and an indicator for an economics major. We use robust, Huber-White standard errors, clustered by session. The breakdown of reported majors was 32N/7E in the 51% treatment, 30N/13E in the 70% treatment, and 22N/19E in the 90% treatment. We observe 718 obfuscated messages for economics majors and 1589 for non-economics majors.

For non-economics majors, each of the null hypotheses $\ell^{51}(\omega_1) = \ell^{70}(\omega_1) = \ell^{90}(\omega_1)$

	Economics majors			Non-economics majors		
Precision p	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$
51%	0.857	0.707	0.150	0.822	0.817	0.005
70%	0.964	0.529	0.435	0.841	0.730	0.112
90%	0.933	0.589	0.344	0.928	0.649	0.280

Table 5: The estimated state-dependent success probabilities of Receivers per treatment, and excess-success probabilities, for economics and other majors.

and $\ell^{51}(\omega_2) = \ell^{70}(\omega_2) = \ell^{90}(\omega_2)$ is rejected (p-values 0.0001 and 0.0000, respectively), and the data clearly support conditions (*iii*) and (*iv*). For (*iv*), the point estimates of the excess-success probabilities strictly increase with p, and the differences are statistically significant, with p-values of 0.0000 for each pairwise comparison involving 90% and a p-value of 0.0001 for the 51% to 70% comparison. For (*iii*), the slack in the inequality is 0.3099 for the 51% treatment (p-value 0.0000), 0.1080 for the 70% treatment (p-value 0.0000), and 0.0002 for the 90% treatment (p-value 0.9904).

For economics majors, the null hypotheses that $\ell^{51}(\omega_1) = \ell^{70}(\omega_1) = \ell^{90}(\omega_1)$ and $\ell^{51}(\omega_2) = \ell^{70}(\omega_2) = \ell^{90}(\omega_2)$ cannot be rejected (p-values 0.1901 and 0.2383, respectively). As for condition (*iv*), the point estimates for the excess-success probabilities violate monotonicity in going from the 70% treatment to the 90% treatment. None of the pairwise differences is statistically significant, with p-values of 0.0907 when comparing the 51% and 70% treatments, 0.2612 when comparing the 51% and 90% treatments, and 0.3922 when comparing the 70% and 90% treatments. As for the optimality requirement (*iii*), the left-hand side exceeds the right-hand side by 0.2737 for the 51% treatment (p-value 0.0000) and by 0.1340 for the 70% treatment (p-value 0.0000), and is smaller than the right-hand side by 0.0011 for the 90% treatment (p-value 0.9563).

We now contrast the estimates for economics and other majors. In the 90% treatment, neither $\ell^{90}(\omega_1)$ nor $\ell^{90}(\omega_2)$ is statistically different across majors (p-values 0.9089 and $p = 0.2413$, respectively); and similarly for the 51% treatment, though the p-value for $\ell^{51}(\omega_2)$ is marginal at 0.0511 (the other p-value is 0.7566). The main difference is in the 70% treatment, where the differences in each of $\ell^{70}(\omega_1)$ and $\ell^{70}(\omega_2)$ across majors are statistically significant (both p-values are 0.0000). Without conditioning on the realized state, non-economics majors have slightly higher estimated expected success rates than economics majors, in all but the 70% treatment: the difference in the expected probability of success for an obfuscated message is 0.0362 in the 51% treatment (p-value 0.2217), -0.0260 in the 70% treatment (p-value 0.0671), and 0.0013 in the 90% treatment (p-value 0.9718).

B.2 By gender

We estimate success probabilities for subjects who self-report as women (W) and those who self-report as men (M), using a logistic regression that interacts the keyword, the treatment, and an indicator for gender. We use robust, Huber-White standard errors, clustered by session. The breakdown of reported genders per treatment is 19W/18M in the 51% treatment, 29W/14M in the 70% treatment, and 28W/13M in the 90% treatment. We observe 1429 obfuscated messages for women and 844 for men.

	Men			Women		
Precision p	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$	$\ell^p(\omega_1)$	$\ell^p(\omega_2)$	$\ell^p(\omega_1) - \ell^p(\omega_2)$
51%	0.767	0.689	0.078	0.889	0.867	0.021
70%	0.896	0.656	0.240	0.871	0.669	0.202
90%	0.965	0.750	0.215	0.913	0.588	0.325

Table 6: The estimated state-dependent success probabilities of Receivers per treatment, and excess-success probabilities, for economics and other majors.

For women, the estimated excess-success probabilities strictly increase, as required by condition (*iv*); and for each pair of treatments, these estimates are statistically different from each other at all significance levels (p-values 0.0000). For the 51% and 70% treatments, the basic optimality condition (*iii*) holds with slack, which is statistically different from zero (p-values 0.0000). For the 90% treatment, the inequality is reversed with the probability loss equal to 0.019, which is small but statistically significant (p-value 0.0294). For men, the point estimates strictly satisfy the basic optimality conditions (*iii*), with the slack statistically different from zero at all significance levels (p-values 0.0000) for the 51% and 70% treatments, and the slack of 0.0435 probability points not statistically different from zero for the 90% treatment (p-value 0.0727). We also see non-monotonicity in the excess-success probabilities going from the 70% to the 90% treatments, though the difference is not statistically significant (p-value 0.7135); the differences are statistically significant between the 51% and 70% treatments (p-value 0.0198) and between the 51% and 90% treatments (p-value 0.0000). For men, the null that success rates are the same for ω_2 across the treatments cannot be rejected at the 5% level (p-value 0.0871) but the corresponding null is rejected for ω_1 (p-value 0.0000). For women, we cannot reject that the success rates in ω_1 are all the same (p-value 0.1069) but we can reject at all significance levels that the success rates in ω_2 are the same (p-value 0.0000).

We next compare, for each treatment p and state ω_i, the estimates $\ell^p(\omega_i)$ by gender. For the 70% treatment, neither $\ell^{70}(\omega_1)$ nor $\ell^{70}(\omega_2)$ is statistically different across gender at any conventional significance levels, with p-values of 0.5739 and 0.6638, respectively. For the 51% treatment, there are statistically significant differences for both $\ell^{51}(\omega_1)$ and $\ell^{51}(\omega_2)$ across genders, with p-values of 0.0000 in both cases. In the 90% treatment, the estimates of $\ell^{90}(\omega_1)$ are not statistically different across genders (p-value 0.1059) but the estimates of $\ell^{90}(\omega_2)$ are (p-value 0.0002).

We next consider differences in expected payoffs across genders, given equilibrium beliefs. In the 51% treatment, where females have higher conditional success rates in each state, the women's expected success rate is higher by 0.149 probability points, which is significant at all levels (p-value of 0.0000). In the 70% treatment, where the conditional success rates are not different, the 0.014 difference in expected success probability of men over women is indeed not statistically significant in (p-value 0.5391). In the 90% treatment, the men's expected success rate is higher by 0.063, which is marginally insignificant (p-value 0.0578).

Welcome to this decision-making experiment!

Please silence and put away electronic devices. Please do not talk with other participants.

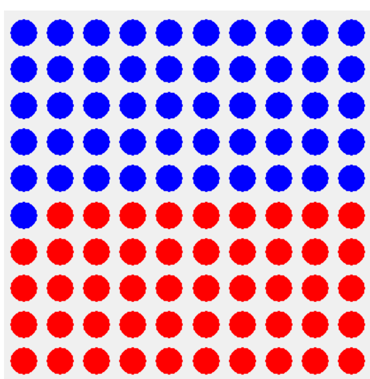
Instructions

You will receive a $10 show-up fee, and will be able to earn more. The exact amount earned will depend on chance and choices made during the experiment.

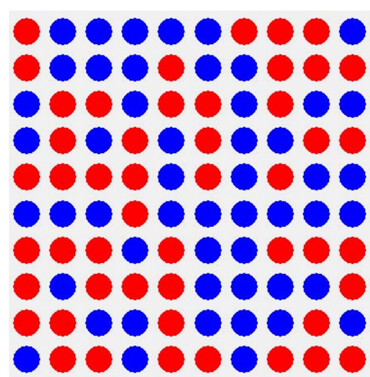
Throughout this experiment, participants will have the opportunity to play what we call `Sender-Receiver games'. We will explain shortly how you'll be able to input your decisions as Senders and Receivers, and how participants will be paired to play these games. First, we describe the game itself.

Sender-Receiver game

The computer randomly draws one of two possible keywords, either `Red' or `Blue', with a 50% chance of each. The Receiver guesses the keyword after seeing a message with 100 balls (of blue and red color). When the keyword is Red, the message contains 51 red balls and 49 blue balls (i.e., red is the majority color). When the keyword is Blue, the message contains 51 blue balls and 49 red balls (i.e., blue is the majority color). However, the balls displayed in the message can either be arranged by color, or in random order (see the two examples below for the keyword Blue).



Arranged by Color



In Random Order

How balls are displayed is impacted by the Sender's decision in the following way. He or she chooses between making it more likely (70% chance) that balls are arranged by color, or making it more likely (70% chance) that balls are arranged in random order. The Sender can make his or her choice depend on whether the keyword is Red or Blue.

Payoffs of a particular Sender-Receiver match will be computed as follows. The Sender gets $15 if the Receiver guesses that the keyword is Red, and $0 if the Receiver guesses Blue. The Receiver gets $15 for guessing the keyword correctly, and $0 otherwise.

How will the Sender-Receiver games be played?

To allow everyone to have a chance to play as a Sender and a chance to play as a Receiver, the session is split into two phases: a Sending phase and a Receiving phase. Decisions made during the Sending phase will be used to determine messages in the Receiving phase of the experiment.

(a) Sending Phase:

In this phase, each participant acts as a Sender, and decides whether balls in the Receiver's message are more likely (70% chance) to be arranged by color, or more likely (70%) to be displayed randomly. You'll be asked to make this decision twice, first for the keyword Blue, and second for the keyword Red.

(b) Receiving Phase:

In this phase, each participant acts as a Receiver. You will be matched 40 times, each time with a randomly drawn Sender other than yourself.

In each match, the computer randomly draws the keyword (with a 50% chance of Blue and a 50% chance of Red), and implements the Sender's decision for the drawn keyword. That is, the computer uses the relative likelihood the Sender picked for that keyword, to display a message with balls arranged either by color, or randomly. You'll be asked to guess the keyword in each match.

What happens at the end of the experiment?

Once the Receiving Phase is complete, there will be a short and optional exit survey. Your participation is voluntary and does not affect your payoff.

At the end of the experiment, the computer randomly picks a role for you (Sender or Receiver) and randomly chooses one match in which you played that role. You will receive your payoff from that match in addition to the $10 show-up fee. All identities remain anonymous. No one will learn what role you played or what payoff you earned.

We are almost ready to start the experiment. Before doing so, there will be a short quiz to check your understanding of some key features of the experiment, as well as a chance to familiarize yourself with the interface.

Welcome

Thank you for participating in this experiment.

Before starting the experiment, there will be a short quiz to check comprehension of the instructions, as well as an opportunity to familiarize yourself with the interface.

Feel free to consult your instruction sheet, both during the quiz and the actual experiment.

Start

Comprehension questions

Question 1

Suppose that the Receiver guesses that the keyword is Red, in a match where the keyword is actually Blue.

Not including the showup fee, what would be the Sender's payoff from this match (in dollars)?

Not including the showup fee, what would be the Receiver's payoff from this match (in dollars)?

Continue

Question 2

Suppose that the Receiver guesses that the keyword is Blue, in a match where the keyword is actually Blue.

Not including the showup fee, what would be the Sender's payoff from this match (in dollars)?

Not including the showup fee, what would be the Receiver's payoff from this match (in dollars)?

Continue

Question 3

Suppose that the Receiver guesses that the keyword is Red, in a match where the keyword is actually Red.

Not including the showup fee, what would be the Sender's payoff from this match (in dollars)?

Not including the showup fee, what would be the Receiver's payoff from this match (in dollars)?

Continue

Question 4

Suppose that the Receiver guesses that the keyword is Blue, in a match where the keyword is actually Red.

Not including the showup fee, what would be the Sender's payoff from this match (in dollars)?

Not including the showup fee, what would be the Receiver's payoff from this match (in dollars)?

Continue

Practice

Everyone has now finished the quiz. Before starting the experiment, you will now have the opportunity to familiarize yourself with the interface.

In the next screen, you will see a practice screen for the Sending Phase. To make a selection for each keyword, simply click on your desired buttons. Your selections will be highlighted.

The interface allows you to change your selections as many times as you like, as long as you have not yet clicked the Submit button. You will not be able to click Submit until you have made selections.

Since the next screen is only for practice, your choices on it have no bearing whatsoever on the actual experiment.

Start

Practice Screen for the Sending Phase

Please make a messaging decision for each of the two possible keywords. Your selections will be highlighted. To finalize them, press the submit button below.

When the keyword is Blue, how would you like the balls to be arranged?

By color, with probability 0.7,
and randomly otherwise.

Randomly, with probability 0.7,
and by color otherwise.

When the keyword is Red, how would you like the balls to be arranged?

By color, with probability 0.7,
and randomly otherwise.

Randomly, with probability 0.7,
and by color otherwise.

Submit

Practice

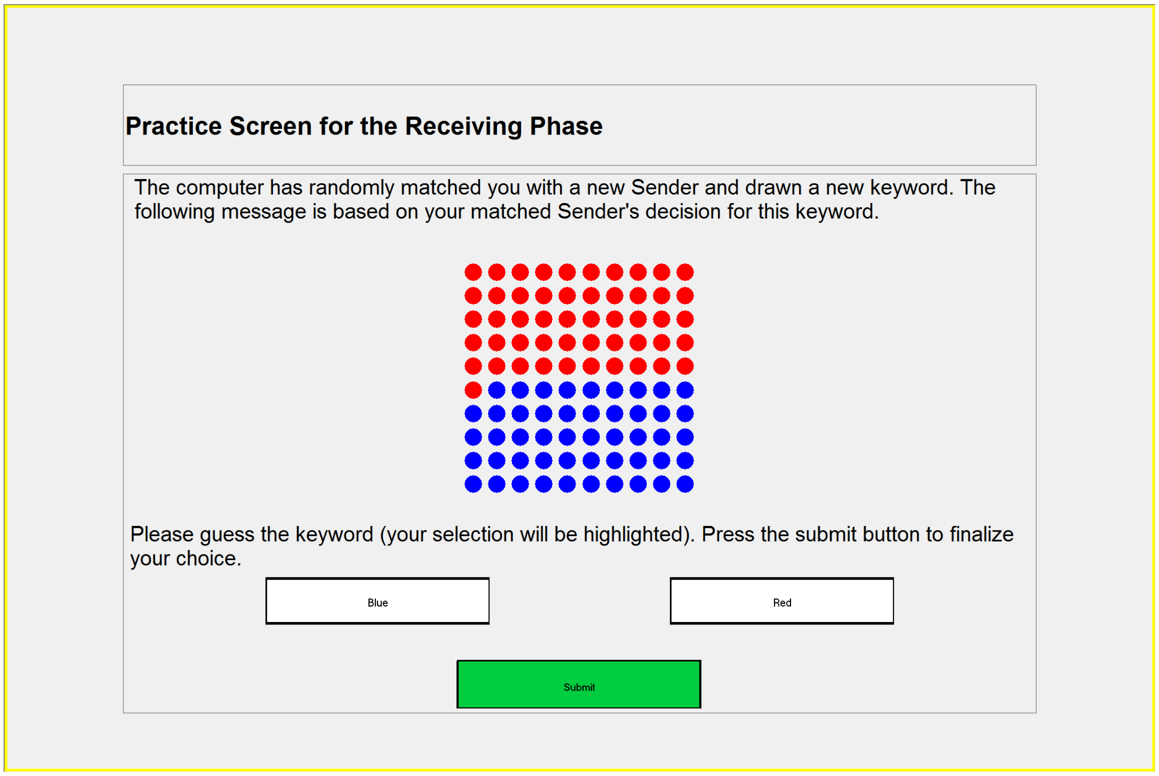
Next, you will see two example screens for the Receiving phase.

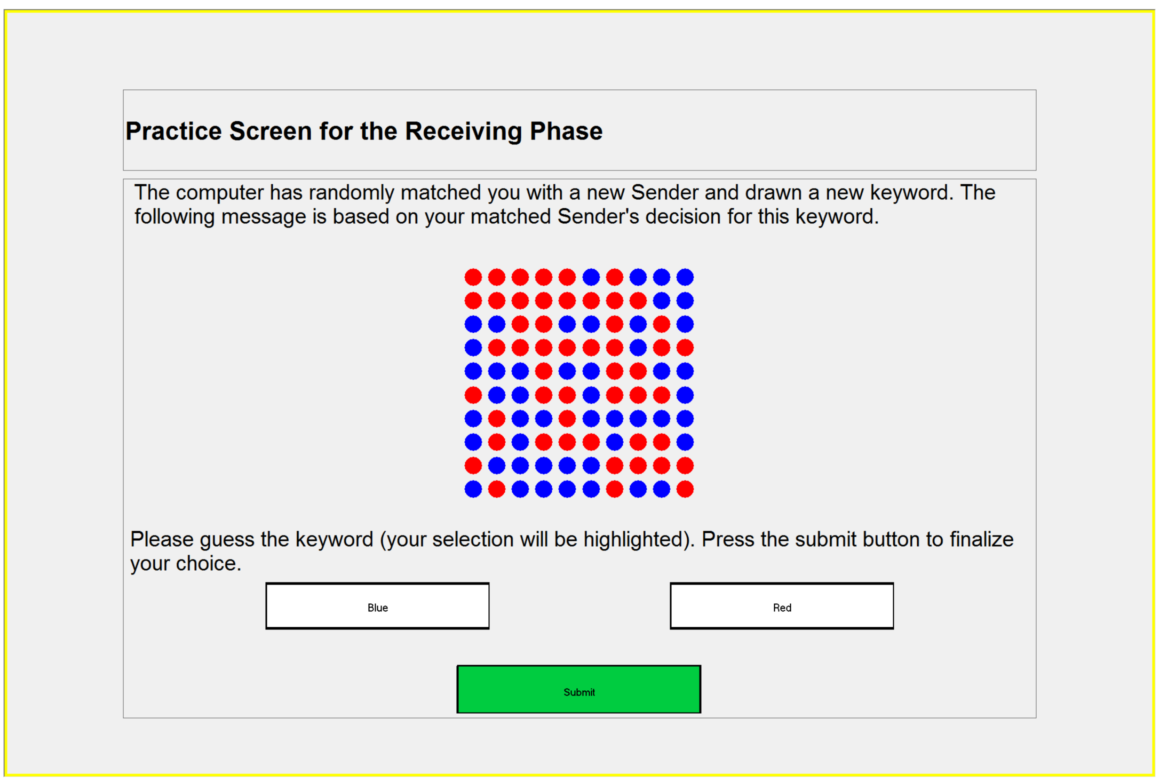
To guess the keyword on each screen, simply click on your desired button. Your selection will be highlighted in the color corresponding to your guess.

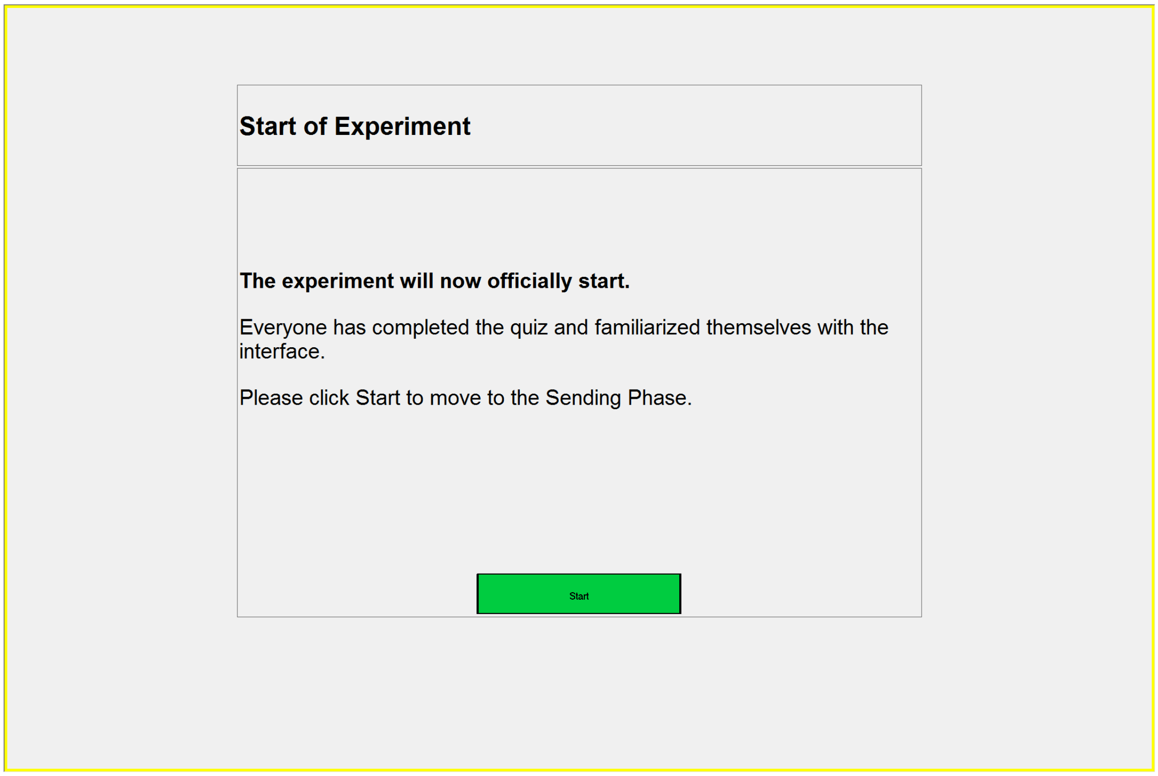
The interface allows you to change your guess on a screen as many times as you like, as long as you have not yet clicked the Submit button to move to the next screen. You will not be able to click Submit until you have made a guess on the screen.

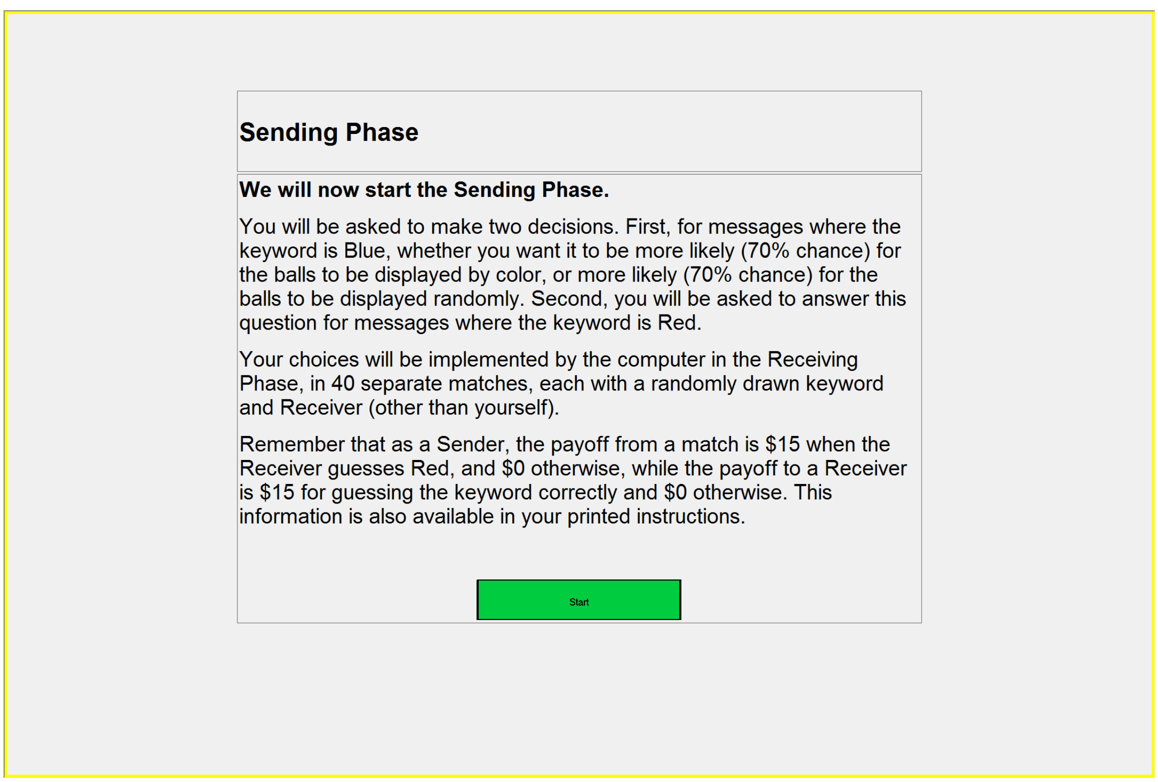
Since the next two screens are only for practice, your choices on them will have no bearing whatsoever on the actual experiment.

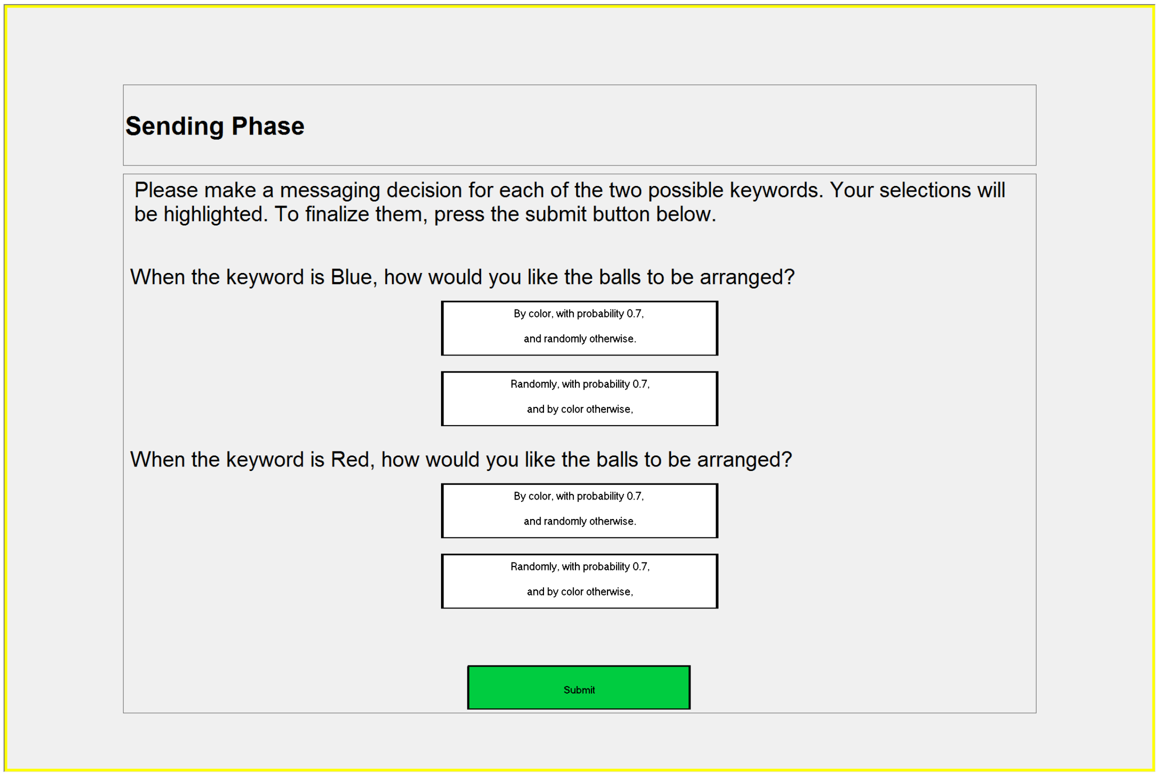
Start

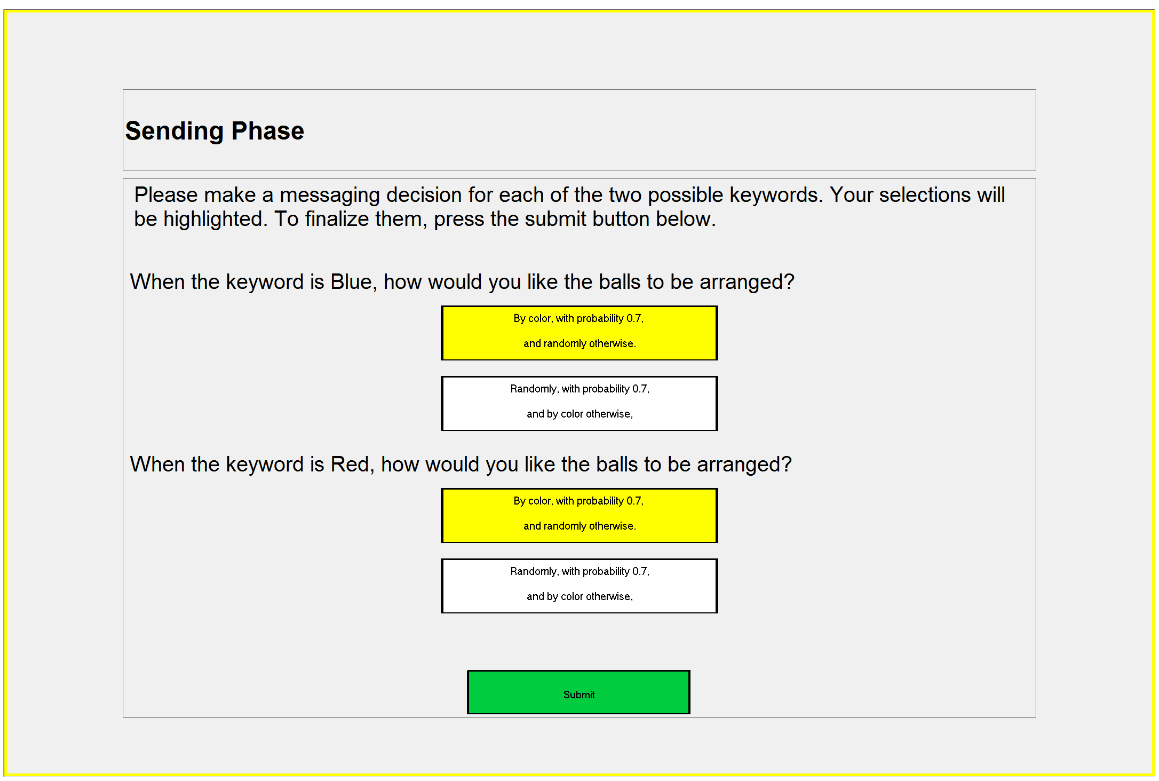

Practice Screen for the Receiving Phase
The computer has randomly matched you with a new Sender and drawn a new keyword. The following message is based on your matched Sender's decision for this keyword.
Please guess the keyword (your selection will be highlighted). Press the submit button to finalize your choice.
Blue
Red
Submit


Practice Screen for the Receiving Phase
The computer has randomly matched you with a new Sender and drawn a new keyword. The following message is based on your matched Sender's decision for this keyword.
Please guess the keyword (your selection will be highlighted). Press the submit button to finalize your choice.
Blue
Red
Submit


Start of Experiment
The experiment will now officially start.
Everyone has completed the quiz and familiarized themselves with the interface.
Please click Start to move to the Sending Phase.
Start


Sending Phase
We will now start the Sending Phase.
You will be asked to make two decisions. First, for messages where the keyword is Blue, whether you want it to be more likely (70% chance) for the balls to be displayed by color, or more likely (70% chance) for the balls to be displayed randomly. Second, you will be asked to answer this question for messages where the keyword is Red.
Your choices will be implemented by the computer in the Receiving Phase, in 40 separate matches, each with a randomly drawn keyword and Receiver (other than yourself).
Remember that as a Sender, the payoff from a match is $15 when the Receiver guesses Red, and $0 otherwise, while the payoff to a Receiver is $15 for guessing the keyword correctly and $0 otherwise. This information is also available in your printed instructions.
Start


Sending Phase
Please make a messaging decision for each of the two possible keywords. Your selections will be highlighted. To finalize them, press the submit button below.
When the keyword is Blue, how would you like the balls to be arranged?
By color, with probability 0.7, and randomly otherwise.
Randomly, with probability 0.7, and by color otherwise,
When the keyword is Red, how would you like the balls to be arranged?
By color, with probability 0.7, and randomly otherwise.
Randomly, with probability 0.7, and by color otherwise,
Submit


Sending Phase
Please make a messaging decision for each of the two possible keywords. Your selections will be highlighted. To finalize them, press the submit button below.
When the keyword is Blue, how would you like the balls to be arranged?
By color, with probability 0.7, and randomly otherwise.
Randomly, with probability 0.7, and by color otherwise,
When the keyword is Red, how would you like the balls to be arranged?
By color, with probability 0.7, and randomly otherwise.
Randomly, with probability 0.7, and by color otherwise,
Submit

Receiving Phase

Thank you. We will now start the Receiving Phase.

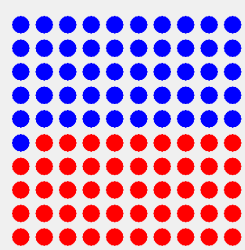
The computer will now match you 40 times with randomly drawn Senders (other than yourself). For each match, the computer randomly draws a new keyword and displays a message based on that Sender's decision for this keyword. You will be asked to guess the keyword in each match.

Remember that as a Receiver, the payoff from a match is $15 for guessing the keyword correctly and $0 otherwise, while the payoff to a Sender is $15 when the Receiver guesses Red, and $0 otherwise. This information is also available in your printed instructions.

Start

Match 1 of 40

The computer has randomly matched you with a new Sender and drawn a new keyword. The following message is based on your matched Sender's decision for this keyword.



Please guess the keyword (your selection will be highlighted). Press the submit button to finalize your choice.

Blue | Red

Submit

Receiving Phase

Please wait while the computer draws your next Sender and keyword.

Outcome of the experiment

Thank you for your participation.

The computer has randomly determined that you will be paid for your match number 23 in your role as a Sender. Including the showup fee of $10, your total payoff for this experiment is $25.

Click the OK button for the experimenter to receive this information and prepare payments. Please wait while payments are being prepared. In the meantime, we would appreciate your participation in the optional exit survey on the next page.

OK

1. What is your major (concentration)?

2. What is your gender?

3. How many semesters of college or university (not including this semester) have you completed?

4. How many economics courses have you taken?

5. What was your SAT math score?

6. Have you participated in an economic experiment before? (Yes/No)

Continue

7. How did you decide what to do in the Sending Phase?

8. How did you decide what to do in the Receiving Phase?

9. What did you expect other people would do when they were Senders?

10. What did you expect other people would do when they were Receivers?

11. Feel free to leave any comment on this experiment below:

Continue