Strategies of Disclosure Timing

Rivin Perinchery*

October 17, 2022

Abstract

This paper extends the signaling and information design literature by introducing and allowing variation in *disclosure timing* as a component of the disclosure and signal itself. Many models assume that disclosure timing is not a relevant signal component, an assumption which may be unreasonable in some environments as shown by Konovalov and Krajbich (2020). We relax that assumption in our environment. Building on the model of Spence (1973), we consider a sender with perfect information who reports the state, but can utilize an imperfect test (in the spirit of Kamenica and Gentzkow (2011)) when they observe that the state is bad to attempt generating a report of 'good'. However, conducting the test requires time which implies that disclosure time can reveal the origin of a report to the receiver (and thus inform them of the state when the test's utilization is state-dependent). Our model is the first to allow senders to directly choose their disclosure timing endogenously. Here, we model interaction with delayed testing (and costly 'quick testing') and solve for two 'pooling' equilibria. We identify one equilibrium which improves the senders payoff ex ante and thus is expected to dominate. Consistent with theoretical benchmarks, we find evidence of pooling in our treatment sessions. However, the pooling strategy that senders choose is weakly dominated. This finding suggests that senders understand the importance of strategically manipulating their time to prevent unintended information disclosures, but are missing a pivotal insight in maintaining persuasion given the cost structure.

JEL Codes: C72, C91, D01, D82, D83, M21

*Department of Economics, Florida State University

The author gratefully acknowledges financial support from the John and Hallie Quinn Foundation to conduct this research. The guidance and direction of Luke Boosey in many discussions is acknowledged with immense gratitude. The author also wishes to thank David Cooper, Marie Claire Villeval, John Hamman, R. Vijay Krishna, R. Mark Isaac, Matthew Gentry, Levi Soborowicz, Eugenio Verrina, Benoit Tarroux, Joe Ballard, Thomas Frye, Bart Lipman, and participants of the Graduate Student Brownbag Workshop, XSFS Workshop, and Third Year Paper Conference, respectively, at FSU for their helpful comments and conversations. Lastly, the author extends his gratitude to the graduate students at FSU who assisted in testing the experimental program. This research was conducted with approval from the IRB at FSU (STUDY00003118). The author has no conflicts of interest to declare.

"I came in here planning on taking a lot of time" - Vrishank Subramani[1]

1 Introduction

When the availability of certain signals is state-dependent, what strategies does a sender utilize to manipulate the revelation of private information to a receiver she wishes to persuade? More precisely, when some signals take longer to obtain than others depending on the state, how can a sender manipulate her signal *disclosure/reporting time* to control the revelation of her private information? Consider an entrepreneur who has invented a product which is either *good* or *bad* and is seeking an investment. Suppose that initially the investor is pessimistic about the product's quality and is unwilling to invest. Naturally, the entrepreneur seeks to persuade the investor that the product is good and worthy of investment regardless of its quality while the investor prefers to invest only if the product is indeed good. To persuade the investor, the entrepreneur must present *evidence* that the product is good. When the product is good, she always has this evidence and can present it. When the product is not good, she does not have any evidence to present. However, the entrepreneur has a team of scientists working for her that can conduct extended trials in an effort to produce evidence for her. The entrepreneur and her scientists are subject to strict audit and cannot fabricate evidence. However, by random chance they may obtain evidence *even though the product is bad*.

If scientist error rates are 'low', the presentation of evidence will induce investment by the investor. Intuitively, although the investor knows that the team of scientists is working against him when the product is bad, he improves his outcome since he'll always invest when the product is good (as opposed to never investing). Notice that the investor knows that extended trials are always an option for the entrepreneur, but *for any individual piece of evidence* he cannot determine whether it originated from the entrepreneur or the extended trial. Thus, for the entrepreneur to persuade the investor she must be careful not to reveal that information to him.

By the nature of 'extended' trial, extra time is required for the team of scientists to attempt to gather evidence. Suppose that the entrepreneur creates many products, whose success rate is independent of each other. Consider when the entrepreneur finds that the product is good and she thus has evidence. After her sales pitch, she can send the evidence to the investor that evening. Now consider when the entrepreneur finds that the product is bad and does not have evidence. After her sales pitch she can either admit that she does not have evidence, or she can request extended trials. If she obtains evidence, after one week she submits it to the investor. Notice here that a sophisticated investor *should*

[1] A contestant on the Netflix game show "Bullsh*t" describing his strategy for when to submit answers. On this show contestants answer trivia questions, learn whether they were right or wrong, and then attempt to persuade other contestants (challengers) that they correctly answered the trivia question.

not invest when he receives evidence after a week because although the evidence is not identified as originating from an extended trial, he can *infer* from the disclosure time that the entrepreneur did not have evidence and thus the product is bad.

A sophisticated entrepreneur will recognize the revelation of information through disclosure timing and act to prevent it. One strategy she may consider is to pay a bonus to the team of scientists to conduct the extended trials very quickly, so that if they find evidence she can report it by the end of the day. By doing this, the disclosure timing is now *uninformative*. However, she can do strictly better than this by *strategically delaying* her reporting when her product is good! The critical observation here is that the absolute amount of time the entrepreneur takes is not informative. Rather, it is the large *variance* of disclosure speeds between the two states which is informative. Thus, when the product is good, if the entrepreneur *holds her tongue* and waits to report until one week has passed now her disclosure timing is uninformative. Thus, she can preserve slow reporting when the product is bad and she does not have evidence. Through strategic delay, she strictly improves her profit since she obtains the same result ex ante without incurring the costs of requesting quick extended trials. By slowing herself down when the state is good, she *rescues* herself when the state is bad!

Exploring the effect of introducing disclosure timing and evaluating the strategies of disclosure timing that emerge as a result will be the focus of this paper. We propose a signalling model where two 'pooling' equilibria emerge. However, we will show that one equilibria produces strictly larger profits and thus dominates. We then conduct an experiment, varying the visibility of disclosure timings and characterizing the strategies that emerge under these varying visibility conditions.

2 Literature Review

This paper builds on the seminal works of Spence (1973) and Kamenica and Gentzkow (2011) (Hereafter, KG 2011). Specifically, this paper allows the time of the KG 2011 test to vary across state, this induces a variation of the Spence signalling model where 'testing speed' emerges as the 'costly signal'. However, this enables senders to reduce or eliminate that variance either through paying for quicker tests, or through strategic manipulation of disclosure times across states.

Cooper, Krajbich and Noussair (2019) provide an excellent review of the literature on response time in experimental economics to date. They note that other investigations consider eye-tracking, mouse-tracking, facial expression, chat, and other methods to explore choice processes. They show how response time has been highlighted in the literature both in a pure computational sense (calculated decisions vs intuitive decisions) and in a choice process sense (strength of preference). Response time theory has focused on those *receiving private signals* and their response time to decision but has not explored deeply how other agents should interpret those results and strategically respond. More recently, Konovalov and Krajbich (2020) study response time as a strategic variable, but do not theoretically allow agents to directly choose their response time. Instead, they model response time as a random walk with drift where agents can strategically influence the response time but cannot directly choose it. Frydman and Krajbich (2021) study information cascades in the laboratory and vary the visibility of response times. They find that response times are associated with private information *and* subjects respond to predecessors' response times.

Bergemann and Morris (2019) provide a superb examination of the information design literature to date. Indeed, many extensions of the original KG 2011 problem have been explored theoretically. However, there are no papers which consider the addition of a *indirect* signal such as time. While time can *reveal* information, the channel we operate in is an *indirect* revelation channel. Furthermore, they highlight a number of papers exploring dynamic information design. While our paper concerns timing, our model does not treat time in the same way as these dynamic models and our approach is more consistent with a static approach. Instead, our paper contributes a hybrid model of Bayesian Persuasion in the presence of an indirect signal structure to the literature. Our paper contributes to the growing experimental literature on Bayesian Persuasion. Frechette, Lizzeri and Perego (2021) study an environment where experimental subjects can choose to commit to an information structure (with flexible ability to choose) under varying verifiability conditions, while Nguyen (2017) lets senders choose from a set of possible information design schemes. We explore a simpler environment where we provide an information structure and enable subjects to use it if they choose. The added richness we provide is the additional indirect signaling layer on top of the experimental information design which has not been done to date.

The literature on the effects of the response time of an agent on other agents in a strategic game has been sparse. This study will build on the literature by exploring the effects how agents respond to the established 'reliable' relationship between deception and decision time (even if its directionality is not completely understood). This paper also will avoid that pitfalls of 'strategic truth telling' effects from the Gneezy (2005) (As observed by Sutter (2009)) procedure by instead implementing the procedure from Belot and VanDeVen (2017). For ease of subject understanding, a variation of the 'stars' procedure by Tergiman and Villeval (2021) was implemented.

3 A Simple Case

To provide fundamental intuition to the nature of the problem, we begin with the analysis of a simple case. The state of the world is $\Theta \in \{(G)ood, (B)ad\}$. There are two agents, a sender and receiver, who share a prior belief $Pr(B) = \frac{2}{3}$. For exposition, we assume that agents have monotonically increasing utility in money. The receiver does not observe the state, and gets \$20 by correctly guessing the state, \$0 otherwise. Notice that due to the receiver's pessimistic prior, ex ante he maximizes his expected payoff by always choosing to guess that the state is bad when he does not have any information. The sender is perfectly informed and receives \$20 when the receiver guesses that the state is good and \$0 otherwise, independent of the true state. Naturally, the sender wishes to *persuade* the receiver to guess that the state is good. The sender chooses whether to report the state to the receiver or conduct a test and report the test results to the receiver. Crucially, *the sender cannot misreport the state.* The test is imperfect. Specifically, when the state is good, the test perfectly identifies it and reports such. However, when the state is bad, the test erroneously reports that the state is good with probability $\frac{1}{4}$, known to both the sender and the receiver. *Reports from the test are indistinguishable to reports from the sender.* The test costs \$1. Notice that the cost of the test is less than the expected return from persuasion in the bad state ($\frac{1}{4} * 20 = 5$). Thus, it follows that the sender's optimal action is to always test when the state is bad, and report the state when the state is good. Notice also that the testing error is 'sufficiently low' to still be persuasive to the receiver. Specifically, observe that $\frac{1}{4} \leq \frac{1}{2}$, the frequency ratio of the good state to the bad state. Thus, the receiver maximizes his payoff by following the report instead of ignoring it (and always guessing the state is B) *even in the presence of noisy/erroneous testing*. Under these conditions, it follows that the equilibrium is that the sender always chooses to conduct the test when the state is bad and report the state when it is good. The receiver always guesses what the test reports.

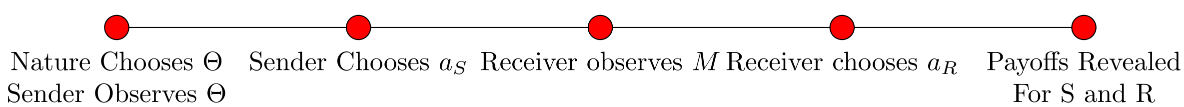
Now consider that the test also requires some time, 20 seconds, to complete. Suppose that the sender decides to report that the state is good when he receives the information (ie at 0 seconds), and reports the test results as soon as he gets them at 20 seconds. This cannot form a Nash Equilibrium since the receiver should only follow immediate reports of a good state and guess the state is bad if he only receives a report after 20 seconds. Intuitively, although the report does not explicitly indicate that a test was conducted, the receiver can *infer* from the relatively long disclosure time that the report originated from a test which was conducted due to the state being bad. Thus, to preserve the equilibrium the sender must choose his timing in a way that is not state dependent as this will be informative to the receiver of the state. Suppose that he has an option to pay \$3 for an instantaneous test. Notice that the cost is low enough that he still is better off paying for an instantaneous test to persuade the receiver (since \$3 < \$5, the expected return of persuasion). Should the sender chose to pay this premium, conduct a 'quick' test, and report instantaneously, the equilibrium is restored. The receiver now receives all reports instantaneously and cannot infer the origin of a report from its timing.

However, there exists another sender strategy which still induces a Nash Equilibrium, and is ex ante strictly preferred by the sender. Although the sender *can* report instantaneously when the state is good, he is not *required* to. Instead, he can choose to *strategically* *delay* himself and report slowly (in this case, at 20 seconds). The benefit of choosing to report slowly when the state is good, is that it allows him to conduct the cheaper slow test when the state is bad and the receiver still is unable to infer the source of the report (since all reports arrive at 20 seconds). Thus, while it is still cost effective for a sender to purchase 'quick' testing, it is not profit maximizing since the 'speed' of the testing does not increase the persuasive ability of the test. The strategic delay strategy allows the sender to receive the same persuasive ability of the test as he does in a world without test timing, but allows him to spend less on testing than if he followed a 'quick' testing strategy.

4 Baseline Model

4.1 Cheap Talk

This section will follow Crawford and Sobel (1982). This game will have two players, a sender (S) and a receiver (R), who play this game for some commonly known number of periods, $T \in \mathbb{N}$. In each period, nature moves first and chooses $\Theta \in \{G, B\}$. R and S share a prior belief $\mu = Pr(\Theta = B)$. The sender can perfectly observe Θ while the receiver cannot. The sender then chooses an action $a_S \in \{g_S, b_S\}$, which translates to a message/report, $M \in \{g, b\}$, which is directly observed by R. R then chooses $a_R \in \{g_R, b_R\}$. Explicitly, the timeline of events is as follows



Without loss of generality, the payoff for S in any period is given by

$$\Pi_S = \beta \mathbb{1}(g_R)$$

Where $\beta > 0$ can be interpreted as the "strength of the incentive". This follows because the sender's payoff is state independent and depends only on a_R.

The receiver's payoff in any period is given by

	G	B
g_R	σ_1	γ_2
b_R	γ_1	σ_2

Where $\sigma_i > \gamma_i$ for $i = \{1, 2\}$. Naturally, R maximizes his payoff by correctly identifying the state. That is, R maximizes his payoff by choosing $a_R = g_R$ when $\Theta = G$ and by choosing $a_R = b_R$ when $\Theta = B$. Notice that this payoff is independent of M. Intuitively, 'correctly identifying the state' would map to a corresponding action (invest in new technology vs invest in current technology) in the natural world. For simplicity, we will assume that R and S have risk neutral utility. Thus, the expected payoff (which will be the expected utility) of R for each action in a period is given by

$$\Pi_{g_R} = \mu(\gamma_2) + (1-\mu)(\sigma_1)$$

$$\Pi_{b_R} = \mu(\sigma_2) + (1-\mu)(\gamma_1)$$

Consider for a moment when $M = \emptyset$. That is, R does not observe a message/report from S. In this case he will always choose $a_R = b_R$ if $\Pi_{g_R} < \Pi_{b_R}$, and $a_R = g_R$ otherwise.

Assumption 1. $\mu > \frac{\sigma_1 - \gamma_1}{(\sigma_1-\gamma_1)+(\sigma_2-\gamma_2)}$

Notice that in order to ensure that S is sufficiently incentivized to provide any information to R, we must require that R's payoff from choosing g_R ex ante is lower than his payoff from guessing b_R ex ante. Rearrange the expected payoff condition and obtain our requirement that

$$\mu > \frac{\sigma_1 - \gamma_1}{(\sigma_1-\gamma_1)+(\sigma_2-\gamma_2)}$$

Intuitively, as the 'correct guess' premium when $\Theta = G$ is increasing, we'll require that the receiver be more pessimistic about that probability that the state is indeed good so that his expected profit maximizing choice is to always choose b_R. Notice that when the correct guess 'premiums' are equal in both states then the condition collapses to the simple case constraint $\mu > \frac{1}{2}$.

Lemma 2. *For any choice of parameters, the sole Nash equilibrium is always* $(a_S = g_S, a_R = b_R)$ *(babbling)*

By construction, notice that whenever $\Theta = G$, it is incentive compatible for S to honestly report the state of the world. Thus, we expect $a_S = g_S$ for all such periods. When $\Theta = B$, honest disclosure results in a payoff of 0 for that period for the sender. However, if the sender dishonestly discloses and the receiver follows his report, then he receives β. Notice that in the last period without a future, it always follows that the sender chooses g_S regardless of Θ. Using backwards induction, the sender always chooses $a_S = g_S$ and the receiver always chooses $a_R = b_R$.

4.2 Persuasion

Continuing with our constraint on μ, now consider an alternate reality where the sender *cannot dishonestly disclose*. That is, when $\Theta = G$ then the sender cannot choose $a_S = b_S$ and when $\Theta = B$ then the sender cannot choose $a_S = g_S$. However, the sender can instead choose to conduct a test and *report the results of the test as his message* (In the spirit of KG 2011). When he chooses this action, we will denote it as $a_S = \lambda_S$. Thus, the action space of the sender when $\Theta = G$ is $A_{SG} = \{g_S, \lambda_S\}$ and the action space of the sender when $\Theta = B$ is $A_{SB} = \{b_S, \lambda_S\}$. Crucially, once the sender elects to conduct the test he is bound to report the results, whatever they may be. Thus, when $a_S = \lambda_S$ the test chooses M with the following probabilities

$$Pr(g|G) = 1 \qquad\qquad Pr(g|B) = \lambda$$

$$Pr(b|G) = 0 \qquad\qquad Pr(b|B) = 1 - \lambda$$

Intuitively, when the state is good the test will always correctly indicate that the state is good. When the state is bad, the test correctly identifies the state with probability $1-\lambda$ and incorrectly identifies the state with probability λ. Crucially, the receiver cannot distinguish when a test is conducted. IE, $M = g$ from $a_S = \lambda_S$ is indistinguishable from $M = g$ following $a_S = g_S$. That is, a_S is entirely unobserved by R. Only M is observed.

Assumption 3. $\lambda \leq \frac{(1-\mu)(\sigma_1-\gamma_1)}{\mu(\sigma_2-\gamma_2)}$

Lemma 4. *The sole Nash equilibrium is* $(a_S = \lambda_S, a_R = M_R)$

Suppose the sender always chooses to implement the test. With a perfect test ($\lambda = 0$), the receiver always prefers to 'follow the report'. That is, he prefers $a_R = M_R$. This is true for sufficiently small λ. However, there exists a threshold for which the test is 'too noisy' and the receiver reverts to $a_R = b_R$ in all periods. For the sender, choosing $a_S = \lambda_S$ is always incentive compatible as long as λ is sufficiently small so that the receiver will follow the report. To obtain this threshold, notice the incentive compatibility constraint of following the report is given by

$$\Pi_\lambda = (1-\mu)(\sigma_1) + \mu[(1-\lambda)\sigma_2 + \lambda\gamma_1] \geq (1-\mu)(\gamma_1) + \mu\sigma_2 = \Pi_{b_R}$$

Rearrange and obtain our constraint on the size of the testing error

$$\lambda \leq \frac{(1-\mu)(\sigma_1 - \gamma_1)}{\mu(\sigma_2 - \gamma_2)}$$

Intuitively, as the 'correct guess' premium when $\Theta = G$ increases, the receiver is more willing to tolerate error to obtain that larger payoff. Similarly, as the 'correct guess' premium when $\Theta = B$ increases, the receiver is less willing to tolerate error since choosing b_R in all periods becomes increasingly attractive. Furthermore, as the receiver becomes more pessimistic he demands a more reliable test to persuade him to choose g_R when the test reports g instead of just choosing b_R for all periods. Thus, under these conditions there is an equilibrium where the sender chooses to honestly report when $\Theta = G$ and conduct a test when $\Theta = B$. The receiver always follows the report of the sender.

5 Model with Costly Testing Time

5.1 General Setup

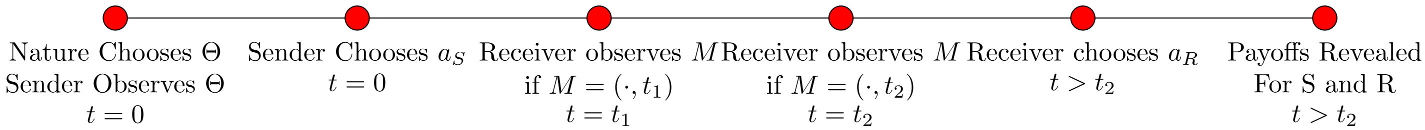
As before, each period nature chooses $\Theta \in \{G, B\}$ and the sender (S) will receive a perfectly informative signal. Additionally, impose the assumption that dishonest disclosure is prohibited and the sender can only choose between honest reporting or conducting some test. Building on the model of Spence (1973), now consider t, the time to conduct a test (or simply, 'testing time'). For simplicity, assume $t \in \{t_1, t_2\}$, where $0 < t_1 < t_2$. Intuitively, t_1 represents a 'quick' test and t_2 represents a 'slow' test. Naturally, conducting a 'quick' test is costlier than a 'slow' test. Thus, we obtain the following cost structure

$$\left\{ \begin{array}{l} C(\Theta_S, t_1) = C(\Theta_S, t_2) = 0 \\ \\ C(\lambda_S, t_1) > C(\lambda_S, t_2) > C(\Theta_S, t_2) \end{array} \right\}$$

Where $C(\lambda_S, \cdot)$ is the cost of conducting a test for some choice of t. $C(\Theta_S, \cdot)$ are the costs of honest disclosure for a choice of t, in the respective state. Notice that honest disclosure is always free regardless of state.

Assumption 5. $\beta \geq \frac{C(\lambda_S, t_2)}{\lambda}$

Intuitively, the assumption is a constraint to ensuring that testing (in general) is cheap enough that the sender is willing to choose $a_S = \lambda_S$. Note that we have not made any assumption on the price of 'quick' testing. Notice also that now the sender also chooses how quickly they wish to report. Thus, the sender's action now is defined as a 2-tuple $a_S = A_{S\Theta} \times t$. Abusing notation, now $M = M \times t = \{g, b\} \times t$. Explicitly, the structure of the game is as follows



5.2 Benchmark 1: The Naïve Receiver

Lemma 6. *The strategies $a_S = (g_S, t_1)$ when $\Theta = G$, $a_S = (\lambda_S, t_2)$ when $\Theta = B$, and $a_R = M_R$ form a Nash Equilibrium when receivers do not observe t*

Suppose that the receiver is naïve and they interpret $(\cdot, t_1) = (\cdot, t_2)$. That is, they are not sensitive to disclosure time and/or they do not observe it. How should the sender respond? Assume $\beta \geq \frac{C(\lambda_S, t_1)}{\lambda}$, then the sender will always choose to conduct the test. Now, the sender simply needs to choose which cost structure to follow. Given that $C(\lambda_S, t_1) > C(\lambda_S, t_2)$ and the assumption that the receiver is not sensitive to the senders choice of t, it immediately follows that when $\Theta = B$, the sender chooses slow testing, $a_S = (\lambda, t_2)$. Notice that this is also true if $\beta \geq \frac{C(\lambda_S, t_2)}{\lambda}$, thus this is the only assumption we will require (relax $\beta \geq \frac{C(\lambda_S, t_1)}{\lambda}$).

Thus, in the 'Naïve Equilibrium' when $\Theta = G$, the sender will always chose $a_S = (g_S, t_1)$ (He is indifferent between choosing t_1 and t_2). When $\Theta = B$, the sender chooses $a_S = (\lambda_S, t_2)$. The receiver will always follow the report of the sender ($a_R = M_R$).

5.3 Benchmark 2: The Sophisticated Receiver

Proposition 7. *The following strategies form a Nash Equilibrium*

$a_S = (g_S, t_1)$ *when* $\Theta = G$
$a_S = (\lambda_S, t_1)$ *when* $\Theta = B$ *if* $\beta \geq \frac{C(\lambda_S, t_1)}{\lambda}$
$a_S = (b_S, t_1)$ *when* $\Theta = B$ *if* $\beta < \frac{C(\lambda_S, t_1)}{\lambda}$
$a_R = M_R$ *when* $M = (\cdot, t_1)$

Now suppose that the receiver is sophisticated[2], and responsive to t. The 'Naïve Equillibrium' is no longer stable. This follows from the separation of signals conditional on the state. Explicitly, the sophisticated receiver infers that $M = (g_S, t_2)$ was obtained from a test conducted when $\Theta = B$ and thus *will not* follow the sender's report in this state. Naturally, it follows that since the receiver will not follow the message then the sender should not be willing to exert any cost to conduct the slow test.

[2] A discussion of sender uncertainty over sophistication of receivers can be found in the Appendix (A.1)

Thus, the sender faces a new choice. Intuitively, he can either bear the cost of conducting the test quickly ($a_S = (\lambda_S, t_2)$) so that the message obtained from the test, $M = (g_S, t_1)$, is only partially informative as before *or* he can decide not to exert any cost to signal and instead choose to honestly disclose, $a_S = (b_S, t_1)$. His optimal choice will depend on β. If $\beta \geq \frac{C(\lambda_S, t_1)}{\lambda}$, then the sender still benefits from conducting quick tests to induce a partially informative report as before. Alternatively, if $\frac{C(\lambda_S, t_1)}{\lambda} > \beta$ then the sender would rather not bear the cost of quick testing and will instead choose costless honest disclosure, $a_S = (b_S, t_1)$.

Thus, with sophisticated receivers we have a new equilibrium. Naturally, we'll call this the 'Quick Testing Equilibrium'. If $\Theta = G$, then the sender chooses $a_S = (g_S, t_1)$. If $\Theta = B$ and $\beta > \frac{C(\lambda_S, t_1)}{\lambda}$ then the sender chooses $a_S = (\lambda_S, t_1)$. Finally, if $\Theta = B$ and $\beta < \frac{C(\lambda_S, t_1)}{\lambda}$ then the sender will choose $a_S = (b_S, t_1)$. As before, the receiver will always follow the report of the sender, $a_R = M_R$.

AN ALTERNATIVE APPROACH: STRATEGIC DELAY

Proposition 8. *The strategies $a_S = (g_S, t_2)$ when $\Theta = G$, $a_S = (\lambda_S, t_2)$ when $\Theta = B$, and $a_R = M_R$ if $[[M = (\cdot, t_1) \; | M = (\cdot, t_2)] \; \forall \; T]$ form a Nash Equilibrium*

In the previous equilibira, when $\Theta = G$, the sender always chooses $a_S = (g_S, t_1)$. This followed intuitively from the fact that the sending g_S when $\Theta = G$ is immediately incentive compatible and costless. However, what if the sender chooses $a_S = (g_S, t_2)$ when $\Theta = G$? This gives rise to a new equilibrium, which we will call the 'Strategic Delay Equilibrium'.

Formally, when $\Theta = B$ the sender would prefer to conduct the slow test, $a_S = (\lambda_S, t_2)$, credibly as opposed to conducting quick test, $a_S = (\lambda_S, t_1)$, given our assumed cost structure. However, with our earlier assumption this falls apart as the sophisticated receiver will recognize that $M = (g_S, t_2) \implies \Theta = B$. Relaxing that assumption, consider if the sender chooses to send $a_S = (g_S, t_2)$ when $\Theta = G$. Then, if the sender also chooses to conduct the slow test when $\Theta = B$ the report $M = (g_S, t_2)$ is partially informative as it were before!

This will form the 'Strategic Delay Equilibrium'. Here, the sender will choose $a_S = (g_S, t_2)$ when $\Theta = G$. When $\Theta = B$, the sender conduct the slow test, $a_S = (\lambda_S, t_2)$. With this structure, the receiver will always follow the report of the sender. Notice that despite the fact that the sender faces no cost to respond quickly in the good state, in this equilibrium the sender *deliberately* delays disclosure in the good state to preserve their ability to conduct tests in future bad states when they'll want to credibly report g_S slowly! Thus, fast reporting disappears even when it is costless to do so!

6 Experimental design and implementation

6.1 Choice of Structure

The game design was chosen to best represent a signalling game with disclosure constraints (regulation) and imperfect testing in the natural world similar to the exposition in the introduction. When $\Theta = G$, the model assumes that all signals are costless. This was chosen to represent the strength of the incentive compatibility in the natural world. In much of the modelling it was assumed that (g_S, t_1) is chosen when $\theta = g$. Again this follows from the intuition that senders will have an 'immediate gut reaction' in the natural world and signal g_S. However, we do also consider relaxing this assumption considering the fact that senders may 'hold their tongue' to credibly allow themselves to conduct tests slowly when $\Theta = B$. *In the experiment, senders will not be constrained.*

Theoretically, risk aversion was relaxed here and agents are assumed to be risk neutral. While this is an unrealistic assumption, it seems unlikely that risk aversion would interact with disclosure timing since the 'lottery' can only improve outcomes for the sender. Thus, any effects from risk aversion should be constant across models theoretically and in the laboratory and should not cause analytic bias or distortions.

The Belot and VanDeVen (2017) procedure was elected over the Gneezy (2005) procedure to minimize the possibility of 'strategic truth telling'. Specifically the latter procedure involves a persuasion message 'Your payoff is higher for choice A than B' or the complement. Thus, senders could choose to send an honest message in hopes that receivers would 'overreact' and choose the opposite state which pays the sender more. This is not present in the Belot and VanDeVen (2017) framework as the message choice is either the state is green (good) or the state is red (bad), so honest disclosure of red cannot be 'strategic' as the receiver knows that if the state is green the sender is better off by informing them of that directly. Thus the receiver should believe that if he sees a message of red that the state is indeed red. This structure maximizes strategic consideration when the state is red and minimizes it when the state is green as desired. Finally, an adaptation of the 'Star/No Star' presentation of states from Tergiman and Villeval (2021) was implemented to enhance subject understanding.

6.2 Procedures

Seven sessions were conducted at Florida State University (FSU) in 2022. 82 FSU students were recruited via ORSEE Greiner (2015). Two types of sessions were conducted. *Delayed Release* (DR) sessions ($N = 40$) served as our control session, while *Time Shown* (TS) sessions ($N = 42$) served as our main treatment.

In all sessions, subjects played the game twice, with each game lasting 8 periods for a total of 16 periods. Pairs were randomly matched and retained throughout each game. Thus, subjects were re-matched once. Roles also were initially randomly selected and retained *throughout the entire session* (both games) to capture if behavior converges in any meaningful way with experience. It also seems natural given that roles in the natural world do not often change with regards to signalling interactions. The roles were sender (S) or receiver (R). States were labelled 'Star' and 'Triangle'. Nature chose 'Triangle' with positive probability 66.66%. The sender was then shown the card which nature had chosen. if the card had a star, he then could choose 'Report Star' at any time. If the card had a triangle he could choose 'Report Triangle' at any time. Regardless of the type of card, the sender could also choose choose to pay 1 ECU to conduct a test which had an error rate of 25% if the card had a triangle on it and had an error rate of 0% if the card had a star on it. The test required 20 seconds to process. Note, after the test is conducted the sender still needed to 'submit' the test results. IE, observing a message at 20 seconds did not imply that the report did not come from a test. To avoid lie aversion, the test was presented with the following exposition

"The test is not perfect. If the card has a star, then the test will *always* report "star". However, if the card has a triangle on it, then there is a 25% chance that the test will *incorrectly* report "star" and a 75% chance that the test will report "triangle". In other words, when your card is triangle the test will report inaccurately 25% of the time. The accuracy of the test is fixed and cannot be manipulated by any subject, and the accuracy of the test does not change if it is conducted quickly."

Reports from senders' honest disclosure are identical in structure to reports from tests from the perspective of the receiver. Senders also had an option to pay 3 ECU for a 'quick test' which had the same error rate as the standard test but gave senders results *instantaneously*. There was no free form chat so that sender actions were the only form of information receivers have about the state (except for their shared prior). The receiver chooses 'Star' or 'Triangle' after receiving the report from the sender. The receiver had an unlimited amount of time to choose an action once they received the report, but after 20 seconds they did see a "Please Make a Decision" prompt. Once an action was chosen, the receiver learned the card's symbol. If the card was 'Star' and the receiver chose 'Star', then he received 20 ECUs. Similarly, If the card was 'Triangle' and the receiver chose 'Triangle', he received 20 ECUs. Else, the receiver earned 0 ECUs. The sender received 20 ECUs if the receiver chose 'Star' and 0 ECUs otherwise. Additionally, subjects were reinforced that a report of 'No Star' necessarily means that the card is indeed 'Triangle' whereas a report of 'Star' could come from a card with a 'Star' or a card with a 'Triangle'. Subjects knew all payoffs in all states and knew that senders were committed to reporting test results once a test is conducted.

Delayed release (DR) eliminates any strategic consideration from disclosure time. Specifically, regardless of when a subject chose their action and submitted their report, reports were only viewable to receivers after 50 seconds[3]. After the entire time constraint has passed then *all* decisions were released to the receivers. Thus, receivers cannot infer how long it took senders to send signals. Receivers saw "Please wait for all senders to choose their action" to reinforce to receivers that decision time is not being revealed to them. Thus senders have no incentive to alter their disclosure time.

In the *Time Shown* (TS) sessions receivers received signals as soon as they were sent by senders. Additionally, receivers saw "Please wait for your sender to choose their action" to reinforce to receivers that they could directly infer the disclosure time. On subject screens they saw "The report was received at (Disclosure Time) seconds."[4]

In both sessions the clock[5] shown to subjects was *ascending*, highlighting how much time had elapsed. This was done because subjects were not given a time limit, and thus 'time remaining' was not a main construct in this experiment. After 35 seconds had elapsed, senders did see a prompt to "Please make a decision." if they had not chosen an action yet. If the sender chose to test, after 50 seconds had elapsed senders were prompted "Please Make a Decision". It is unlikely that this "constraint" influenced subjects, as the average disclosure time was less than 16 seconds

To prevent bankrupcy, senders were endowed with 5 ECUs per round. Receivers also had the opportunity to earn tokens for a separate 'Home Fund' which was not available during the experiment. 'Home Fund' tokens were earned while receivers were waiting for messages from senders. Specifically, receivers were shown at most 5 multiple choice two digit addition problem while they waited for their sender[6]. Each correctly solved problem earned them 1 ECU. The implementation of this task is to preserve anonymity. In the absence of this task, subjects may have been able to deduce who their sender was in treatment sessions. In the presence of the task, receivers were unable to distinguish if someone was clicking their mouse as a receiver or as a sender.

Conductors utilized a standardized script for instructions. Launching the session in Z-Tree, subjects were given a short quiz prior to beginning the experiment.[7] The goal of this quiz was to ensure that subjects understand the payoffs for themselves and the other

[3]Because time was not binding, it was possible that a sender took more than 50 seconds. In that case, *none* of the receivers would see their corresponding report until the slowest sender sent their report. In other words, receivers would be able to infer the speed of the *slowest* sender (but they would not know if their sender was the slowest). This happened in 3 periods across all control sessions (64 periods total).

[4]In order to prevent subjects sending signals quickly to move through the experiment faster, all rounds were a minimum of 50 seconds, regardless of when signals were sent

[5]It is a standard practice in Ztree experiments to show subject a clock/timer

[6]In the control (DR), this amounted to each receiver having exactly 50 seconds. In order to prevent socially motivated delay (ie, delay in order to allow partners more time to work on math), in the treatment receivers were given 50 seconds total to answer math questions. If the report was received before 50 seconds had elapsed, the receiver would be given the opportunity to finish their math questions after they guessed the card's state until 50 seconds *on math questions* had elapsed

[7]The Payoff Quiz can be found in the Appendix (D)

subject depending on actions and outcomes. Subjects were told that in order to advance all questions had to be answered correctly. They were also told that this quiz was just to ensure that they understood the structure and that they should ask questions if they are confused. If there were any incorrect answers, they would see the following prompt at the top of their screen in bold red 'At least one of the questions is incorrect. Please try again. If you are stuck, please raise your hand and an experimenter will come assist you"

During the instructions, subjects observed a conductor in the role of S observing a triangle card, choosing a slow test, and obtaining a report of 'Star' (20 seconds later). This was done to ensure subject understand that the timing is indeed occurring in real time and not simply exposition. Additionally, it is worth noting that the problem senders face arguably has a 'CRT' structure. Indeed, the immediate intuitive answer is 'Test quickly to avoid detection' while the profit maximizing answer is 'reveal slowly to credibly test slowly' and is less intuitive. As such, subjects were also asked to answer 5 CRT questions at the end of the experiment, with each correct answer earning subjects an additional 1 ECU. Finally, a demographic survey was conducted.

Senders and receivers were always given a full history which included the card's symbol, the report, the receiver's guess, and their round payoffs (including endowments). In the Time Shown treatment, the full history of senders and receivers also included the time the report was received. All sessions lasted 45 minutes. Subjects earned \$1 per every 100 ECUs.

7 Preliminary Hypotheses

H1: Signal time will be significantly larger in the bad state than the good state in the control (Validity of Naïve Equilibrium)

When the state is bad, agents need to decide whether to honestly report or to pursue a test. This decision in and of itself requires more time than when the state is good (and the only choice is honest reporting). Furthermore, if the agent elects to conduct a test there is no incentive for him to test quickly since all results are released at the end of the decision period.

H2: In the treatment, As the variance in disclosure time when the message is g_S increases, the likelihood of a receiver choosing b_R is increasing in decision time (Sophistication of Receivers)

This hypothesis directly follows from the first. If testing times vary, larger disclosure times should be correlated with Θ, thus the disclosure time is informative and should not be ignored. Subjects should recognize the correlation between delays and bad states and adjust their posteriors accordingly over time.

H3: The difference in the disclosure time when $\Theta = G$ and $\Theta = B$ in the treatment will be significantly different than the same difference in the control (Validity of Pooling Bayesian Nash Equilibria)

This hypothesis follows directly from the theoretical model and from the second assumption. In the presence of receiver sophistication, separation cannot be an equilibrium and one of the pooling equilibria are expected to emerge.

H3a: In the treatment, H3 will hold & the mean disclosure time when $\Theta = B$ in the treatment will not significantly differ from t_1 (Validity of Quick Testing Equilibrium)

This hypothesis follows directly from the theoretical model. This hypothesis supports the notion that senders signal quickly when the state is good and thus they must pay for quick testing when the state is bad if they wish to test credibly/without detection.

H3b: In the treatment, H3 will hold & the mean disclosure time when $\Theta = B$ in the treatment will not significantly differ from t_2 (Validity of Strategic Delay Equilibrium)

This hypothesis follows directly from the theoretical model. This hypothesis supports relaxing the notion of senders signalling quickly when $\Theta = G$ and instead allows them to 'hold their tongue' so that they can test slowly and send the result g_S when $\Theta = B$ credibly/without detection.

8 Results

8.1 Summary Statistics

We begin with summary statistics over the entire sample

Table 1: Summary Statistics

N=82 (Subjects)

Variable	Mean	Std. Dev.	Min	Max
Bad Card	.6905488	.4624437	0	1
Test Error	.2538631	.4357014	0	1
DT	15.63196	12.54752	1.875	53.844
GuessTime	7.202521	5.512681	1.063	35.719
CRT Score	1.188262	1.122714	0	4
Report	.5411585	.4984931	0	1
SlowTest	.3689024	.4826914	0	1
QuickTest	.089939	.2862035	0	1
ECU Earnings	269.33	95.55505	84	400

As expected by construction, the state is bad just over 66% of the time and the testing error when the state is bad is just above 25%. The average CRT Scores are relatively low, with the average subject roughly only getting one question correct. It is worth noting that no subject got all 5 CRT questions right. Disclosure times and actions vary considerably. For better understanding, we report the key summary statistics by card symbol[8] and treatment.

Table 2: Summary Statistics by Card Symbol & Treatment

Delayed Release (DR, Control)

N=40 (Subjects)

Card Symbol = Star (Good)

Variable	Mean	Std. Dev.	Min	Max
DT	8.56675	9.854657	2.125	53.844
Report	.94	.2380828	0	1
SlowTest	.06	.2380828	0	1
QuickTest	0	0	0	0

Card Symbol = Triangle (Bad)

Variable	Mean	Std. Dev.	Min	Max
DT	17.88611	12.50531	1.984	51.579
Report	.5090909	.5004864	0	1
SlowTest	.4772727	.5000518	0	1
QuickTest	.0136364	.1161079	0	1

Time Shown (TS, Treatment)

N=42 (Subjects)

Card Symbol = Star (Good)

Variable	Mean	Std. Dev.	Min	Max
DT	7.811107	8.457534	1.875	38.047
Report	.9126214	.283077	0	1
SlowTest	.0776699	.2683034	0	1
QuickTest	.0097087	.0982923	0	1

Card Symbol = Triangle (Bad)

Variable	Mean	Std. Dev.	Min	Max
DT	19.99313	12.2764	2.156	45.359
Report	.2360515	.4251106	0	1
SlowTest	.527897	.4997577	0	1
QuickTest	.2360515	.4251106	0	1

Separating by state and treatment, we obtain a clearer picture. As expected, the utilization of the Quick Test is near 0 regardless of the state in the Control, suggesting that subjects correctly understood that the Quick Test provided no additional value relative to the Slow Test. Across treatments, surprisingly, subjects choose to conduct a test (of any speed) when the card's symbol is star roughly 7.4% of the time. It is worth noting

[8] Recall that the sender perfectly observes these

that one subject accounts for 26.7% of these unexpected tests[9]. We note that 73.3% of these behaviors occur within the first 8 periods, suggesting that the prevalence of these behaviors is decreasing in experience[10].

By treatment and state, we summarize sender actions and disclosure times in Figures 1 and 2, respectively.

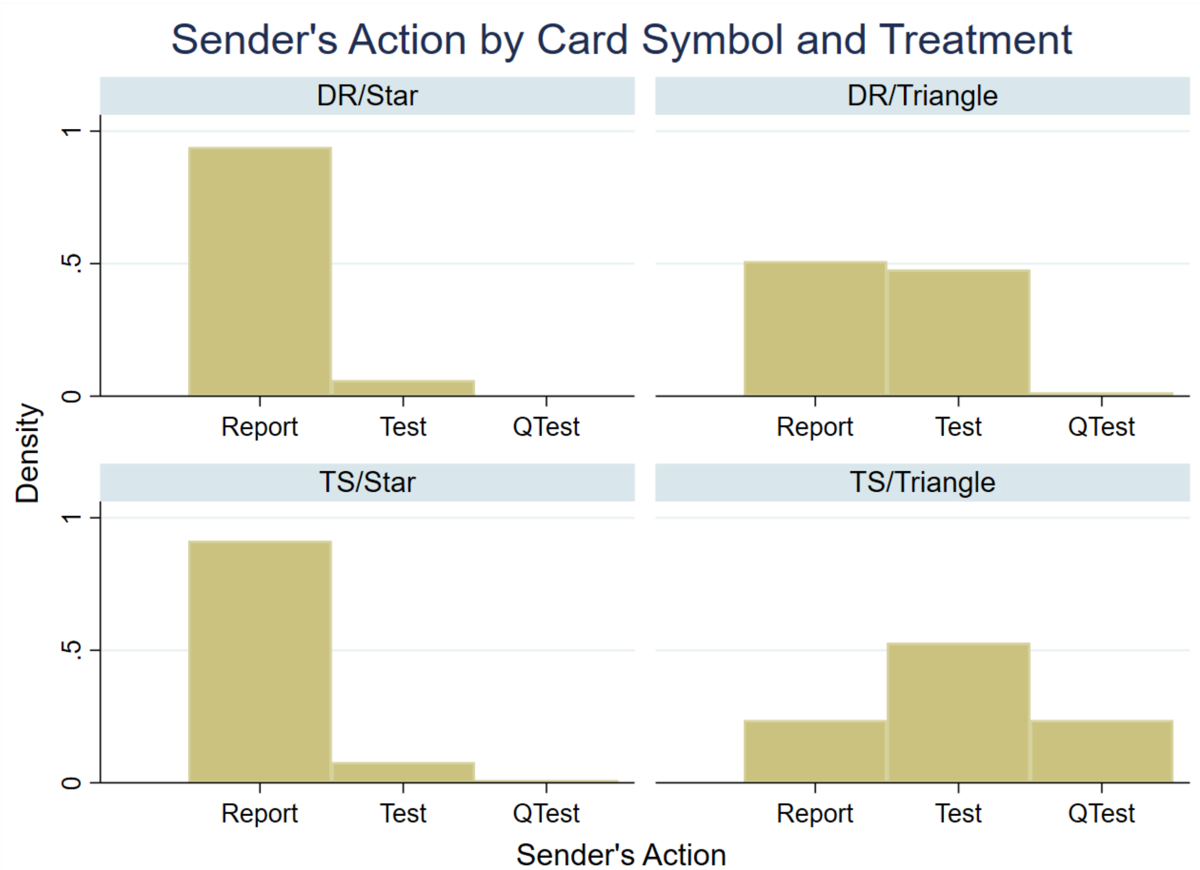


Figure 1

[9]In total, 9 subjects made a decision to conduct a test when the card's symbol is Star at least once, or 21.95% of senders

[10]Formally, regressing Action-Tester (A binary equal to 1 if the action is Slow Test or Quick Test) on the subset of observations where the card's symbol was star, the coefficient on Period is statistically significantly negative, with p=.088

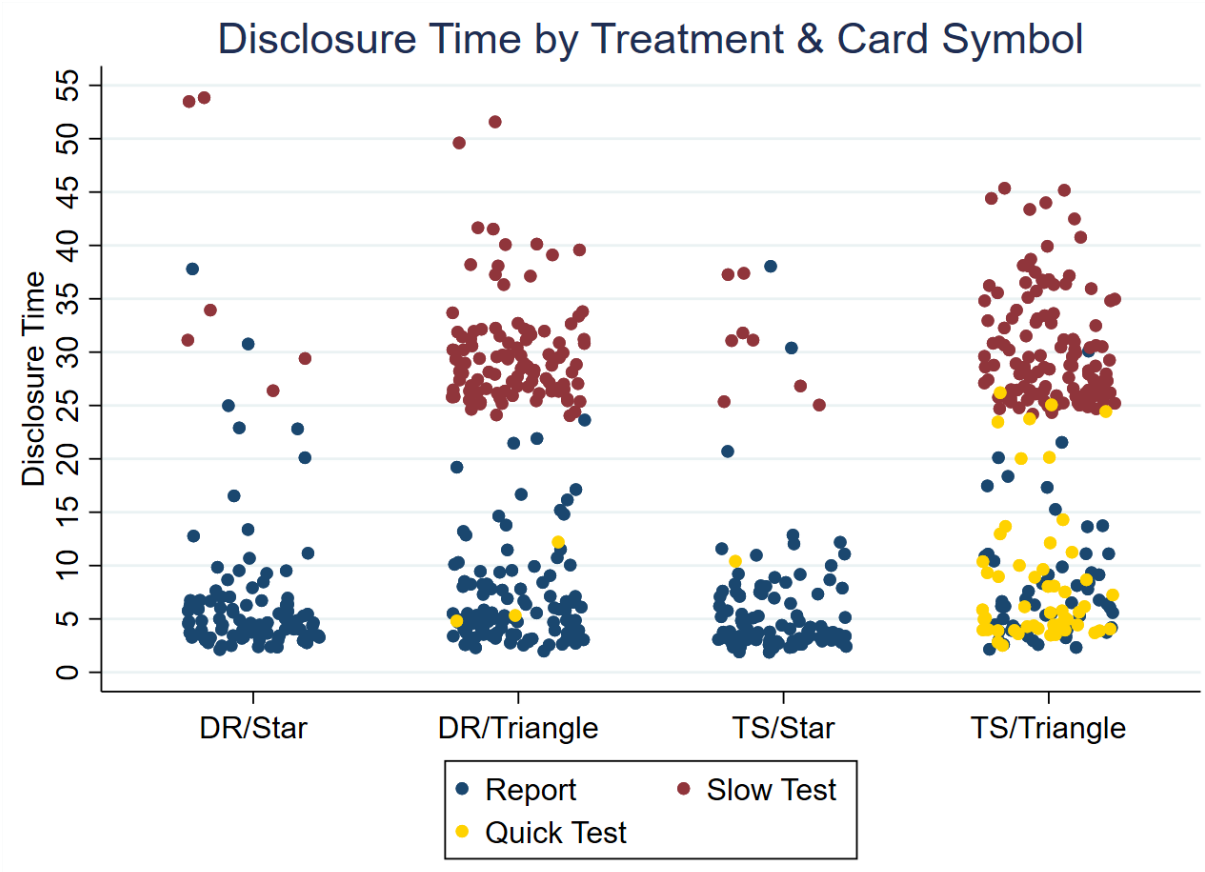


Figure 2

8.2 Sender Behavior

Now considering our hypotheses in turn, **we find support for H1**. Indeed, across all the DR sessions the mean disclosure time when the card has a star on it (good state) is 8.57, while the mean disclosure time when the card has a triangle on it (bad state) is 17.89. As expected, the mean disclosure time in the bad state is significantly higher than in the good state[11]

To explore our additional hypotheses, we calculated the *Mean* Disclosure Time for each individual sender when the card had a star on it, and when the card had a triangle on it. Using those calculations we obtained the *Difference* between the means[12]. In treatment sessions, we expect that strategic senders should minimize this statistic while naive senders will not optimize on it.

[11]Wilcoxon rank-sum test, z = -6.322, p=.00001

[12]We utilize the Absolute Value transformation to keep all values positive

Table 3: DTDifference

	DR	TS	Difference
Mean	13.00305	13.09559	-0.09254
Std. Dev.	8.474238	7.940787	0.533451
Min	.1188667	.2937984	-0.1749317
Max	32.92471	28.46857	4.45614
N (Games)	38	39	-1

Upon initial inspection, it appears that the treatment had no effect on DTDifference. However, it is also crucial to notice that many senders exhibited "truthtelling" behavior. IE, when the card's symbol was triangle these senders simply reported the symbol instead of utilizing any test option. This truthtelling could arise for a number of reasons. Candidate explanations are preferences for "truthtelling" and "over-reaction" to a failure (if the test fails to persuade the receiver). These behavior profiles are outside of the scope of this study, and interfere with the analysis of the strategies of interest. To see this, notice that if a sender is always reporting honestly (quickly, since there is no reason to delay), then that subject's DTDifference will be very small. Thus, they appear as if they are strategically manipulating their disclosure timing to hide testing behavior, despite the fact that they aren't testing often at all! To remedy this, we calculate the proportion of the time where senders choose to simply report that the card has a triangle on it when observing a triangle (as opposed to attempting to persuade the receiver with a test). Focusing on senders who report truthfully when the state is bad less than 75% of the time ($PRB <= .75$), the pattern that emerges is more consistent with our third hypothesis.

Table 4: DTDifference, PRB <= 75%

	DR	TS	Difference
Mean	17.28576	14.35128	2.93448
Std. Dev.	6.255923	7.390234	-1.134311
Min	6.230249	.2937984	5.9364506
Max	32.92471	28.46857	4.45614
N (Games)	25	35	-10

To further understand sender behavior, we introduce a complete typology over sender behaviors. These types are

1. Honest Sender (PercentReportBad >= .75[13])
2. Naïve Sender (DTDifference > 6.6[14] & not (Honest Sender))

[13]This threshold was chosen as it is the 90th percentile of PercentReportBad in the treatment group
[14]This threshold was chosen as it is 1 standard deviation of DTDifference in the sample

3. Strategic Sender
    (a) Quick Sender (DTDifference <= 6.6, MeanDT-Triangle < 20[15] & not (Honest Sender))
    (b) Slow Sender (DTDifference <= 6.6, MeanDT-Triangle >= 20 & not (Honest Sender))

Intuitively, a Naïve sender does not consider time in his optimization problem or strategic behavior. Thus, the quick test provides no gain in utility relative to slow testing, and there is no reason not to immediately report the state in the good state. Thus, we expect the difference in their Disclosure Time to be large. Conversely, we categorize Strategic senders as those senders whose DTDifference is relatively low. From there, we characterize the mechanism through which Strategic senders minimize DTDifference (through Quick Testing or Strategic Delay). The summary stats for these new statistics are reported, by game and treatment, in turn below.

Table 4: Sender Types by Treatment & Game
Delayed Release (DR, Control)

Game 1

Variable	Mean	Std. Dev.
HonestSender	.4	.4914361
NaiveSender	.6	.4914361
QuickSender	0	0
SlowSender	0	0

Game 2

Variable	Mean	Std. Dev.
HonestSender	.35	.4784672
NaiveSender	.65	.4784672
QuickSender	0	0
SlowSender	0	0

Time Shown (TS, Treatment)

Game 1

Variable	Mean	Std. Dev.
HonestSender	.1428571	.3509732
NaiveSender	.7142857	.4531045
QuickSender	.0952381	.2944211
SlowSender	.047619	.2135955

Game 2

Variable	Mean	Std. Dev.
HonestSender	.0952381	.2944211
NaiveSender	.7619048	.427191
QuickSender	.0952381	.2944211
SlowSender	.047619	.2135955

Unsurprisingly, there are no Strategic Senders in our control group. Upon initial inspection, it is counter-intuitive that there are more Naive Senders in the treatment than the control. However, it is important to recognize that the DR sessions had a much higher proportion of Honest Senders than the TS sessions[16]. If we assume that these

[15]This threshold was chosen as it is impossible to conduct a test slowly and report in less than 20 seconds by construction

[16]Significant at the 10% level for Wilcoxon Rank-Sum tests within and across games

Honest Senders would have acted Naively, then the intuitive result appears that there are less Naive Senders in the treatment than in the control[17].

However, we proceed without making any assumptions on the unobserved behavior of Honest Senders. **we find support for H3**. Indeed, across the TS sessions and both games 14.29% of senders are strategic, significantly more than in the control[18]. Thus, we do find evidence of senders strategically manipulating their disclosure times. We also find that across both games there are significantly more 'Quick Testers' than in the control[19]. However, we are unable to draw any conclusions regarding which of the described strategies (Quick Testing or Strategic Delay), is more dominant in this environment.

8.3 Receiver Behavior

Given that strategically manipulating ones time can be more expensive than not, and the prevalence of naïve behavior, it is of interest to determine if receiver's exhibited sophistication around disclosure times.

To explore receiver behavior, we analyze how often they 'follow the report' when the state is good. Indeed, as expected this number is higher in the control than in the treatment. Unsurprisingly, receivers paired with Honest Senders also follow the report more than those paired with Naive or Strategic Senders. Interestingly, Strategic senders only do marginally better than naive senders on average, suggesting some degree of naïevity by receivers or over optimism regarding the signal.

9 Conclusion

Response time is a relevant signalling component which has been largely unexplored as a strategic variable. We build a model which, to the best of our knowledge, is the first to consider state-dependent implementation of a KG (2011) style 'test' which maintains optimal posteriors while introducing variation due to testing times. We find support for our model of strategic disclosure timing. Specifically, we observe that disclosure timing is heavily state-dependent when there is no incentive to manipulate disclosure timing, but less state-dependent when timing is a visible signal. We observe that senders are willing to pay (rationally) high costs to strategically manipulate their disclosure timing. Thus, we find support for our 'Quick-Testing' Equilibrium. This implies that subjects 'first-order'

[17] Significant at a 5% level for two-tailed t-tests within and across games

[18] Wilcoxon rank-sum test, z = -2.468, p=.0136

[19] Wilcoxon rank-sum test, z=-1.989, p=.0457

understand the importance of strategically manipulating their disclosure timing but are failing to recognize the opportunity to strategically delay in the good state to increase their profits through slow testing in the bad state.

These results are vulnerable to potential experimenter demand effects since we explicitly inform receivers of the sender's disclosure time on their screen. However, we give no instruction or demand for how subjects should utilize this information, nor do we give any indication to them that this is special information. For robustness, future sessions may explore the effect of less explicit reports of disclosure timing to observe the effect of *salience* on strategies of disclosure timing. Of interest also is the robustness of receiver sophistication, we may explore this in future sessions where receivers are shown disclosure times from the control (where strategic time manipulation should not be present, and subjects know this) and explore how receivers respond to information and disclosure times.

While there is evidence of senders playing a strategic delay strategy, this has not been observed widely. Further research into why senders don't play this strategy and if certain environments facilitate or induce them to play this strategy is of interest. Additionally, a very large number of senders decided to play 'honestly', especially in the control. Further exploration of why senders exhibit these truthtelling behaviors and if the treatment induced these behaviors is also of interest.

References

Belot, Michele, and Jeroen VanDeVen. 2017. "How private is private information? The ability to spot deception in an economic game." *Experimental Economics*, 20: 19–43.

Bergemann, Dick, and Stephen Morris. 2019. "Information Design: A Unified Perspective." *Journal of Economic Literature*, 57: 44–95.

Cooper, David, Ian Krajbich, and Charles Noussair. 2019. "Choice-Process Data in Experimental Economics." *Journal of the Economic Sciences Association*, 5: 1–13.

Crawford, Vincent, and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica*, 50: 1431–1451.

Frechette, Guillaume, Alessandro Lizzeri, and Jacopo Perego. 2021. "Rules and Commitment in Communication: An Experimental Analysis." *Electronic Journal*, 1–71.

Frydman, Cary, and Ian Krajbich. 2021. "Using Response Times to Infer Others' Beliefs: An Application to Information Cascades." *Management Science*, in press: 1–67.

Gneezy, Uri. 2005. "Deception: The Role of Consequences." *American Economic Review*, 95(1): 384–394.

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

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian Persuasion." *American Economic Review*, 101.

Konovalov, Arkady, and Ian Krajbich. 2020. "Decision Times Reveal Private Information in Strategic Settings: Evidence from Bargaining Experiments." *In Press*.

Nguyen, Quyen. 2017. "Bayesian Persuasion: Evidence from the Laboratory." *Electronic Journal*.

Spence, Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics*, 355–374.

Sutter, Matthias. 2009. "Decption Through Telling the Truth?! Experimental Evidence from Individuals and Teams." *The Economic Journal*, 47–60.

Tergiman, Chloe, and Marie Claire Villeval. 2021. "The Way People Lie in Markets: Detectable vs. Deniable Lies." *Management Science (RR)*, 1–61.

A Extensions

A.1 Mixtures of Sophisticated and Naïve Types

Suppose that the receiver's type is unknown to the sender. That is, the sender faces a mixture of sophisticated types or naïve type. Specifically, let ϕ be the probability of the sender encountering a naïve type[20]. Assume once more that when the state is good, the sender chooses $a_S = (g_S, t_1)$. if the sender chooses the slow test, $a_S = (\lambda_S, t_2)$ when $\Theta = B$, the sender's report will only be followed by a naïve type of receiver. If a sophisticated type sees $M = (g_S, t_2)$ They will recognize that $\Theta = B$ and not follow the sender's report. Thus, the sender needs to determine if it is worth the high cost to test quickly given the probability of facing a naïve type. Formally, the sender should choose $a_S = (\lambda_S, t_2)$ if

$$\phi\beta\lambda - C(\lambda_S, t_2) \geq \beta\lambda - C(\lambda_S, t_1)$$

Where the left-hand expression is the expected utility of the sender conducting the slow test, λ, t_2, and the right-hand expression if the expected utility of the sender conducting the quick test, λ, t_1. Rearrange and observe that the sender conducts the slow test if

$$\phi \geq 1 - \frac{\rho}{\beta\lambda}$$

Where $\rho = C(\lambda_S, t_1) - C(\lambda_S, t_2)$, the "speed premium" for faster testing. Naturally, as quick testing becomes more expensive, the likelihood threshold of naïve types required for the sender to choose the cheaper action, $a_S = (\lambda_S, t_2)$, is decreasing. Furthermore, as β increases, the likelihood threshold of naïve types required for the sender to choose to conduct the slow test, $a_S = (\lambda_S, t_2)$, is also increasing. Intuitively, as the "incentive" becomes stronger, the sender increasingly prefers to pay the cost associated with fast signalling and ensure all receivers follow the message. Likewise, as the test becomes noisier (λ increasing), the liklihood threshold of naïve types required for the sender to conduct the slow test is increasing. Intuitively, provided that the receiver will follow the test, increasing λ increases the probability of getting a positive payoff which increases the value of quick testing.

From these observations we will observe a final equilibrium, the "Mixed R Type Equilibrium". By assumption, when $\Theta = G$, senders will choose $a_S = (g_S, t_1)$. When

[20]Alternatively, one can consider the single receiver case from before, where ϕ is the probability of the receiver being sophisticated in any given period.

$\Theta = B$, and $\phi \geq 1 - \frac{\rho}{\beta}$ the sender will choose to conduct the slow test, $a_S = (\lambda_S, t_2)$. Otherwise, when $\Theta = B$ the sender will conduct the quick test, $a_S = (\lambda_S, t_1)$. If the sender chooses to conduct the slow test when $\Theta = B$, sophisticated receivers will not follow the senders report when observing $(M = g_S, t_2)$. If the sender chooses to conduct the quick test when $\theta = B$, then all receiver types will follow the senders report.

It is worth noting here the existence of another equilibrium. Namely, the equilibrium that arises if we relax the assumption that the sender chooses $a_S = (g_S, t_1)$ when $\Theta = G$. Relaxing this assumption, then the sender should choose $a_S = (g_S, t_2)$ when $\Theta = G$ and thus they can choose to test slowly and both types of receivers will follow the report. IE, even with a mixture of receiver types relaxing this assumption leads us to the 'Strategic Delay Equilibrium' once again. In other words *The strategic delay equilibrium is robust to uncertainty over sophistication of receivers, and holds in the naïve case as well.*

B Proofs

B.1 Non-existence of other equilibria

Theorem 9. *There are no equilibria other than the ones described when* $\mu > \frac{\sigma_1 - \gamma_1}{(\sigma_1 - \gamma_1) + (\sigma_2 - \gamma_2)}$, $\lambda \leq \frac{(1-\mu)(\sigma_1 - \gamma_1)}{\mu(\sigma_2 - \gamma_2)}$, *and* $\beta\lambda \geq C(\lambda, t_2)$

Proof.

Case 1: Full Revelation Equilibria

First, recall that there does exist full revelation if $C(\lambda, t_1)$ is sufficiently high.

Suppose that another 'full revelation' equilibrium exists. That is, suppose that a 'full revelation' equilibrium exists when $C(\lambda, t_1) \leq \beta\lambda$

It immediately follows that the receiver follows the senders advice when there is full revelation. However, since this is the case then there is a a profitable one-shot deviation where the sender chooses to implement a test for any $\lambda > 0$.[21]. Thus, this contradicts the assumption.

Case 2: Separation Equilibria

Suppose there exists a 'separation equilibria' (IE, one where t perfectly predicts Θ)

[21] Notice here that receiver punishment is not possible, since deviation from following the report is a 'non-credible threat'

Notice then that there is a profitable one-shot deviation for the receiver, who chooses to identify the correct state instead of following the report (when there is conflicting information) based on t. Thus, this contradicts the assumption.

Case 3: Mixture Equilibria

Suppose there exists some equilibria where at least one of the players mixes.

Notice that, by construction, 'mixing' when $\Theta = G$ is degenerate since the test provides the same report as honest reporting at some positive cost.

Thus, consider when the sender mixes when $\Theta = B$. The receiver can either mix only when $M = g$, only when $M = b$, or always. By choice of λ, without sender mixing the receiver always prefers g_R or is exactly indifferent between g_R and b_R when $M = g$. Thus, if the sender mixes he increases the receiver's preference for g_R (which is already the dominant strategy) while reducing his own payoff (by decreasing the amount of times the report returns g since he is increasing the amount of honest disclosures).

Finally, notice that the receiver will never mix when $M = b$. Since $M = b \implies \Theta = B$ he strictly reduces his payoff by choosing $a = g_R$ with any positive probability when observing b.

Thus, the assumption is contradicted.

Finally, by exhaustion conclude that indeed no other equilibria exist for the parameters described as desired.

□

C Experiment Instructions

Subjects were given a handout of the following instructions[22] [23], while the experimenter[24] read the instructions verbally.

Welcome

Welcome. This is an experiment in the economics of decision-making. Florida State University has provided funds for this research. You will be paid for your participation in

[22]Some formatting is minimally different than that which subjects saw. All content presented is identical, including any **bold**, *italic*, or underlined text. Original copies available upon request.

[23]DR in BLACK, TS identical to DR, except for changes in OLIVE GREEN, RT is identical to TS, except for changes in RED

[24]The same experimenter was present for all sessions

the experiment. The exact amount you will be paid will depend on your and/or others' decisions. All payments for this experiment are denominated in Experimental Currency Units (ECUs). 1 ECU is worth $.01 (1 cent). Your payment will consist of the amount of ECUs you accumulate, plus a $5 participation fee. You will be paid privately, either via Venmo, check, or cash (USD $) at the conclusion of the experiment.

If at any time you have a question, please raise your hand and an experimenter will assist you. Please do *not* talk, exclaim, or try to communicate with other participants during the experiment. At this time, please put away all outside materials (such as cellphones, books, bags, notebooks, etc.). Participants violating the rules will be asked to leave the experiment and will *not* be paid. At no point during today's experiment will an experimenter deceive you.

At the conclusion of these instructions, we will be conducting a brief quiz. Everyone will need to answer all questions correctly before we can begin the experiment, so please pay close attention to these instructions. You may refer to these instructions at any time during the quiz and during the experiment.

Stages and Rounds

This experiment consists of 2 stages, with 8 rounds in each stage. You will be assigned to one of two roles, *Sender* or *Receiver*. Regardless of your decisions, each round will last at least 50 seconds. You will be assigned to your role in the beginning of stage one, and this will be your role *for both stages*. IE, you will stay in the same role for the entire experiment. You will be randomly assigned a partner in the beginning of a stage, and they will be your partner for the duration of that stage. It is not possible to play with the same partner in both stages. IE, you will never play with your Stage 1 Partner in Stage 2, and it is never possible for the partner you play with in Stage 2 to be someone you have played with before.

Playing a Round of the Game

Drawing Cards

The experimental software is programmed with a virtual deck of 3 cards. Each card has a symbol on it. Two of the cards have a triangle symbol, while the other card has a star symbol. At the beginning of each round, one card will be randomly drawn. At the end of each round, the card is 'returned' to the virtual deck and the deck is 'shuffled'. Consequently, each card draw is *independent* across rounds. A card drawn in one round will not influence the card drawn in any other round. Thus, there is a (2/3), or approximately 66.67%, chance that the card will have a triangle on it, and a (1/3), or approximately 33.33% chance that the card will have a star on it.

What does the SENDER do?

At the beginning of each round, each sender will be *endowed* with 5 ECUs. The senders will move first, and they will directly observe the symbol on the card. They will then send a report to the receiver with information about the card's symbol. After receiving the report, the receiver will guess the symbol on the card. Details on how the senders' reports are produced are provided further below.

What does the RECEIVER do?

Before receiving the report, receivers will have 50 seconds to answer up to 5 multiple choice math questions, with correct answers earning them an additional 1 ECU per correct answer. **Once an answer is selected, it will be recorded and the next question will be shown, Receivers cannot go back, or change their answers to the math questions**. Their screen will look like this.

[[IMAGE]]

Once the report is received, the receiver will guess whether the card's symbol is star or triangle.

How are Your Payoffs Determined?

If your role is Receiver, you will earn 20 ECUs if you correctly guess the symbol on the card. If your guess is incorrect, you earn nothing (0 ECUs).

If your role is Sender, you will earn 20 ECUs if the Receiver guesses the card's symbol is "Star" *regardless of the symbol that is on the card*. If the Receiver guesses the card's symbol is "Triangle", then you will earn 0 ECUs (again, regardless of the symbol that is on the card).

How is the SENDER'S REPORT produced?

Upon observing the symbol of the card, the sender can choose to either *report the symbol, conduct a test, or conduct a test quickly*. Below is how a sender's screen will appear. In this case, the card's symbol is TRIANGLE.

[[IMAGE]]

If the sender chooses to *report the symbol*, then the symbol reported will be the symbol of the card.

Conducting a Test

Instead of reporting the symbol, the sender may choose to conduct a test and report the results of the test. The test is not perfect. If the card has a star, then the test will *always* report "star". However, if the card has a triangle on it, then there is a 25%

chance that the test will *incorrectly* report “star” and a 75% chance that the test will report “triangle”. In other words, when the card has a triangle, the test will report inaccurately 25% of the time. *Conducting a test* costs 1 ECU, and takes 20 seconds to produce results. However, a sender may instead choose to pay 3 ECUs and *conduct a test quickly*, producing results immediately. **The accuracy of the test is fixed and cannot be manipulated by any subject,and the accuracy of the test does not change if it is conducted quickly**.

If a sender chooses to conduct a test (regularly), they will see the following screen while the test is being conducted.

[[IMAGE]]

Notice that the sender now has 4 ECUs, since the test cost 1 ECU to conduct

Below is the test results screen. Notice that once a test is conducted and the sender obtains results, *the sender still needs to report the results to the receiver, this does not happen automatically!*

Regardless of when a sender sends a report, receivers will not receive the report *until at least 50 seconds has elapsed*. Once the report is received, the receiver’s screen will look like this (please see next page).

[[THIS REPLACES THE TEXT ABOVE]]Once a sender sends a report, receivers will receive it immediately. Once the report is received, the receiver’s screen will look like this (please see next page)

[[IMAGE]]

The report does not tell the receiver what the sender’s action is, but only whether the card is reported to have a “Star” or “Triangle” on it. *The reported symbol may or may not be the card’s symbol*. The receiver will then guess what symbol they think is on the card.

[[INSERT]] *If the report is received before 50 seconds has elapsed, once they have guessed the card’s symbol they will be allowed to resume math questions until a total of 50 seconds on math questions has elapsed (or they answer all 5 questions).Thus, regardless of when a report is sent, all receivers will have up to 50 seconds to answer math questions.*

We will now demonstrate each of these sender actions on the overhead above[25].

[25] At this point, the experimenter clicked through screens showing how clicking “report” simply reported the card’s symbol, while conducting a test would lead subjects to a “testing in progress” screen for 20 seconds (which the experimenter allowed to pass in real time) before reaching a result screen, while conducting a ’quick test’ would lead subjects to the results screen immediately

Payoffs & Feedback

Each round will conclude with a summary of what happened that round. Senders will see the card symbol, their action, [[INSERT]] when the report was received, the reported symbol, the receivers guess, and their payoff. Receivers will see the card symbol, the reported symbol, [[INSERT]] when the report was received their guess, and their payoff. As stated above, senders will earn an additional 20 ECUs if the receiver guesses that the card has a star on it, and 0 otherwise. Sender's reported payoffs will include any ECUs from their endowment that round, after deducting testing costs (if any). Receivers will earn 20 ECUs if they correctly guess the symbol of the card, and 0 otherwise. Note that Receiver earnings from the math questions *will not* be reported in this feedback, but instead will be added to your take-home payment at the conclusion of the experiment.

Stage 3

Following stages 1 and 2, we will conduct stage 3. In this Stage you will be asked a series of questions. Only NUMERIC answers will be accepted (no letters, no unit labels). You will have 30 seconds to answer each questions. Once you type in an answer, you **must push submit for it to be counted, if you do not push submit before 30 seconds expires, your answer will not be counted**. Additionally, **once you submit an answer you cannot go back, and you cannot change your answer**. Each correctly answered question will earn you an additional 1 ECU. After Stage 3, you will see your earnings from the experiment. Afterwards, you will be asked to answer some demographic and payment questions.

Confidentiality and Payment

At no point in time will we identify any of the people you interact with in the experiment nor will you be identified to them. In other words, the actions you take will remain confidential. At the end of the experiment you will be paid your accumulated earnings along with your participation fee of $5. You will be paid privately, and we will not disclose your payoff to the other participants in the experiment. *Please do not leave the experiment unless dismissed by the experimenter*.

Reference Guide

On the next page, you'll find a summary of the structure of the game. Please feel free to refer to it, and the rest of these instructions, during any part of the payoff quiz or the experiment. You may detach pages of these instructions. Unless there are questions, we will begin the payoff quiz shortly.

——————PAGE BREAK——————

Reference Guide

Probability the Card's Symbol is a Star = 33.33% (1/3) Probability the Card's Symbol is a Triangle = 66.67% (2/3)

The Sender's Payoffs (Not Including Their Endowment or Test Cost Deductions)

	Receiver Guesses Star	Receiver Guesses Triangle
Card's Symbol is Star	20 ECUs	0 ECUs
Card's Symbol is Triangle	20 ECUs	0 ECUs

The Receiver's Payoffs (Not Including Earnings from Math Question Payments)

	Guess Star	Guess Triangle
Card's Symbol is Star	20 ECUs	0 ECUs
Card's Symbol is Triangle	0 ECUs	20 ECUs

Probabilities of Testing Outcomes

	Probability the Test Reports Star	Probability the TestReports Triangle
Card's Symbol is Star	100%	0%
Card's Symbol is Triangle	25%	75%

*****The Probabilities of Testing Outcomes are the same when testing quickly*****

1. When the card's symbol is **STAR**, a test will always report Star, and never report Triangle.
2. When the card's symbol is **TRIANGLE**, there is a 35% chance a test reports star, and a 75% chance the test reports triangle
3. The above 2 statements apply when conducting tests quickly

Test Costs and Testing Times

	Cost	Testing Time
Test	1 ECUs	20 Seconds
Test Quickly	3 ECUs	0 Seconds

D Payoff Quiz Questions

1. True of False: Receiving a Report of **STAR** means that the card's symbol must be star (Correct Answer: FALSE)
2. True of False: When the card has a **Triangle** on it, the test will always report TRIANGLE (Correct Answer: FALSE)
3. True of False: When the card has a **STAR** on it, the test will always report STAR (Correct Answer: TRUE)
4. True or False: If the card has a TRIANGLE on it, and the receiver guesses that the card has a STAR. The sender and receiver's payoff that round will be 0. (Correct Answer: FALSE)
5. True of False: The quick test is more accurate than the regular test (Correct Answer: FALSE)

E Cognitive Response Test (CRT) Questions

1. How many cubic feet of dirt are there in a hole that is 2' deep x 2' wide x 2' long? IN CUBIC FEET ______________ (Intuitive Answer: 8, Correct Answer: 0)
2. A man buys a pig for \$40, sells it for \$50, buys it back for \$60, and sells it finally for \$70. How much money has he made? IN DOLLARS ______________ (Intuitive Answer: 10, Correct Answer: 20)
3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 56 days to cover the entire lake, how long would it take for the patch to cover half of the lake? IN DAYS ______________ (Intuitive Answer: 28, Correct Answer: 55)
4. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? IN MINUTES ______________ (Intuitive Answer: 100, Correct Answer: 5)
5. A farmer makes 4 piles of hay in one corner of a field and 5 other piles in another corner. If he merges them, how many piles will he have? ______________ (Intuitive Answer: 9, Correct Answer: 1)