STRATEGIC INATTENTION IN THE SIR PHILIP SIDNEY GAME

By Mark Whitmeyer[1]

Infamously, the presence of honest communication in a signaling environment may be difficult to reconcile with small signaling costs or a low degree of common interest between sender and receiver. This paper posits that one mechanism through which such communication can arise is through inattention on the part of the receiver, which allows for honest communication in settings where, should the receiver be fully attentive, honest communication would be impossible. We explore this idea through the Sir Philip Sidney game in detail and show that some degree of inattention is always weakly better for the receiver, and may be strictly better. Moreover, some inattention may be a Pareto improvement and leave the sender no worse off. We compare limited attention to Lachmann and Bergstrom's (1998) notion of a signaling medium and show that the receiver-optimal degree of inattention is equivalent to the receiver-optimal choice of medium.

Keywords: Costly Signaling, Strategic Inattention, Handicap Theory, Information Design

[1]Department of Economics, University of Texas at Austin

Email: mark.whitmeyer@utexas.edu. Thanks to Rosemary Hopcroft, Cooper Howes, Ryan Oprea, Jan Schlupp, Joseph Whitmeyer, and Thomas Wiseman. All remaining errors are my own.

date: January 7, 2020

1. Introduction

> I have only one eye–I have a right to be blind sometimes.
> ...I really do not see the signal.
>
> Admiral Horatio Lord Nelson

The handicap principle is an important notion in signaling games. Put simply, this principle states that order to facilitate meaningful communication in situations in which there are conflicts of interest, a cost is necessary. One standard setting for formally investigating this principle is the discrete Sir Philip Sidney Game (Maynard Smith [13]). In the classic formulation of this game, two players interact, a sender and a receiver. The sender's type is uncertain: he is either healthy (with probability $1 - \mu$) or needy (with probability μ). The sender is the first mover, and may choose to either cry out and incur a cost of $c > 0$, or stay silent and incur no cost. Following this action (henceforth referred to as a *signal*) by the sender, the receiver observes the signal before choosing whether to donate a resource and incur a cost of $d > 0$ or do nothing and incur no cost.

Should a sender receive a donation, his probability of survival is 1 regardless of his type. On the other hand, if a sender does not receive a donation then his probability of survival is $1-a$ if he is needy and $1-b$ if he is healthy, where $a > b$. In addition, there is a relatedness parameter $k \in [0, 1]$ that captures the degree of common interest between the sender and the receiver–under any vector of strategies each player receives k times the payoff of the other player plus his or her own payoff.

The purpose of this paper is to explore *strategic inattention* in this setting. That is, in recent work, Whitmeyer (2019) [15] highlights that *full transparency* is not generally optimal for the receiver in signaling games, and takes steps to characterize the receiver-optimal degree of transparency in those games. In particular, less than full transparency may beget informative equilibria in situations where there is little to no meaningful communication under full transparency.[1]

In the Sir Philip Sidney game, such a breakdown in communication occurs for certain

[1] This has been noted by Lachmann and Bergstrom (1998) [10] in the context of the continuous Sir Philip Sidney game.

regions of the parameters corresponding to a low cost of crying out, a low degree of relatedness, and a low cost of donation d. However, not all is lost. We embed the game into a slightly larger game, one with an additional (first) stage in which the receiver chooses a level of attentiveness in the signaling game. She chooses a probability x: in the signaling portion of the game she observes the sender's choice of signal with probability x and does not with probability $1 - x$, which probabilities are independent of the signal choice of the sender.

Remarkably, some degree of inattention is always (weakly) optimal for the receiver, and in certain regions of the parameter space the receiver is strictly better off when she is partially inattentive. Inattention is helpful in the following two ways. First, it provides a lower bound on the set of equilibrium payoffs of the game: since complete inattention may be chosen in the initial stage, at equilibrium the receiver can do no worse than the unique equilibrium payoff in the signaling game given complete inattention, where the two types of sender both remain silent. Second, and perhaps more compellingly, there is an interval of the attention parameter in which a separating equilibria may manifest, despite the non-existence of such an equilibrium in the game with full attention. In short, inattentiveness enhances communication.

1.1. Related Work

There is a substantial literature in theoretical biology, philosophy, and economics exploring signaling games, commencing with Lewis (1969) [12]. In the same era, Zahavi (1975) [19] published his seminal work on handicaps in biology, and this notion was later incorporated in signaling games themselves.

The discrete Sir Philip Sidney Game (Maynard Smith (1991) [13]) has become one of the paradigmatic settings for investigating the handicap principal in signaling games. Numerous other papers investing this and other closely related games have ensued, and if ever an area could be termed burgeoning it is this one. A list of recent works includes Bergstrom and Lachmann (1998) [2], Huttegger and Zollman (2010) [6], Huttegger, Skyrms, Tarrès, and Wagner (2014) [5]. Also pertinent is the survey article on

the handicap principle, Számadó (2011) [14].

In recent work, Zollman, Bergstrom and Huttegger (2013) [20] highlight the fact that empirically, "researchers have not always been able to find substantial signal costs associated with putative costly signal systems–despite evidence that these systems do convey, at least, some information among individuals with conflicting interests", and ask, "What then, are we to make of empirical situations in which signals appear to be informative even without the high costs required by costly signaling models?" As they mention, other works have illustrated–see e.g. Lachmann, Számadó, and Bergstrom (2001)[11]–that this issue may be ameliorated by recognizing that costly signals need not be sent on the equilibrium path, and that it is the high cost of sending a (deviating) signal that keeps the sender types honest. In addition, they forward an alternate resolution of the issue: there are also partially informative equilibria (in which players mix), which may be sustained despite low or non-existent costs. Hence, in the same vein, we explore a third possibility: that limited attention may explain the existence of honest signaling, even with low costs.

The paper closest in spirit to this one is Lachmann and Bergstrom (1998) [10]. There, the authors allow for perceptual error on the part of the receiver and introduce the notion of a *medium*, which distorts the signals observed by the receiver. They illustrate that different media may beget different equilibria, and that some media may even foster honest communication impossible in other media. In this paper we endogenize the medium by making it a choice of the receiver. Moreover, we restrict the set of media the receiver can choose to those of a specific sort: those which correspond to inattention. However, as we show in Theorem 3.9, this is no restriction, the receiver-optimal equilibrium under her optimal choice of attention remains supreme even were she able to choose any medium, however complex. Other papers that allow for perceptual error, or noise, include Johnstone and Grafen (1992) [9], Johnstone (1994, 1998) [7, 8], Lachmann, Számadó, and Bergstrom (2001)[11], and Wiley (2013, 2017) [17, 18].

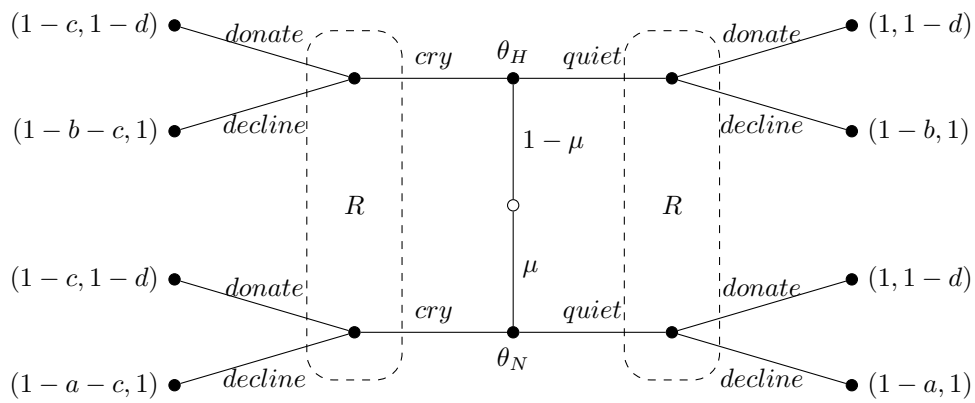

FIG 1. *The Sir Philip Sidney Game*

2. The Classic Sir Philip Sidney Game

We begin by revisiting the (more-or-less standard) Sir Philip Sidney game (where we allow the cost, c, to equal 0), which game is depicted in Figure 1. There are two players, a sender (he) and a receiver (she); and the sender is one of two types, healthy or needy: $\Theta = \{\theta_H, \theta_N\}$. The sender's type is his private information, about which both sender and receiver share a common prior, $\mu \coloneqq \Pr(\Theta = \theta_N)$.

After being informed of his type, the sender chooses to either cry out (*cry*) or stay quiet (*quiet*). The receiver observes the sender's choice of signal (but not his type), updates her belief about the sender's type based on her prior belief and the equilibrium strategies and elects to either donate a resource (*donate*) or refuse to donate (*decline*). We impose that $a > b$ and that a, b, c, and d take values in the interval $[0, 1]$. There is also a relatedness parameter $k \in [0, 1]$: after each outcome, a player receives his own payoff plus k times the payoff of the other player.

Throughout, we impose the following conditions:

$$a > \frac{d}{k} > b$$

and

$$a > b > dk + c$$

The first assumption ensures that if the receiver is (sufficiently) confident that the sender is healthy then she strictly prefers not to donate and if she is sufficiently confident that

the sender is needy then she strictly prefers to donate. The second assumption eliminates any separating equilibria.

In addition, we define

$$\hat{d} \coloneqq k\left(\mu a + (1-\mu)\, b\right)$$

which will thankfully save some room on the manuscript. We describe the equilibrium in the signaling game as a four-tuple $(\cdot, \cdot; \cdot, \cdot)$, where the first entry corresponds to the strategy of θ_H, the second entry to the strategy of θ_N, the third entry to the response of the receiver to *quiet*, and the fourth entry to the response of the receiver to *cry*. In the case of pooling equilibria (equilibria in which both types of sender choose the same signal), we leave the response of the receiver to an off-path signal as $\cdot$ when there may be multiple responses that would sustain an equilibrium.

We have the following result, which follows from our parametric assumptions:

Lemma 2.1. *There exist no separating equilibria.*

Proof. Standard see e.g. Bergstrom and Lachmann (1997) [1]. ■

There do; however, exist pooling equilibria, both those in which both sender types choose *cry* and those in which both sender types choose *quiet*. Note that the pooling equilibrium in which both sender types choose *cry* requires that the receiver's belief upon observing *quiet* (an off-path action) be such that she would at least (weakly) prefer to choose *decline* rather than *donate*. Moreover, the pooling equilibrium in which both sender types choose *quiet* and to which the receiver responds with *decline* also requires that the receiver's belief upon observing *cry* (an off-path action) be such that she would at least (weakly) prefer to choose *decline* rather than *donate*. In some sense, this is less convincing of an equilibrium: shouldn't the needy bird be more likely to cry out?[2]

The other pooling equilibrium, that in which both sender types choose *quiet* and the receiver responds with *donate* makes no restrictions on the receiver's off-path beliefs and is in that sense quite strong. Formally,

[2] *Cf.* the "Intuitive Criterion" as developed in Cho and Kreps (1987) [3].

LEMMA 2.2. *There exist pooling equilibria:* $(cry, cry; \cdot, decline)$ *is never an equilibrium. If* $d \leq \hat{d}$ *then* $(cry, cry; \cdot, donate)$ *is an equilibrium, and* $(quiet, quiet; donate, \cdot)$ *is an equilibrium. If* $d \geq \hat{d}$*, then* $(quiet, quiet; decline, \cdot)$ *is an equilibrium.*

PROOF. Standard, see e.g. Bergstrom and Lachmann (1997) [1]. ■

There also exist equilibria in which players mix. We omit those equilibria since the receiver's payoff is higher under the best pure-strategy (pooling) equilibrium than under one in which at least one sender mixes. To see this observe that there cannot be an equilibrium in which each message is played and a different action is strictly preferred after each since there would be a sender type who could deviate profitably to the message that is followed by *donate*. Thus, all equilibria must be ones where the receiver is (weakly) willing to do the same action following any message. However, of those equilibria, the best for the receiver is the one in which the sender types pool on *quiet*.

Accordingly, if $d \geq \hat{d}$ then the receiver optimal equilibrium is $(quiet, quiet; decline, \cdot)$, which yields her a payoff of

$$V = 1 + k\,(1 - b) - k\mu(a - b) \tag{1}$$

and if $d \leq \hat{d}$ then the receiver optimal equilibrium is $(quiet, quiet; donate, \cdot)$, which yields her a payoff of

$$V = 1 - d + k \tag{2}$$

Note that we may rewrite the condition $d \geq \hat{d}$ as

$$\mu \leq \frac{d - bk}{k(a - b)}$$

and so we can write the receiver-optimal payoff as a function of the belief

$$V(\mu) = \begin{cases} 1 + k\,(1 - b) - k\mu(a - b), & 0 \leq \mu \leq \frac{d-bk}{k(a-b)} \\ 1 - d + k, & \frac{d-bk}{k(a-b)} \leq \mu \leq 1 \end{cases}$$

Evidently, V is convex in μ. An interesting ramification of this is that any (free) *ex ante* information about the sender's type benefits the receiver. This is not generally true in

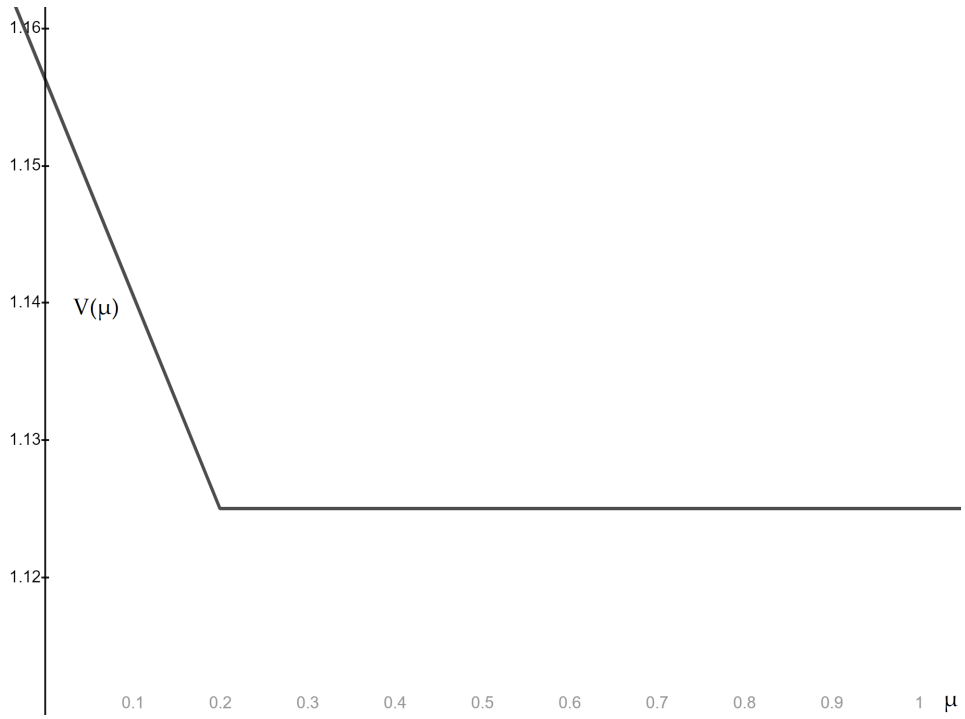

FIG 2. *Receiver's Payoff, No Inattention*

communication games even if there are just two types and two receiver actions (when the receiver's payoff is directly affected by the sender's message).[3] However, it is true here: any free information benefits the receiver.

The receiver's optimal equilibrium payoff is depicted in Figure 2 for the following values of the parameters: $a = 1$, $b = 3/8$, $c = 11/64$, $d = 1/8$, and $k = 1/4$.

3. Strategic Inattention

We now explore the notion that full attention may not be generally optimal for the receiver. We modify the game by introducing an initial stage in which the receiver chooses and commits to a level of attention. This is modeled in the following parsimonious fashion: in the first stage the receiver chooses an **Attention Parameter**. x, which is

[3]To see more discussion of this in communication games see Whitmeyer (2019) [15] and Whitmeyer 2019 [16]

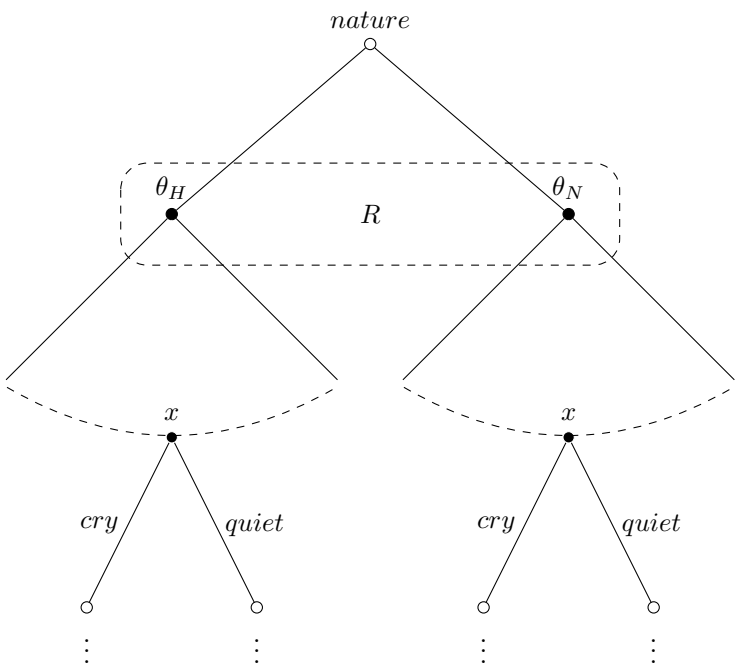

FIG 3. *Choosing an Attention Parameter*

publicly observable,[4] and then the signaling game proceeds in the standard manner. See Figure 3.

The attention parameter is straightforward: the receiver simply observes the sender's choice of signal with probability x and does not observe the choice with probability $1 - x$. If the receiver does not observe the signal then she simply makes the optimal choice given her information, which is merely her prior. Hence, she chooses *decline* if and only if $d \geq \hat{d}$ and *donate* otherwise.

The game has two stages: first, the receiver chooses x; then second, the signaling game occurs under parameter x. Hence, we search for pure-strategy[5] subgame perfect equilibria through backward induction.

[4]Alternatively, we can think of x as being an exogenous primitive of the model, and thus our endeavor may be reinterpreted as searching for the optimal level of attention for the receiver.

[5]This is justified in part because there must be a receiver-optimal equilibrium in pure strategies. This follows due to the fact that the optimal level of attention is equivalent to the receiver-optimal medium, as detailed in Section 3.3; and [15] establishes that in that problem it is without loss of generality to restrict the sender types to pure strategies.

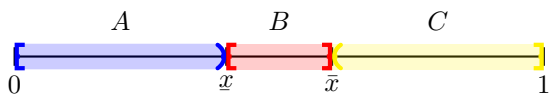

FIG 4. *The Critical Regions of x*

3.1. The Signaling Subgame

Let us examine the subgame that consists of the signaling game for a given choice of attention parameter x. As we discover, there are two critical cutoff beliefs of x, $\underline{x}$ and $\bar{x}$, which divide the range of possible values of x into three regions, depicted in Figure 4. Explicitly,

$$\underline{x} := \frac{c}{a - dk}, \qquad \bar{x} := \frac{c}{b - dk}$$
$$A := [0, \underline{x}), \qquad B := [\underline{x}, \bar{x}], \qquad C := (\bar{x}, 1]$$

Our first result highlights that, in contrast to the first section of this paper, in which there did not exist separating equilibria, other values of x may beget separation. ,

LEMMA 3.1. *If $x \in B$ then there exists a separating equilibrium in which θ_H chooses quiet and θ_N chooses cry. These equilibria yield payoffs to the receiver of*

$$V(x) = 1 + k(1 - b) - k\mu(a + c - b) + \mu x(ak - d) \tag{3}$$

for $d \geq \hat{d}$, and

$$V(x) = 1 - d + k(1 - c\mu) + (1 - \mu)x(d - bk) \tag{4}$$

for $d \leq \hat{d}$. The level of attention that maximizes the receiver's payoff is

$$x^* = \bar{x} = \frac{c}{b - dk}$$

PROOF. Proof is left to Appendix A.1. ■

The intuition behind this result is simple. By being strategically inattentive (choosing x within B), the receiver can lessen the incentive of either type to deviate and mimic the other. The receiver-optimal attention parameter x^* is that which leaves type θ_H

indifferent between separating and deviating to mimic θ_N. The receiver would like to pay the maximal amount of attention such that she is sufficiently inattentive for the sender types to separate.

Furthermore, when $d \geq \hat{d}$,

$$V(x^*) = 1 + k(1-b) - k\mu(a+c-b) + \frac{\mu c(ak-d)}{b-dk}$$

and when $d \leq \hat{d}$,

$$V(x^*) = 1 - d + k(1-c\mu) + \frac{(1-\mu)c(d-bk)}{b-dk}$$

Of course, there exist pooling equilibria as well. *Viz*,

Lemma 3.2. *There does not an exist an $x \in [0,1]$ such that $(cry, cry; \cdot, decline)$ is an equilibrium. Moreover,*

1. *For $d \geq \hat{d}$, $(quiet, quiet; decline, \cdot)$ is an equilibrium. The receiver's resulting payoff is given in Expression 1.*
2. *For $d \leq \hat{d}$, $(quiet, quiet; donate, \cdot)$ is an equilibrium. $(cry, cry; \cdot, donate)$ is also an equilibrium if and only if $x \geq \bar{x}$. The receiver's resulting payoff is given in Expression 2.*

Proof. Proof is left to Appendix A.2. ■

The receiver-optimal pooling equilibria are those in which the sender types remain silent, and those equilibria are optimal (among all equilibria pooling or otherwise) for all $x \in A$ and $x \in C$. Note that there is only an equilibrium in which the sender types both choose *cry* if both d and x are sufficiently high. In particular, for any d, if x is sufficiently low then there is no equilibrium in which the sender types both *cry*.

Next, the following lemma establishes that if $x \in B$, it is possible that the separating equilibrium is optimal. To wit,

Lemma 3.3. *Let $d \geq \hat{d}$ and $x \in B$. Then the equilibrium that maximizes the receiver's payoff is $(quiet, cry; decline, donate)$ if and only if $x \geq kc/(ak-d)$. Otherwise it is $(quiet, quiet; decline, \cdot)$.*

Let $d \leq \hat{d}$ and $x \in B$. Then the equilibrium that maximizes the receiver's payoff is $(quiet, cry; decline, donate)$ if and only if $x \geq (kc\mu)/\left((1-\mu)(d-bk)\right)$. Otherwise it is $(quiet, quiet; donate, \cdot)$.

Proof. First, let $d \geq \hat{d}$. Using the receiver's payoff from the pooling equilibrium (Expression 1) and her payoff from the separating equilibrium (Expression 3), we have

$$\begin{gathered} V^{sep}(x) \geq V^{pool} \\ 1 + k(1-b) - k\mu(a+c-b) + \mu x(ak-d) \geq 1 + k(1-b) - k\mu(a-b) \end{gathered} \tag{5}$$

Or,

$$x \geq \frac{kc}{ak-d} \tag{6}$$

Note also that if $x = x^* = c/(b-dk)$ then this simplifies to $(ak-d) \geq k(b-dk)$.

Second, let $d \leq \hat{d}$. Then,

$$\begin{gathered} V^{sep}(x) \geq V^{pool} \\ 1 - d + k(1-c\mu) + x(1-\mu)(d-bk) \geq 1 - d + k \end{gathered} \tag{7}$$

Or,

$$x \geq \frac{kc\mu}{(1-\mu)(d-bk)} \tag{8}$$

If $x = x^*$ then this simplifies to $(1-\mu)(d-bk) \geq \mu k(b-dk)$. ■

Pausing briefly to look at the cutoffs above which separation is better, we see that the right-hand side of Inequality 6 is increasing in both c and d and decreasing in k and a. Hence, small[6] decreases in the signaling and/or donation costs enlarge the set of attention parameters x such that separation is better for the receiver than pooling. Analogously, small increases in the relatedness parameter and/or the cost suffered by the needy type have the same effect.

The right-hand side of Inequality 8 is also increasing in c and decreasing in k, and so the same intuition holds. However, it is now increasing in b and in μ and decreasing in

[6]The modifier "small" is required since we have already imposed several conditions on the values that the parameters may take.

d. It is easy to see why it should be increasing in μ: as the proportion of needy types increases, the uninformativeness of pooling is not as harmful to the receiver (recall that since $d \leq \hat{d}$ the receiver is donating). Similar reasoning explains the relationship with d: as d increases, pooling becomes more costly since donation itself is more costly. As b increases it becomes harder to elicit separation (from the high type), which reduces the receiver's benefit from the separating equilibrium.

Before proceeding onward let us revisit x^* and the properties of the optimal attention level. Directly, we have

Lemma 3.4. *x^* is decreasing in b and increasing in c, d, and k.*

This result seems reasonable: as signaling costs, donation costs, and relatedness increase, so does the amount of attention paid by the receiver at the optimum. This optimal attention level decreases, on the other hand, as the high type's loss from not receiving a donation increases. Thus, an increase in a parameter that helps drive communication (separation), or a decrease in the parameter that helps thwart communication, increases the optimal level of attention.

3.2. The First Stage

Armed with our analysis from the preceding subsection, we may now characterize the subgame perfect equilibria of the game. We introduce the following conditions

Condition 3.5. $d \geq \hat{d}$ and Inequality 6 holds.

Condition 3.6. $d \leq \hat{d}$ and Inequality 8 holds.

At long last, we characterize the subgame perfect equilibria of this game:

Theorem 3.7. *Suppose there exists some $\hat{x} \in B$ such that either Condition 3.5 or Condition 3.6 holds. Then, there exists a collection of subgame perfect equilibria consisting of a choice of $x \geq \hat{x}$ in the first stage, and $(quiet, cry, decline, donate)$ in the signaling portion of the game. The receiver optimal subgame perfect equilibria is that in which $x = x^* = c/(b - dk)$.*

Proof. This follows from Lemma 3.3. Subgame perfection is obtained since the sender types are willing to pool following any deviation by the receiver. ■

Next, we see that there are always subgame perfect equilibria in which no information is transmitted.

Theorem 3.8. *There is always a collection of subgame perfect equilibria consisting of any choice of x in the first stage, and $(quiet, quiet, decline, \cdot)$ for $d \geq \hat{d}$, and $(quiet, quiet, donate, \cdot)$ for $d \leq \hat{d}$ in the signaling portion. If neither Condition 3.5 nor Condition 3.6 holds then this collection is unique.*

Proof. This follows from the fact that the equilibria in which both sender types are quiet are the pooling equilibrium that (uniquely) maximize the receiver's payoff. Since these are the unique equilibrium should the receiver choose any $x \in A$, this must be the equilibrium played for any x since otherwise the receiver would have a profitable deviation in the initial stage to an $x \in A$. ■

This pair of theorems illustrates the two main effects of allowing the receiver to choose her level of attention initially. First, limited attention yields separating equilibria even when such equilibria could not exist under full attention. That is, honest communication manifests in a scenario in which the conflict between the receiver and the sender would usually be too great. Second, enabling the receiver to choose her level of attention ensures that the equilibrium played in the signaling portion of the game is relatively "good" for the receiver (either best or second-best) and provides a lower bound for the receiver's payoff.

3.3. Inattention Corresponds to the Receiver-Optimal Medium

Recall that in Bergstrom and Lachmann (1997) [10] the authors introduce the notion of a imperfect medium, which distorts signals. That is, the medium defines a conditional probability distribution of perceived signals dependent on which signals are actually sent. Here, we show that the optimal level of inattention, is equivalent to the best-possible medium for the receiver. To wit,

THEOREM 3.9. *V^{med} be the receiver's payoff for the receiver-optimal equilibrium under the best-possible medium for the receiver. Then there exists an attention parameter x such that $V(x) = V^{med}$.*

If either Condition 3.5 or Condition 3.6 holds then the optimal parameter is $x^* = c/(b - dk)$. If neither holds then any parameter $x \in [0, 1]$ is optimal.

PROOF. We wish to choose a medium in order to maximize V, which we will then show coincides with the receiver's payoff under inattention. From Whitmeyer (2019) [15] it is without loss of generality to restrict our attention to pure strategies of the sender types. Moreover, Theorem 3.8 in [15] establishes that we need only consider a (relaxed) *commitment* problem for the receiver. That is, suppose that the receiver can commit to choosing *donate* with probability p and *decline* with probability $1 - p$ following *cry*; and *donate* with probability q and *decline* with probability $1 - q$ following *quiet*. The receiver solves the following optimization problem,

$$\max_{p,q} \{V\}$$

subject to

$$q\,(1 + (1 - d)\,k) + (1 - q)\,(1 - b + k) \geq p\,(1 - c + (1 - d)\,k) + (1 - p)\,(1 - b - c + k) \tag{$IC1$}$$

and

$$p\,(1 - c + (1 - d)\,k) + (1 - p)\,(1 - a - c + k) \geq q\,(1 + (1 - d)\,k) + (1 - q)\,(1 - a + k) \tag{$IC2$}$$

where

$$\begin{aligned} V = (1 - \mu)\,[q\,(1 - d + k) + (1 - q)\,(1 + (1 - b)\,k)] \\ + \mu\,[p\,(1 - d + (1 - c)k) + (1 - p)\,(1 + (1 - a - c)\,k)] \end{aligned}$$

This optimization problem is easy to solve and yields $q = 0$ and $p = c/(b - dk)$ for $d \geq \hat{d}$, and $q = 1 - c/(b - dk)$ and $p = 1$ for $d \leq \hat{d}$. Substituting these into the value function, we obtain $V\,(x^*)$. If either Condition 3.5 or Condition 3.6 holds, then this maximizes

the receiver's payoff, and as illustrated in Whitmeyer (2019) [15], since p and q solve the commitment problem, this must be the solution to the problem of choosing an optimal medium.[7]

On the other hand, if neither condition holds then the result is trivial. The receiver-optimal equilibrium is one in which both types pool on *quiet* and is attainable under any x. ■

As we did earlier, the receiver's payoff can be written as a function of the prior, V_I (where I stands for inattention). By Whitmeyer (2019) [15], the receiver's payoff when she may choose the optimal medium is convex in μ. Thus, here, since inattention corresponds to the optimal medium, V_I must be convex. That is, any (free) *ex ante* information about the sender's type must benefit the receiver at least weakly. Substituting the parameter values used *supra*[8] the receiver's payoff function is, explicitly:

$$V_I(\mu) = \begin{cases} \frac{296}{256} - \frac{35}{256}\mu, & 0 \leq \mu \leq \frac{1}{5} \\ \frac{73}{64} - \frac{15}{256}\mu, & \frac{1}{5} \leq \mu \leq \frac{4}{15} \\ \frac{9}{8}, & \frac{4}{15} \leq \mu \leq 1 \end{cases}$$

This function is depicted in Figure 5. The receiver's payoff without inattention (from Figure 2) is the dashed black curve and her payoff with optimal inattention is the green curve.

4. Discussion

The primary goal of this paper was to illustrate the counter-intuitive fact that limited attention may enhance honest communication in situations of conflict. The main results suggest several qualitative and hence testable implications. Broadly, these results imply that in situations in which there are severe conflicts of interest between the sender and receiver, we should expect inattentive receivers, even if there is no cost (of either an

[7] See Whitmeyer (2019) [15] for a more in depth exposition of this concept.

[8] $a = 1$, $b = 3/8$, $c = 11/64$, $d = 1/8$, and $k = 1/4$.

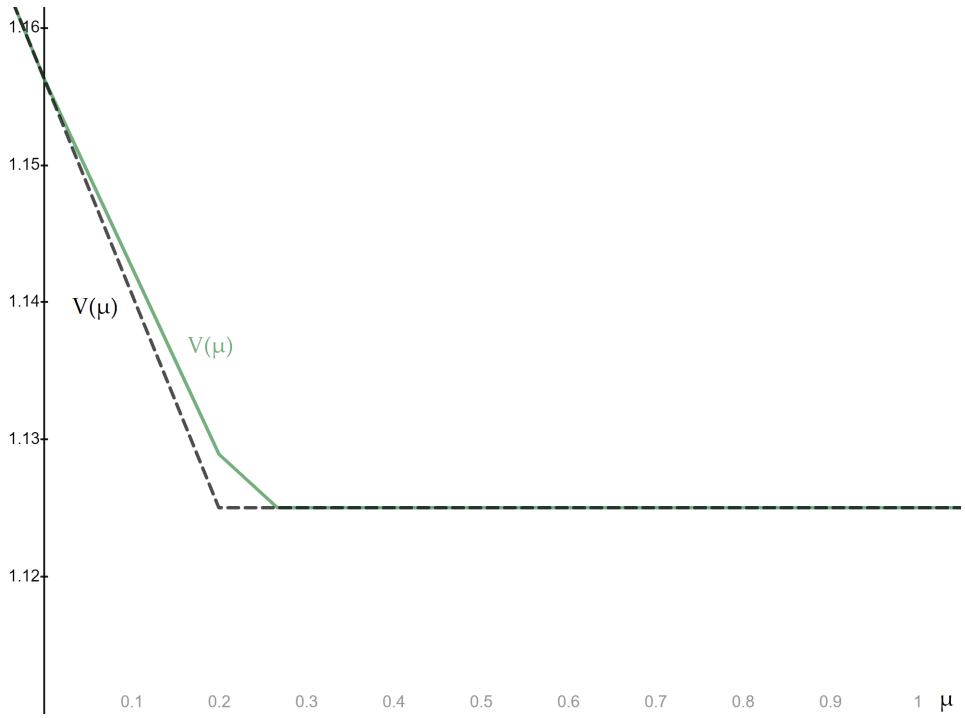

Fig 5. *Receiver's Payoff, Inattention*

explicit or implicit sort) of attention. Moreover, such inattention may beget honest communication even in the absence of communication costs, or if such costs are surprisingly low.

Explicitly, this manuscript's solution for the optimal attention level, x^*, provides the following testable predictions. In an environment in which receiver attentiveness is somehow a choice variable, we should expect to see the receiver pay more attention when she and the sender are more closely related or when signaling is costlier. Analogously, if the receiver's donation cost is higher or if the benefit to the healthy sender from receiving a donation is lower, we should again expect the receiver to me more attentive.

In addition, in these situations where communication is more difficult to sustain, we may also expect to see pooling equilibria in which the sender types choose the least cost signal, which signal the receiver may not even observe. The possibility of limited attention greatly strengthens the least cost pooling equilibria, since a receiver could always default to that by being very inattentive.

Another interpretation of this paper is to see it as highlighting the benefits of inatten-

tion to the receiver. That is, this work forwards an idea of the ilk, "if she were (involuntarily) inattentive, she would be better off".[9] This fits in well with the formulation of the initial stage in which the receiver chooses her attention level (which is observed by the sender). In other words, there is commitment to this level of attention that corresponds to the question, "if the receiver could choose her attention level in the signaling game, what would she choose?" It is important to note that without this commitment, the separating result would vanish and indeed the only equilibria that would remain would be the one in which both types of sender remain quiet. Of course, the separating equilibrium would reappear should the model be altered further so that attention is costly for the receiver, say increasing in x. Such a model is beyond the scope of this work, and perhaps merits further attention elsewhere.

Finally, note that the solution concept used throughout this work is a refinement of the standard Nash Equilibrium, the Subgame Perfect Equilibrium, and not Evolutionary Stability. However, note that if x is fixed and in the interior of B, then there is a separating equilibrium that is strict, and thus must therefore be an Evolutionary Stable Strategy in the symmetrized game (see e.g. Cressman 2003 [4]). It might be interesting to explore the dynamic properties of the scenario from this manuscript in greater detail.

References

[1] Carl T. Bergstrom and Michael Lachmann. Signalling among relatives. i. is costly signalling too costly? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352(1353):609–617, 1997.

[2] Carl T. Bergstrom and Michael Lachmann. Signaling among relatives. iii. talk is cheap. *Proceedings of the National Academy of Sciences*, 95(9):5100–5105, 1998.

[3] In-Koo Cho and David M. Kreps. Signaling games and stable equilibria. *The Quarterly Journal of Economics*, 102(2):179–221, 1987.

[4] Ross Cressman. *Evolutionary Dynamics and Extensive Form Games.* MIT Press, Cambridge, MA, 2003.

[9] Note that, as is mentioned at the end of Section 3.2, for $d \geq \hat{d}$, the best separating equilibrium for the receiver would leave the sender no worse off as well.

[5] Simon Huttegger, Brian Skyrms, Pierre Tarrès, and Elliott Wagner. Some dynamics of signaling games. *Proceedings of the National Academy of Sciences*, 111:10873–10880, 2014.

[6] Simon M. Huttegger and Kevin J. S. Zollman. Dynamic stability and basins of attraction in the sir philip sidney game. *Proceedings of the Royal Society B*, 277, 2010.

[7] Rufus A. Johnstone. Honest signalling, perceptual error and the evolution of 'all-or-nothing' displays. *Proceedings of the Royal Society B*, 256, 1994.

[8] Rufus A. Johnstone. Efficacy and honesty in communication between relatives. *The American Naturalist*, 152(1):45–58, 1998.

[9] Rufus A. Johnstone and Alan Grafen. Error-prone signalling. *Proceedings of the Royal Society B*, 248, 1992.

[10] Michael Lachmann and Carl T. Bergstrom. Signalling among relatives: ii. beyond the tower of babel. *Theoretical Population Biology*, 54(2):146 – 160, 1998.

[11] Michael Lachmann, Szabolcs Számadó, and Carl T. Bergstrom. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, 98(23):13189–13194, 2001.

[12] David K. Lewis. *Convention. A Philosophical Study.* Harvard University Press, Harvard, 1969.

[13] John Maynard Smith. Honest signalling: the philip sidney game. *Animal Behavior*, 42:1034–1035, 1991.

[14] Szabolcs Számadó. The cost of honesty and the fallacy of the handicap principle. *Animal Behaviour*, 81(1):3 – 10, 2011.

[15] Mark Whitmeyer. Bayesian elicitation. *Mimeo*, January 2019. URL https://whitmeyerhome.files.wordpress.com/2020/01/bayesian_elicitation_jmpv24.pdf.

[16] Mark Whitmeyer. In communication games, when does ex ante fact-finding benefit the receiver. *Mimeo*, January 2019. URL https://whitmeyerhome.files.wordpress.com/2020/01/in_communication_games__when_does_ex_ante_fact_finding_benefit_the_receiver__v18.pdf.

[17] R. Haven Wiley. A receiver-signaler equilibrium in the evolution of communication in noise. *Behaviour*, 150(9/10):9570993, 2013.

[18] R. Haven Wiley. How noise determines the evolution of communication. *Animal Behaviour*, 124: 307 – 313, 2017.

[19] Amotz Zahavi. Mate selection—a selection for a handicap. *Journal of Theoretical Biology*, 53(1): 205 – 214, 1975.

[20] Kevin J. S. Zollman, Carl T. Bergstrom, and Simon M. Huttegger. Between cheap and costly signals: the evolution of partially honest communication. *Proceedings of the Royal Society B*, 280, 2013.

APPENDIX A: OMITTED PROOFS

Here, we present the proofs of selected results contained in the paper.

A.1. Lemma 3.1 Proof

Proof. Suppose the two types of sender separate: θ_H chooses *quiet* and θ_N chooses *cry*. First, let $d \geq \hat{d}$, so that *decline* is the receiver's response to not observing. Hence, the receiver's payoff is

$$\begin{aligned} V &= (1-\mu)(1+(1-b)k) + \mu(x(1-d+(1-c)k) + (1-x)(1+(1-a-c)k)) \\ &= 1 + k(1-b) - k\mu(a+c-b) + \mu x(ak-d) \end{aligned}$$

which is increasing in x. The sender types' incentive constraints are

$$1 - b + k \geq x(1-c+(1-d)k) + (1-x)(1-b-c+k)$$

or $c \geq (b-dk)x$ for θ_H, and

$$x(1-c+(1-d)k) + (1-x)(1-a-c+k) \geq 1-a+k$$

or $(a-dk)x \geq c$ for θ_N. Thus, any x that satisfies

$$\frac{c}{b-dk} \geq x \geq \frac{c}{a-dk} \tag{A1}$$

begets a separating equilibrium of this form. Since $a > b > dk + c$, both $a - dk$ and $b - dk$ are greater than c. Hence,

$$x^* = \frac{c}{b-dk} \tag{A2}$$

and so

$$V(x^*) = 1 + k(1-b) - k\mu(a+c-b) + \frac{\mu c(ak-d)}{b-dk} \tag{A3}$$

When $d \leq \hat{d}$, we simply proceed in the same manner and obtain the same interval and x^* as in Inequality A1 and Equation A2, respectively. The resulting payoff for the receiver is

$$V(x^*) = 1 - d + k(1-c\mu) + \frac{(1-\mu)c(d-bk)}{b-dk} \tag{A4}$$

■

A.2. Lemma 3.2 Proof

Proof. First, it is obvious that there is no equilibrium in which the two types of sender pool on *cry* if *cry* is followed by *decline*.

Next, let $d \geq \hat{d}$ and consider $(quiet, quiet; decline, \cdot)$. It is clear that the off-path belief that leaves the equilibrium in greatest jeopardy is that which insists the receiver prefer *donate* upon observing *cry*. Under this "worst" case scenario, the incentive constraint for θ_H is

$$1 - b + k \geq x\,(1 - c + (1 - d)k) + (1 - x)\,(1 - b - c + k)$$

Or, $c/(b-dk) \geq x$. Analogously, the incentive constraint for θ_N reduces to $c/(a-dk) \geq x$.

Thus, if $x \leq c/(a-dk)$ then regardless of the receiver's off-path beliefs, $(quiet, quiet; decline, \cdot)$ is an equilibrium. If x is above this threshold, then it is clear that an off-path belief that results in the receiver (weakly) preferring *decline* upon observing cry is required.

Now, let $d \leq \hat{d}$. The receiver's optimal action should she choose not to observe a signal is *donate*. First, we explore whether there is an equilibrium in which the two types of sender pool on *cry*. For θ_H we have

$$1 - c + (1 - d)k \geq x\,(1 - b + k) + (1 - x)\,(1 + (1 - d)\,k)$$

which holds if and only if $x \geq c/(b - dk)$. For θ_N we have

$$1 - c + (1 - d)k \geq x\,(1 - a + k) + (1 - x)\,(1 + (1 - d)\,k)$$

which holds if and only if $x \geq c/(a - dk)$. Note that here we have assigned the receiver's off-path belief to be such that *decline* is a (weak) best response to *quiet*. This is clearly necessary for the existence of this equilibrium, irrespective of x.

Finally, suppose the two types of sender pool on *quiet*. Again, it is clear that this is an equilibrium, regardless of x or the off-path beliefs.

■