Evolution of Honest Signaling by Social Punishment

David Catteeuw
Artificial Intelligence Lab
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
dcatteeu@vub.ac.be

The Anh Han
Artificial Intelligence Lab
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
h.anh@ai.vub.ac.be

Bernard Manderick
Artificial Intelligence Lab
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
bmanderi@vub.ac.be

ABSTRACT

When facing dishonest behavior of any form, individuals may choose to punish in order to enhance future honesty from others, even if it is costly for the punishers. Such behavior can be found ubiquitously in human and animal communications, suggesting that it may play an important role in the evolution of honest signaling or reliable communication. By applying Evolutionary Game Theory to the Philip Sidney game, we provide a computational model to investigate whether costly punishment can be a viable strategy for the evolution of honest signaling. We identify four different forms of dishonesty, and study how punishing them affects the level of honesty in the final outcome of evolutionary dynamics. Our results show that punishing those that lie can significantly boost honest signaling when conflicts are moderate and signals are cheap or cost-free. It hence provides an important alternative to the well-known Handicap Principle, which states that honest signaling can evolve only if signals are sufficiently costly for their senders. Furthermore, punishing greedy responses promotes honest signaling if conflicts of interest are high and signals are costly. Lastly, punishing timid or worried individuals does not lead to a clear improvement of honesty.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent Systems*

General Terms

Theory

Keywords

Philip Sidney game, punishment, honest signaling, evolutionary game theory

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.

ACM 978-1-4503-2662-9/14/07 ...$15.00.
http://dx.doi.org/10.1145/2576768.2598312.

1. INTRODUCTION

Communication is undoubtedly among the most important processes in human society as well as in the societies of many other species [39, 25]. It enables a smooth and efficient coordination among participants of a joint venture. Diverse examples of such behavior can be found in nature, such as the invention of human languages [28], the honey bees' waggle dance to indicate the location of a food source [41], and alarm calls in vervet monkeys [34]. In artificial multiagent systems, communication enables teamwork [2], such as in playing robotic soccer [4], in cooperative hunting [21], and wireless sensors that communicate to find an optimal route via which all data is sent to the sink node [45].

Despite the important benefits of communication, when there are conflicts of interests between communicating individuals, dishonesty may emerge leading to the breakdown of the communication systems [1, 8]. Conflicts of interest provide incentives to lie, and when most communication becomes unreliable, no one will respond to it and when that happens, it is best not to communicate since it is ineffective while being costly to the communicator. Hence, the (evolutionary) puzzle here is, why are communication systems so abundant in different societies even though conflicts of interests are largely unavoidable [24, 27, 39]?

This question of explaining the evolution of communication (also called evolution of honest signaling) has been actively investigated in several fields of research, most notably in biology, economics, and evolutionary linguistics [27, 32, 37]. The main and most influential explanation is the Handicap Principle [43, 44, 11], saying that when it is sufficiently costly to initiate the communication (or to send a signal), honest communication (signaling) can evolve and remains stable. However, it has been shown that there are communication systems where it is quite cheap to communicate, and several alternatives to the Handicap Principle have been proposed (see the surveys by Számadó [38] and Zollman [46]). For instance, other equilibria (called partial pooling and hybrid equilibria) exist where communication is slightly less informative—but still honest—and a lot less costly [23, 19]. Another alternative is based on the assumption (supported by examples) that it is more costly to the dishonest than honest signalers to send a signal [18, 24] (see Section 2 for further discussion). Differently, here we study how punishing dishonest signaling behavior can provide a pathway for the evolution of honest signaling. Note that a large body of evidence shows that in human and many other species' societies, individuals are willing to pay a personal cost in order to punish those that infringe their interests, such as

the non-contributors of a public good or dishonest individuals in communication [7, 9, 12]. Punishment has been shown to promote cooperation in different scenarios, including the Prisoner's Dilemma and the Public Goods Game, e.g. in [15, 36, 16]. However, little effort has been provided to clarify the role of punishment in the emergence and stability of honest signaling or reliable communication systems.

In order to study this problem, we apply Evolutionary Game Theory methods in finite populations [30, 35] to an extension of the Philip Sidney game [26]. It is the standard game-theoretical framework to study honest communication in biology and economics, which best illustrates the Handicap Principle: when interests conflict, signals must be costly in order to maintain honesty [43, 44, 11]. We extend the Philip Sidney game to study four distinct types of punishment which target the four possible deviations from honest signaling, namely the liars, the timid, the greedy and the worried individuals (see definitions in Section 3.2). We show how each of them affects the level of honesty in the final outcome of the evolutionary dynamics.

The rest of the paper is structured as follows. Section 2 discusses relevant literature of honest signaling and punishment in different contexts. Section 3 describes our model of punishment within the context of the Philip Sidney game, and gives methods to analyze the model. Section 4 then shows our numerical results. Finally, in Section 5 we discuss the obtained results and describe some future directions.

2. RELATED WORK

The most influential explanation for the evolution of honest signaling is the Handicap Principle: costly signals can maintain honesty even if interests conflict [43]. Grafen [11] proved this mathematically but two issues remain: signal costs in nature are often smaller than predicted, and the signal cost makes other non-signaling equilibria (also called pooling equilibria) often more attractive to both senders and receivers [38]. Several alternatives to the Handicap Principle have been proposed that allow honesty with cheap or cost-free signals (see the surveys by Számadó [38] and Zollman [46]). A first alternative explanation is that *partially* honest communication can be an equilibrium at a much lower cost than completely honest communication. These are so-called partial pooling equilibria and hybrid equilibria. In a partial pooling equilibrium, communication is only partially honest, because the same signal is used in many different but closely related states [23]. In a hybrid equilibrium, senders may now and then signal dishonestly and receivers may sometimes ignore the sender [19]. Another alternative defines signals that are cheap when honest, but costly when dishonest. Hurd [18] and Lachmann et al. [24] give examples of such cost functions and prove that they maintain honesty. Several mechanisms that generate such cost functions have been suggested, including social punishment, but they remain implicit in these models [38, 46]. Instead, we explicitly analyze whether costly punishment can be a *viable* strategy for the evolution of honest communication systems.

There is a great body of literature, theoretical and experimental, showing the promoting role of costly punishment in the evolution of cooperation. Such theoretical models and experimental works are usually framed in the context of the Prisoner's Dilemma and the Public Good Games [15, 36, 7, 9, 12, 16]. To the best of our knowledge, this is the first effort to provide a computational model analyzing the role of costly punishment in the evolution of reliable communication systems, particularly within the context of the Philip Sidney game. As will be seen, punishment in the context of the Philip Sidney game (or signaling games in general [32]) is much richer and more complex to analyze. In the context of the study of the evolution of cooperation, punishment is usually carried out against those that defect or do not help others, which is similar to the punishment of greedy individuals in the current context. Signaling games exhibit the possibility to punish other types of dishonest behavior, including liars, timid and worried individuals. We discuss these differences into more detail in Section 5.

Last but not least, it is undoubtedly important to mention the large body of work on agent communication in multiagent systems research [2, 22, 42, 33]. One main concern of these works is to design and implement efficient and reliable communication systems in multiagent systems so as to allow coordination and teamwork [2], for example in wireless sensor networks [45] and playing robotic soccer [4], and also to collaboratively assess trust and reputation of others [22, 33]. In contrast, the current work studies how communication systems can remain reliable while it may be advantageous to act dishonestly, with the help of social punishment targeting different forms of dishonesty in signaling. As will be seen, the results from this work provide important insights into the design of more reliable communication systems in concrete domains, namely which forms of dishonesty should be targeted with punishment and under which circumstances—in order to achieve a high level of honesty.

3. MODEL AND METHODS

3.1 The Philip Sidney game

The Philip Sidney Game [26] is a signaling game: a two-player extensive-form game with asymmetric information (Figure 1). The first player, Sender, may be in one of two states: healthy or needy, with probability p and $1-p$, respectively. In each state, he may either signal at cost c ($0 \leq c \leq 1$) or stay quiet (at no cost). The second player, Receiver, does not know Sender's true state, but observes whether or not Sender signals and must then decide to donate his resource to Sender, or to keep it for himself.

We assume that Sender and Receiver are related such that Receiver may benefit from donating his resource to Sender. This is the principle of *inclusive fitness* [13]. Each player's payoff is the sum of his own survival probability plus a fraction r of the other player's survival probability (the benefit of survival is normalized), where r is the players' relatedness coefficient.

The survival probabilities depend on the outcome of the game (Table 1). Receiver is sure to survive if he keeps the resource, but if he donates the resource he is only sure to survive with probability $S < 1$. If Sender receives the resource he is sure to survive, otherwise his survival probability depends on his initial state: if he is needy, he will die; if he is healthy, he will survive with probability $V < 1$. As already mentioned, signaling is costly: Sender's survival probability is decreased by a factor $(1-c)$ if he signals.

The Philip Sidney game has sixteen pure strategies which are represented by a couple (S_i, D_j) (see Table 2). The first element (S_i) determines what to do when in the role of Sender, which can be either: always signal (S_A), never signal $(S_\emptyset)$, signal only when healthy (S_H), or signal only when

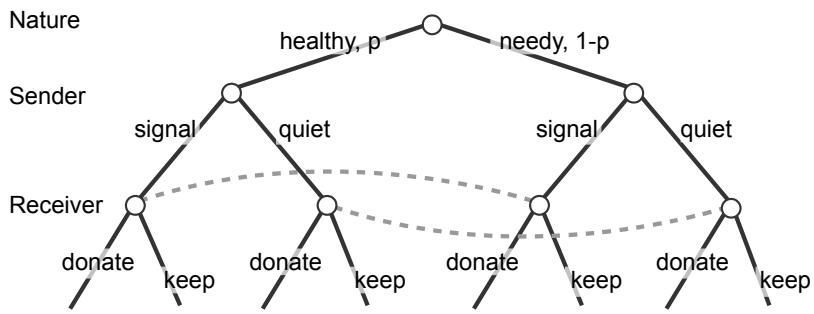

Figure 1: Philip Sidney game.

Table 1: Sender's and Receiver's survival probabilities for all possible outcomes of the Philip Sidney game.

	healthy		needy	
	donate	keep	donate	keep
quiet	$1, S$	$V, 1$	$1, S$	$0, 1$
signal	$1-c, S$	$V(1-c), 1$	$1-c, S$	$0, 1$

Table 2: Abbreviation and meaning of the four Sender and four Receiver strategies.

	strategy
S_A	always signal
$S_\emptyset$	never signal
S_N	signal only when needy
S_H	signal only when healthy
D_A	always donate
$D_\emptyset$	never donate
D_S	donate only when signal
D_Q	donate only when quiet

needy (S_N). The second element (D_j) determines what to do when in the role of Receiver, which can be either: always donate (D_A), never donate ($D_\emptyset$), donate only when signal (D_S), or donate only when quiet (D_Q). For example, the honest signaling strategy—signal only when needy, donate only when signal—is represented by (S_N, D_S). The payoffs of these pure strategies are calculated assuming that each strategy is equally likely to be in the role of Receiver or of Sender.

Note that players do not need to be related genetically. The relatedness factor r may simply represent the level of common interest [3]. If $r = 1$, interests are perfectly aligned. For smaller r the conflict of interest increases. Furthermore, depending on the other parameters of the game, there may be a conflict when Sender is needy, when Sender is healthy, in both cases, or never. We say there is a conflict if, knowing Sender's state, both players would prefer other outcomes, whether Sender signals or not. Concretely, there is a conflict when Sender is healthy if $r(1-V) < 1-S$ and $r(1-S) < (1-c)(1-V)$ [5]. Similarly, there is a conflict when Sender is needy if $r < 1-S$ and $r(1-S) < (1-c)$. Note that conflicts do not depend on the probability of Sender being healthy p [26]. If there is no conflict, we expect communication to emerge only if this is best for both players. When there is a full conflict, we expect no communication will emerge. The most interesting region is the one where there is a partial conflict [5].

3.2 Model

We consider new honest signaling strategies that punish dishonest behavior. We distinguish four different such strategies which target each of the possible deviations from honest signaling, namely punishment of lying (P_L), of greedy (P_G), of timid (P_T), and of worried opponents (P_W). An individual is lying (or simply a liar) if he signals when healthy; greedy if he keeps when Sender signals; timid if he remains quiet when needy; and worried if he donates when Sender is quiet.

Note that a strategist can be punished on several occasions. For example, the strategy 'signal only when healthy, donate only when quiet', (S_H, D_Q), is lying, greedy, timid, and worried simultaneously, albeit in different cases. When adopting this strategy, a healthy Sender is deemed a liar while a needy Sender a timid one.

In the new game, a player's survival probability is decreased by d if he punishes his opponent, and decreased by e if he is punished. As before, we apply inclusive fitness with relatedness factor r to the survival probabilities of the players. Punishing your opponent always lowers your own payoff even if the cost to punish $d = 0$, due to the relatedness to your opponent ($r > 0$) and decreasing his survival probability decreases your payoff. For example, if Sender punishes Receiver, Sender earns $u_S = (v_S - d) + r(v_R - e)$ and Receiver $u_R = (v_R - e) + r(v_S - d)$, where Senders and Receiver survival probability, v_S and v_R, depend on the game's outcome as shown in Table 1. If Receiver punishes Sender, the costs d and e are swapped. If no one punishes, the costs d and e are left out (as in the original Philip Sidney game).

While greedy and worried strategists are always revealed and hence are always punished by a strategy such as P_{GW}, liars and timid are only revealed with probability $q < 1$. Hence, they may sometimes (with probability $1 - q$) escape punishment. Table 3 summarizes all parameters of the Philip Sidney game with punishment.

3.3 Evolution in Finite Populations

The numerical analysis in the following section is based on methods of Evolutionary Game Theory methods for finite populations [30, 20]. In such a setting, individuals' payoff represents their *fitness* or social *success*, and evolutionary dynamics is shaped by social learning [17, 35], whereby the

Table 3: Parameters of the Philip Sidney game with punishment.

	domain	meaning
p	$(0,1)$	probability of *Sender* being healthy
q	$[0,1]$	probability that the true state of *Sender* is revealed
r	$[0,1]$	relatedness factor
S	$(0,1)$	*Receiver*'s survival probability when donating resource
V	$(0,1)$	*Sender*'s survival probability when healthy but not getting resource
c	$[0,1]$	signal cost
d	$[0,1]$	cost payed to punish the opponent
e	$[0,1]$	cost incurred when punished

most successful individuals will tend to be imitated more often by others. In the current work, social learning is modeled adopting the so-called pairwise comparison rule [40]. It assumes that an individual A with fitness f_A adopts the strategy of another individual B with fitness f_B with probability given by the Fermi function, $(1+\exp(-\beta(f_B-f_A)))^{-1}$. The parameter β represents the 'imitation strength' or 'intensity of selection', i.e., how strongly the individuals decide to imitate based on fitness comparison. For $\beta=0$, we obtain the limit of neutral or random drift—the imitation decision is random. For large β, imitation becomes increasingly deterministic.

In the absence of mutations or exploration, the end states of evolution are inevitably monomorphic (i.e all individuals in the population adopt the same strategy): once such a state is reached, it cannot be escaped through imitation. We thus further assume that, with a certain mutation probability, an individual switches randomly to a different strategy without imitating another individual. In the limit of small mutation rates, the behavioral dynamics can be conveniently described by a Markov Chain, where each state represents a monomorphic population, whereas the transition probabilities are given by the fixation probability of a single mutant [20, 15, 36]. The resulting Markov Chain has a stationary distribution, which characterizes the average time the population spends in each of these monomorphic end states.

Let N be the size of the population. Suppose there are at most two strategies in the population, say, k individuals using strategy A ($0 \le k \le N$) and $(N-k)$ individuals using strategy B. Thus, the (average) payoff of the individual that uses A and B can be written as follows, respectively,

$$\begin{aligned}\Pi_A(k) &= \frac{(k-1)\pi_{A,A}+(N-k)\pi_{A,B}}{N-1},\\ \Pi_B(k) &= \frac{k\pi_{B,A}+(N-k-1)\pi_{B,B}}{N-1},\end{aligned} \tag{1}$$

where $\pi_{X,Y}$ is the payoff an X-strategist obtains in an interaction with a Y-strategist.

The probability to change the number k of individuals using strategy A by ± 1 in each time step can be written as

$$T^{\pm}(k) = \frac{N-k}{N}\frac{k}{N}\left[1+\exp(\mp\beta[\Pi_A(k)-\Pi_B(k)])\right]^{-1}. \tag{2}$$

The fixation probability of a single mutant with a strategy A in a population of $(N-1)$ individuals using B is given by [40, 10]

$$\rho_{B,A} = \left(1+\sum_{i=1}^{N-1}\prod_{j=1}^{i}\frac{T^-(j)}{T^+(j)}\right)^{-1}. \tag{3}$$

In the limit of neutral selection (i.e. $\beta=0$), $\rho_{B,A}$ equals the inverse of the population size, $1/N$.

Considering a set of strategies $\{1,...,Z\}$, these fixation probabilities determine a transition matrix $M=\{T_{ij}\}_{i,j=1}^{Z}$, with $T_{ij,j\neq i}=\rho_{ji}/(Z-1)$ and $T_{ii}=1-\sum_{j=1,j\neq i}^{Z}T_{ij}$, of a Markov Chain. The normalized eigenvector associated with the eigenvalue 1 of the transposed of M provides the stationary distribution described above [10, 20], describing the relative time the population spends adopting each of the strategies.

Invasion and Replacement in Finite Populations.

We describe some key concepts of Evolutionary Game Theory in finite populations that we use to discuss the results presented in the next section. In finite populations, evolutionary stability is defined differently from the traditional concept of evolutionarily stable strategies [35]. It takes into account the population size N and is based on the dynamics described above. A strategy is evolutionarily stable in finite populations if selection (i) opposes any mutant strategy invading the population, and (ii) opposes any mutant replacing the entire population [30]. Strategy B is resistant to *invasion* if in a population where $N-1$ individuals adopt strategy B and one individual adopts the (mutant) strategy A, the fitness of the single mutant A is lower than that of the other individuals (the Bs): $\Pi_A(1)<\Pi_B(1)$ for all $A\neq B$. Together with Equation (1) this implies that for all $A\neq B$: $(N-1)\pi_{A,B}<\pi_{B,A}+(N-2)\pi_{B,B}$. Strategy B is resistant to *replacement* if the probability that any mutant A fixates in a population of individuals adopting B is smaller than the neutral fixation probability: $\rho_{B,A}<1/N$ for all $A\neq B$. If selection pressure is weak ($\beta \ll 1$), we simply have: $\pi_{A,A}(N-2)+\pi_{A,B}(2N-1)<\pi_{B,A}(N+1)+\pi_{B,B}(2N-4)$.

4. RESULTS

To study the effect of punishing dishonest behavior on the evolution of honest signaling, we consider a finite population of the sixteen (pure) strategies of the Philip Sidney game together with one of the punishment strategies described above. Then, we numerically compute stationary distributions of the strategies using the methods described in Section 3.3 for different configurations of the parameters.

4.1 Punishing Dishonest Signalers: Liars and Timid

Punishing liars. First, let us consider the honest signaling strategy that punishes those who signal when healthy: P_L. Figure 2a shows the frequency of honest signaling (i.e. the sum of the frequencies of P_L and of the pure honest signaling strategy) as a function of the signal cost c, and the relatedness level r. For a wide range of c and r, there is a high level of honest signaling. Interestingly, that occurs even for low values of c, which is not the case in the original Philip Sidney game [6, 43, 11]. So, punishment of liars can provide an alternative for the Handicap Principle, which states that,

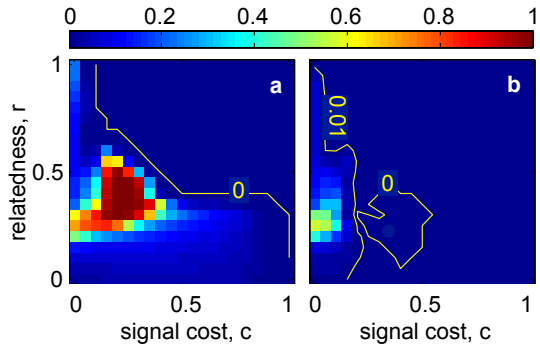

Figure 2: ***a)*** **Total frequency of honest signaling in the Philip Sidney game with punishment of liars (P_L).** ***b)*** **Increase in frequency of honest signaling when P_L is present compared to when it is absent (being replaced with another pure honest signaling strategy). For small signal costs there is a clear increase in the frequency of honest signaling. There is also a region where punishment has a slightly negative effect ($\approx -10^{-3}$ inside the 0-contour line). Both figures were obtained for population size $N = 100$, selection pressure $\beta = 5$, $p = q = d = 1/2$, $S = V = 4/5$, $e = 1$, and varying signal cost c and relatedness level r.**

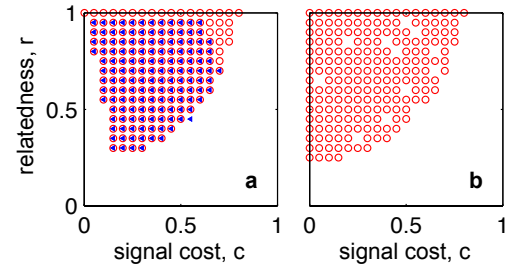

Figure 3: Red circles indicate resistance to invasion by any mutant strategy, blue triangles indicate resistance to replacement by any mutant. ***a)*** **Honest signaling is evolutionarily stable in the colored area, but requires a minimum signal cost ($c > 0$).** ***b)*** **Honest signaling with punishment of liars is not evolutionarily stable, since it can always be replaced by honest signaling (without punishment) through random drift. But it is resistant to invasion for small signal costs. We used the same parameters as in Figure 2.**

under conflict of interest, honest signaling cannot evolve or has low frequency for low-cost signaling systems.

The effect of punishing liars on the level of honest signaling is clearer in Figure 2b, where we plot the increase in frequency of honest signaling when P_L is present in the population compared to the case where P_L is replaced by another pure honest signaling strategy. Irrespective of the relatedness level r, when signal cost c is small enough, the presence of P_L always has a positive effect on the level of honest signaling compared to when it is absent.

For higher signal costs c, the cost of punishment and honesty may not outweigh the benefits. Figure 2b shows a region where punishment has no effect and even slightly decreases the level of honest signaling (approximately -10^{-3}).

This interesting observation can be explained as follows. P_L is effective for small signal costs c since it is resistant to invasion by any mutant in that region whereas honest signaling without punishment can be invaded by mutants that 'always signal,' see Figure 3. In the original Philip Sidney game, honest signaling can be stable only if the signal cost $c > 0$. In the new model, honest signaling with punishment cannot be stable, since it can always be replaced by honest signaling without punishment through random drift. But except for the pure honest signaling strategy, punishment is stable even for cost-free signals, $c = 0$.

Moreover, punishment directly affects liars and indirectly affects other strategies that can replace liars. Figure 4 shows the Markov chain and the transition probabilities of the Philip Sidney game with P_L and without P_L (ignoring the P_L state with the arrows to and out of it). The strategies 'always signal, donate when signal,' (S_A, D_S), and 'always signal, never donate,' $(S_A, D_\emptyset)$, can replace honest signaling, but not honest signaling with punishment (P_L). P_L can only be replaced through random drift by honest signaling itself. As such, lying is less frequent when the punisher P_L is present. For example, the frequency of 'always signal, never donate,' $(S_A, D_\emptyset)$, and 'signal when healthy, donate when quiet,' (S_H, D_Q), dropped below the averaged frequency (i.e. $1/Z$ where Z is the number of strategies). Since liars are less frequent, other strategies that replace liars are also affected. The frequency of 'never signal, never donate,' $(S_\emptyset, D_\emptyset)$, for example, also dropped below the averaged frequency.

Although we call 'signal when healthy, donate when quiet' a liar, it is also honest in some sense. This strategy simply has the meaning of the signals 'signal' and 'quiet' reversed. When the signal cost c is very small (for example $c = 0.01$ as in Figure 4), this strategy is only slightly worse than honest signaling. When the signal cost $c = 0$, both strategies behave symmetrically and perform equally well. When signals have different costs, individuals have a common preference of when to signal what, and end up using the same signal under the same circumstances: the cost creates the meaning. When signals are cost-free there is no a priori common preference for one signal or the other. This does not only makes lying cheap, but also creates two equally valid honest signaling strategies: one where 'signal' means 'needy' and 'quiet' means 'healthy', and the other where 'signal' means 'healthy' and 'quiet' means 'needy'. Punishment deters lying, but can also teach others the preferred meanings.

These results show that costly, social punishment is indeed an alternative explanation for honest signaling. Signals do not need to be costly, as the Handicap Principle suggests; the cost may be paid by liars and punishers. This cost deters liars and so, is rarely paid.

Punishing timid signalers. A different form of deviation from honest signaling occurs when the signalers are 'so timid' that they do not signal even when they are needy. Our results show that punishing this behavior exhibits no clear improvement and its effect is significantly smaller than punishing the liars (Figure 6). The effect is mostly neutral, and even slightly negative for lower relatedness (dark blue region in Figure 5c). That is because in that case timid

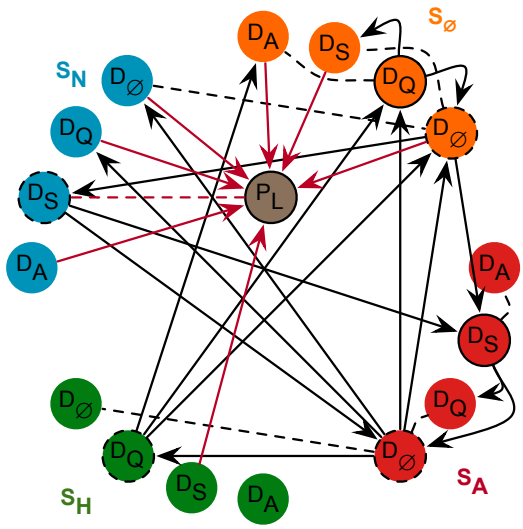

Figure 4: Transition probabilities and stationary distribution for all strategies of the Philip Sidney game (see methods described in Section 3.3 and strategy names in Table 2). Punishment of liars can only be replaced by honest signaling, (S_N, D_S), through random drift. Parameter values are $p = q = d = 1/2$, $S = V = 4/5$, $e = 1$, $c = 0.01$, and $r = 0.3$. Black borders indicate strategies which have above the average frequency $(1/Z)$; *dashed* black borders indicate strategies which have no longer above average frequency after the introduction of the punishment strategy (P_L). For clarity, we only draw transitions departing from strategies with above average frequency and transitions to the punishment strategy (P_L). Dashed lines indicate random drift.

strategies have very low frequencies, hence the effort to punish them is not worth the cost. Although for more efficient punishment (i.e. e/d and q are greater) a small improvement is observed for high signal cost c and low relatedness r, its overall average effect is negative (see Figure 6).

4.2 Punishing Dishonest Responders: Greedy and Worried

When being in the role of a signaler, the player can choose to punish those that deviate from the honest or helpful response, namely those that keep when the sender signals for help (greedy, P_G) and those that donate when the sender is quiet (worried, P_W).

Punishing greedy individuals improves honest signaling for low relatedness if signals are costly ($c > 0.2$ and $r < 0.3$). This region characterizes high conflicts of interest, where greed pays off in the original Philip Sidney game [5]. It may also decrease the frequency of honest signaling for small signal costs (blue region near $(c, r) = (0.1, 0.3)$ in Figure 5a).

Punishing worried individuals slightly improves honest signaling if there is a full conflict or no conflict at all. The improvement is too small to be visible in Figures 5b and 6. Its effect is mostly neutral and even negative when there is a conflict only when the sender is healthy ($0.2 < r < 1 - c$).

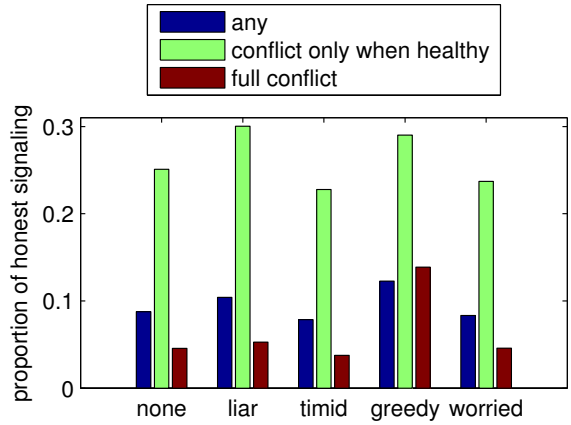

Figure 6: Frequency of honest signaling for different forms of punishment averaged over different regions of conflict as described in Section 3.1. The baseline model (labeled 'none') includes two honest signaling strategies without punishment. Punishing liars is most effective when there is a partial conflict when healthy. Punishing greedy is very effective when there is a full conflict. Punishing worried helps slightly when there is a full conflict. Punishing timid individuals is, on average, counterproductive for the evolution of honest signaling. We used the same parameter values as in Figure 2.

Note that greedy and worried strategists can always be detected, whereas lying and timid ones cannot (in all our experiments lying and timid behaviors are detected with $q = 50\%$ chance). We may expect that punishing greedy and worried is more effective than punishing liars, but this is not the case.

Finally, to have a better understanding on the overall effects of different forms of punishment, we measure the average (total) frequency of honest signaling in different regions of conflict, as described in Section 3.1. Figure 6 shows, varying r and c, the result for the entire region ('any'), the region where there is a conflict only when Sender is healthy, and the region of full conflict. The baseline mode replaces punishment strategies with another pure honest signaling strategy, and is labeled 'none'. Punishing liars is most effective when there is a partial conflict when healthy. Punishing greedy is very effective when there is a full conflict. Punishing worried is slightly effective when there is a full conflict and otherwise not. Punishing timid individuals shows no improvement on average. Additional analysis shows that these results are robust for varying probability q of detecting liars and timid, effectiveness of punishment e/d, and intensity of selection β.

5. DISCUSSION

Within the context of the Philip Sidney game, we studied the effects of punishing four different forms of dishonest behavior in communication—the lying, timid, greedy, and worried individuals—on the level of honest signaling. We have shown these distinct forms of punishment are not all beneficial for promoting the evolution of honest signaling or reliable communication, and when they are beneficial, it

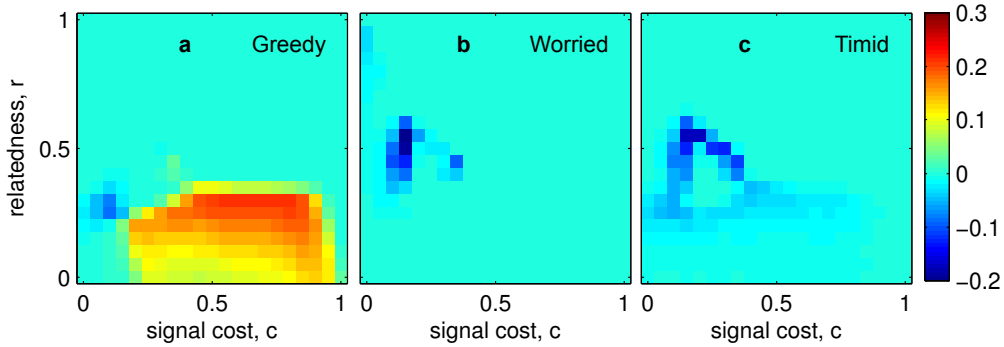

Figure 5: Increase in frequency of honest signaling when punishing *a)* greedy individuals (P_G), *b)* worried individuals (P_W), and *c)* timid individuals (P_T). We used the same parameter values as in Figure 2.

occurs in different regions of the signal costs and relatedness levels. Punishing liars increases the frequency of honest signaling when the signals are cheap and even cost-free. We deem this an important form of punishment because it clearly suggests an alternative for the Handicap Principle—the main and most influential explanation for the evolution of communication. Punishing greedy individuals increases the frequency of honest signaling when conflicts of interest are high (i.e. relatedness level is low) and signals are sufficiently costly. Finally and surprisingly, punishing timid or worried individuals is mostly counterproductive for the evolution of honest signaling. They do not lead to any clear improvement in general, and even result in an overall decrease of the frequency of honest signaling.

We have not yet taken into account anti-social (punishing honest signalers) and spiteful punishment (punishing everyone) [16]. As shown in the context of the study of the evolution of cooperation [31, 16], these types of punishment may destroy any benefit provided by the *social* punishment targeting dishonest signalers. However, as described in some recent works, antisocial and spiteful punishment may be avoided if reputation effects are taken into account [16] or prior agreements regarding posterior consequences of not honoring them are made [14]. In future work, we will analyze whether these mechanisms can still deal with anti-social and spiteful punishment in the context of the evolution of honest signaling. We expect they do, as there is evidence from nature and human societies showing that anti-social and spiteful punishment may be unlikely in this context [29].

In contrast to the analysis of honest signaling, there have been a great number of theoretical and experimental studies analyzing the role of costly punishment in the evolution of cooperative behavior, within the contexts of the Prisoner's Dilemma and the Public Goods Game (Section 2). It has been shown that, in general, social punishment (i.e. punishing those that defect or do not contribute to the Public Good) promotes the evolution of cooperation. However, in the context of the Philip Sidney (signaling) game, that is not generally the case. Herein punishment (even only social) is more complex, showing diverse possibilities which result in different outcomes. Defectors are similar to the greedy individuals herein, but the other types of dishonest behavior (liars, timid and worried) are not present in the Prisoner's Dilemma and the Public Goods Game.

In short, our work demonstrates that punishing dishonest behavior in communication promotes the evolution of honest signaling or reliable communication in several contexts, leading to higher levels of honesty. It provides a richer and more complex framework for the study of the evolutionary roles of punishment than in the context of cooperation. More effort and attention is required to further clarify the roles of punishment in the evolution of communication.

6. ACKNOWLEDGMENTS

The Anh Han acknowledges the support provided by the F.W.O. Belgium (postdoctoral fellowship 05_05 TR 7142).

7. REFERENCES

[1] G. A. Akerlof. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

[2] S. Barrett, N. Agmon, N. Hazon, S. Kraus, and P. Stone. Communicating with Unknown Teammates. In S. Devlin, D. Hennes, and E. Howly, editors, *Proceedings of the 13th Adaptive and Learning Agents workshop*, pages 46–52, Saint Paul, MN, USA, 2013.

[3] C. T. Bergstrom and M. Lachmann. Signalling among relatives. I. Is costly signalling too costly? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 352(1353):609–617, May 1997.

[4] C. Castelpietra, L. Iocchi, D. Nardi, M. Piaggio, A. Scalzo, and A. Sgorbissa. Coordination among heterogeneous robotic soccer players. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1385–1390. IEEE, 2000.

[5] D. Catteeuw and B. Manderick. Honesty and deception in populations of selfish, adaptive individuals. *The Knowledge Engineering Review*, 31(2):1–15, 2016.

[6] D. Catteeuw, B. Manderick, and T. A. Han. Evolutionary stability of honest signaling in finite populations. In *Proceedings of IEEE Congress on Evolutionary Computation*, pages 2864 – 2870. IEEE Computer Society, 2013.

[7] T. Clutton-Brock and G. Parker. Punishment in animal societies. *Nature*, 373:209–216, 1995.

[8] M. S. Dawkins and T. Guilford. The corruption of honest signalling. *Animal Behaviour*, 41(5):865–873, 1991.

[9] E. Fehr and S. Gächter. Altruistic punishment in humans. *Nature*, 415:137–140, 2002.

[10] D. Fudenberg and L. A. Imhof. Imitation processes with small mutations. *Journal of Economic Theory*, 131:251–262, 2005.

[11] A. Grafen. Biological signals as handicaps. *Journal of Theoretical Biology*, 144(4):517–46, June 1990.

[12] F. Guala. Reciprocity: Weak or strong? what punishment experiments do (and do not) demonstrate. *Behavioral and Brain Sciences*, 35(1):1, 2012.

[13] W. D. Hamilton. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7(1):1–52, 1964.

[14] T. A. Han, L. M. Pereira, F. C. Santos, and T. Lenaerts. Good agreements make good friends. *Scientific reports*, 3(2695), 2013.

[15] C. Hauert, A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund. Via freedom to coercion: the emergence of costly punishment. *Science*, 316(5833):1905–7, June 2007.

[16] C. Hilbe and A. Traulsen. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Scientific reports*, 2, 2012.

[17] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.

[18] P. L. Hurd. Communication in discrete action-response games. *Journal of Theoretical Biology*, 174(2):217–222, May 1995.

[19] S. M. Huttegger and K. J. S. Zollman. Dynamic stability and basins of attraction in the Sir Philip Sidney game. *Proceedings of the Royal Society B: Biological Sciences*, 277(1689):1915–1922, June 2010.

[20] L. A. Imhof, D. Fudenberg, and M. A. Nowak. Evolutionary cycles of cooperation and defection. *Proceedings of the National Academy of Sciences of the United States of America*, 102:10797–10800, 2005.

[21] K. C. Jim and C. L. Giles. Talking helps: evolving communicating agents for the predator-prey pursuit problem. *Artificial life*, 6(3):237–54, Jan. 2000.

[22] A. Koster, J. Sabater-Mir, and M. Schorlemmer. Personalizing communication about trust. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 517–524, 2012.

[23] M. Lachmann and C. T. Bergstrom. Signalling among Relatives II. Beyond the Tower of Babel. *Theoretical Population Biology*, 54:146–160, 1998.

[24] M. Lachmann, S. Szamado, and C. T. Bergstrom. Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23):13189–13194, Nov. 2001.

[25] D. K. Lewis. *Convention: A Philosophical Study*. Harvard University Press, Cambridge, 1969.

[26] J. Maynard Smith. Honest signalling: The Philip Sidney game. *Animal Behaviour*, 42:1034–1035, 1991.

[27] J. Maynard Smith and D. Harper. *Animal signals*. Oxford University Press, Oxford, UK, 2003.

[28] J. Maynard Smith and E. Szathmary. *The major transitions in evolution*. Oxford University Press, Oxford, UK, 1997.

[29] A. P. Moller. Social control of deception among signalling house sparrows Passer domesticus. *Behavioral Ecology and Sociobiology*, 20(5):307–11, 1987.

[30] M. A. Nowak, A. Sasaki, C. Taylor, and D. Fudenberg. Emergence of cooperation and evolutionary stability in finite populations. *Nature*, 428:646–650, 2004.

[31] D. G. Rand and M. A. Nowak. The evolution of antisocial punishment in optional public goods games. *Nature Communications*, 2:434, 2011.

[32] J. G. Riley. Silver signals: Twenty-five years of screening and signaling. *Journal of Economic Literature*, 39(2):432–478, 2001.

[33] E. Serrano, M. Rovatsos, and J. Botia. A qualitative reputation system for multiagent systems with protocol-based communication. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pages 307–314, 2012.

[34] R. M. Seyfarth, D. L. Cheney, and P. Marler. Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, 28(4):1070–1094, Nov. 1980.

[35] K. Sigmund. *The Calculus of Selfishness*. Princeton University Press, 2010.

[36] K. Sigmund, H. De Silva, A. Traulsen, and C. Hauert. Social learning promotes institutions for governing the commons. *Nature*, 466(7308):861–3, Aug. 2010.

[37] B. Skyrms. *Signaling: Evolution, Learning and Information*. Oxford University Press, New York, 2010.

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

[39] M. Tomasello. *Origins of Human Communications*. MIT Press, 2008.

[40] A. Traulsen, M. A. Nowak, and J. Pacheco. Stochastic dynamics of invasion and fixation. *Physical Review E*, 74(1):011909, July 2006.

[41] K. Von Frisch. *The dance language and orientation of bees*. Harvard University Press, Cambridge, MA, USA, 1967.

[42] S. A. Williamson, E. H. Gerding, and N. R. Jennings. Reward shaping for valuing communications during multi-agent coordination. In *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems*, pages 641–648, 2009.

[43] A. Zahavi. Mate selection - a selection for a handicap. *Journal of Theoretical Biology*, 53:205–214, Sept. 1975.

[44] A. Zahavi. The cost of honesty (further remarks on the handicap principle). *Journal of Theoretical Biology*, 67(3):603–5, Aug. 1977.

[45] W. Zhuang and M. Ismail. Cooperation in wireless communication networks. *Wireless Communications, IEEE*, (April):10–20, 2012.

[46] K. J. S. Zollman. Finding alternatives to handicap theory. *Biological Theory*.