Putting oneself in everybody’s shoes - Pleasing enables indirect reciprocity under private assessments

Marcus Krellner\textsuperscript{1,2} and The Anh Han\textsuperscript{1}

\textsuperscript{1}Teesside University, UK
\textsuperscript{2}TU Dresden, Germany
m.krellner@tees.ac.uk, t.han@tees.ac.uk

Abstract

Indirect reciprocity is an important mechanism for promoting cooperation among self-interested agents. Simplified, it means you help me, therefore somebody else will help you (in contrast to direct reciprocity: you help me, therefore I will help you). Indirect reciprocity can be achieved via reputation and norms. Strategies relying on these principles can maintain high levels of cooperation and remain stable against invasion, even in the presence of errors. However, this is only the case if the reputation of an agent is modeled as a shared public opinion. If agents have private opinions and hence can disagree if somebody is good or bad, even rare errors can cause cooperation to break apart. This paper examines a novel approach to overcome this private information problem, where agents act in accordance to others’ expectations of their behavior (i.e. pleasing them) instead of being guided by their own, private assessment. As such, a pleasing agent can achieve better reputations than previously considered strategies when there is disagreement in the population. Our analysis shows that pleasing significantly improves stability as well as cooperativeness. It is effective even if only the opinions of few other individuals are considered and when it bears additional costs.

INTRODUCTION

Cooperation is already considered central to complex life (Nowak, 2006; Perc et al., 2017) and may prove central to artificial systems as well (Langton, 1997). But cooperation relies on a handful of mechanisms that can enable it. One of these is indirect reciprocity (Rand and Nowak, 2013). It stands out, because it does not require any relatedness or other structural order between individuals, nor does it require repeated interactions. This makes it especially relevant for modern global economies, where human or artificial agents often engage in non-repeated interactions. Indeed, real-world applications of indirect reciprocity, particularly, reputation-based systems, are prevalent in e-commerce (Jiang and Li, 2007), socio-technical systems (Andras et al., 2018) and artificial societies (Conte and Paolucci, 2002).

Indirect reciprocity is usually modeled using self-interested agents playing the donation game: a random agent (donor) is selected to pay a personal cost $c$ to grant benefit $b$ to another randomly selected agent (recipient). The benefit is assumed to be bigger than the cost, i.e. $b > c$. Hence, it is best for the population collectively if every agent decides to take the cost, so that the sum of all wealth would increase by $b - c > 0$ with each interaction. However, individually preferred choice for each agent is to avoid the cost (i.e. defect), making the game a social dilemma. Thus, to maintain cooperation and prevent defection, agents may use strategies that are based on reputations and norms (Nowak and Sigmund, 1998a; Ohtsuki and Iwasa, 2006).

The reputation of the recipient determines how the potential donor should act (pay the cost or not). Norms determine what reputations the acting agent will earn. A simple version of such a strategy may state: If the reputation of the recipient is good, then donate, otherwise defect; and: if an agent donates, he earns a good reputation, otherwise a bad one (Nowak and Sigmund, 1998b). It was shown that such simple strategies, whose norms only consider actions, cannot maintain stable cooperation. With these norms, agents will earn a bad reputation if they do not donate to those with a bad reputation. This is known as the problem of justified punishment (Panchanathan and Boyd, 2003).

This problem can be solved using norms that also consider the current reputation of the recipient (so called second-order norms) or even the reputation of the donor itself (third-order norms) (Ohtsuki and Iwasa, 2004, 2006). Each possible scenario (a good/bad donor cooperates/defects against a good/bad recipient), linked to a resulting reputation (good/bad), leads to 256 possible norm combinations. An exhaustive search for successful norms showed that only eight can reliably maintain cooperation (Table 1). They are collectively referred to as the “leading-eight” but some are mentioned separately in the literature, e.g. L1 “standing” (Leimar and Hammerstein, 2001; Panchanathan and Boyd, 2003), L6 “stern judging” (Pacheco et al., 2006; Santos et al., 2018) and L7 “staying” (Okada et al., 2017; Sasaki et al., 2017). Standing was one of the first strategies to address the problem of justified punishment (see above), whereas stern judging and staying have been suggested as the best indirect reciprocity norms in recent years.

Most previous studies on the subject, however, assumed a
Table 1: Norms and action rules of the leading-eight. They show many similarities that often confirm with intuitive moral judgement, e.g. if a donor is currently considered bad and he cooperates with a good recipient, he can earn a good reputation, i.e. being forgiven. C = cooperates, D = defects, G = good reputation, B = bad reputation

In this paper, we investigate a novel approach to coping with private assessment, called ‘pleasing’. Pleasing rests on two simple rationales: i) agents are being aware that others may have distinct opinions and ii) agents want to maximize the chance of getting a good reputation (so as to obtain a better payoff). For instance, if a donor wants to defect, but the majority of players think he should cooperate (Figure 1), he will cooperate to please them.

In this scenario, information sharing is not a collective act. Instead only the current donor is actively seeking information. Which seems almost natural for him to do, since this information is viable for his future reputation. This enables one to avoid a major problem associated with information sharing: who is responsible and ready to pay the cost of the information distribution (Suzuki and Kimura, 2013)? It is natural to assume that the direct beneficiary of some infor-
mation would be willing to make effort to gather it. The
donor could even decide to compensate others for the infor-
mation they share, depending on how much additional cost
pleasing is worth to him.

Is it possible to solve the problems of private assessment
with ‘pleasing’? Judging from moral intuition, pleasing may
seem an undesirable behavior. People who are too easily
persuaded by the majority might be looked down upon. And
it is easy to imagine situations where pleasing would cause
harm to society or individuals; when one goes along with
punishing people who do not deserve it. But are such acts
the misfire of an otherwise good mechanism that prevents
us more often from punishing people we ourselves mistaken-
ly believe to deserve it? And besides humans, how should
artificial agents interact to maximize overall productivity?
As a first step to answer such questions, we study pleasing
from a strict evolutionary game theory perspective (Sig-
mund, 2016).

The remainder of the paper is structured as follows. We
next describe our models and methods. We then show how
pleasing improves the evolution of cooperation under private
assessment when it is not restraint. Afterwards we analyze
how robust pleasing is if only a few other agents (instead of
all possible agents) are pleased and when pleasing bears an
additional cost. Finally, we further discuss the implications
of our findings while pointing out limitations and future di-
rections.

MODELS AND METHODS

We test pleasing for stability and cooperativeness. We run
agent-based simulations under private, noisy and incom-
plete information, adopting a similar setting as in Hilbe
et al. (2018), for a clear and convenient comparison. Agents
play either unconditional strategies, i.e. always cooperate
(All-C) or always defect (All-D), or a conditional strategy
based on norms (leading-eight: L-X). In addition to the
strategies described in the literature (Table 1), L-X play-
ners may try to please other L-X agents. We analyze sta-
bility as the time spent in the state of homogeneous L-X
population under the constraint of rare mutations (select-
mutation equilibrium: SME) (Sigmund, 2016; Fudenberg
and Imhof, 2006). We subsequently compute the average co-
operation rate in the evolving population overtime (denoted
by Copop), which we use to indicate the success of indirect
reciprocity.

Model and Agent-based Simulation Setup

We consider a well-mixed population of size $N$. Reputa-
tions have two possible states, 1 (good) or 0 (bad), and players can
choose between two actions, cooperate or defect. Agents
have either an unconditional strategy (All-C, i.e. always co-
operate, or All-D, i.e. always defect) or a conditional one.
Conditional strategies are defined by a vector pair $(\alpha, \beta)$ that
represent their assessment and action rules. The assessment
rule $\alpha$ has eight entries, one for each combination of three
situational features: action and current reputations of donor
and recipient. An entry is equal to 1 if the resulting reputa-
tion is good, otherwise it is 0. The action rule $\beta$ corresponds
to the four possible combinations of reputations of the donor
and recipient. An entry is 1 if the agent will cooperate in that
situation and 0 if he will defect. As in Hilbe et al. (2018),
we will focus on the ‘leading-eight’ strategies L-X (Ohtsuki
and Iwasa, 2006) given in Table 1.

Reputation Dynamics

We simulate the reputation dynamics for players with at
most two fixed strategies. The state of reputation is given
by the $N \times N$ image matrix $M(t)$ of the population at time $t$
(Uchida, 2010). An entry $m_{i,j}(t)$ is equal to 1 if player $i$
has a good opinion of player $j$, and 0 otherwise. Initially all en-
tries are set to 1, a state of homogeneously good reputation.
As described in Hilbe et al. (2018), other initial states only
change results for the extreme opposite, homogeneously bad
reputations, and even then, only for L7.

Each time step in a simulation consists of three parts.
First, a donor $do$ and a recipient $re$ are drawn at random from
the population. The donor then decides whether to coopera-
te. Unconditional players always act the same, whereas the
standard leading-eight players decide by their action rule $\beta$
and their opinions of themselves $m_{do,do}(t)$ and of the recipi-
ent $m_{do,re}(t)$. For pleasing leading-eight players see below.
In the third step, reputations are updated due to observa-
tions.

The donor and the recipient always observe the interaction,
whereas other players independently observe it with proba-
bility $q$. Any observation (even by the focal donor or
recipient) is independently altered due to error with proba-
ibility $\epsilon$ towards the opposite action (e.g. cooperation instead
of defection). Each player who has observed the interaction
updates her opinion of the donor according to her assessment
rules, the action she observed and her current opinions
(i.e. even the donor updates her opinion of herself). The re-
result is the image matrix $M(t + 1)$. This process is repeated for
$10^6$ rounds in all the results shown below.

Pleasing

Pleasing L-X players differ from the original L-X in two
ways. They share their opinion with others and they decide
how to act based on the opinion of the majority when being
chosen as a donor. Let $n_w$ be the number of players the
donor wants to please and $N_p$ the number of pleasing L-X
players in the current population. The donor will consider
the opinions $\lambda$ of $n_p$ pleasing players, where $do \notin \lambda$ and $n_p$
is the minimum of $n_w$ and $N_p - 1$. The average opinions
$\hat{m}_{do}$ and $\hat{m}_{re}$ are given by

$$
\hat{m}_y = \frac{\sum_{i=\lambda} (m_{i,y}) + m_{do,y}}{n_p + 1}.
$$

(1)
The majority opinion

\[ m_{maj,y} = \begin{cases} 0, & \text{if } \hat{m}_y < 0.5 \\ 1, & \text{if } \hat{m}_y > 0.5 \\ \text{randomly 0 or 1, otherwise} \end{cases} \] (2)

The pleasing donor will act according to her action rule as if the majority opinion was her own. Afterwards observers assess the action as before (Note, the averaged opinions are not used again, especially not to update the opinions of the donor. All-C and All-D do not have opinions and do not share information.)

**Payoffs**

Based on simulations we compute the average rate of cooperation \( x_{A,B} \) of agents with strategy \( A \) towards agents with strategy \( B \). We only consider rounds after an initial transition period, i.e. we average only the second half of all rounds. If two strategies are present in a population of size \( N \), strategy \( A \) is played by \( k \) agents and strategy \( B \) by \( N - k \). The average payoff per interaction in which the agent is involved is given by the benefits she receives minus the cost.

\[ \hat{\pi}_A = \frac{x_{A,A}(N - k - 1) + x_{B,A}k}{N - 1} - \frac{A,B}{N - 1}c. \]

For some analysis we also subtract a cost \( c_p \) for pleasing strategies, which gives the payoff for a leading-eighth strategy that adopts pleasing \( A_p \) with \( \hat{\pi}_{A_p} = \hat{\pi}_A - c_p \).

**Evolutionary Dynamics**

We assume that players may consider updating their strategy according to their accumulated payoffs. We model this process using Evolutionary Game Theory methods for finite populations (Traulsen et al., 2006). Particularly, in each step of the evolutionary process, two players, \( i \) and \( j \), are randomly selected from the population for update via pairwise comparison: \( i \) will adopt the strategy of \( j \) with a probability given by the Fermi function

\[ p_{i,j} = \left( 1 + e^{-s(\hat{\pi}_j - \hat{\pi}_i)} \right)^{-1}. \] (3)

The parameter \( s \) defines the ‘imitation strength’, i.e. how much \( p \) depends on the payoff difference. For \( s = 0 \) imitation will be entirely random and as \( s \) approaches infinity the imitation process becomes more deterministic. We adopt here \( s = 1 \) for convenient comparison to Hilbe et al. (2018).

We want to test the stability of L-X strategies against All-C and All-D. Thus, for each L-X we consider populations with players that either play the L-X strategy at hand or All-C or All-D. We further assume for simplicity and convenience the limit of rare mutations (Wu et al., 2012). A mutation is an event where a player adopts any of the three strategies at random. In the limit of rare mutation, the population will almost always be in a homogeneous state with a single strategy present, since these states are absorbing and cannot be escaped by imitation alone. If a population consists of only strategy \( B \) and a new mutant of strategy \( A \) arrives, the population will subsequently either reach a homogeneous state of \( A \) or \( B \) and there will be at most these two strategies present at during the transition. The probability that \( A \) will take over, i.e. fixate, is denoted by \( \rho_{B,A} \). It is determined through the probabilities of incremental steps towards more players of \( A, T_{A,A} \), or the reverse step \( T_{A,A}^- \), as follows.

Let \( k(1 \leq k \leq N - 1) \) be the number of individuals using strategy \( A \) and \( (N - k) \) the number for strategy \( B \), in the population, then

\[ T^+(k) = \frac{N - k}{N} \frac{k}{N} \rho_{A,B}. \] (4)

Then,

\[ \rho_{B,A} = \left( 1 + \sum_{i=1}^{N-1} \prod_{i=1}^i \frac{T^-(j)}{T^+(j)} \right)^{-1}. \] (5)

These fixation probabilities determine the Markov Chain that describes the evolutionary transitions between the homogeneous states of \( q \) strategy, with a transition matrix \( Z \) of size \( q \times q \). The entry \( z_{A,B} = \rho_{B,A}/(q - 1) \) for \( A \neq B \) and \( z_{A,A} = 1 - \sum_{i=1}^q z_{j,A} \). The normalized eigenvector with eigenvalue 1 of the transposed matrix of \( Z \) provides the selection-mutation equilibrium \( \sigma \) for each considered strategy. It describes, in the long run, how often the population will spend in a homogeneous state of the corresponding strategy (Fudenberg and Imhof, 2006).

**Evolution of Cooperation**

Evolutionary success of L-X is not a guarantor for high cooperation rates. Strategies may succeed with little or no cooperation (e.g. L8). We therefore compute the cooperation rates within each homogeneous population. That is either in the state of All-C, All-D or one of the L-X strategies. We then obtain the average cooperation \( \bar{X} \) (Copop) of the whole system \( \bar{X} = \sum_{i=1}^q \sigma_i x_i \), where \( \sigma_i \) is the frequency of strategy \( i \) at the selection-mutation equilibrium and \( x_{i,i} \) the cooperation rate in the homogeneous state of strategy \( i \) (which is one for All-C and zero for All-D).

**RESULTS**

**Improvements by Pleasing**

We first investigate pleasing L-X strategies when it is not constraint (i.e. players consider the opinion of all other pleasing players) for varying the rate of perception error \( c \) and benefit of cooperation \( b \) (fixing \( c = 1 \)). Results for Copop are shown in Figure 2. The abundance of warm colors in the right columns (i.e. with pleasing) compared to the left shows the great potential of the new pleasing strategies. Copop is increased in wide parameter spaces for most L-X. With pleasing, higher benefits always increase cooperation rates and high Copop is even possible for frequent errors.
L1, 2, 6 & 7 utilize pleasing the best within the leading-eight. High benefits and low error rates (upper right quadrant) seem generally best for cooperation, compared to low benefits and high error rates (lower left quadrant).

The improvements are the result of both increased stability and cooperation within leading-eight. Cooperativeness in homogeneous populations (Coho) of pleasing leading eight are in fact mostly perfect for $\epsilon < 0.2$. The differences in the resulting cooperation rates (Copop) were mainly caused by differences in stability. The only exception, as in previous studies (Hilbe et al., 2018), was L8, which failed to have any cooperation with itself and was therefore exactly as stable as All-D. Stern judging (L6) is also different in that it stays almost perfectly stable for increasing error rates but groves steadily less cooperative.

The observation rate, the probability that a non-involved player perceives what the donor does, was previously reported to be of little effect when any errors are present (Brandt and Sigmund (2004), compare also with Hilbe et al. (2018)). The same is true for pleasing, except if zero observation rate is approached.

**Efficiency of Pleasing**

Enquiring all other opinions is likely to be expensive (e.g. calories, time or opportunity costs). We therefore analyze cooperation rates if players would try to please smaller numbers of agents. Figure 3 shows an example for some commonly used parameters, but our additional analysis show that the qualitative results are robust for other values of $c$ and $b$. As the number of pleased players increases, Copop often increases steeply, then levels of. That means, that pleasing a subset can be as effective as pleasing the entire population. Even pleasing just a few agents is enough for L1 (standing) and L7 (staying) to reach their full potential. L6 (stern judging) on the other hand appears to have a threshold after which it rapidly reaches total stability (SME = 1), but each additional agent that it pleases makes it more cooperative.

We also consider the cost of pleasing itself (Figure 4). Namely, pleasing players have to pay a cost $c_p$ each time they are selected as donors. This cost represents the time and energy they spend obtaining others’ opinions, but could also include compensation for some third-party agents from whom they gather information. Increasing the cost significantly decreases Copop with a single exception; L6 (stern judging) was not affected by small costs. Since costs do not change the cooperativeness in homogeneous states (Coho), the changes in Copop are entirely a result of lowered stability against the unconditional strategies, which are not affected by costs. As described above, L6 had the highest SME and was stable enough to not lose any ground, even when baring considerable costs for pleasing. Additional analysis with different parameters support this finding, whereby higher benefits enable L6 to endure even higher costs.

**DISCUSSION**

**Explanations**

We expected pleasing to benefit reputation-based strategies such as the leading-eight. They rely on information and pleasing uses more information for each decision, compared to the original leading-eight. However, pleasing players use this information not to update their own opinions and to directly reduce disagreement. Instead, they use it to selfishly improve their own reputations. So, it may come as a surprise that this strategy increases both stability and cooperativeness. In the study that discovered the leading-eight strategies, all successful norms had a single best set of action rules (Ohtsuki and Iwasa, 2006). They can all be described as follows: defect, unless cooperation would give you a better reputation than defection. In that sense, these strategies were acting selfish already. Selfishness can lead to stable cooperation.

Since pleasing is a better strategy to optimize one’s reputation than the original leading-eight, better reputations and hence more cooperation within L-X players come not as a surprise. However, pleasing has three less direct effects. It does also raise the reputation of All-C and lowers the reputation of All-D. As a result, pleasing strategies fare much better against All-D, but lose some advantages over All-C. This causes the loss of stability for L2 for some parameter ranges. In all other cases however, it is a net win in SME.

The changes in reputation cause a further effect. Pleasing shifts reputations away from mixed (half bad, half good opinions) towards good for L-X and All-C and towards bad for All-D. As a consequence, there is less possibility for disagreement. If reputation is unanimously good (or bad), there cannot be any disagreements. Increasing reputation that was already above 50% good tends to decrease the number of disagreements between agents. Disagreement was an important cause of bad reputation for L-X in other studies with private assessment (Brandt and Sigmund, 2004; Hilbe et al., 2018). Hence, pleasing likely leads to a positive feedback to improve reputation of L-X players even further.

Third, pleasing affected the actions of agents directly and caused more cooperation for the same amount of good reputation. Whereas an agent with 60% good reputation would receive a benefit 60% percent of the time from original leading eight agents, it would receive it 100% of the time from pleasing agents (if they please all other L-X in the population). A random opinion would be good 60% of the time, but the majority opinion is always good. The same principle causes the opposite effect for agents with bad reputations (i.e. in most cases All-D) which earned less benefits than for the same reputation under the old strategies. These polarization of cooperation and defection seems to have helped the stability of L-X players as well as increasing their cooperation with each other.
Copop in the parameter space of $b$ and $\epsilon$

Figure 2: Effect of pleasing on the evolution of cooperation (Copop), across benefit $b$ (x-axis) and error rate $\epsilon$ (y-axis). First and third columns show Copop with original leading-eight (L-X); second and fourth columns show L-X with pleasing. $N = 50$, $n_w = 50$ (i.e. pleasing all L-X players in the population), $q = 0.9$.

**General effect**: Pleasing greatly facilitates the evolution of cooperation in large parameter regions. Copop is lowered by pleasing only in a small area of low error rate and low benefits for a single strategy, L2. For these parameters, pleasing slightly decreases stability of the L2 (i.e. SME) and hence the resulting cooperation (Copop).

**Parameter space**: The original leading-eight were somewhat successful only if error rates and benefits were both relatively low (upper left quadrant). More frequent errors but also higher benefits decrease their stability (due to invasion by All-C). With pleasing, higher benefits always increase cooperation rates instead. Now, high benefits can also offset the negative effects of high error rates (lower right quadrant). Even if error occur as often as every third time, there are still a few cases that can maintain more than 50% cooperation.

**Different strategies**: Originally only three of the leading-eight were somewhat successful (L1 "standing", L2 & L7 "staying"), showing small parameter ranges that allow for at least 50% cooperation. With pleasing, this is true for all but one strategy, L8. Although relatively stable, L8 showed no cooperation, with or without pleasing. Four strategies (L1, L6 “stern judging” & L7, less so L2) can maintain more than 70% cooperation for large parameter ranges.
Relevance, Limitations and Future Work

Pleasing is a novel information sharing approach to solve the problem of private assessment that does not aim to synchronize opinions. It can lead to reliable indirect reciprocity even for the hardest conditions studied to far, when information is private, noisy and incomplete. It is therefore definitely a strategy worth of future consideration. The current work is our first step to understand the consequences and relevance of pleasing.

Next, we aim to study how pleasing fares against other strategies besides the unconditional ones that were investigated here. Particularly, we will address several candidates that could challenge the new strategy. For example, players that want to rely on the opinions of others but don’t want to share their own opinions to avoid costs. Such strategies were pointed out as a problem of information sharing in general (Suzuki and Kimura, 2013). As mentioned, a solution could be to pay players for their information.

Another problem could be lies or irrelevant opinions. Note, that we do not assume that pleasing L-X can differentiate between L-X and unconditional strategies. If those strategies would share opinions, a pleasing player would consider it valuable information. He may fail to recognize the majority opinion of L-X players and therefore may fail to please them. Pleasing unconditional players instead does not yield any benefit, since they treat him independently of his actions. However, lying can take many different forms. Our preliminary simulations show for example, that All-D players, who claim everybody is bad, actually benefit pleasing players. The full range of lying strategies needs to be investigated in the future.

Is pleasing realistic? It seems highly impracticable to literally run around and ask what others think in a situation where one is supposed to help someone, especially if help is required fast. However, pleasing could be done in real life by other means. Instead of asking right beforehand, one could ask habitually and have the information ready at hand when needed. This way the information might be outdated, but a small decrease in accuracy should not diminish the effect of pleasing all together, comparable to pleasing a subset of players. Secondly, instead of asking, opinions could be inferred via theory of mind or cognitive abilities, see e.g. Han et al. (2012); Han (2013). Similar capabilities were already shown to improve indirect reciprocity, but when attributed to the observer instead of the donor (Radzvilavicius et al., 2019).

A minor limitation of the current study is that the pleasing described here will not always find the action with the best possible reputation. Rather, it is a cheap heuristic to find a very good action but could be improved. Instead of using the average opinions (see Methods) to decide for an action, it could use each opinion of each agent itself to infer the consequences of each action for each pleased agent and then chose the action that grants the best possible reputation. Considering this more sophisticated approach could push the limits of pleasing even further.

This paper may also contribute to the question, which indirect reciprocity strategy is the best. Two strategies may have been endorsed the most; stern judging (L6, e.g.
Pacheco et al. (2006); Santos et al. (2018)) and staying (L7, e.g. Okada et al. (2017); Sasaki et al. (2017)). Our study is in line with these findings. Both strategies do also very well when they are combined with pleasing. Stern judging can achieve total dominance against All-C and All-D for a large range of parameters (the rate of perception error $\epsilon$ and the benefit of cooperation $b$) and is also robust against additional costs of pleasing. The staying strategy on the other hand remains stable even under hard conditions (high $\epsilon$ and low $b$) and shows more cooperation in homogeneous states.

But, these two strategies are not the only candidates. Standing (L1), one of the first strategies ever to be investigated (Leimar and Hammerstein, 2001; Panchanathan and Boyd, 2003), does also very well, sometimes better than both L6 and L7. These results seem to suggest that, to achieve optimal indirect reciprocity, different situations require different strategies. On an abstract level, this may reflect how different societies enabled cooperation with different norms, without a single one being strictly better than the others.

**Conclusion**

In summary, we have shown that pleasing can significantly improve stability and cooperativeness of the leading-eight. The improvement is especially large for high benefit to cost ratios and high rates of perception error. The evolution of cooperation is only reduced by pleasing for a single strategy (L2) and a narrow parameter range. Pleasing is effective even when only a small subset of other players are being pleased and when it bears additional costs. Standing (L1), stern judging (L6), staying (L7) and L2, are particularly successful when combining with pleasing.

In short, our investigation has provided novel insights into the importance and success of indirect reciprocity as a mechanism for promoting cooperation among self-interested agents in non-repeated interactions, for the hard and more realistic conditions where information is private, noisy and incomplete.

**Acknowledgements**

This work is supported by Future of Life Institute (grant RFP2-154).

**References**


Han, T. A. (2013). Intention recognition, commitment and their roles in the evolution of cooperation: From Artificial intelligence techniques to evolutionary game theory models. Springer SAPERE.


