Ecology

ISBE
International Society for Behavioral Ecology

Original Article

Pairs of cleaner fish prolong interaction duration with client reef fish by increasing service quality

Simon Gingins[a,b] and Redouan Bshary[a]

[a]Institute of Biology, University of Neuchâtel, Emile-Argand 11, 2000 Neuchâtel, Switzerland and
[b]Department of Ecology and Evolution, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland

Few biological examples of cooperation seem to precisely fit the assumptions of an iterated prisoner's dilemma. In an attempt to increase biological validity, one model altered the assumption that cooperating is an all-or-nothing decision to a situation where benefits are a function of interaction duration, which in turn is a function of the levels of cooperation. A potential application involves pairs of cleaner fish coinspecting a client fish. In this mutualism, clients visit cleaners to have ectoparasites removed but a conflict of interest exists, as cleaners prefer to eat client mucus, which constitutes cheating. As large clients often flee in response to a cleaner cheating, pair inspections lead to a dilemma: the cheater obtains the benefit while both cleaners share the cost of the client leaving. The model predicts that pairs of cleaners behave more cooperatively toward reef fish clients than when inspecting alone, to entice clients to profit from the increased parasite removal rate and keep interaction duration almost constant. Here, we present field experiments that first replicate results that pairs behave indeed more cooperatively than when inspecting alone and second show that levels of cooperation quantitatively predict the duration of cleaning interactions. We also found that several additional variables may affect the duration of cleaning interactions, such as a client's willingness to interact with a cleaner, identity of interaction terminator, and the presence of bystanders. In conclusion, introducing benefits as a function of interaction duration into the prisoner's dilemma framework provides a biologically relevant framework to study cooperation.

Key words: cooperation, game theory, *Labroides dimidiatus*, marginal value theorem, reciprocity.

INTRODUCTION

Of the many forms of cooperation, reciprocity has attracted particular interest because it appears to be highly vulnerable to exploitation by cheaters. In reciprocity, each act is an investment, that is, an act that reduces the immediate payoff of the actor and increases the immediate payoff of the recipient (Trivers 1971; Bshary and Bergmüller 2008). Hence, there is a strong temptation to cheat, as usually illustrated with the well-known prisoner's dilemma (PD) game, which in its iterated form (IPD) has been the standard framework to model reciprocity (Axelrod and Hamilton 1981). However, empiricists have found it challenging to provide evidence for the reciprocal strategies predicted to solve the IPD. A variety of potential examples exist (Raihani and Bshary 2011), including cases with obvious biological relevance, like predator inspection (Milinski 1987), allogrooming in primates (Seyfarth and Cheney 1984), mobbing of predators (Krams et al. 2007), and laboratory setups that

explicitly construct an IPD (St-Pierre et al. 2009; Schneeberger et al. 2012). Nevertheless, other forms of cooperation appear to be far more abundant (Leimar and Hammerstein 2010; Archetti et al. 2011; Bshary and Bronstein 2011) and the examples listed above are often contested. Some authors suggest that reciprocity is cognitively too demanding for most animals (Stevens and Hauser 2004; Connor 2010), some doubt the existence of the IPD payoff structure in real-life examples (Noë 2006; Clutton-Brock 2009), and others doubt that animals use the solutions proposed by theoreticians, that is, simple but precise counting strategies like the variations of tit for tat (Brosnan and de Waal 2002; Schino and Aureli 2009).

Here, we explore a recent extension of the IPD framework that may partly resolve several of the criticisms. Following West et al. (2007), our aim is not to defend the IPD as a whole, but to show the usefulness of biologically informed models that may otherwise be quite similar to the IPD framework. In the standard IPD, the actual behavioral act is best described with pressing 1 of 2 possible buttons. The behavior, therefore, has no time dimension and no subtlety: cooperating is an all-or-nothing decision. Although studies explored

---

the effects of variable investment in an IPD (Barrett et al. 2000; Roberts and Renwick 2003; Barclay and Willer 2007), the decision was always discrete and lacked a time component. However, a time dimension is a feature of many cooperative interactions. For example, the inspection of a predator by small fishes consists of repeated saltatory advances (Pitcher et al. 1986), mobbing of a predator by birds can vary in both duration and intensity (Krams et al. 2007), and lions take time to approach potential intruders (Heinsohn and Packer 1995). The time dimension gives individuals the opportunity to monitor the behavior of partners and to respond accordingly within an interaction: cooperative female lions approach speakers simulating a territorial intrusion more slowly when paired with a laggard versus a cooperative individual (Heinsohn and Packer 1995). Thus, adding a time dimension to a PD-like interaction may change the strategic options of players or even reduce cognitive constraints, as responses are immediate rather than delayed (Brosnan et al. 2010).

The model to be tested

A relevant model explored the scenario that pairs of service providers interact simultaneously with the same client (Bshary et al. 2008). The assumptions of the model were explicitly based on marine cleaning mutualism involving the cleaner wrasse *Labroides dimidiatus* (the service provider), which feeds on the ectoparasites of other reef fishes (the clients). A conflict of interest exists because cleaners prefer to eat client mucus, which constitutes cheating from the client fish perspective (Grutter and Bshary 2003). Cleaners take some risks when they cheat; for example, visitor clients with access to several cleaning stations may flee in response and, in addition, seek a different cleaner for their next inspection (Bshary and Schäffer 2002). Being left prematurely imposes opportunity costs on cleaners, as they have to wait longer for a new client to invite inspection. The model takes both the benefits of cheating and the opportunity costs into account. Also, the model makes the reasonable assumption that the potential gains for cleaners from eating parasites are a function of diminishing returns as the longer an interaction continues, the more the parasites on the client are depleted. With these assumptions, a single cleaner's solution to the optimal foraging problem fits the logic of the marginal value theorem (Charnov 1976; Bshary et al. 2008). The situation becomes more complex if 2 cleaners inspect the same client simultaneously. When taking a bite of mucus, the cleaner does not only cheat the client fish but also jeopardizes foraging opportunities for its partner cleaner: only the cheating cleaner gains an extra benefit from its deceitful act although both cleaners pay the associated costs. In the perspective of the cleaners, pair cleaning can thus be viewed as a dilemma between increasing foraging opportunities and obtaining the extra benefit from exploitation before the partner does. Under these circumstances, 2 evolutionarily stable strategies (ESS) emerge: either to cheat immediately or a cooperative solution that fits again the logic of the marginal value theorem. Interestingly, pair inspections lead to higher levels of cooperation (service quality) than when a cleaner inspects alone (Bshary et al. 2008). The reason is as follows: as long as cleaner service quality remains stable, increased parasite removal rates during pair inspections will cause cheating to occur earlier as well and, hence, lead to opportunity costs due to prolonged waiting time for the next client. The only way to reduce these opportunity costs is to improve service quality as interaction duration is a simple function of cleaner cheating rates in the model (Bshary et al. 2008). The model, thus, predicts that as long as 2 cleaners do not hinder each other while searching for parasites, the service quality is predicted to almost double,

leading to almost the same duration as in singleton interactions. It is, however, possible that 2 cleaners inspecting a client simultaneously would interfere with each other, for example, when both are simultaneously inspecting the same body part of the client. As a result, the parasite search efficiency for each cleaner might be lower in a pair situation than in a singleton situation. According to the model, interference in search efficiency would cause a reduction in service quality but, nevertheless, prolong interactions further as it slows down foraging (Bshary et al. 2008). Although interference is thus of empirical importance, the key conclusion for the current article is that the model's predictions fit the logic of an IPD as only the cheater obtains the extra benefits, whereas both cleaners share the costs of the client leaving in response, leading to a cooperative and a noncooperative ESS. The main difference to the standard IPD is that the interaction has a time dimension and possible levels of cooperation are continuous from cheating immediately to never cheating (for similar assumptions, see Sherratt and Roberts 1998; Killingback et al. 1999; Johnstone and Bshary 2002). In summary, for a cooperative solution to the dilemma of cleaning in pairs, the model predicts that 1) service quality should almost double when clients interact with pairs versus singletons and 2) duration of interactions should be very similar between pairs and singletons, albeit slightly shorter for pairs.

Existing evidence

Currently, there is observational and experimental evidence that cleaner wrasse indeed seem to almost double their service quality when inspecting in pairs (Bshary et al. 2008). With real clients, cheating becomes visible to the observer by client jolts in response to cleaner mouth contact (Bshary and Grutter 2002). Client jolt rates in the field were very similar between interactions involving 1 cleaner and interactions involving 2 cleaners (Bshary et al. 2008). Given that cooperating involves feeding against preference, lab experiments using other food sources revealed that cleaners eat almost the double amount of nonpreferred food prior to eating a preferred item when feeding in pairs than when feeding alone (Bshary et al. 2008). In both natural observations and experiments, the increase in service quality was mainly due to changes in female behavior (Raihani et al. 2010), as the large male partners would punish females for cheating with aggressive chasing (Raihani et al. 2010). Thus, data on cheating frequencies conform to the predictions of the model while the asymmetry between partners was not foreseen.

Our study

Here, we test the second crucial prediction of the model that increased service quality leads to prolonged interactions, which has not yet been investigated. The predicted duration of an interaction is a simple function of cleaner cheating rates. Cheating rates, in turn, are an expression of cleaner service quality in the case of singleton inspections, whereas for pair inspections, one has to additionally take into account the possibility that interference between cleaners reduces search efficiency. Reduced search efficiency reduces the slope of the cumulative benefit function of parasite removal and, hence, should lead to an increased cheating frequency (equivalent to a reduced stay in a patch of lower food abundance; Charnov 1976). Finally, asymmetry between partners with respect to the ability to punish a cheating cleaner is one of the several real-life variables the model does not take into account. We, therefore, investigated various additional variables. First and most importantly, cleaners are known to manipulate client decisions

by providing tactile stimulation, which consists of cleaners touching the client with their pelvic fins on specific body parts in a way that is incompatible with foraging (Bshary and Würth 2001). Because clients appear to benefit from tactile stimulation (Soares et al. 2011), cleaners may induce clients into staying longer at the cleaning station by providing tactile stimulation, and 2 cleaners may be better at this than a single one. If this were the case, we expected that pairs of cleaners would prolong interactions beyond the model's prediction. Furthermore, when addressing the effect of pair inspections on interaction duration, one must be particularly careful with confounding variables. The presence of bystanders can enhance cooperative behavior through image scoring (Pinto et al. 2011). The presence of several clients may lead to frequent switching of cleaners from one client to another (personal observation). This causes pair interactions to be of longer duration than singleton interactions as whenever one cleaner switches to another client, this would terminate the interaction in singleton inspections but not in pair inspections. Another potential bias can be caused by interactions with passing clients ("passing-by interactions" hereafter)—fish passing by the cleaning station that are unwilling to interact. These short interactions usually involve only 1 cleaner. Finally, singleton interactions are increasingly likely to turn into a paired interaction the longer they last, as the likelihood that the other partner joins the interaction increases with duration. As a result, singleton interactions at pair stations are most likely to be observed if the duration is short.

In conclusion, to adequately test the model, we investigated the effect of all variables introduced above on cleaning duration. Data collection involved the temporary removal of 1 cleaner, so we could observe nonconfounded singleton interactions at a pair station. We hypothesized that the model's predictions regarding the link between client jolt frequency and interaction duration should be most closely matched for interactions sought (no passing-by interactions) and terminated by the client (one of the model's assumptions),

in the absence of bystanders (to avoid the above-mentioned biases). Furthermore, taking interference between coinspecting cleaner partners into account will yield the best model prediction concerning client jolt rates when interacting with a singleton relative to a pair of cleaners.

METHODS
Interaction duration, jolt rates, and tactile stimulation

Data were collected in an ~1.3-km strip along the coast at Marsa Bareika, Ras Mohamed National Park, Egypt (27°47′21″N, 34°13′17″E). The site consists of a discontinuous wall-like fringing reef of a maximum depth of 5 m following the coastline and a sandy slope with patch reefs extending toward the bay. Observations were made between October and November 2009 by S.G. either on the fringing reef or on patch reefs (between 1 and 14 m of depth).

In order to assess differences in behavior between pair and singleton inspections by *L. dimidiatus*, the following 4 observation sessions were conducted on 11 pairs of cleaners (Figure 1). Each session was conducted on a different day and always followed the same order: during session (1), the pair was observed for 1 h. Observations were performed using a client's perspective; thus, some interactions involved both cleaners, whereas others involved only the male or only the female. During session (2), one of the cleaners was caught in order to cause singleton interactions while controlling for identity and site. Catching was performed regardless of sex, using either 2 landing nets alone or a combination of landing nets and a barrier net placed at the site. Once removed, the individual was kept in a net bag for 45 min while observations on the remaining individual were conducted. A catching event can be compared with a predation event. Therefore, the behavior of the individual that was left on the site could be altered by the stress caused by the capture event. Observations were conducted for

Sessions	Description	N cleaners	Situations
1 3 4	The pair was observed for 1 h		**Paired** Both cleaners interacting with the client for more than 50% of interaction time **Solo Natural** Only one cleaner interacting with the client
2	One of the cleaners was caught and kept in a net bag for 45 min while observations were made on its partner		**Solo Experimental (1)** One cleaner present on site
	After 45 min the captured cleaner was released and observations were made for 30 minutes		**Solo Experimental (2)** Both cleaners present, but only one active. Due to the stress of the recent capture event, the released cleaner swims around but does not engage in pair inspections.

Figure 1
Field experiments. Four different observation sessions were made on 11 pairs of cleaner wrasse *Labroides dimidiatus*. Each session took place on a different day and always followed the same order (1–4). The solo experimental situations 1 and 2 showed no significant differences and were grouped under "solo experimental" for further analysis (see Results).

30 min immediately after the captured individual was released back to the site to assess whether the behavior of the focal individual was altered. Once the removed individual was replaced, it was expected to stay in the vicinity of its partner without participating in cleaning interactions, due to the stress of removal. This situation makes it possible to observe singleton interactions by the focal individual (the one observed during the previous 45 min) while its partner is present but not cleaning. During each of the sessions (3) and (4), the pair was observed for 1 h in the same way as in session (1). The analysis focused on 11 pairs for which the catching and removal events worked as planned. For these 11 pairs, 8 males and 3 females were caught during the capture sessions.

Clients categorized as nonpredatory visitors, defined as nonpredatory clients with large enough territories to have access to several cleaning stations (Bshary and Côté 2008), were the focus of field observations. Along with predatory visitors, these fish are considered attractive clients because cleaners actively seek them out and give them priority over the smaller resident clients. Focusing on interactions with this category of clients is thus more representative of the cleaners' choices. In addition, both cleaners will rarely inspect smaller and/or unattractive clients. Therefore, it seemed relevant to focus only on categories of clients that have similar chances to be inspected by pairs or singletons, as it was intended to make comparisons without controlling for client species identity. As *L. dimidiatus* is known to provide predators with a better service than nonpredatory clients (Bshary and Côté 2008), interactions with these visitors were not included in data collection. The data were collected on interactions with 16 different species of nonpredatory visitors, with the majority consisting of *scarines* (parrotfish) and *mullidae* (goatfish).

The following information was noted on a Plexiglas plate for every interaction:

- Client species according to Lieske and Myers (1996).
- Duration of the interaction (in seconds) from the clients' perspective.
- Initiation of the interaction. The client initiated if it posed (stopped swimming and remained stationary in a more or less upright position, maintaining its position by agitating its pelvic fins only) prior to physical contact. Otherwise, the cleaner was scored to initiate.
- Termination of the interaction. Whoever initiated movement away from the partner terminated the interaction.
- Number of cleaners. We scored pair inspection if both cleaners inspected simultaneously for more than 50% of the total interaction time, based on the observer's estimation. Singleton inspections consisted of 1 cleaner inspecting 100% of total interaction time. In these cases, the sex was noted when recognized. Interactions with less than 50% simultaneous inspection were discarded from the analyses.
- Tactile stimulation occurring or not.
- Number of jolts.
- Number of attractive clients within 1 m (bystanders) at the end of the interaction.

Interference

In order to determine whether pairs of cleaners interfere while cleaning together, data on cleaner nibbling rate (number of nibbles/second) were collected on 7 *L. dimidiatus* pairs along the coast of Dahab, Egypt, in May and November 2010. The analysis on interference focused on 126 interactions (totaling 1762 s), 51 of

them documented by S.G. and 75 of them documented by R.B. The following protocol was used: pairs were localized and observed for undefined amounts of time. Depending on the depth, observers were either snorkeling or SCUBA diving. During observations, the following information was recorded: pair identity, observer identity, sex of the focal individual, presence or absence of the partner during the interaction, client species, interaction duration, and number of nibbles (mouth contacts) the focal individual took on the client during the interaction. Data were collected on the same category of clients (nonpredatory visitors) as for interaction duration, jolt rate, and tactile stimulation.

Data analysis

We distinguished between "pair inspection" (both cleaners inspect the same client >50% of the total interaction time) and "singleton inspection" (1 cleaner inspects a client 100% of the total interaction time). Within singleton inspections, we further distinguished between "solo natural" (singleton inspection under natural conditions; i.e., the partner is nearby and active) and the 2 experimental situations: "solo experimental 1" (singleton interactions when the partner has been removed) and "'olo experimental 2" (singleton interactions when the partner had returned but was not yet cleaning again due to the stress of the recent capture).

The complete data set for duration, tactile stimulation, and jolt rates includes 872 (totaling 17 884 s) interactions distributed among the different observation sessions and the 11 different pairs. In order to control for potentially confounding variables, a subset of the data including only interactions terminated by the client, in absence of a bystander and excluding passing-by interaction had been created. Passing-by interactions were defined as follows: cleaner initiates, client terminates, no jolts, and lasting less than 10 s. In addition, each of the 3 potentially confounding variables was also removed from the data set independently to assess which one had most impact on the duration of interactions. The subset controlling for confounding variables contains 291 interactions (totaling 7633 s).

For singleton inspections, the effect of sex on duration was tested with a linear mixed model (LMM) and on tactile stimulation with a generalized linear mixed model (GLMM), with sex as the explanatory variable and pair identity as grouping factor. For jolt rates, the differences between sexes were tested with Wilcoxon signed-rank tests when data for singleton inspection for both sexes were available in every replicate (solo natural situation and overall analysis). When data for singleton inspection for only 1 sex were available in every replicate (solo experimental situations), Wilcoxon rank sum tests (equivalent to Mann–Whitney U) were performed.

Differences between situations in terms of duration of interactions were tested using a LMM with the situation as explanatory variable and pair identity as grouping factor. Differences in terms of tactile stimulation were tested using a GLMM. Pairwise comparisons between treatments were obtained with Tukey contrasts.

Jolt frequencies were calculated as the sum of jolts divided by the sum of interactions durations (in seconds) for 1 pair in 1 situation. Thus, for each pair in each situation, a single value of jolt frequency was used for further testing. Friedman tests were performed to detect overall differences between situations in terms of jolt frequencies. Duration and jolt rates are directly dependent in the model. We thus calculated both the jolt rate ratio and the duration ratio between paired and singleton interactions (using the median values) to document this relationship. For singleton interactions, values from the experimental situations were used because they are the most representative of the behavior of a singleton cleaner.

Interference was investigated through nibbling rates (number of nibbles/second), where each nibble is a cleaner mouth contact with the client's body surface. Differences between observers were tested using a Wilcoxon rank sum test (equivalent to Mann–Whitney U). To illustrate the interferences that might happen while both cleaners are cleaning simultaneously, nibbling rates were calculated for each individual in each situation (solo or paired) as follows: sum of nibbles across all interactions/sum of duration across all interactions.

All statistics were performed using the program R 2.14.1 (R Core Team 2013). For the LMMs, the R package lme4 (Bates and Maechler 2009) was used; Tukey contrasts were obtained with the R package "multcomp" (Hothorn et al. 2008).

RESULTS

A first analysis was performed to investigate differences between the 2 experimental situations: partner temporarily removed by the experimenter and partner present but inactive due to the stress of the capture (solo experimental 1 and 2 in Methods). There were no significant differences regarding durations of interactions (LMM fit by restricted maximum likelihood (REML): complete data: $z = 1.009$, $P = 0.313$; subset: $z = 1.312$, $P = 0.189$), occurrence of tactile stimulation (GLMM fit by Laplace—complete data: $z = 0.916$, $P = 0.36$; subset: $z = 0.175$, $P = 0.861$), and jolt rates (complete data: $V = 47$, $P = 0.240$; subset: $V = 19$, $P = 0.945$). Therefore, these 2 situations were grouped under the "solo experimental" category for further analysis.

Within singleton interactions, there were no significant differences between sexes in terms of duration in this study (LMM fit by REML—complete data: $z = 0.201$, $P = 0.841$; solo experimental: $z = 0.693$, $P = 0.488$; solo natural: $z = 0.943$, $P = 0.346$). Regarding tactile stimulation, females gave tactile stimulation significantly more often than males (GLMM fit by Laplace: $z = 2.869$, $P = 0.004$). However, when solo natural and solo experimental situations are considered separately, the difference is only significant for the solo natural situation (GLMM fit by Laplace: solo experimental: $z = 0.848$, $P = 0.397$; solo natural: $z = 3.218$, $P = 0.001$). Finally, males and females did not differ significantly with regard to jolt rates (complete data: $V = 33$, $P = 1$; solo experimental: $W = 5$, $P = 0.194$; solo natural: $V = 31$, $P = 0.7598$). We, therefore, pooled male and female solo data for the analyses, ending with 3 situations: pair inspection ("paired"), experimentally induced singleton inspection ("solo experimental"), and solo inspection under natural conditions ("solo natural").

Full data set

Contrary to the model's predictions, paired interactions were significantly longer than singleton interactions (LMM fit by REML: paired vs. solo natural: $z = -8.319$, $P < 0.001$; paired VS solo experimental: $z = -4.685$, $P < 0.001$) (Figure 2A). Among singleton interactions, solo experimental interactions were significantly longer than solo natural interactions (LMM fit by REML, $z = -3.855$, $P < 0.001$).

Jolt rates did not differ significantly among situations (Friedman chi square = 0.546, degrees of freedom [df] = 2, $P = 0.761$) (Figure 3A). Tactile stimulation appeared to be given more often in paired than in singleton interactions (GLMM fit by Laplace—paired VS solo experimental: $z = -4.594$, $P < 0.001$; paired VS solo natural: $z = -4.189$, $P < 0.001$). Among singleton interactions, there were no significant differences between the "solo experimental" and the "solo natural" situation (GLMM fit by Laplace, $z = 0.037$, $P = 0.999$) in terms of tactile stimulation (Figure 4A).

Subset of data controlling for passing-by interactions, presence of bystander, and client termination of the interaction

We first independently removed data on passing-by interactions, data with bystanders present, or data on interactions terminated by the cleaner. No single removal altered the conclusion from the general data set that pair inspections lasted longer than singleton inspections (LMM fit by REML, all $z < -2.7$, all $P < 0.02$, Figure 2B–D). It is only when all 3 variables were controlled for that the results on duration differed notably from the complete data set (Figure 2E). In this case, the solo experimental situation yielded interaction durations similar to pair inspection (LMM fit by REML, $z = -0.430$, $P = 0.902$), and both situations yielded significantly longer interaction times than the solo natural condition (LMM fit by REML—paired VS solo natural: $z = -3.304$, $P = 0.003$; solo experimental VS solo natural: $z = -2.475$, $P = 0.035$).

In the subset of the data, jolt rates did not differ significantly among situations (Friedman chi square = 2, df = 2, $P = 0.368$) (Figure 3B). Pairs provided tactile stimulation significantly more often than singletons in the natural situation, but this difference did not appear when compared with singletons in the experimental situation (GLMM fit by Laplace—paired VS solo natural: $z = -2.456$, $P = 0.037$; paired VS solo experimental: $z = -2.053$, $P = 0.098$) (Figure 4B). Finally, solo experimental and solo natural conditions did not yield significant differences in terms of tactile stimulation (GLMM fit by Laplace, $z = -0.502$, $P = 0.869$).
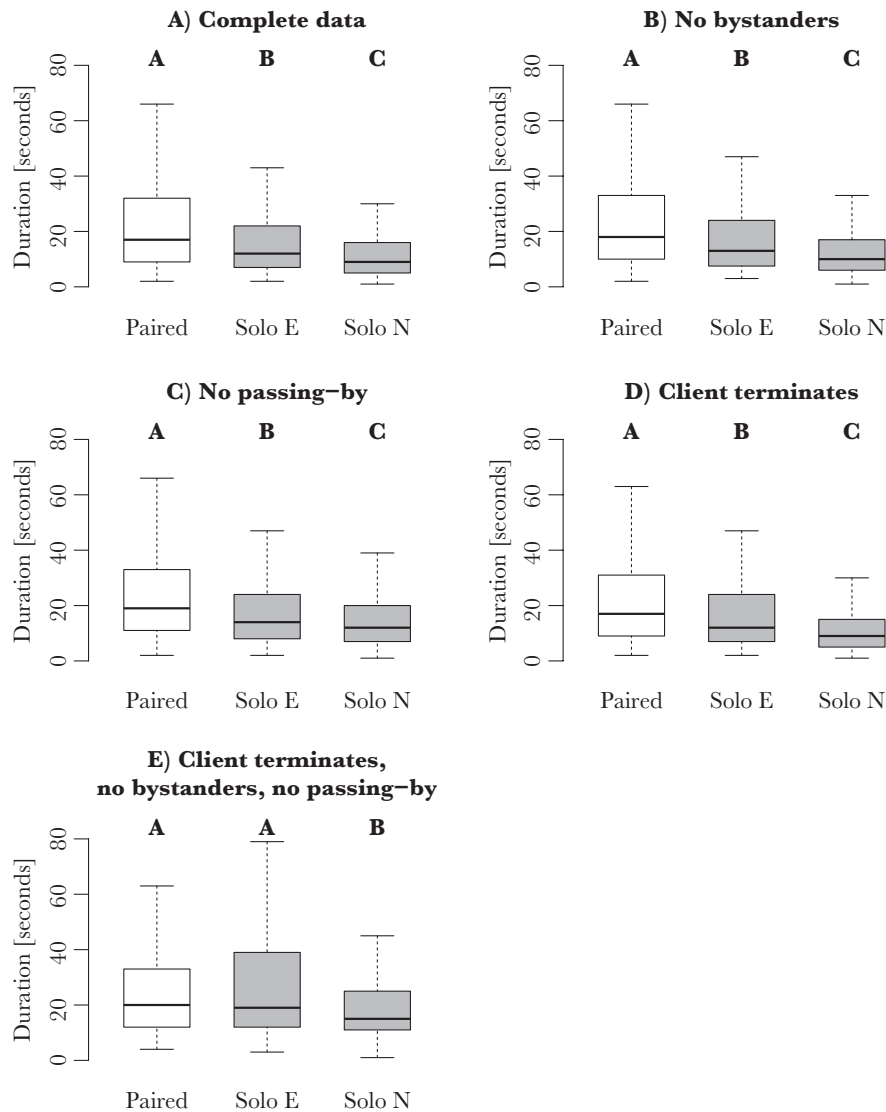
Interference

There was no significant difference between observers ($W = 1837$, P-value = 0.708) and hence data were combined. Given that we are primarily interested in a quantitative interference factor rather than in the question whether nibbling rates differ significantly between sexes and/or situation, we present median values as a basis for the calculation of the interference factor. When alone, females nibbled at clients 0.42 times per second, and males nibbled 0.44 times per second. When cleaning together, nibbling rates dropped to 0.38 for females and 0.40 for males. Thus, pairs appeared to be slightly less than twice as efficient as singletons: 0.78 nibbles per second compared with a mean of 0.43 nibbles per second, that is, 1.81 times more efficient.

Fitting our results into the model

When using only interactions that fitted the model's assumptions (client is initially willing to interact, it terminates the interaction, there are no bystanders, the cleaner is alone because the partner was experimentally removed), clients jolted on average 0.91 times as often with pairs as with singletons. Therefore, the model predicts that pair inspections should on average last 1.1 times longer than singleton interactions, which is close to the observed value of 1.05 times. On the other hand, the model predicts that any interference between coinspecting cleaners will translate linearly into increased cheating rates and hence reduced interaction duration relative to singleton inspections (Bshary et al. 2008 supplements). In our case, data suggest an efficiency loss of 10%, which translates into a predicted 10% reduction in inspection duration due to a correspondingly increased cheating frequency. Thus, we observed a certain mismatch between the predicted jolt ratio pair/singleton of 1.1 and the observed ratio of 0.91.

DISCUSSION

We had asked whether integrating a time component in an IPD-like game yields predictions concerning the link between levels of

A) Complete data

A   **B**   **C**

Duration [seconds]

Paired   Solo E   Solo N

B) No bystanders

A   **B**   **C**

Duration [seconds]

Paired   Solo E   Solo N

C) No passing−by

A   **B**   **C**

Duration [seconds]

Paired   Solo E   Solo N

D) Client terminates

A   **B**   **C**

Duration [seconds]

Paired   Solo E   Solo N

E) Client terminates, no bystanders, no passing−by

A   **A**   **B**

Duration [seconds]

Paired   Solo E   Solo N

Figure 2
Duration of interactions. Box-and-whisker plots for pair inspections (white) and singleton inspections (gray) at 11 pair cleaning stations for the complete data set (A) as well as subsets controlling for confounding variables (B–E). Interactions documented under natural conditions either involved the 2 cleaners (paired) or only 1 cleaner (Solo N: solo natural), whereas singleton interactions in absence of a partner were induced by temporarily removing 1 cleaner (Solo E: solo experimental). Different letters at the top indicate pairwise significant differences.

cooperation and interaction duration that quantitatively match observations on cleaner wrasse–client interactions. We found indeed a very good quantitative match between the model's predictions concerning the comparison between singleton and pair inspections regarding the link between client jolt rate and inspection duration, as long as confounding variables are controlled for.

The test of the model

We could reproduce results from a previous study that had found that client jolt rates did not differ between singleton and pair inspections (Bshary et al. 2008). Separate data on the frequency of cleaners touching clients with their mouths suggest a 10% loss in individual cleaning efficiency during pair inspections, which together with the data on jolt rates implies that cleaners are quite close to doubling their service quality when inspecting as a pair. Note, however, that this conclusion is still preliminary as we cannot exclude that our experimental manipulation somehow affected

the cleaning behavior of singletons and hence their client's jolt rate. Nevertheless, on a qualitative level, our results bring additional strong support to the model's prediction that the dilemma of cleaning in pairs appears to be solved through an increase in cooperation levels. On a quantitative level, cleaners were relatively more cooperative during pair inspections than predicted by the model. Although it is possible that the mismatch is simply noise or due to the effects of our manipulation, there is also a functional explanation. The model does not specify the strategies that lead to cooperative outcomes, but the mathematical analyses assume that both partners adopt the same reduced cheating frequency during pair inspections (Bshary et al. 2008 supplements). This assumption contrasts with the various empirical studies that have shown that the increase in service quality during pair inspections is largely due to females in response to male punishment (Raihani et al. 2010 2011 2012). Thus, future modeling should test whether asymmetric punishment and resulting differences in cheating frequencies may yield
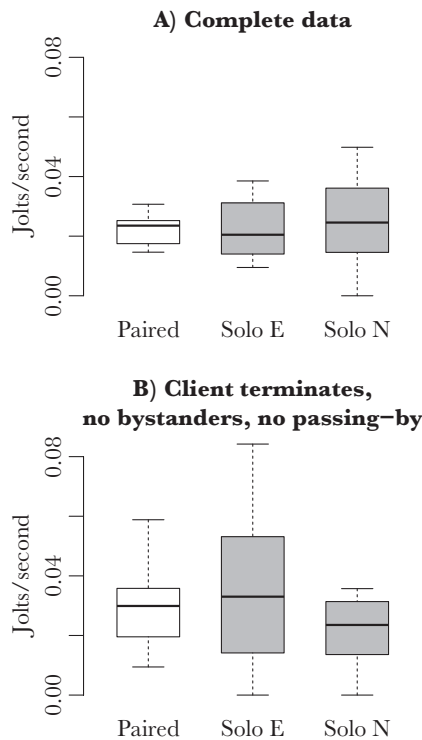
Figure 3

Jolts per seconds (sum of jolts/sum of interaction durations). Box-and-whisker plots for pair inspections (white) and singleton inspections (gray) at 11 pair cleaning stations, for the complete data set (A) as well as a subset controlling for confounding variables (B). Interactions documented under natural conditions either involved the 2 cleaners (paired) or only 1 cleaner (Solo N: solo natural), whereas singleton interactions in absence of a partner were induced by temporarily removing 1 cleaner (Solo E: solo experimental). There were no significant differences between groups.

cooperative outcomes that push client jolt frequency during pair inspections below jolt frequencies during singleton inspections.

With respect to our main question, we note that observed relative interaction durations between singleton and pair inspections are almost perfectly predicted by their relative jolt frequencies. It thus appears that jolt rates alone explain interaction durations well in this study and that cleaner pairs can indeed induce clients to prolong interaction duration by reducing cheating rates, as predicted by the model (Bshary et al. 2008). As a consequence of this quantitative fit, we infer that tactile stimulation has little effect on the duration of interactions that fit the model's parameters. Tactile stimulation was the 1 variable where its impact on interactions could not be predicted. Tactile stimulation reduces cortisol levels in clients (Soares et al. 2011) and, thus, is likely to have a positive effect on client fitness. Although cleaners use tactile stimulation to alter client decisions in regard to interaction duration (Bshary and Würth 2001), we propose that instead of prolonging interactions with cooperating clients, tactile stimulation is most effective in making passing clients slow down for an inspection and hence mainly prolongs interactions under these restricted conditions (Bshary and Würth 2001).

Model assumptions and real cleaning interactions

Our results show that the model only captures the outcome of the subset of data that conforms to the model's assumptions, which in our data set represents one third of the interactions. The inclusion
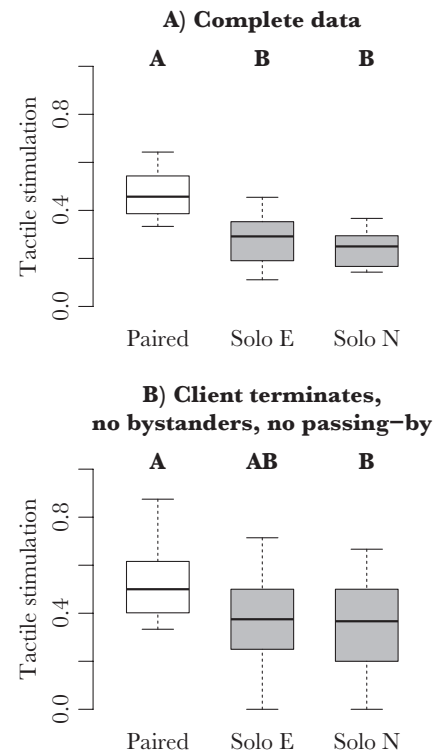

Figure 4

Tactile stimulation (proportion of interactions with tactile stimulation). Box-and-whisker plots for pair inspections (white) and singleton inspections (gray) at 11 pair cleaning stations for the complete data set (A) as well as a subset controlling for confounding variables (B). Interactions documented under natural conditions either involved the 2 cleaners (paired) or only 1 cleaner (Solo N: solo natural), whereas singleton interactions in absence of a partner were induced by temporarily removing 1 cleaner (Solo E: solo experimental). Different letters at the top indicate pairwise significant differences.

of situations in which clients are apparently unwilling to interact, or in which bystanders are present, or in which cleaners terminate the interactions yields a data set that does not fulfill all the model's assumptions and indeed produces the unexpected result that pair inspections last significantly longer than singleton inspections. Thus, to understand the full complexity of the link between cheating rates and interaction duration in cleaning interactions, it would probably be best to develop several models that explore each parameter of interest separately. Another key result is that only experimental introduction of singleton inspection due to removal of the partner yielded data as predicted by the model. It, thus, appears that short interactions at pair stations are mainly performed by a singleton cleaner and that longer interactions are more likely to involve both cleaners. As a result, correlational data on singleton inspections in the presence of a partner are indeed highly flawed and not useful for an evaluation of the model.

Methodological concerns

We acknowledge a few issues that might have had some effect on the data. First, as already mentioned, the data on client jolt rate and interaction duration were collected separately from the data on the rates with which cleaner touch clients with their mouth. Season and/or location might have affected the results in addition to the already discussed possibility that removing a partner may affect this

variable as well. In addition, our data are not properly counterbalanced in time, but the experimental manipulation was invariably conducted after the first control observation and before the remaining 2 control observations, a design that prevents to investigate whether sequence effects might exist. Finally, the observer was not blind to the general predictions of the model. Nevertheless, we note that the key variable of interest in this article—interaction duration—does not allow for large observer biases because the start is clearly defined (first physical touch) while termination may sometimes leave room for 1–2 s variance (when a client swims off while the cleaner tries to follow). Hence, we are confident that the main conclusions are robust.

Prolonged interactions

IPD models are typically built in a way that cooperating and defecting can be described as pressing a C or a D button, which is also the way laboratory experiments with humans have been run in the past and indeed are still run (Fehl et al. 2012; Gracia-Lázaro et al. 2012). However, real-life interactions typically have a measurable duration, which adds an important layer of complexity. The behavior of a partner can be monitored and own behavior fine-tuned in response, as that occurs during predator inspection in fish (Milinski 1987), territorial defense in lions (Heinsohn and Packer 1995), or predator mobbing in birds (Krams et al. 2007). The potential dynamics within an interaction needs to be explored. In particular, one has to determine the payoffs as a function of duration. The model by Bshary et al. (2008) was inspired by the optimal foraging problem in marine cleaning mutualism, and hence, the logic of the marginal value theorem was integrated in the model, which yielded indeed very precise quantitative predictions about relative levels of cooperation and resulting interaction duration during singleton and pair inspections. For cleaning mutualism, our data thus strongly suggest that the marginal value theorem has a strong explanatory power. In other examples, a sigmoid or even step function may be more appropriate. During predator inspection, for example, stopping the approach early may yield few net benefits, whereas further approach may offer exponential increase in net benefits until the increase in danger causes further approach to yield diminishing returns (Noë 2006). In the lion case, not showing up at the border for defense will yield zero benefits. In the case of birds mobbing predators, one has to evaluate how the duration of mobbing affects a predator's subsequent decisions. Apart from the payoff function of single interactions, one also needs to assess how behavior in a current interaction affects the time interval until the next interaction starts. Optimal decisions within the marginal value theorem framework depend on a link between interaction duration and the time interval to the next interaction, as is the case in cleaning mutualisms. This link is less likely to exist in other cases. Nevertheless, we note that general abstract cooperation models are abound (Lehmann and Keller 2006; Nowak 2012). To further advance our understanding of cooperation, biologically informed modeling is essential.

FUNDING

Handling editor: Shinichi Nakagawa

REFERENCES

Archetti M, Scheuring I, Hoffman M, Frederickson ME, Pierce NE, Yu DW. 2011. Economic game theory for mutualism and cooperation. Ecol Lett. 14:1300–1312.

Axelrod R, Hamilton WD. 1981. The evolution of cooperation. Science. 211:1390–1396.

Barclay P, Willer R. 2007. Partner choice creates competitive altruism in humans. Proc R Soc B. 274:749–753.

Barrett L, Henzi SP, Weingrill T, Lycett JE, Hill RA. 2000. Female baboons do not raise the stakes but they give as good as they get. Anim Behav. 59:763–770.

Bates D, Maechler M. 2009. lme4: linear mixed-effects models using S4 classes. Available from: http://cran.r-project.org/web/packages/lme4/index.html

Brosnan SF, Salwiczek L, Bshary R. 2010. The interplay of cognition and cooperation. Phil Trans R Soc Lond B. 365:2699–2710.

Brosnan SF, de Waal FBM. 2002. A proximate perspective on reciprocal altruism. Hum Nat. 13:129–152.

Bshary R, Bergmüller R. 2008. Distinguishing four fundamental approaches to the evolution of helping. J Evol Biol. 21:405–420.

Bshary R, Bronstein JL. 2011. A general scheme to predict partner control mechanisms in pairwise cooperative interactions between unrelated individuals. Ethology. 117:271–283.

Bshary R, Côté IM. 2008. New perspectives on marine cleaning mutualism. In: Magnhagen C, Braithwaite VA, Forsgren E, Kapoor BG, editors. Fish behaviour. Enfield (NH): Science Publishers. p. 563–592.

Bshary R, Grutter AS. 2002. Experimental evidence that partner choice is a driving force in the payoff distribution among cooperators or mutualists: the cleaner fish case. Ecol Lett. 5:130–136.

Bshary R, Grutter AS, Willener AS, Leimar O. 2008. Pairs of cooperating cleaner fish provide better service quality than singletons. Nature. 455:964–966.

Bshary R, Schäffer D. 2002. Choosy reef fish select cleaner fish that provide high-quality service. Anim Behav. 63:557–564.

Bshary R, Würth M. 2001. Cleaner fish *Labroides dimidiatus* manipulate client reef fish by providing tactile stimulation. Proc R Soc Lond B. 268:1495–1501.

Charnov EL. 1976. Optimal foraging, the marginal value theorem. Theor Popul Biol. 9:129–136.

Clutton-Brock T. 2009. Cooperation between non-kin in animal societies. Nature. 462:51–57.

Connor RC. 2010. Cooperation beyond the dyad: on simple models and a complex society. Phil Trans R Soc Lond B. 365:2687–2697.

Fehl K, Sommerfeld RD, Semmann D, Krambeck HJ, Milinski M. 2012. I dare you to punish me—vendettas in games of cooperation. PLoS One. 7:e45093.

Gracia-Lázaro C, Ferrer A, Ruiz G, Tarancón A, Cuesta JA, Sánchez A, Moreno Y. 2012. Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. Proc Natl Acad Sci USA. 109:12922–12926.

Grutter AS, Bshary R. 2003. Cleaner wrasse prefer client mucus: support for partner control mechanisms in cleaning interactions. Proc R Soc Lond B. 270:S242–S244.

Heinsohn R, Packer C. 1995. Complex cooperative strategies in group-territorial African lions. Science. 269:1260–1262.

Hothorn T, Bretz F, Westfall P. 2008. Simultaneous inference in general parametric models. Biom J. 50:346–363.

Johnstone RA, Bshary R. 2002. From parasitism to mutualism: partner control in asymmetric interactions. Ecol Lett. 5:634–639.

Killingback T, Doebeli M, Knowlton N. 1999. Variable investment, the Continuous Prisoner's Dilemma, and the origin of cooperation. Proc R Soc Lond B. 266:1723–1728.

Krams I, Krama T, Igaune K, Mänd R. 2007. Experimental evidence of reciprocal altruism in the pied flycatcher. Behav Ecol Sociobiol. 62:599–605.

Lehmann L, Keller L. 2006. The evolution of cooperation and altruism—a general framework and a classification of models. J Evol Biol. 19:1365–1376.

Leimar O, Hammerstein P. 2010. Cooperation for direct fitness benefits. Phil Trans R Soc Lond B Biol Sci. 365:2619–2626.

Lieske E, Myers RF. 1996. Coral reef fishes: Caribbean, Indian Ocean, and Pacific Ocean : including the Red Sea. Princeton (NJ): Princeton University Press.

Milinski M. 1987. TIT FOR TAT in sticklebacks and the evolution of cooperation. Nature. 325:433–435.

Noë R. 2006. Cooperation experiments: coordination through communication versus acting apart together. Anim Behav. 71:1–18.

Nowak MA. 2012. Evolving cooperation. J Theor Biol. 299:1–8.

Pinto A, Oates J, Grutter A, Bshary R. 2011. Cleaner wrasses *Labroides dimidiatus* are more cooperative in the presence of an audience. Curr Biol. 21:1140–1144.

Pitcher TJ, Green DA, Magurran AE. 1986. Dicing with death: predator inspection behaviour in minnow shoals. J Fish Biol. 28:439–448.

R Core Team. 2013. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: http://www.R-project.org/

Raihani NJ, Bshary R. 2011. Resolving the iterated prisoner's dilemma: theory and reality. J Evol Biol. 24:1628–1639.

Raihani NJ, Grutter AS, Bshary R. 2010. Punishers benefit from third-party punishment in fish. Science. 327:171.

Raihani NJ, Grutter AS, Bshary R. 2012. Female cleaner fish cooperate more with unfamiliar males. Proc R Soc Lond B. 279:2479–2486.

Raihani NJ, Pinto AI, Grutter AS, Wismer S, Bshary R. 2011. Male cleaner wrasses adjust punishment of female partners according to the stakes. Proc R Soc B. 279:365–370.

Roberts G, Renwick JS. 2003. The development of cooperative relationships: an experiment. Proc R Soc Lond B Biol Sci. 270:2279–2283.

Schino G, Aureli F. 2009. Chapter 2 reciprocal altruism in primates: partner choice, cognition, and emotions. In: Advances in the study of behavior. Vol. 39. San Diego (CA): Academic Press. p. 45–69.

Schneeberger K, Dietz M, Taborsky M. 2012. Reciprocal cooperation between unrelated rats depends on cost to donor and benefit to recipient. BMC Evol Biol. 12:41.

Seyfarth RM, Cheney DL. 1984. Grooming, alliances and reciprocal altruism in vervet monkeys. Nature. 308:541–543.

Sherratt TN, Roberts G. 1998. The evolution of generosity and choosiness in cooperative exchanges. J Theor Biol. 193:167–177.

Soares MC, Oliveira RF, Ros AF, Grutter AS, Bshary R. 2011. Tactile stimulation lowers stress in fish. Nat Commun. 2:534.

Stevens JR, Hauser MD. 2004. Why be nice? Psychological constraints on the evolution of cooperation. Trends Cogn Sci. 8:60–65.

St-Pierre A, Larose K, Dubois F. 2009. Long-term social bonds promote cooperation in the iterated Prisoner's Dilemma. Proc R Soc Lond B. 276:4223–4228.

Trivers RL. 1971. The evolution of reciprocal altruism. Q Rev Biol. 46:35–57.

West SA, Griffin AS, Gardner A. 2007. Evolutionary explanations for cooperation. Curr Biol. 17:R661–R672.