Juvenile cleaner fish can socially learn the consequences of cheating

Noa Truskanov, Yasmin Emery & Redouan Bshary

Social learning is often proposed as an important driver of the evolution of human cooperation. In this view, cooperation in other species might be restricted because it mostly relies on individually learned or innate behaviours. Here, we show that juvenile cleaner fish (Labroides dimidiatus) can learn socially about cheating consequences in an experimental paradigm that mimics cleaners' cooperative interactions with client fish. Juvenile cleaners that had observed adults interacting with model clients learned to (1) behave more cooperatively after observing clients fleeing in response to cheating; (2) prefer clients that were tolerant to cheating; but (3) did not copy adults' arbitrary feeding preferences. These results confirm that social learning can play an active role in the development of cooperative strategies in a non-human animal. They further show that negative responses to cheating can potentially shape the reputation of cheated individuals, influencing cooperation dynamics in interaction networks.
Social interactions often involve conflicts of interest between participants\textsuperscript{1–3}, providing opportunities for both cooperation and defection. These interactions frequently occur in the presence of observing individuals, who can gain valuable information about the costs and benefits associated with different strategies, and the nature of potential interaction partners\textsuperscript{4–8}. While social learning is often suggested to be vital to the evolution of large-scale cooperation in human societies\textsuperscript{9–10}, its role in shaping cooperation dynamics in other species is currently unclear\textsuperscript{11,12}. Although cooperation and social learning are widespread in nature\textsuperscript{12–15}, evidence for the use of social learning by non-human animals in cooperative contexts is extremely limited\textsuperscript{16,17}. This, in turn, often leads to the assumption that social learning about cooperative behaviours is a uniquely human feature\textsuperscript{11,18}.

Studies aiming to test the potential influence of social learning on cooperation in other animals, must bear in mind that even in humans, the links between the two are ambiguous and highly contentious\textsuperscript{5,7–9,19–23}. Observation of the social behaviour of others may provide individuals with information on different aspects of cooperative interactions, including the behavioural strategies being employed, their prevalence, associated payoffs and consequences. Social learning can thus promote either more helping, or exploitation, depending on the information being used and the biases that underlie its acquisition (also termed social learning strategies, i.e., what and from whom to learn, and under which circumstances\textsuperscript{24,25}). Identifying the pathways of information transmission is crucial for determining the impact of social learning on cooperative interactions and their evolutionary stability.

Learning socially to adjust cooperative behaviours can be especially beneficial to animals living in complex social environments, involving multiple interaction partners, or different potential behavioural strategies. In such cases, individuals are faced with the challenge of tracking the characteristics of partners and identifying the behavioural strategies that are expected to be beneficial in interactions with them. While individual learning provides a more direct way to assess the consequences of different social strategies, it can be costly, time-consuming and lead to errors in cases in which personal experience is limited\textsuperscript{24–26}. Taking into account social information, gained by observing the interactions of others, can help individuals bypass these problems, and increase the probability that they will discover the optimal social strategies in different interaction contexts.

A model system in which such learning might be particularly beneficial is the mutualistic interactions of the bluestreak cleaner wrasse (\textit{Labroides dimidiatus}) with a great diversity of so-called ‘client’ reef fish. Cleaners offer clients cleaning services in which they remove the clients’ ectoparasites\textsuperscript{27}. However, a conflict of interests arises, as cleaners prefer to feed on clients’ protective mucus, which is costly to the clients and constitutes cheating\textsuperscript{28}. Clients use various control mechanisms (responses that reduce cheaters’ payoffs\textsuperscript{29}) in response to cleaners’ cheating: they can either leave the cleaning station, or terminate the interaction by chasing the cleaners, which functions as punishment\textsuperscript{30}. These responses vary between different client species, and are important for enforcing cooperation, causing the cleaners to eat against their preference\textsuperscript{31}. Given that cleaners interact with a range of clients that differ in abundance, parasite load, mucus quality and responsiveness to cheating\textsuperscript{32–34}, observing conspecifics interact with clients may provide cleaners with useful information about when and with whom to cooperate. Such information may be especially relevant to young cleaners, who tend to spatially overlap with adults, but may have had less opportunities to gain relevant personal experience as their client assembly shifts along their ontogeny\textsuperscript{35,36}.

Here, we experimentally test the ability of juvenile cleaners to learn socially from observing the interactions of adults with unfamiliar model clients. To simulate the key features of natural cleaner–client interactions, we use a well-established experimental paradigm that has already contributed substantially to the study of cooperation (e.g., \textsuperscript{37,38}). We present cleaners with plexiglass plates (model clients) containing both prawn and fish flake items on their surface. As cleaners generally prefer prawn to fish flakes, eating flake items requires them to eat against this preference, and corresponds to behaving cooperatively by eating ectoparasites in the wild. Eating a preferred prawn item, on the other hand, represents eating client mucus, which constitutes cheating in natural interactions (also see\textsuperscript{37,38}). The juveniles can observe interactions in which adult cleaners are presented with model clients and are free to choose which items to consume, but their choices would elicit pre-defined responses (the plates are attached to levers, enabling the experimenter to control their movement). Behaving ‘cooperatively’ by eating against preference causes the plate to remain available for longer, whereas ‘cheating’ by consuming a preferred prawn item elicits an evasive response by the plate, equivalent to clients’ reactions to cheating in the wild and prohibiting further food consumption.

Using this setup, we conduct three experiments in which we test whether young cleaners can learn socially about different aspects of cooperative interactions with clients. We test whether the cleaners: (1) behave more cooperatively by eating more against their preference, following observation of adults interacting with model clients that flee in response to cheating (the consumption of preferred food); (2) prefer to service model clients that were observed responding more favourably to conspecifics’ cheating and (3) copy adults’ arbitrary, non-meaningful, model client preferences. We find that juvenile cleaners can learn socially to adjust their cooperative behaviour and client choice to model clients’ responses, but do not copy adults’ arbitrary preferences. Our results thus show that social learning about the consequences of cheating can shape strategic behaviour in young cleaner fish, and suggest that social learning may play an active role in the maintenance of cooperative interactions in natural model systems.

Results

Socially learning to feed against preference. In the first experiment, we tested whether naïve juveniles would socially learn to eat against preference after observing adults interacting with plates that leave in response to the consumption of preferred food. In the observer treatment, juveniles could observe the adult interacting with the plate, and the plates’ subsequent responses (rapid ‘fleeing’ when prawn is consumed). In the control treatment, no interaction between the adult and the plate could occur, but the plates were rapidly removed after being presented for a similar amount of time (Fig. 1). For all tested individuals, we first measured their baseline preference for prawn, and then tested their feeding choices following observation in a set of 15 consecutive tests, in which plate responses were similar to those of the observation phase. The design thus allowed subjects to combine the social information with personal experience obtained during the tests, a situation that is likely to resemble juvenile cleaners’ opportunities for social learning under natural conditions.

Our results reveal clear evidence for social learning: juvenile cleaners that could observe adults interacting with simulated clients, ate more against their preference than individuals in the control treatment who did not observe the social interaction itself (LM: social observation: estimate ± SE = 0.339 ± 0.114, N = 20, F = 8.796, df = 1, P = 0.009, effect size (Cohen’s \(d\)) = 1.322,
Fig. 1 Experimental setup of experiments 1 and 2. During the demonstration phase (a, b), the cleaners were separated via a clear partition (juvenile depicted on the left side, and as smaller compared to the adult). In each trial, a plate with food facing the juvenile was introduced behind a second clear partition that made it inaccessible for the adult. a In the observer treatments, the clear partition between the demonstrator and the plate was removed, and the demonstrator was allowed to interact with the plate. b In the control treatments, the plate remained confined and no interaction could occur. c During individual testing, an opaque barrier separated the young and adult cleaners. The juveniles were initially confined and could then approach the plates presented. The plates would follow the same response rules they exhibited during training.

Fig. 2). Thus, the potential for individual experience in the tests was not sufficient to override the advantage of being initially exposed to social information. Flake palatability (i.e., the experiment was conducted in two cohorts differing in the palatability of flake items) did not seem to affect the cleaners’ feeding choices (LM: flake palatability: estimate ± SE = 0.12 ± 0.114, N = 20, F = 1.097, d = 1, P = 0.309). Importantly, the results of this experiment indicate that social observation can affect cleaners’ feeding in ways that correspond to behaving more cooperatively under natural conditions.

Socially learning about partner responsiveness. In the second experiment, we tested whether juvenile cleaners can use social learning to assess client responsiveness and prioritise plates that are tolerant to cleaners’ consumption of preferred food over plates that negatively respond to it (after verifying that the subjects could also learn this through individual learning, for further details see Supplementary Note 1 and Supplementary Fig. 2). Each juvenile cleaner experienced four different treatments (treatment order counterbalanced): in the two ‘observer’ treatments, juveniles could observe an adult interacting separately with a tolerant and a responsive plate, differing in colours and patterns. In these interactions, the responsive plate would respond to cleaners’ ‘cheating’ by either leaving swiftly (fleeing) or chasing the cleaner for ~3 s (punishing). In the two control treatments, the adults were blocked from interacting with the plates by a transparent barrier but both observation time and plate movement (gentle leaving/fast fleeing/simulated punishing) were matched to the observer treatments (Fig. 1). At the end of each treatment, the juveniles’ preferences were tested in a set of ten choice tests in which the chosen plate’s responses to the consumption of prawn would match the ones used during the demonstration phase. Thus, like in experiment 1, the juveniles could potentially combine the social information with personal experience gained during this test phase.

We found that social observation influenced juveniles’ model client choice (GLMM: social observation: estimate ± SE = 0.345 ± 0.208, N = 19, χ² = 15.665, d = 1, P < 0.0001, Fig. 3) regardless of plate reaction type (GLMM: effect of plate response type: estimate ± SE = −0.256 ± 0.207, N = 19, χ² = 0.011, d = 1, P = 0.915; plate response type × social observation: estimate ± SE = 0.497 ± 0.297, N = 19, χ² = 2.789, d = 1, P = 0.095; R² = 0.04). Observing adult cleaners interacting with the plates, diverted juveniles’ preferences towards tolerant model clients: only following trials of the ‘observer’ treatments the cleaners significantly preferred the tolerant plates, whereas after experiencing control conditions, their preferences remained at chance level (GLHT comparing the juveniles’ preferences to a no preference null hypothesis: observation of fleeing, z = 2.472, P = 0.027; observation of punishing, z = 3.889, P = 0.0002; control fleeing, z = 0.275, P = 0.783; control punishing, z = −1.378, P = 0.168, Fig. 3). As in experiment 1, the effect of social learning in this experiment remained pronounced despite the potential for individual learning during the tests. Taken together, our results confirm that observer juvenile cleaners can extract information about plates’ reactions that correspond to clients’ partner control mechanisms in nature and use this information to guide their choices.

Observation of adults exhibiting arbitrary preferences. In the third experiment, we tested whether juvenile cleaners would copy whichever behaviours they happen to observe, by exposing them to social information regarding adult cleaners’ arbitrary feeding
Fig. 2 Socially learned adjustment of a cooperative foraging decision. The effects of social observation on the extent to which juvenile cleaners ate against preference in the test phase of experiment 1. Dashed line marks the juveniles’ predicted score based on their initial feeding preferences. Values above and below 0 indicate that the fish ate more and less flame items, respectively, than predicted according to their baseline food preferences. Observer treatment, in which the juveniles had witnessed the plates in response to adults’ consumption of preferred prawn, is marked in grey (N = 10). Control treatment, in which the adults could not interact with the plates, is marked in white (N = 10). Boxplots show the median and interquartile range, whiskers denote 1.5× interquartile range and dots mark individual data points. Source data are provided as a Source Data file.

Discussion

Our results show the ability of a non-human animal to learn socially about behaviour in an experimental paradigm that mimics cooperating/cheating in interspecific social interactions. Observation of adult cleaners’ interactions facilitated the behaviour adjustment of juvenile cleaner fish: first, they socially learned to eat against preference in order to prolong interactions with model clients, an adjustment that would lead to behaving cooperatively in natural interactions with client fish. Second, they socially learned to strategically pick partners that yielded higher payoffs through the absence of evasive actions. Our results further suggest that information about the consequences of interactions was more salient to the juvenile cleaners than information about demonstrator’s choices or feeding behaviour. This exhibition of payoff-based social learning in a social game, indicates that learning from observation can promote cooperative behaviour in self-serving ways.

Learning from observation about cooperation and cheating is often treated as a replication process, in which strategies are being copied by naïve individuals (e.g.,5,7,20). However, as the information encoded in cooperative interactions is multifaceted, social learning can be directed at different aspects of the interaction and elicit responses that do not necessarily match the behaviour being observed. In our experiments, the fish did not copy the demonstrators, but rather extracted information about the negative consequences associated with their behaviour. In experiments 1 and 3, the juveniles did not copy adults’ behaviour or take their preferences into account, even when this was indeed feasible (for further discussion of the potential explanations for this result, see Supplementary Discussion section). In experiment 2, copying preferences. We allowed the juveniles to observe an adult repeatedly choosing between two simultaneously presented plate types that only differed in an arbitrary manner (their colour). The juveniles were divided into two treatments, whose goal was to generate variation in demonstrated preferences: in the ‘preferring demonstrator’ treatment, the demonstrators had been previously trained to clearly prefer one of the plates (plate role counterbalanced between observers). In the ‘indifferent demonstrator’ treatment they had been trained to approach the plates at random. Each observer could see the demonstrator’s choices and the time it spent interacting with chosen plates (Fig. 4a). We then presented the observers with ten choice trials in which both plates were equally rewarding. If adults’ preferences substantially influence the choices of juveniles, we would expect these preferences to be copied. Contrary to this expectation, juveniles’ preferences were not affected by those of the demonstrators and did not differ between the two treatment groups (GLM: demonstrator preference: estimate ± SE = −3.2368 ± 3.019, N = 19, χ² = 0.854, df = 1, P = 0.358; treatment group: estimate ± SE = −0.817 ± 1.056, df = 1, P = 0.438, Fig. 4b). This suggests that although juvenile cleaners can use social information to learn about ecologically relevant cues pertaining to their complex social environment (as shown in our previous experiments), arbitrary effects induced by demonstrator’s choices, or interaction time with a client, are unlikely to significantly affect their behaviour. Another result that points in this direction, is a finding in our first experiment, that demonstrators’ feeding choices (number of flakes consumed) did not seem to affect the observers’ subsequent levels of feeding against preference (LM: demonstrator’s average flake consumption: estimate ± SE = −0.115 ± 0.226, N = 10, F = 0.261, df = 1, P = 0.625. Supplementary Fig. 1). Thus, it appears that the cleaners learned socially about the outcomes of the interaction rather than merely copying the demonstrators’ behaviour.
was simply not possible—the juveniles did not see an adult choosing between a tolerant and an evasive plate, but rather, saw the consequences of interactions with both client types separately, and then had to make a choice based on this information.

In studies on the effects of social learning on cooperation in humans, payoff based social learning caused cooperation to deteriorate22,23. This may be expected in interactions with a prisoner’s dilemma pay-off matrix, where cheating is dominant over cooperating. In our experiments, in which payoff based social learning promoted cooperation, behaving cooperatively towards model clients led to overall higher payoffs. Taken together, these results indicate that the effect of payoff based social learning on cooperation may not be exclusively detrimental. Instead, the payoff matrix underlying interactions is likely to dictate whether observation will lead to cooperation or to cheating. According to this logic, observation of cheating can induce cooperative behaviour if it leads to a more favourable outcome, which can occur when partner control mechanisms offset the benefits associated with cooperative and cheating strategies2,23.

Can social learning about pay-off outcomes influence cooperation in other non-human species? The current lack of evidence for social learning of cooperative strategies11 makes it hard to make concrete predictions. However, as animals learn socially in a variety of contexts, and can extract information from seeing other types of interactions15,39,40 (see Supplementary Discussion for specific examples), we believe that social learning about cooperation is more prevalent than currently recognised. Nevertheless, social learning about effective cooperative or cheating strategies may also be constrained by difficulties associated with tracking payoff related information. Studies of payoff based social learning in non-cooperative contexts, indicate that success in learning about differential payoffs can be variable at both the inter- and intraspecific levels41−44. In experiments with chimpanzees and vervet monkeys, evidence for the use of social information for learning about more favourable outcomes is rather limited42−44, while the evidence for nine-spined sticklebacks is strong41. Whether social learning about cooperation and cheating leads to higher payoffs, and the precise impact that this will have on interaction dynamics, will therefore depend on the content of the information that animals manage to extract from observing others. This can further differ between contexts and individuals45, as indicated by the variation in humans’ use of social learning strategies in social dilemmas21,46,47.

Theoretical models on reputational effects and their influence on cooperation have largely focused on how helping can induce higher cooperation levels48−50. However, negative reactions to cheating can also have reputational effects51. Such a reputation can promote cooperative behaviour in bystanders (as in the case of punishment in humans52,53), but may also bare negative consequences: in humans, bystanders reward helpers more than punishers54,55 and punishers pay to hide their punishment from observers in lab-based economic games56. Here, we find that observation of plates responding negatively to the behaviour of adults affects the subsequent behaviour of juvenile cleaners towards these plates. Observation of plates’ negative reactions incited the cleaners to behave more cooperatively (a response that would benefit real clients), but also to avoid responsive plates when alternative options were available (a response that would negatively affect real clients). It is hence plausible that real clients, as at least equally salient interaction partners, gain reputation based on their responses to cheating and that such reputation affects interaction dynamics in a natural model system. Using partner control mechanisms to respond negatively to cheating can thus act as a double-edged sword, favourably increasing cooperating in eavesdroppers, but reducing the chance of being chosen as interaction partner.

Methods

Subjects and housing. This research was conducted in March−April 2018, at the Lizard Island Research Station, Australia. A total of 40 fish participated in the experiments, 20 juvenile and 20 adult cleaner wrasse, distinguishable by their distinct colour pattern: juveniles are mostly blue with a black stripe, while the blue becomes whiter in adults. There was no overlap between age classes with respect to body length (juvenile cleaners: average = 4.45 ± 0.64 SD cm; adult cleaners: average = 7.725 ± 0.56 SD cm). The cleaners were captured in the reefs surrounding Lizard Island at least 24 days prior to the beginning of the experiments and housed in the lab in separate aquaria with a constant flow of running seawater. Each fish was provided with a polyvinyl chloride tube for shelter. Before the experiments began, the cleaners were fed daily with mashed prawn that was smeared on plexiglass plates of varying colours and sizes. The fish were returned to their original habitat at the end of the experiments.

Three experiments were conducted consecutively, in two cohorts, each consisting of 9−10 pairs of young and adult cleaners (N=20 cleaner pairs in experiment 1, and N=19 in the following experiments due to the death of a juvenile in the first cohort). Prior to the beginning of the experiments, each pair of...
cleaners was placed in a glass aquarium, separated into two compartments by a clear partition. An additional opaque barrier was used to prevent visual contact between cohorts, we included ‘cheating’ in the wild28. In some cases, when the cleaner encountered a total of one prawn item, the experimenter chased the cleaner with the plate, the equivalent of a client punishment.

In each phase, the cleaners were presented with plates containing both food types that did not respond to prawn consumption of the second prawn item, then led to the removal of the plate from the aquarium until the next prawn is consumed. The immediate removal of the plate for 60 s, following which the plate would be reintroduced into the aquarium until the next prawn consumption of preferred prawn. The two phases were pronounced despite the potential for individual learning, and not because of it.

**Experimental setup.** In experiments 1 and 2, plexiglass plates were used as surrogates for clients, mucus was replaced with mashed prawn and ectoparasites with plates. Both food types were used to create discrete food items that were placed on the plates. Flake items were marked with an additional white stripe on their left side, to distinguish them from the plates used earlier, in the pre-experimental training phase. Each plate contained three flake and three prawn items, the location of which was randomly varied between trials. In the treatment group (N = 10), the consumption of a prawn item by the demonstrator led to the immediate removal of the plate from the aquarium. In the control group, the demonstrator prohibited the adult from interacting with the plate (Fig. 1b). The duration of plate presentation (the time the plates stayed in the aquarium) was matched between treatment and control groups: each control individual was paired to an individual from the observer treatment, and in each trial round, the control plate was removed at the same time as the experimental plate had been removed in response to the demonstrator eating a prawn item in the paired trial.

At the end of the demonstration phase, we conducted 15 test trials (5 on the day of demonstration, and 10 during the following day), in which the juveniles were presented with the same plates, and the plates followed the same response rule that was used during demonstrations. In each of these tests, we measured the amount of prawn items that were consumed by the cleaners prior to the consumption of prawn (the consumption of a food item was indicated by cleaners’ attachment of the mouth to the plate at the item’s location). This design enabled the juveniles to both acquire information by observing adults, and gain personal experience by interacting with the plates during the test phase. Prior to the test, we observed that when juveniles’ use of social information likely involves getting direct feedback. It further fits the notion that social learning is a biasing of individual learning by social stimuli rather than a completely distinct process (also see refs. 14, 57–62). Note, however, that in itself, individual learning would not be able to account for differences between observer and control treatments. If nothing were to occur, social learning would be expected to diminish any potential differences between these treatments, as it allows individuals of the control groups to fill any knowledge gaps created by the lack of exposure to social information. Thus, any significant effects of social observation on juvenile performance, would be pronounced despite the potential for individual learning, and not because of it.

At the end of the experiment, we calculated for each fish an eating against preference score. In order to achieve this, we first calculated the cleaner’s predicted feeding score: the number of flake items that it is expected to target prior to the consumption of prawn, on plates containing three flake and three prawn items. We used a derivation of a formula developed by Guguis and Bahary33 for the exact same purpose:

\[
\text{Predicted feeding score} = \frac{3(1 - x)}{2x + 1}
\]

The formula combines the initial feeding preferences of the fish and the probabilities of eating preferred and less preferred items in a set of sequential choices. The feeding preference score measured prior to the beginning of the experiment (proportion of prawn consumed in the tests—see details above) is denoted by x. The probability of eating a flake item is denoted by 1 – x. This probability changes each time a flake item is consumed, as its consumption increases the proportion of prawn in the remaining items. The formula takes into account these changes in probabilities and generates for each fish a predicted feeding score: the number of flake items it is expected to eat prior to the consumption of prawn in a single test session. This score was then subtracted from the average number of flake items that the fish actually consumed in the tests (measured—predicted), thus generating a feeding against preference score. When the feeding against preference score is larger than 0, it means that the fish ate more flake items than predicted, whereas, when the feeding against preference score is lower than 0, the fish performed below the expected value.

**Experimental setup.** In experiments 1 and 2, juveniles (N = 19) participated in two consecutive experimental phases: an individual learning phase and a social learning phase, in which they were repeatedly confronted with the need to choose between plates that differ in their responsiveness to cleaners’ consumption of preferred prawn. The two phases were conducted as paired in the individual learning phase and in the social learning phase (see further details in the description of each experimental phase below). Each of the fish participated in all treatments (a within subject design), and treatment order was counterbalanced between cleaners within each phase. Overall, during this experiment the juvenile cleaners encountered a total of eight different treatments (two in the individual learning phase and four in the social learning phase; see details above).

In the individual learning phase, the cleaners were exposed to presentations of prawn consumption (see Fig. 3). Four food items, two prawn items and two flake items, were placed on equal distant dots in the middle of each plate. Flake concentration was 40% for
the juvenile cleaners, and 20% in adult demonstrations. Food locations were counterbalanced, so that individuals were exposed to a variety of food location combinations throughout trials. One treatment provided juveniles with the plates directly during the training phase while being visually isolated from the adults in the neighbouring compartments (Fig. 1c). Treatment order was counterbalanced between individuals. In each treatment, the cleaners were presented with two plates. In a subset of trials each cleaner was allowed to choose between the two plates. Plates served as the reward and were taken away to eat as much as they wanted (‘tolerant to cheating’), gently allowing the aquarium when the cleaner is done feeding, and a plate that responds in a negative way if and when the cleaner eats a preferred prawn item (‘responsive to cheating’). Plate responses varied between treatments; in the ‘feeding’ treatment, the response placed in the aquarium when the cleaner is done feeding, and a plate that responds in a negative way if and when the cleaner eats a preferred prawn item (‘responsive to cheating’). Plate responses varied between treatments; in the ‘feeding’ treatment, the response

During pre-experimental training, the barrier between the demonstrator and observer compartments was opaque, preventing juvenile cleaners from observing the feeding choices of the adult cleaner. Observers were trained individually to find food on the back of plexiglass plates by swimming behind the plate. The adult cleaners participated in at least 24 training sessions (some of them received more training, depending on their achievements, see further below), involving a choice between the two plate types. In each session, the two plates were placed in the aquarium, visually isolating the two compartments was introduced, and the juvenile cleaners were presented with a set of ten simultaneous choice tests between the two plates. In these tests, one prawn item was located at the back of each plate, making the plates of equal value. During demonstrations, the plate not chosen was removed from the aquarium before the juveniles could feed off it.

 experiment 1, we were interested in whether juvenile cleaners adjusted the extent to which they ate against preference following observation. We tested the juveniles’ feeding adjustment by fitting a linear model (model A, Supplementary Table 1) in which their calculated eating against preference score served as the response variable. Treatment group (observers vs. control) and plate palatability (feeding behaviour was run in two blocks, differing the plate palatability of items) were added as fixed predictors. The formula syntax of this model was the following: eating against preference score ~ treatment group + plate palatability. We checked the model’s diagnostics, normality of residuals and homogeneity of variances visually, by using residual plots and q-qplots, and statistically, by using Shuro-Wilk's tests. In addition, we were also interested in whether the adults' feeding choices in the demonstration phase affected the feeding adjustment of juveniles in the observer treatment (control individuals did not observe the adults consuming food and were therefore not included in this analysis). To that aim, we fitted a linear model (model B, Supplementary Table 1) in which the juveniles' eating against preference score served as the response variable. Demonstrators’ average plate consumption during observation and plate palatability served as predictors. The formula syntax

Experiment 3: observing demonstrators’ arbitrary preferences. In experiment 3, the juvenile cleaners were allowed to observe demonstrations in which the adult would choose between two plates types. The juveniles’ preference for either of the plates was then examined in a set of choice tests. The two plates used in this experiment were monochromatic plates (red and green, size 5 × 8 cm) with food items placed on dots drawn on their back. This ensured that the cleaners would not be able to see the food itself before making a choice, thus forcing them to base their choices on the colour cues rather than just readily approaching the visible food. Placing the food on the back of the plates further allowed us to pre-train the demonstrator to choose between the two plates, rather than the other plate was empty of food. In the ‘non-prefering’ demonstrator treatment, the goal of the training was to cause the demonstrators to approach the plates in a random manner. The adults were thus trained with plates that were equally rewarding, both containing one prawn item.

In some cases, we provided the adults with additional pre-training sessions. In the ‘non-prefering’ treatment, this occurred when demonstrators showed a strong preference towards one of the plates (chose repeatedly the same plate type in three sequential trials). These demonstrators would then receive some trials in which the preferred plate did not offer any food to reduce the preference. In cases in which the demonstrators nevertheless continued to exhibit a bias towards one of the plates, the experimenters would allow the cleaners to explore both plates, rather than immediately removing the non-chosen (and rewarding) plate. In the ‘preference’ treatment, additional trials were added if demonstrators showed weak preferences at the end of the pre-training (non-significant preference in the 24 sessions, according to a binomial test). This was conducted in order to maximise the chance that the adults would indeed show some preferences in the experiments. As in previous experiments, the experiment was divided into two phases: social observation and testing. In the social observation phase, the juveniles were provided with the opportunity to witness numerous demonstrations of plate presentations (14 in the first cohort and 20 in the second cohort). Food was placed on the back of the plates (and thus not visible to the observers), and its allocation was similar to that of the adults’ initial training period: 2 vs. 0 items in the preferring demonstrator treatment (N = 10), and 1 vs. 1 in the non-prefering treatment (N = 9). At the beginning of each trial, the demonstrator was confined to the side adjacent to the observer, and the two plates were placed in the aquarium. The barrier confining the demonstrator was then removed, allowing it to swim towards the plates. Following a choice of one of the plates, the plate not chosen was retracted. The remaining plate was removed from the aquarium only after the interaction ended and the cleaner swam away from it. This could take quite some time, as the cleaners would often provide the plates with tactile stimulation, a ‘massaging’ behaviour that is part of their service to clients in the reef. Note that the demonstrators’ choices in this phase, indicate that our initial training reached its goal, and generated variation in demonstrated preferences as well as clear differences in the demonstrations of the two treatment groups (Wilcoxon rank sum: n = 19, W = 8.5, P = 0.003). Compare the x-axis values of grey and white dots in Fig. 4.

At the end of the demonstration phase, the opaque barrier separating the two compartments was introduced, and the juvenile cleaners were presented with a set of ten simultaneous choice tests between the two plates. In these tests, one prawn item was located at the back of each plate, making the plates of equal value. During demonstrations, the plate not chosen was removed from the aquarium before the juveniles could feed off it. Statistical analysis. Statistical analyses were conducted in the statistical software R version 3.5.1. The results of the different experiments were analysed using linear mixed models, general linear models, and general linear mixed models, where applicable. Continuous predictors were all standardised using the function scale() from the base package in R language. Reported P values were extracted using the function Anova() from the car package. Summary outcomes from all the fitted models are available in Supplementary Table 1. In experiment 1, we were interested in whether juvenile cleaners adjusted the extent to which they ate against preference following observation. We tested the juveniles’ feeding adjustment by fitting a linear model (model A, Supplementary Table 1) in which their calculated eating against preference score served as the response variable. Treatment group (observers vs. control) and plate palatability (feeding behaviour was run in two blocks, differing the plate palatability of items) were added as fixed predictors. The formula syntax of this model was the following: eating against preference score ~ treatment group + plate palatability. We checked the model’s diagnostics, normality of residuals and homogeneity of variances visually, by using residual plots and q-qplots, and statistically, by using Shuro-Wilk's tests. In addition, we were also interested in whether the adults’ feeding choices in the demonstration phase affected the feeding adjustment of juveniles in the observer treatment (control individuals did not observe the adults consuming food and were therefore not included in this analysis). To that aim, we fitted a linear model (model B, Supplementary Table 1) in which the juveniles’ eating against preference score served as the response variable. Demonstrators’ average plate consumption during observation and plate palatability served as predictors. The formula syntax
of this model was the following: eating against preference score – demonstrator’s
flue consumption + flue palatability. Tests of models’ assumptions were similar
to those described above (model A).
In experiment 2, we tested the individual learning and social learning phases
separately (N = 19 in each phase), by fitting two generalised linear mixed models
(GLMM with a binomial distribution, function glmer() from the
lme4 package in R
language). In the individual learning model, the response variable was the cleaners’
binary choices between responsive and tolerant plates in the test phase, plate response
type was fitted as a fixed predictor and cleaner identity as a random factor (see
Supplementary Note 1). The formula syntax of this model was the following: cleaner’s
choice – plate response type + (cleaner identity as a random factor). In the social
learning model (model C, Supplementary Table 1), cleaners’ binary choices between
responsive and tolerant plates was again the response variable. Social observation,
response type and the interaction between them were fitted as fixed predictors, and
cleaner identity was fitted as a random factor. The formula syntax of this model was
cleaner’s choice – social observation * plate response type + (cleaner identity as a random
factor). For both models, we checked the models’ diagnostics using residual plots,
and the normality of residuals using random factors using qqplots and
Shapiro–Wilk tests. R squared for the mixed models was obtained using the function
r.squaredGLMM(), from the
lme4 package in R
language. In addition, we measured the
performance of the fish in each treatment, by testing whether their preferences
deviated from those expected by random choice (i.e., whether the proportion of
tolerant plate choice in each treatment differed from 0.5). This was done using general
linear hypotheses tests (GLHT) for the estimates of the mixed models, and P values
were adjusted using the Holm method.
In experiment 3, we tested whether treatment group and/or demonstrators’
exhibited preferences affected the subsequent plate preferences of the observer
juvenile cleaners. We fitted a generalised linear model (GLM) with a quasibinomial
distribution; a distribution that controlled for the overdispersion of the residuals
(by using an additional scale parameter). In this model, observers’ choices of
demonstrators’ preferred and less preferred plates in the tests was the response
variable, while demonstrators’ preferences (proportion of choices of the more
preference of each demonstrator during the demonstration phase) and
treatment group were the predictors. The formula syntax of this model was the
following: cleaner’s choices of demonstrator’s preferred vs. less preferred plates ±
treatment group + demonstrator’s preference. We checked the models’ diagnostics
(homogeneity of variance, residual normality and potential violations of linearity)
using residual plots.

Ethical note. This study was performed in accordance with the guideline of the
Animal ethics committee, Queensland, Australia (approval number: CA 2018/01/
1156).

Reporting summary. Further information on research design is available in
the Nature Research Reporting Summary linked to this article.

Data availability
The data that support the findings of this study are deposited in Figshare data repository:
https://doi.org/10.6084/m9.figshare.8068280.v1. Source data for Figs 2–4 and
Supplementary Figs 1 and 2 are provided in a Source Data file.

Received: 17 June 2019; Accepted: 29 January 2020;
Published online: 03 March 2020

References
1. Axelrod, R. & Hamilton, W. D. The evolution of cooperation. Science 211,
1390–1396 (1981).
2. Bhary, B. & Bergmueller, R. Distinguishing four fundamental approaches to
the evolution of helping. J. Evolut. Biol. 21, 405–420 (2008).
3. Leimar, O. & Hammerstein, P. Cooperation for direct fitness benefits. Philos.
Trans. R. Soc. B 365, 287–296 (2010).
4. Pollock, G. & Dugatkin, L. A. Reciprocity and the emergence of reputation. J.
Theor. Biol. 159, 25–37 (1992).
5. Lehmann, L., Foster, K. R., Borenstein, E. & Feldman, M. W. Social and
individual learning of helping in humans and other species. Trends Ecol.
Evol. 23, 664–671 (2008).
6. Easton, R. L. Social eavesdropping and the evolution of conditional cooperation
and cheating strategies. Philos. Trans. R. Soc. B 365, 2675–2686 (2010).
7. Boyd R, Richerson P. J. Culture and the Evolutionary Process. (University of
Chicago Press, 1985).
8. Henrich, J. Cultural group selection, coevolutionary processes and large-scale
cooperation. J. Econ. Behav. Organ. 53, 3–35 (2004).
9. Boyd R, Richerson P. J. Culture and the evolution of human cooperation. Philos.
Trans. R. Soc. B 364, 3281–3288 (2009).
10. Richerson P., et al. Cultural group selection plays an essential role in
explaining human cooperation: a sketch of the evidence. Behav. Brain Sci. 39,
520 (2016).
11. Melis, A. P. & Semmann, D. How is human cooperation different? Philos.
Trans. R. Soc. B 365, 2663–2674 (2010).
12. Dugatkin L. A. Cooperation among Animals: an Evolutionary Perspective.
(Oxford University Press, 1997).
13. Hoppitt, W. & Laland, K. N. Social processes influencing learning in animals:
A survey of the evidence. Adv. Stud. Behav. 38, 105–165 (2008).
14. Cialdieri, N. et al. Selective and contiguous prosocial resource donation
in capuchin monkeys, chimpanzees and humans. Sci. Rep. 5, 7631 (2015).
15. Lehmann, L., Foster, K. R., Borenstein, E. & Feldman, M. W. Social and
individual learning of helping in humans and other species. Trends Ecol.
Evol. 14, 49–53 (1999).
16. Boyd, R. & Richerson, P. J. Social evolution of helping. Evolutionary
perspectives on the coevolution of human and animal cooperation. Proc. Natl
Acad. Sci. USA 112, 2912–2917 (2015).
17. Lamba, S. Social learning in cooperative dilemmas. Proc. R. Soc. B 281,
20140147 (2014).
18. van den Berg, P., Molleman, L. & Weissing, F. J. Focus on the success of others
leads to selfish behavior. Proc. Natl Acad. Sci. USA 112, 2912–2917 (2015).
19. Burton-Chellew, M. N., En Mouden, C. & West, S. A. Social learning and the
demise of costly cooperation in humans. Proc. R. Soc. B 284, 20170067 (2017).
20. Laland, K. N. Social learning strategies. Learn. Behav. 32, 4–14 (2004).
21. Kendal, R. L. et al. Social learning strategies: bridge-building between fields. Trends
Cogn. Sci. 22, 651–665 (2018).
22. Kendal, R. L., Coolen, I., van Bergen, Y. & Laland, K. N. Trade-offs in the
adaptive use of social and associational learning. Adv. Stud. Behav. 35, 333–379
(2005).
23. Cote, I. M. Evolution and ecology of cleaning symbioses in the sea. Oceanogr.
Mar. Biol. 38, 311–355 (2000).
24. Grutter, A. S. & Bhary, R. Cleaner wrasse prefer client mucus: support for
partner control mechanisms in cleaning interactions. Proc. R. Soc. B 270,
S242–S244 (2003).
25. Bhary, B. & Bronstein, J. L. A general scheme to predict partner control
mechanisms in pairwise cooperative interactions between unrelated individuals.
Ethology 117, 271–283 (2011).
26. Bhary, B. & Grutter, A. S. Asymmetric cheating opportunities and partner
control in a cleaner fish mutualism. Anim. Behav. 63, 547–555 (2002).
27. Bhary, R. & Grutter, A. S. Punishment and partner switching cause
cooperative behaviour in a cleaner fish mutualism. Biol. Lett. 1, 396–399 (2005).
28. Grutter, A. S. Relationship between cleaning rates and ectoparasite loads in
coral reef fishes. Mar. Ecol. Prog. Ser. 118, 51–58 (1995).
29. Arnaud-C, Cote, I. M. & Morand, S. Why clean and be cleaned? The
importance of client ectoparasites and mucus in a marine cleaning symbiosis.
Behav. Ecol. Sociobiol. 51, 1–7 (2001).
30. Bhary, R. The cleaner fish market. (eds Noël, R., van Hooff, J. A. R. M. &
Hammerstein, P.) In: Economics in Nature (Cambridge University Press,
2001).
31. Adam, T. C. & Horii, S. S. Patterns of resource-use and competition for
mutualistic partners between two species of obligate cleaner fish. Coral Reefs
31, 1149–1154 (2012).
32. Triki, Z. et al. Biological market effects predict cleaner fish strategic
sophistication. Behav. Ecol. 30, 1548–1557 (2019).
33. Bhary, R. & Grutter, A. S. Image scoring and cooperation in a cleaner fish
mutualism. Nature 411, 975–978 (2006).
34. Raihani, N. J., Grutter, A. S. & Bhary, R. Punishers benefit from third-party
punishment in fish. Science 327, 171–171 (2010).
35. Danchin, E., Giradille, L. A., Valone, T. J. & Wagner, R. H. Public information
from nosy neighbors to cultural evolution. Science 305, 487–491 (2004).
36. McGregor P. K. Animal Communication Networks. (Cambridge University
Press, 2005).
37. Kendal, R. L., Rendell, L. & Laland, K. N. Nine-spined stickleback
display a hill-climbing social learning strategy. Behav. Ecol. 20, 238–244 (2009).
38. Vale, G. L. et al. Testing differential use of payoff-biased social learning
strategies in children and chimpanzees. Proc. R. Soc. B 284, 20171573 (2017).
39. van Leeuwen, E. J. C. & Call, J. Conservatism and “copy-if-better” in
chimpanzees (Pan troglodytes). Anim. Cogn. 20, 575–579 (2017).
44. Bono, A. E. et al. Payoff-and sex-biased social learning interact in a wild primate population. *Curr. Biol.* 28, 2800–2805. e2804 (2018).

45. Mesoudi, A., Chang, L., Dall, S. R. X. & Thornton, A. The evolution of individual and cultural variation in social learning. *Trends Ecol. Evol.* 31, 215–225 (2016).

46. Molleman, L., Van den Berg, P. & Weissing, F. J. Consistent individual differences in human social learning strategies. *Nat. Commun.* 5, 3570 (2014).

47. Molleman, L. & Gächter, S. Societal background influences social learning in cooperative decision making. *Evol. Hum. Behav.* 39, 547–555 (2018).

48. Kandori, M. Social norms and community enforcement. *Rev. Econ. Stud.* 59, 63–80 (1992).

49. Lotem, A., Fishman, M. A. & Stone, L. From reciprocity to unconditional altruism through signalling benefits. *Proc. R. Soc. Lond. Ser. B* 270, 199–205 (2003).

50. Ohtsuki, H. & Iwasa, Y. Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation. *J. Theor. Biol.* 244, 518–531 (2007).

51. dos Santos, M., Rankin, D. J. & Wedekind, C. The evolution of punishment through reputation. *Proc. R. Soc. B* 278, 371–377 (2011).

52. Barclay, P. Reputational benefits for altruistic punishment. *Evol. Hum. Behav.* 27, 325–344 (2006).

53. dos Santos, M., Rankin, D. J. & Wedekind, C. Human cooperation based on punishment reputation. *Evolution* 67, 2446–2450 (2013).

54. Raihani, N. J. & Bshary, R. Third-party punishers are rewarded, but third-party helpers even more so. *Evolution* 69, 993–1003 (2015).

55. Raihani, N. J. & Bshary, R. The reputation of punishers. *Trends Ecol. Evol.* 30, 98–103 (2015).

56. Rockenbach, B. & Milinski, M. To qualify as a social partner, humans hide severe punishment, although their observed cooperativeness is decisive. *Proc. Natl Acad. Sci. USA* 108, 18307–18312 (2011).

57. Heyes, C. What’s social about social learning? *J. Comp. Psychol.* 126, 193–202 (2012).

58. Dawson, E. H., Avargues-Weber, A., Chittka, L. & Leadbeater, E. Learning by observation emerges from simple associations in an insect model. *Curr. Biol.* 23, 727–730 (2013).

59. Truskanov, N. & Lotem, A. The importance of active search for effective social learning: an experimental test in young passerines. *Anim. Behav.* 108, 165–173 (2015).

60. Leadbeater, E. What evolves in the evolution of social learning? *J. Zool.* 295, 4–11 (2015).

61. Truskanov, N. & Lotem, A. Trial-and-error copying of demonstrated actions reveals how fledglings learn to ‘imitate’ their mothers. *Proc. R. Soc. B* 284, 20162744 (2017).

62. Truskanov, N. & Prat, Y. Cultural transmission in an ever-changing world: trial-and-error copying may be more robust than precise imitation. *Philos. Trans. R. Soc. B* 373, 20170050 (2018).

63. Gingins, S. & Bshary, R. The cleaner wrasse outperforms other labrids in ecologically relevant contexts, but not in spatial discrimination. *Anim. Behav.* 115, 145–155 (2016).

64. Bshary, R. & Wurth, M. Cleaner fish *Labroides dimidiatus* manipulate client reef fish by providing tactile stimulation. *Proc. R. Soc. B* 268, 1495–1501 (2001).

65. R Core Team. *R*: A Language and Environment for Statistical Computing (2016).

66. Bates, D., Machler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48 (2015).

67. Truskanov, N., Emery, Y., Bshary, R. Juvenile cleaner fish can socially learn the consequences of cheating. https://doi.org/10.6084/m9.figshare.8068280.v1 (2020).

Acknowledgements
We thank the friendly staff of Lizard Island Research Station for their help and support in the field; R. Slobodeanu for his help with statistical analysis; M. Burton-Chellew, O. Kolodny, A. Moran, and Y. Prat for helpful comments and discussion. This work was funded by SNF grant 310030B_173334 / 1 awarded to R.B., and a Swiss Government Excellence Scholarship (FCS) awarded to N.T.

Author contributions
N.T. and R.B. designed the study. N.T. and Y.E. carried out data collection. N.T. analysed the data and led the writing. All authors discussed the results and commented on the paper at all stages.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-14712-3.

Correspondence and requests for materials should be addressed to N.T.

Peer review information Nature Communications thanks Culum Brown, Rachel Kendal and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020