

## Do cleaner fish learn to feed against their preference in a reverse reward contingency task?

Evin Danisman · Redouan Bshary · Ralph Bergmüller

**Abstract** The ability to control impulsive behaviour has been studied in animals with a standard test in which subjects need to choose the smaller of two food items in order to receive the larger one (reverse reward contingency). As a variety of mammals that have been tested so far (mostly primates) have great difficulties to solve the task, it has been proposed that it is generally cognitively demanding. However, according to an ecological approach to cognition, a species' ability to solve the task should not depend on its general cognitive abilities but on whether its ecology causes selective pressure on the ability to restrain foraging behaviour. We tested this hypothesis using the cleaner wrasse (*Labroides dimidiatus*), a fish species that feeds against its preference in nature when engaging in cleaning interactions with so called 'client fish'. None of the eight tested individuals learned to choose a non-preferred item after 200 trials. In a subsequent test, one subject learned to respond correctly in a large or none contingency task (only the choice of the small food was rewarded). After a short re-experience treatment, this individual learned to solve the reverse reward task after 30 trials. In conclusion, we did not find support for the general idea that interactions with clients prepared cleaners to quickly solve a reverse reward test. However, the results suggest that the potential to solve a reverse reward contingency may not be restricted to mammals but could be present also in a fish species in which the problem of choosing a non-preferred food over a preferred one is an ever present challenge in nature.

**Keywords** Reverse reward contingency · Impulse control · Ecological approach to cognition · Cleaner fish · Animal cooperation · Learning

### Introduction

The ability to control impulsive behaviour in order to acquire an overall long term benefit has received a large amount of interest in studies on human and non-human animal behaviour (Mischel et al. 1989; Genty et al. 2004; Abeyesinghe et al. 2005; Murray et al. 2005; Genty and Roeder 2006; Wittmann and Paulus 2008). Impulse control may be advantageous in order to withhold an action that would have deleterious consequences in the future. Individuals may profit from impulse control behaviour by avoiding unbeneficial temporal discounting (Stevens and Hauser 2004; Abeyesinghe et al. 2005; Stevens et al. 2005; Wittmann and Paulus 2008) or when feeding against a preference (Genty and Roeder 2006; Wittmann and Paulus 2008). The ability to control impulsive behaviour should be particularly important in cooperative interactions, i.e. when investments in others yield a net gain in the future compared to not investing (Stevens et al. 2005; Bergmüller et al. 2007a; West et al. 2007; Bshary and Bergmüller 2008).

Studies on impulse control in humans have aimed at investigating the age at which children are able to choose a delayed large reward over an immediate small reward (Russell et al. 1991). This ability has been suggested to be critical for the development of an awareness of self and has hence been linked to self-control in humans. As such self control appears to be crucial for the quality of human interactions, many non-human species that have been studied in this context have been primates, possibly reflecting an anthropocentric approach that seeks to understand human cognitive abilities

E. Danisman · R. Bshary · R. Bergmüller (✉)  
Department of Biology, University of Neuchâtel,  
Emile-Argand 11, Case Postale 158, 2009 Neuchâtel, Switzerland  
e-mail: ralph.bergmueller@unine.ch

by studying closely related species (Bshary et al. 2002; Bshary et al. 2007). In contrast, the ecological approach to cognition proposes that we should find certain cognitive capabilities including impulse control in those species that need to solve similar problems in their natural environment (Shettleworth 1994; Kamil 1998; Bshary et al. 2002).

A paradigm that has been widely used to study impulse control across species has been the so called ‘reverse reward contingency task’. In this task, a subject is typically shown two different quantities of food. When the subject reaches or points towards one quantity, the non-selected quantity is given as the reward. Many species such as chimpanzees, *Pan troglodytes* (Boysen and Berntson 1995), Japanese macaques, *Macaca fuscata* (Silberberg and Fujita 1996), squirrel monkeys, *Saimiri sciureus* (Anderson et al. 2000), cottontop tamarins, *Saguinus oedipus* (Kralik et al. 2002), rhesus macaques, *Macaca mulatta* (Murray et al. 2005) and two species of prosimians, black and brown lemurs, *Eulemur fulvus* and *E. macaco* (Genty et al. 2004) showed an initial preference for the larger quantity of food. Such discrimination between two quantities of food could also be shown in red-backed salamanders (*Plethodon cinereus*) (Uller et al. 2003). However, these species did not succeed in spontaneously selecting the smaller reward in order to receive the larger one. In a few other species, subjects managed to maximise their reward by choosing the smaller reward. Orang-utans, *Pongo pygmaeus* (Shumaker et al. 2001) did so, but without showing an initial preference for the large reward. Rhesus macaques learned to consistently select the smaller reward but only after 180–2,800 trials, while in sea lions (*Zalophus californianus*) 3 of 4 individuals succeeded just after 80–150 trials (Genty and Roeder 2006). Mangabeys (*Cercocebus torquatus lunulatus*) mastered the test without procedural modifications (e.g. such as the large-or-none contingency, see below) and were able to transfer this ability to novel pairs of arrays with different quantities of food (Albiach-Serrano et al. 2007).

Given that other complex cognitive abilities including numerical skills, symbolic representation, and other higher-order information processing capacities had already been shown in chimpanzees, failure in the reverse reward test was especially surprising in the initial study in chimpanzees (Boysen and Berntson 1995). As a result, many subsequent studies focused on developing procedures that could overcome the problem of failure in this task. In some species that failed to choose the non-preferred reward in order to obtain the preferred one, cues such as Arabic numerals for chimpanzees and colour cues for cottontop tamarins associated with the two quantities of food enabled the species to succeed (Boysen and Berntson 1995; Kralik et al. 2002). Differences in performance when testing the subjects with food only or with Arabic numerals have been interpreted to result from a conflict between an associative disposition to select the smaller reward in order to obtain the large one

and the intrinsic features of the food reward, which could be removed by using symbols (Boysen et al. 1996; Boysen et al. 1999). Also, in a different procedure, the so called large-or-none contingency facilitated subsequent success in the reverse reward contingency. In this procedure, individuals receive the large reward when choosing the small reward, but do not receive any reward when choosing the large reward. Several species succeeded in choosing the small reward in this procedure and continued to do so, even when the reverse reward procedure was reapplied (Silberberg and Fujita 1996; Anderson et al. 2000; Genty et al. 2004). The large-or-non procedure appears to facilitate choosing against a preference because it increases the costs of choosing the large reward which is not rewarded.

Here we investigated the ecological approach to cognition by testing a fish species that is phylogenetically distant from humans, but that has been selected for to feed against its preference in nature. We predict this should translate into the ability to succeed in a reverse reward contingency task. Cleaner wrasses (*Labroides dimidiatus*) feed on parasites but prefer mucus which they obtain by biting so called client fish (Grutter and Bshary 2003). This results in a conflict of interests between cleaners and clients: clients visit cleaners at their stations in order to get their parasites removed, but cleaners prefer the mucus they obtain by biting cleaners. As clients punish non-cooperative (i.e. biting) cleaners by leaving the cleaning station or attacking (punishing) the cleaner (Bshary and Grutter 2002a; Bshary and Schäffer 2002), cleaners adjust their behaviour and feed on the parasites they obtain from clients (Grutter 1996; Bshary and Grutter 2002b). Several laboratory experiments show that cleaners can easily learn to feed against their preference if this allows them to continue with foraging (Bshary and Grutter 2005; Bshary and Grutter 2006). This foraging behaviour results in a cooperative interaction under natural conditions. The cleaners’ ability to feed against their preference suggests that they are capable of a specific form of impulse control. This situation provides the ideal opportunity to test the cleaners for their ability to choose against their preference in a reverse reward contingency task. The ecological approach leads us to predict that cleaner wrasses should be able to choose against their preference in order to get the preferred resource, while the prediction of an anthropocentric approach would be that cleaners will fail the task.

## Methods

The study was carried out in the lab at the University of Neuchâtel, Switzerland. We used eight wild caught *Labroides dimidiatus* that originated from the Philippines (Island Luzon near Legaspi) and were directly imported to Switzerland in February 2006. The fish were kept individually in

aquaria with a size of  $100 \times 50 \times 40$  cm which were half filled with water. Water conditions were kept similar across all tanks as all aquaria were connected via a flow through system that pumped water from a large cleaning tank (a  $160 \times 80 \times 60$  cm tank containing pieces of hard corals (Scleractinia) served as natural filter) into the single tanks. The water was pumped into the experimental tanks and flowed passively back to the cleaning tank. With this system the nitrite concentration was kept to a minimum (always below 0.3 mg/l). Each tank contained an air supply and a commercial aquarium heater (Eheim, Jäger 125 W for 200 l tanks). Water salinity was kept at a specific gravity of  $1.025 \pm 0.005$  at a temperature of  $25 \pm 1^\circ\text{C}$ . The pH ranged between 8.1 and 8.4. Partial water changes were performed monthly with a commercial marine salt water mixture (Aquamedic). Several small polyvinyl pipes (10–15 cm long and 2.5 cm diameter) served as shelter for the fish. The subjects were fed daily while conducting the experiments (mashed shrimp and fish flakes) and after the experiments had been completed in the afternoon (mashed shrimps, krill, mysis, artemia or commercial fish tropical flakes (Tetramin)).

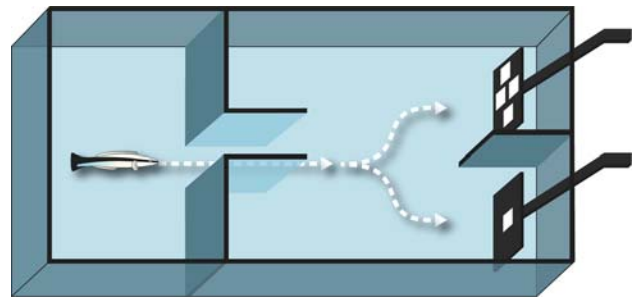
## Experiments

The experiments were carried out in May–August 2006 and were conducted in the aquaria where the fish were kept. Before the start of the experiments, the fish were trained to feed from a Plexiglas plate as a substitute for the client fish that serve as food source in nature. Each aquarium was divided into two compartments which were separated by an opaque partition. The partition could be closed or opened by pulling up the partition (Fig. 1.). All eight subjects were tested individually from 8 a.m. till 5 p.m. Trials were performed 5 times in the morning and 5 times in the afternoon.

### Reverse reward contingency test

In a reverse reward contingency task individuals are given a choice between two rewards, one of them less preferred than the other. As individuals obtain the reward they have not chosen they need to choose against their preference in order to succeed in the task. Before start of the experiments the fish were acclimatised to change between the two compartments and to feed from PVC plates ( $7 \times 10$  cm) with mashed shrimp. These plates were used in the experiment in order to present choices between two quantities or qualities of food.

As we did not know at the onset of the experiments to which differences among the plates or among food types the fish would react best by distinguishing between two types of reward, we performed four different procedures, each one with two subjects (Table 1). In three of the four procedures



**Fig. 1** The experimental aquarium (view from above) was divided into a back compartment (*left*) and the test compartment (*right*) with help of an opaque PVC partition. Two plates with different quantities (indicated by the number of *white patches*) or qualities of food were introduced into the test compartment. The two Plexiglas plates were visually separated from each other with help of a dividing white partition wall in-between the plates. The tested fish could move into the test compartment through a slot followed by a corridor when an opaque partition in front of the corridor (not displayed) at the side of the back compartment was pulled up

the individuals were given the choice between a small reward, i.e. one item with 1 mg of mashed shrimp (the non-preferred reward) or four item patches of one 1 mg of shrimp each (the preferred reward). Mashed prawn is very adhesive and can easily be attached to a Plexiglas plate. Each food item was attached on a separate white coloured patch on the plate (i.e. plates with four rewards items had four white patches, size =  $3 \times 3$  mm) so the fish could distinguish between the possible choices when entering the test compartment. In a fourth procedure the fish were given the choice between 1 mg of mashed shrimps or 1 mg of dried food flakes mixed with water. Before each trial, the subjects were gently moved into the back compartment and the partition was closed (Fig. 1). This way, the subjects could not observe the preparation for the experiment in the choice compartment. Two plates with different quantities or qualities of food were introduced into the test compartment. The two Plexiglas plates were placed on the left and right of the dividing white partition wall. A trial started when the opaque partition was pulled up so the fish could move into the test compartment through a slot. In order to make sure that the fish was swimming in the middle of the aquarium while choosing a plate, the subjects needed to swim through a corridor that was created with help of two Plexiglas partitions (length 10 cm), before they made a choice. The fish chose one of both rewards when it reached the dividing white partition between both plates (i.e. it entered the compartment created by the white partition and the aquarium wall). In the same moment, the plates were exchanged (or the first plate was pulled out to release a second plate behind, see Table 1) in order to reward them against their choice (reverse reward).

The plates were positioned at the left or right side of the aquarium and the position of each plate was randomised

**Table 1** Summary of the four different experimental procedures with two individuals tested per procedure

Procedure	Reward	Plate size	Plates	Individuals
1	1 or 4 items	Equal	First withdrawn	Mi, El
2	1 or 4 items	Equal	Exchanged	Ma, Am
3	1 or 4 items	Small versus large	Exchanged	Bo, Wa
4	Shrimps or flakes	Equal	Exchanged	Ci, Ch

In procedures 1–3 the preferred reward was determined by the number of items on two plates while in procedure 4, shrimps were preferred over flakes. The plate sizes could be equal or different. In order to reward the test subject contrary to what it had chosen, the first plate was withdrawn so the fish had access to the second one in procedure 1. In procedures 2–4, both plates were exchanged so the subjects received the reverse reward

(one plate was placed for a maximum of three successive trials at the same side in order to minimise the risk that the fish would develop a side preference). One session consisting of 10 trials was performed per day. In total, 20 sessions were completed for each individual.

### Large-or-none contingency test

All subjects that did not learn to feed against their preference in the reverse reward test were subject to a large-or-none contingency test (Silberberg and Fujita 1996). The individuals were subjected to a choice between a plate with the preferred reward (large reward or prawn) and another plate with a non preferred reward (small reward or flakes). The individuals received the preferred reward when choosing the non-preferred one or no reward when choosing the preferred reward. This procedure facilitates learning to choose against a preference, because choosing the preferred reward is not reinforced. All other aspects of the procedure were performed in the same way as described above in the reverse reward task. One session consisting of ten trials was conducted once per day with each subject for up to a maximum of 30 sessions or until the individual had learned (i.e. three successive sessions in which the individual was choosing in at least eight of ten trials the plate without reward). Binomial tests were carried out for the first 50 trials and the last 50 trials in order to determine changes in choice behaviour due to the contingency of the task.

### Return to the reverse-reward contingency test

Individuals who were successful in the large-or-none contingency test, were again subjected to a reverse reward contingency task. Before, however, they were first subjected to a re-experience procedure. This was done so the subjects would not simply continue to avoid the preferred item after the large-or-non contingency task. The re-experience was given by presenting the two rewards (the preferred and the non-preferred) on two plates simultaneously so the subjects could experience again the two possible rewards. During the re-experience procedure the fish could feed on both

plates without any interference for 10 times in total. After the re-experience, the fish were once again subjected to the reverse reward contingency task for ten sessions with ten trials per session.

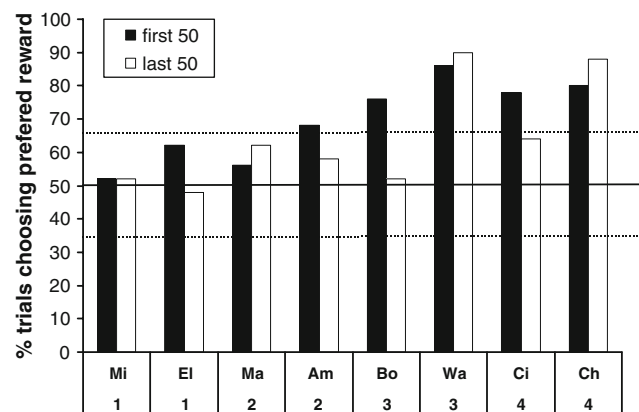
### Statistical analysis

All tests were performed with SPSS version 14.0 and all results reported are two-tailed.

## Results

### Reverse reward contingency

Three out of six individuals showed a significant preference for the plate with the large reward during the first 50 trials (Fig. 2). In detail, when the first of two plates was withdrawn, Mi choose the larger reward in 52% of times (Binomial test,  $x = 26$ ,  $P > 0.05$ ) and El in 62% of times ( $x = 31$ ,  $P > 0.05$ ). When two equal sized plates were exchanged, Ma chose the larger reward in 56% of times ( $x = 28$ ,



**Fig. 2** Percentage of trials in which individuals chose the preferred reward during the first 50 trials (*filled bars*) and during the last 50 trials (*open bars*) in the reverse-reward contingency test. The 50% line corresponds to random choice. Dotted lines at 66% ( $x = 33$ ) and 34% ( $x = 17$ ) indicate threshold values for significant deviation from random. Letters are individual codes for the test subjects. The numbers 1–4 indicate the experimental situation (for details see Table 1)

**Table 2** Initial food choice after different periods of testing

	Mi	El	Ma	Am	Bo	Wa	Ci	Ch
<i>N</i> trials								
50	26	31	28	34*	38**	43**	39**	40**
100	45	64**	55	60°	72**	90**	67**	73**
	Four over one item of prawn				Prawn over flakes			

The table shows the number of times individuals chose a plate with either four items of prawn instead of one, or prawn instead of flakes after 50 and 100 trials. Significance levels for food preferences are depicted as ° $P < 0.1$ , \* $P < 0.05$  or \*\* $P < 0.01$  (binomial tests)

$P > 0.05$ ) and Am in 68% of times ( $x = 34$ ,  $P < 0.001$ ). When two different sized plates were exchanged (the larger plate contained 4 items of shrimp), Bo chose the large plate in 76% ( $x = 38$ ,  $P < 0.001$ ) and Wa in 86% of trials ( $x = 43$ ,  $P < 0.001$ ). Both individuals that were choosing between two different types of food (shrimps or flakes) showed a preference for shrimps. Ci choose shrimps in 78% of times ( $x = 39$ ,  $P < 0.05$ ) and Ch in 80% of times ( $x = 40$ ,  $P < 0.001$ ).

Most individuals who showed a significant preference for one of both plates did so too after 100 trials (Table 2). However, one of these individuals (Am) did not significantly prefer the plate with four rewards after 100 trials and one individual (El) only showed a significant preference for the large reward after 100 trials but not after 50 trials.

During the last 50 of the 200 trials in the reverse reward task, none of the eight individuals significantly fed against a preference (Fig. 2). Two individuals maintained their preference (Wa and Ch) while all other individuals were not choosing differently than would be expected from random choice. In detail, Mi and El feed without preference for one of both plates (Mi choosing the plate with more reward in 52% of cases (Binomial test,  $x = 26$ ,  $P > 0.05$ ) and El in 48% of the trials ( $x = 24$ ,  $P > 0.05$ )). Ma chose the larger reward in 62% of times ( $x = 31$ ,  $P > 0.05$ ) and Am in 58% of the trials ( $x = 29$ ,  $P > 0.05$ ), Bo did not continue to chose the larger reward (52% of choices for the large reward ( $x = 26$ ,  $P > 0.05$ ) but Wa significantly chose the larger reward (90% of choices for the large reward ( $x = 45$ ,  $P < 0.001$ )). Ci chose the plate with shrimp instead of flakes in 64% of trials ( $x = 32$ ,  $P > 0.05$ ) and Ch in 88% of trials during the last 50 trials ( $x = 44$ ,  $P < 0.001$ ).

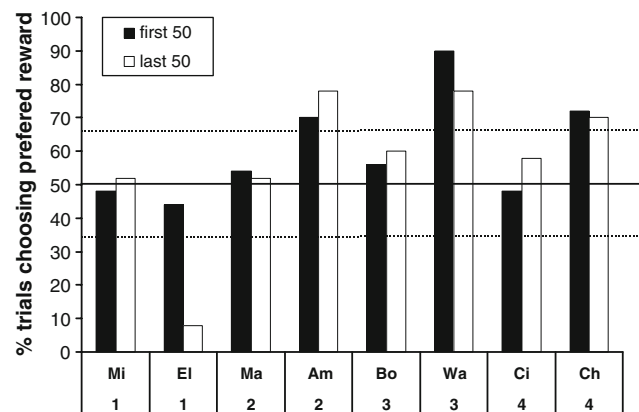
Three individuals developed a side preference. Ma selected the plate on the left side in 42.5% (85) of the trials (Chi-square test,  $N = 200$ ;  $\chi^2 = 4.5$ ;  $DF = 1$ ;  $P < 0.05$ ), Mi in 88% (176) of the trials ( $N = 200$ ;  $\chi^2 = 115.5$ ;  $DF = 1$ ;  $P < 0.001$ ) and Bo in 76% (152) of trials ( $N = 200$ ;  $\chi^2 = 54.1$ ;  $DF = 1$ ;  $P < 0.001$ ). All other individuals did not show a significant side preference.

## Large-or-none contingency

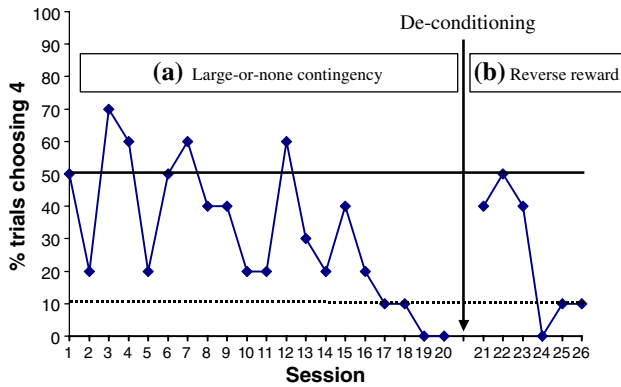
During the first 50 trials in the large or none contingency task, three individuals preferentially chose the plate with the large reward and five did not feed differently than would be expected from random choice (Fig. 3). In detail, Mi (48% of choices for the large reward,  $x = 24$ ), El (44%,  $x = 22$ ), Ma (54%,  $x = 27$ ), Bo (56%,  $x = 28$ ) and Ci (48%,  $x = 24$ ) feed without preference for one of both plates (Binomial tests,  $P > 0.05$ ) while Am (70%,  $x = 35$ ), Wa (90%,  $x = 45$ ) and Ch (72%,  $x = 36$ ) preferred to feed from the plate with the large reward (all results  $P < 0.01$ ).

During the last block of 50 trials one individual (El) was significantly choosing the plate with one food item in order to obtain the large reward (Fig. 3). Three individuals continued to preferentially choose the plate with the large reward while the other four individuals continued to choose randomly. In detail, El was choosing significantly more often the plate without reward (Binomial test, 8% of choices for plate with large reward,  $x = 4$ ,  $p < 0.001$ ). Mi (52%,  $x = 26$ ), Ma (52%,  $x = 26$ ), Bo (60%,  $x = 30$ ) and Ci (58%,  $x = 29$ ) feed without preference for one of both plates (all results  $P > 0.05$ ) while Am (78%,  $x = 39$ ), Wa (78%,  $x = 39$ ) and Ch (70%,  $x = 35$ ) continued to prefer to feed from the plate with the large reward (all results  $P < 0.01$ ).

Three individuals developed a side preference. Ma selected the plate on the left side in 95% (190) of the trials (Chi-square test,  $N = 200$ ;  $\chi^2 = 162.0$ ;  $DF = 1$ ;  $P < 0.001$ ), Mi in 99% (198) of the trials ( $N = 200$ ;  $\chi^2 = 192.1$ ;  $DF = 1$ ;  $P < 0.001$ ), Bo in 65% (130) of the trials ( $N = 200$ ;  $\chi^2 = 18$ ;  $DF = 1$ ;  $P < 0.001$ ).



**Fig. 3** Percentage of trials in which individuals chose the preferred reward during the first 50 trials (filled bars) and during the last 50 trials (open bars) in the large-or-none contingency test. The 50% line corresponds to random choice. Dotted lines at 66% ( $x = 33$ ) and 34% ( $x = 17$ ) indicate threshold values for significant deviation from random. Letters are individual codes for the test subjects. The numbers 1–4 indicate the experimental situation (for details see Table 1)



**Fig. 4** Percentage of trials per session (10 trials per session) El was choosing reward 4 in **a** a large-or-none contingency task and **b** the reverse reward contingency task. A re-experience procedure (de-conditioning) preceded the reverse reward contingency task. The (solid) 50% line corresponds to random choice and all values equal or below the 10% line (dotted line) show choice behaviour that significantly differs from random expectation (binomial test,  $N = 10$ ,  $x = 1$ ,  $P < 0.05$ )

#### Return to reverse reward contingency task

Only one individual (El) succeeded in choosing significantly the non-rewarding plate in order to get access to the large reward and was subsequently again subjected to the reverse reward task. After the re-experience procedure, El chose according to random expectation during the first three sessions and subsequently chose significantly more the small reward in order to obtain a large reward in three successive sessions (Fig. 4).

## Discussion

#### Ecological versus anthropocentric approach

The two main approaches in the cognition literature make opposite predictions as to whether cleaners should be able to solve the reverse-reward contingency task. Scientists who emphasise the role of phylogenetic relationships would predict that fishes are too distantly related to humans (presumably the animal with the most ‘complex’ cognitive abilities) and hence should be incapable of solving the task (anthropocentric approach). In contrast, the ecological approach to cognition predicts that an animal’s ability to solve a specific problem should be tightly linked to its ecological need to solve such a problem. Therefore, cleaners should be able to master a reverse reward preference task. Neither approach can fully explain our results.

In favour of the ecological approach, we provide some evidence suggesting that a fish can solve the reverse reward contingency task. We predicted that cleaners should be able

to solve a reverse reward contingency task, based on the fact that cleaners can control impulsive behaviour in a different situation, i.e. when they feed against their preference in interactions with client reef fish (Grutter and Bshary 2003; Bshary and Grutter 2005; Bshary et al. 2007; Bshary et al. 2008).

In favour of the anthropocentric approach, we note that no individual learned to solve the task within 200 trials (as has been the case for most primate species as well (Murray et al. 2005)). Moreover, while it is not possible to exclude that they would have mastered the task after more trials (as has been shown for rhesus macaques; Murray et al. 2005), only one individual developed a significant preference for the correct solution in the large or none situation, which is in contrast to the high levels of success observed in several primate species (Silberberg and Fujita 1996; Anderson et al. 2000; Genty et al. 2004).

One individual (El) may have transferred its knowledge from the large or none to a subsequent reverse reward test. This may provide some evidence that cleaner fish as a species may have the potential to solve the reverse reward task.

#### Methodological considerations

Previous studies did not attempt to exclude the possibility that individuals just continued with their choice behaviour after the large or none situation without realising a change in the reward condition. Therefore, as far as we are aware, our study has a methodological advantage because we introduced a re-experience procedure between the large or none task and the second reverse reward task. In this procedure the individuals had the possibility to experience the value of both plates, i.e. with one and four rewards before being subjected to the reverse reward test. As the successful individual El started to choose the small reward after 30 trials of initial random choice, it may indeed have learned to solve the reverse reward task after a large or none task. However, we cannot conclude with certainty that El indeed chose against a preference because the short extinction phase only caused random choices rather than a choice of the large reward. Therefore, El might have retained information from the previous large or none procedure, i.e. El might have used the white marks on the plates as a cue where to swim. Future studies should involve a longer re-experience phase than 10 trials only, and make attempts to assure extinction before returning to the reverse reward contingency procedure.

Most of the tested individuals developed a clear preference for four over one items of prawn, or for prawn over flakes. However, two individuals did not develop a significant food preference and El showed a clear preference for four items only after 100 trials. Different levels of acclimatisation to the experimental conditions might explain these

results. Such effects may cause undesired behaviours such as the development of a side preference which could explain the lack of a food preference in two individuals (Ma and Mi) and the finding that individual El showed a clear preference for the large reward only after 100 trials. Although we aimed at minimising the development of side preferences by randomising the plate position and by constructing a corridor so the fish needed to swim in the middle of the tank, three individuals developed a significant side preference, both in the reverse reward and the large or none contingency task. This might have precluded these individuals from succeeding in the task. Future studies should aim at further reducing the factors that favour the development of a side preference, for instance, by using only one plate in the centre of the tank.

Finally, as cleaners in nature appear to feed against a preference due to punishment (either a cheated client may leave the cleaning station or attack the cheating cleaner), future experiments may try to emphasize the punishment aspect within the reverse reward paradigm to make it more similar to the situation under natural conditions.

We assume that the task in our experiment was actually more difficult to solve than the reverse reward task for primates. This is because primates sitting in front of both options only need to reach out for the smaller reward with their hands. As cleaners do not have extremities to indicate their choice, they needed to swim towards one of the two plates with food. Therefore, they needed to swim away from the more attractive item and approach the less attractive item in order to solve the task. This subtle but perhaps important difference might explain why cleaners largely failed in the reverse reward contingency task, while they can easily learn to feed on a non preferred food item from a plate when a preferred item is present on the same plate in 1 cm distance (Bshary and Grutter 2005; Bshary and Grutter 2006). In these earlier experiments, each subject was allowed to continue to feed on a plate with two types of food as long as it continued to feed on the non-preferred item (flakes). However, when the subject fed on a preferred item (prawn) the plate was withdrawn.

Our low total sample size of eight individuals precludes any statistical analysis on whether the four different experimental designs had any influence on the performance of the subjects. We do not see any intuitive explanation why the experimental setup in the conditions of the successful individual (the reward plate was behind the chosen plate, the latter was removed as soon as the fish had chosen) should be more suitable for cleaner fish than the other three alternatives.

Does the reverse reward task test for “self control”?

Initially, studies on impulse control were mainly performed with humans (Mischel et al. 1989; Russell et al. 1991) and

primates (Boysen and Berntson 1995; Boysen et al. 1996; Silberberg and Fujita 1996; Kralik et al. 2002; Genty et al. 2004). Some authors have suggested that success in the reverse reward task may be interpreted as a form of “self control” thereby implying that in order to solve the task, some “awareness of self” needs to be involved (e.g. Genty et al. 2004). In line with this argument, our results suggest that solving a reverse reward contingency task is difficult even for a species with an ecology that selects for the ability to feed against a food preference. On the other hand, the observation that cleaners apparently show some potential to solve the reverse reward task suggests that seemingly complex behaviours might not necessarily require an overly complex explanation such as the concept of self awareness (Bshary et al. 2007). Instead, we propose the ability to solve a reverse reward task may result from an evolved flexibility in learning that originates from selection due to the challenges animals face in their natural environment. The underlying mechanism behind this ability may differ depending on whether primates or fishes are the focus of interest. Furthermore, contrary to what we assumed, the ecology of cleaners may not fit to the requirements of the reverse reward task. Although cleaners often feed on ectoparasites instead of the more preferred mucus, if given the choice under natural conditions, they would usually approach the more attractive client (Bshary 2001). In contrast, predators such as sea lions, may find it particularly easy to approach the smaller reward in such a test situation as the hunting success of many predators relies on their ability to isolate a single prey from its group in order to avoid confusion effects (Landeau and Terborgh 1986; Ruxton et al. 2008). The hunting technique of chasing after isolated prey provides a straightforward ecological explanation for the sea lions’ ability to solve the reverse reward task (Genty and Roeder 2006). For the future, it would therefore be interesting to repeat the test, both in non predators (including primates) and in predators (including fish species), to test whether approaching a less preferred food item is particularly difficult for non-predators compared to predators.

Finally, after the surprising finding that many primates have difficulties in refraining from the larger amount in order to obtain a small one (Boysen and Berntson 1995), many subsequent studies focussed on procedures that help the test subjects to overcome this problem. However, this approach may distract from an important question, i.e. to understand the prevalence and the evolutionary roots of succeeding in such a task in animals. An ecological approach might help to understand why many primate species initially fail in this task despite their highly developed cognitive skills. Ultimately, to understand the prevailing variation in such cognitive abilities we will need comparative studies in which a sufficient number of species with a different ecology can be subjected to the same tasks.

## Impulse control and cooperation

The evolution and maintenance of cooperative behaviour is likely constrained by the cognitive abilities in the species concerned (Stevens and Hauser 2004; Stevens et al. 2005). This is particularly important in case individuals interact in reciprocal interactions, where the act of cooperating reduces the immediate payoffs for the actor and is hence an investment that often only yields benefits in the future (Bergmüller et al. 2007a, b; Bshary and Bergmüller 2008). Under these conditions, animals may require the capacity for individual recognition, low temporal discounting (Stevens and Hauser 2004), book keeping of past interactions with partners, and, as we suggest here, the ability to control impulsive behaviour in reverse reward choice situations. An important future research question in the field of cognition in the context of cooperation is therefore to investigate how the ability to control impulsive behaviour is linked to the ability of low temporal discounting, and how past experience with a partner (positive or negative) influences impulsive behaviour.

**Acknowledgments** We are grateful to Sylvain Bruschweiler for help with producing Fig. 1 and Matthias Erb, Russell Naisbit, and four anonymous referees for helpful comments on the manuscript. We thank Anna Albiach Serrano and the Eco-Ethology group at the University of Neuchâtel for discussion, Jean-Pierre Duvoisin for technical assistance and Ana Pinto for her help with the fish. This study was supported by the Swiss National Science Foundation (to R. Bshary). The experiment was licensed by the Veterinary Commission of the Canton de Neuchâtel.

## References

- Abeyesinghe SM, Nicol CJ, Hartnell SJ, Wathes CM (2005) Can domestic fowl, *Gallus gallus domesticus*, show self-control? *Anim Behav* 70:1–11
- Albiach-Serrano A, Guillen-Salazar F, Call J (2007) Mangabeys (*Cercocebus torquatus lunulatus*) solve the reverse contingency task without a modified procedure. *Anim Cogn* 10:387–396
- Anderson JR, Awazu S, Fujita K (2000) Can squirrel monkeys (*Saimiri sciureus*) learn self-control? A study using food array selection tests and reverse-reward contingency. *J Exp Psychol Anim Behav Process* 26:87–97
- Bergmüller R, Johnstone RA, Russell AF, Bshary R (2007a) Integrating cooperative breeding into theoretical concepts of cooperation. *Behav Process* 76:61–72
- Bergmüller R, Russell AF, Johnstone RA, Bshary R (2007b) On the further integration of cooperative breeding and cooperation theory. *Behav Process* 76:170–181
- Boysen ST, Berntson GG (1995) Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees (*Pan troglodytes*). *J Exp Psychol Anim Behav Process* 21:82–86
- Boysen ST, Berntson GG, Hannan MB, Cacioppo JT (1996) Quantity-based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *J Exp Psychol Anim Behav Process* 22:76–86
- Boysen ST, Mukobi KL, Berntson GG (1999) Overcoming response bias using symbolic representations of number by chimpanzees (*Pan troglodytes*). *Anim Learn Behav* 27:229–235
- Bshary R (2001) The cleaner fish market. In: Noë R, van Hooff JARAM, Hammerstein P (eds) *Economics in nature*. Cambridge University Press, Cambridge pp 146–172
- Bshary R, Bergmüller R (2008) Distinguishing four fundamental approaches to the evolution of helping. *J Evol Biol* 21:405–420
- Bshary R, Grutter AS (2002a) 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 (2002b) Parasite distribution on client reef fish determines cleaner fish foraging patterns. *Mar Ecol Prog Ser* 235:217–222
- Bshary R, Grutter AS (2005) Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Biol Lett* 1:396–399
- Bshary R, Grutter AS (2006) Image scoring and cooperation in a cleaner fish mutualism. *Nature* 441:975–978
- Bshary R, Schäffer D (2002) Choosy reef fish select cleaner fish that provide high-quality service. *Anim Behav* 63:557–564
- Bshary R, Wickler W, Fricke H (2002) Fish cognition: a primate's eye view. *Anim Cogn* 5:1–13
- Bshary R, Salwiczek L, Wickler W (2007) Social cognition in non primates. In: Dunbar R, Barrett L (eds) *The Oxford handbook of evolutionary psychology*. Oxford Handbooks in Psychology, Oxford
- Bshary R, Grutter AS, Willener AST, Leimar O (2008) Pairs of cooperating cleaner fish provide better service quality than singletons. *Nature* 455:964–966
- Genty E, Roeder JJ (2006) Self-control: why should sea lions, *Zalophus californianus*, perform better than primates? *Anim Behav* 72:1241–1247
- Genty E, Palmier C, Roeder JJ (2004) Learning to suppress responses to the larger of two rewards in two species of lemurs, *Eulemur fulvus* and *E. macaco*. *Anim Behav* 67:925–932
- Grutter A (1996) Parasite removal rates by the cleaner wrasse *Labroides dimidiatus*. *Mar Ecol Prog Ser* 130:61–70
- Grutter AS, Bshary R (2003) Cleaner wrasse prefer client mucus: support for partner control mechanisms in cleaning interactions. *Proc R Soc Lond B Biol Sci* 270:S242–S244
- Kamil AC (1998) On the proper definition of cognitive ethology. In: Balda RP, Pepperberg IM, Kamil AC (eds) *Animal cognition in nature*. Academic Press, San Diego, pp 1–28
- Kralik JD, Hauser MD, Zimlicki R (2002) The relationship between problem solving and inhibitory control: cotton-top tamarin (*Saguinus oedipus*) performance on a reversed contingency task. *J Comp Psychol* 116:39–50
- Landeau L, Terborgh J (1986) Oddity and the confusion effect in predation. *Anim Behav* 34:1372–1380
- Mischel W, Shoda Y, Rodriguez ML (1989) Delay of gratification in children. *Science* 244:933–938
- Murray EA, Kralik JD, Wise SP (2005) Learning to inhibit prepotent responses: successful performance by rhesus macaques, *Macaca mulatta*, on the reversed-contingency task. *Anim Behav* 69:991–998
- Russell J, Mauthner N, Sharpe S, Tidswell T (1991) The “windows task” as a measure of strategic deception in preschoolers and autistic subjects. *Br J Dev Psychol* 9:331–349
- Ruxton GD, Jackson AL, Tosh CR (2008) Confusion of predators does not rely on specialist coordinated behavior. *Behav Ecol* 18:590–596
- Shettleworth SJ (1994) Biological approaches to the study of learning. In: Mackintosh NJ (ed) *Animal learning and cognition*. Academic Press, San Diego, pp 185–219
- Shumaker RW, Palkovich AM, Beck BB, Guagnano GA, Morowitz H (2001) Spontaneous use of magnitude discrimination and ordination by the orangutan (*Pongo pygmaeus*). *J Comp Psychol* 115:385–391

- Silberberg A, Fujita K (1996) Pointing at smaller food amounts in an analogue of Boysen and Berntson's (1995) procedure. *J Exp Anal Behav* 66:143–147
- Stevens JR, Hauser MD (2004) Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn Sci* 8:60–65
- Stevens JR, Cushman FA, Hauser MD (2005) Evolving the psychological mechanisms for cooperation. *Annu Rev Ecol Evol Syst* 36:499–518
- Uller C, Jaeger R, Guidry G, Martin C (2003) Salamanders (*Plethodon cinereus*) go for more: rudiments of number in an amphibian. *Anim Cogn* 6:105–112
- West SA, Griffin AS, Gardner A (2007) Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol* 20:415–432
- Wittmann M, Paulus MP (2008) Decision making, impulsivity and time perception. *Trends Cogn Sci* 12:7–12