CrossMark

Rats' acquisition of the ephemeral reward task

Thomas R. Zentall[1] · Jacob P. Case[1] · Jonathon R. Berry[1]

Abstract The ephemeral reward task provides a subject with a choice between two alternatives A and B. If it chooses alternative A, reinforcement follows and the trial is over. If it chooses alternative B, reinforcement follows but the subject can also respond to alternative A which is followed by a second reinforcement. Thus, it would be optimal to choose alternative B. Surprisingly, Salwiczek et al. (PLoS One 7:e49068, 2012. doi:10.1371/journal.pone.00490682012) reported that adult fish (cleaner wrasse) mastered this task within 100 trials, whereas monkeys and apes had great difficulty with it. The authors attributed the species differences to ecological differences in the species foraging experiences. However, Pepperberg and Hartsfield (J Comp Psychol 128:298–306, 2014) found that parrots too learned this task easily. We have found that with a similar task pigeons are not able to learn to choose optimally within 400 trials (Zentall et al. in J Comp Psychol 130:138–144, 2016). In Experiment 1 of the present study, we found that rats did not learn to choose optimally in 840 trials; however, in Experiment 2 we added a prior commitment to the initial choice by increasing delay to reinforcement for the choice response from a single lever press to the first lever press after 20 s (FI20 s). In a comparable amount of training to Experiment 1, the rats learned to choose optimally. Although the use of a prior commitment increases the delay to reinforcement, it appears to reduce impulsive responding which in turn leads to optimal choice.

Keywords Ephemeral choice · Ephemeral reward · Delay of reinforcement · Commitment · Rats

✉ Thomas R. Zentall
  zentall@uky.edu

[1] Department of Psychology, University of Kentucky, Lexington, KY 40506, USA

Introduction

Many tasks have been used to assess the relative learning ability of animals (Bitterman 1965). For example, with the serial reversal task animals are trained on a simple simultaneous discrimination. After acquisition, the discrimination is reversed until the reversal is acquired, and then, it is reversed again and again. The measure of learning flexibility is the extent to which reversal acquisition improves with successive reversals, relative to original learning. Using this procedure, Bitterman found that that different species show different degrees of improvements with successive reversals that correspond to less formal comparisons among species. For example, monkeys show greater improvement than rats, which show greater improvement than pigeons, which show greater improvement than fish (Bitterman 1965). This approach to comparative cognition was quite popular at one time, but there is reason to suspect that it is overly simplistic.

Recently, research using what would appear to be a relatively easy task (Bshary and Grutter 2002; Salwiczek et al. 2012) has shown surprising species differences. With this task, which we have called the ephemeral choice task (Zentall et al. 2016), animals are given a choice between two distinctive stimuli (e.g., red and yellow dishes) on which two identical reinforcers are placed. If, for example, the animal chooses the reinforcer on the red background, it is also allowed to have the reinforcer on the yellow background, but if the animal chooses the reinforcer on the yellow background, the reinforcer and the red background are removed. With this task, the optimal solution is for the subject always to choose the reinforcer on the red background because it results in two reinforcers per trial, whereas if it chooses the reinforcer on the yellow background it would only receive one reinforcer. The surprising finding is that several primate species including capuchin

🖄 Springer

monkeys (*Cebus apella*), orangutans (*Pongo pygmaeus*), and several chimpanzees (*Pan troglodytes*) when trained on this task were not able to learn to choose optimally within 100 trials, whereas adult fish (cleaner wrasse, *Labroides dimidiatus*) appear to acquire it quite easily.

Salwiczek et al. (2012) attributed the difference to species differences in natural foraging behavior. According to the authors, the cleaner wrasse often clean other large fish that live on their reef (permanent clients), but sometimes they clean fish that are temporarily present (ephemeral clients). The authors reasoned that the cleaner fish learn to clean the ephemeral clients first because those clients may not be there if they return at a later time. Thus, the fish may learn to prioritize servicing clients that are less likely to remain on the reef over client fish that will be there when they get back. Primates, however, have not developed such a foraging strategy. This argument was challenged recently by Pepperberg and Hartsfield (2014) who showed that grey parrots (*Psittacus erithacus*), whose foraging behavior should be more like that of the primates than of the cleaner wrasse, can easily acquire the optimal response.

To test the hypothesis that birds may acquire this task differently from other animals, Zentall et al. (2016) conducted a similar experiment with pigeons. Choice of a dried pea on, for example, a blue disk ended the trial and both disks were removed, but choice of a similar dried pea on a yellow disk allowed the pigeon to also eat the pea on the blue disk. Surprisingly, not only did the pigeons not choose the optimal alternative, but they showed a significant preference for the suboptimal alternative, the alternative that fed them once rather than twice. In a follow-up experiment, the authors demonstrated that the preference for the suboptimal alternative could be replicated under operant (automated) conditions. That is, initially choosing a yellow light provided food at a feeder and allowed the pigeon also to peck at the other, blue, light for food, but initially choosing the blue light provided food at the feeder but terminated the trial. Zentall et al. reasoned that the pigeons had a preference for the suboptimal alternative, because although the optimal alternative allowed the pigeons to obtain more food, the difficulty in learning to choose that alternative may have occurred because no matter which alternative was originally selected, all trials ended with a response to the suboptimal alternative. That is, an initial response to yellow was followed by a response to blue and an initial response to blue ended the trial. Thus, there would have been two reinforcements associated with blue but only one associated with yellow. Interestingly, perhaps for the same reason, Prétôt et al. (2016a) using a similar task with rhesus macaques found that four of their eight monkeys initially chose the suboptimal alternative more often than the optimal alternative.

To test the hypothesis that in the earlier research initially animals would receive twice as many reinforcements associated with the suboptimal alternative, Zentall et al. (2016) again gave pigeons a choice between a yellow and a blue alternative, but if the pigeon chose the yellow (optimal) alternative and was fed, the blue alternative changed to red, and now the pigeon could be fed again by pecking the red light. Thus, choice of yellow still resulted in two rewards, but now those trials did not end with pecks to the blue stimulus. Therefore, the contingency for choosing the yellow light (two reinforcements) was the same as in the earlier experiment but now choosing the optimal alternative resulted in reinforcement associated with pecking red instead of blue. The result of that experiment was that the pigeons no longer had a preference for the suboptimal alternative, but they still failed to learn to choose the optimal alternative (i.e., now they were indifferent between the two alternatives).

The purpose of the first experiment was to test the generality of finding by Zentall et al. (2016), which found that pigeons did not readily acquire the ephemeral reward task. However, given that fish (Salwiczek et al. 2012) and parrots (Pepperberg and Hartsfield 2014) readily acquired the optimal choice with this task and rats sometimes show optimal performance under conditions in which pigeons do not (Rayburn-Reeves et al. 2013), we thought it appropriate to test rats with this task.

Although the present experiments were conducted with automated procedures and in earlier research primates, fish, parrots, and pigeons were trained using manual procedures, Zentall et al. (2016) compared the automated procedure with the manual procedure with pigeons and found very similar results with both.

Experiment 1

To test rats on the ephemeral reward task, we gave them a choice between two levers. If they chose the suboptimal lever, they received a food pellet and the trial was over but if they chose the optimal lever, after receiving a food pellet they could also respond to the other lever for an additional pellet. To ensure that idiosyncratic spatial preferences would not be interpreted as optimal or suboptimal learning, for some rats the optimal lever on each trial was signaled by a light above the lever, and for other rats the suboptimal lever on each trial was signaled by a light above the lever. The light over the lever changed location (left or right) randomly from trial to trial.

Method

Subjects

Our subjects were 12 Sprague–Dawley rats. All rats were approximately 1-year-old males maintained at 85% of their

free-feeding weight. The rats had prior experience with autoshaping in an operant chamber to associate a lever, a tone, or a light with reinforcement. They were housed individually in plastic bins and given free access to water and chew sticks. The rat colony room was kept on a 12:12-h light dark cycle. All rats were cared for in accordance with the University of Kentucky Animal Care Guidelines.

Apparatus

The experiment was conducted in a Med Associates (St. Albans, VT) modular operant testing chamber. The front wall of the chamber had a feeding trough in the center that was connected to an automatic pellet dispenser that delivered reinforcement according to the rat's performance. On either side of the trough were two retractable levers, with a green signal light mounted directly above each lever. There was a house light centered in the ceiling of the box, and a microcomputer in an adjacent room controlled the experiment. The side walls on either side of the front wall of the chamber were made of clear Plexiglas, and the right wall contained a hinged door through which the rat could be placed in the chamber. The back wall of the chamber was made of sheet metal.

Procedure

Each trial began with the extension of both levers and the illumination of one of the signal lights. For some rats, the signal light always appeared above the ephemeral lever (the lever associated with the optimal choice, defined as the ephemeral reward). For the remaining rats, the signal light always appeared above the suboptimal lever (counterbalanced over subjects, six rats per group). The side on which the signal light appeared was determined randomly on each trial; thus, it was a visual rather than a spatial discrimination. This is likely to be important because if animals are indifferent between alternatives they often do not choose the alternatives randomly from trial to trial but often choose one side or the other consistently from trial to trial. Thus, if the animals are indifferent, it may appear that there is a bimodal preference for the ephemeral alternative, some animals showing what appears to be a strong preference for one schedule, others showing what appears to be a strong preference for the other (see Smith and Zentall 2016). If the rat selected the suboptimal lever, both levers retracted, the signal light turned off, reinforcement was delivered, and the trial was terminated. A 5-s intertrial interval (ITI) followed, during which the house light was illuminated. If the rat chose optimally (i.e., first selected the ephemeral lever), that lever retracted and reinforcement was delivered, but the suboptimal lever remained extended, allowing the rat to respond to it for reinforcement as well, followed by the 5-s

ITI. All rats received 40 trials per session, for 21 sessions, one session per day, six days a week.

Results

Although at the start of training most of the rats showed a preference for the lever with the light over it, as shown in Fig. 1, the difference between the counterbalancing groups disappeared quickly with training. The only session on which the light optimal group was significantly different from the light suboptimal group was on Session 2, $t(10) = 3.93$, $P = .003$. As suggested by the small error bars, most of the rats developed a strong spatial preference and over the course of the 21 sessions of training, the rats showed little sign of learning to choose the optimal alternative (see Fig. 1). When the data were pooled over the last three sessions of training, they averaged 50.6% choice of the optimal alternative.

Discussion

The results of Experiment 1 suggest that the ephemeral reward task is almost as difficult for rats as it was for pigeons (Zentall et al. 2016), although it should be noted that the pigeons chose the suboptimal alternative significantly below chance, whereas the rats chose it at chance. That is, in spite of the fact that all trials ended with the rats pressing the suboptimal lever, overall, they did not show a preference for the suboptimal lever. The results of this experiment add to the list of species that have great difficulty with this task: pigeons (Zentall et al. 2016), macaque
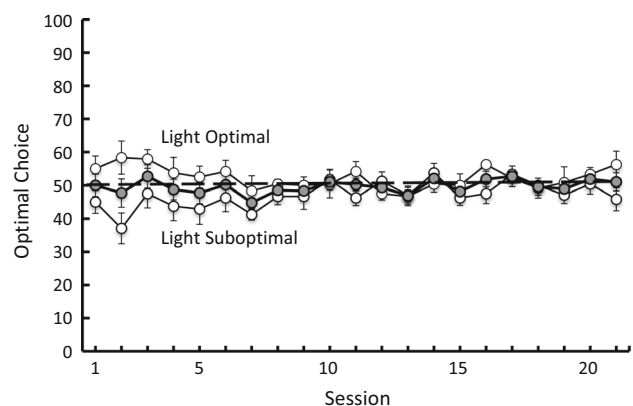
Fig. 1 Experiment 1: optimal choice by rats with a fixed ratio 1 choice response requirement. For the light optimal group, the ephemeral alternative was the lever with the light over it. For the light suboptimal group, the nonephemeral alternative was the lever with the light over it. Mean of the two groups is indicated by the *filled circles*. *Error bars* = ±1 SEM

monkeys (Prétôt et al. 2016b), capuchin monkeys, orangutans, and even chimpanzees (Salwiczek et al. 2012). The difficulty of this task is especially surprising, given that cleaner fish (Salwiczek et al.) and parrots (Pepperberg and Hartsfield 2014) can master it easily.

Several hypotheses have been proposed to account for the species difference. Salwiczek et al. (2012) proposed that the task matched the foraging strategy of the fish but not the primates. They propose that the fish experience the ephemeral nature of client fish that visit the reef, as well as client residents of the reef that can wait to be serviced, whereas the primates do not have that experience. However, the fact that parrots easily acquire the task yet have a foraging strategy that is more similar to that of the primates suggests that foraging strategy alone cannot account for the species differences. Furthermore, parrots like nonhuman primates have fewer young and devote more care to them than fish, they also live longer, have longer periods of maturation, are larger in body size, and they forage over wider distances.

Another suggestion by Salwiczek et al. (2012) for the fishes' success and nonhuman primates' failure is that the fish use their mouth to make their choices, whereas the primates use their hands. As the primates have two hands, in nature they would be able to choose both alternatives simultaneously, whereas the fish would have to choose one at a time. In fact, when monkeys used a "joy stick" to choose between two stimuli presented on a computer monitor, they showed much better acquisition than with the manual procedure (Prétôt et al. 2016a). The success of parrots is consistent with this hypothesis; however, the failure of pigeons to acquire this task suggests that prior experience with choosing one at a time cannot account for the rapid acquisition of the optimal choice by fish and parrots but not by primates, pigeons, and rats.

A different approach to understanding the species differences in acquisition of this task was proposed by Zentall et al. (2016). They proposed that the task can be thought of in the context of delay discounting: rewards that are delayed have less value than those that are immediate (see, e.g., Doyle 2013). If the second reinforcement associated with initial choice of the optimal alternative is sufficiently discounted, the immediacy of the first reinforcement may make it difficult for the animal to discriminate between the differential consequences of the choice. Although given the choice of the optimal alternative, the delay between the rewards is generally quite short (in the manual procedure the pigeon can eat the second reward within 1 s of the first, see Zentall et al. 2016, Exp. 1). Thus, if the pigeons are sufficiently impulsive, it may be sufficient to make the consequences difficult to discriminate.

If the analogy between the ephemeral choice task and delay discounting is correct, one might be able to increase choice of the optimal alternative by using a technique that has been used with delayed discounting to increase choice of the delayed larger reward over the smaller immediate reward (Ainslie 1975; Rachlin and Green 1972). Paradoxically, the procedure used by Rachlin and Green involved forcing the animal to make a choice some time before obtaining the reward. Thus, rather than obtaining the smaller reward immediately, pigeons were forced to choose a few seconds before obtaining it. Like Rachlin and Green (1972), Zentall et al. (2016) conceived of the procedure as making a prior commitment. They proposed that this procedure might increase the pigeons' choice of the optimal alternative. Specifically, Zentall et al. (2016) incorporated the prior commitment procedure into the ephemeral choice task by requiring pigeons to make their initial choice, which then started a fixed-interval 20-s schedule (after choosing one of the alternatives, the first response after 20 s provided the first reward). If they had chosen the optimal alternative, they could then make a single peck to the other alternative to obtain the second reward. The results of that experiment indicated that pigeons that were forced to make a prior commitment acquired the ephemeral reward task, whereas those that were not forced to make a prior commitment did not.

Interestingly, when Prétôt et al. (2016b) modified their task such that the food on the plates was not visible at the time of choice but the cups that covered the food were distinctive, the monkeys learned the task. Covering the food not only delayed the monkeys' access to the food but also removed the food from view, thus likely further reducing impulsivity (see also Boysen et al. 1996).

Another modification of the task that resulted in relatively rapid acquisition was having the monkeys make their choices with a "joy stick" on a computer screen (Prétôt et al. 2016a). Although the authors suggested that this change increased the ecological relevance of the task by reducing the number of "extraneous cues," it is also possible that this change delayed the monkeys' access to the food and removed the food from view as in the Prétôt et al. (2016b) study.

Experiment 2

The purpose of Experiment 2 was to determine whether forcing rats to make a prior commitment, thus delaying access to the reinforcer, would allow them to acquire the optimal response in the ephemeral choice task.

Methods

Subjects and apparatus

Six male Sprague–Dawley rats were used in Experiment 2. They were approximately four months old. The rats were maintained at 85% of their free-feeding weight. They had prior experience with a probability discounting task. Animal care and housing were the same as in Experiment 1.

Apparatus

Experiment 2 was conducted in the same apparatus as Experiment 1.

Procedure

The procedure for Experiment 2 was similar to the procedure for Experiment 1; the critical difference was the schedule of reinforcement associated with the initial choice. In Experiment 2, initial choice required completion of a fixed-interval 20-s schedule (FI20s). If the rat first selected the nonephemeral lever, the ephemeral lever retracted just as it did in Experiment 1, and reinforcement followed the first response to the nonephemeral lever after 20 s, which also started a 5-s ITI with the houselight lit. If the rat's initial choice was the ephemeral lever, again the other lever retracted and the first response after 20 s provided reinforcement and retracted the ephemeral lever. The nonephemeral lever then reemerged, allowing the rat to press it, and a second reinforcement followed a single response (fixed ratio 1; FR1), followed by a 5-s lit ITI. For half of the rats, the green jeweled cue light was illuminated above the ephemeral lever. For the remaining rats, the cue light appeared above the nonephemeral lever. Each session consisted of 40 trials. The location of the cue light varied randomly from trial to trial. Sessions were conducted once a day, six days a week, for 41 sessions.

Results

The rats initially were indifferent between the two alternatives. In fact, they first exceeded 60% choice of the optimal alternative on Session 16, and the first session on which they were significantly different from the rats in Experiment 1 was Session 18 (Mean Optimal Choice = 62.1%, $t(5) = 2.52$, $P = .05$, Cohen's $d = 2.75$, r(effect size) = .75. Choice of the optimal alternative on the last two sessions for which we have data from Experiment 1 (Sessions 20 and 21) was at 69.8% as compared with 50.6% for the rats in Experiment 1 (see Fig. 2). Unfortunately, we only trained the rats in

Experiment 1 for 21 sessions, so comparison of the two groups beyond that point was not possible. Nevertheless, with additional training, the rats in Experiment 2 continued to improve in their choice of the optimal alternative, and their choice of the optimal response was quite similar to that of the pigeons in Zentall et al. (2016). For comparison purposes, we have included the data from Experiment 1 in Fig. 2 as well as the results from the experiment with pigeons (Zentall et al.).

Discussion

The results of Experiment 2 indicate that much like pigeons, delaying the outcome of the rats' initial choice by 20 s facilitated acquisition of the ephemeral reward task. The mechanism by which it facilitated acquisition is likely to be similar to early commitment in the delay discounting task. That is, in the delay discounting task commitment reduces the relative difference in delay between the smaller sooner reward and larger later reward. For example, if the smaller sooner reward can be obtained in 1 s and the larger later in 10 s, the ratio would be 10:1, whereas if the choice must be made 20 s before the smaller sooner the ratio would be $(20 + 10)/(20 + 1) = 30/21$ or 1.42:1, a much smaller difference. In the case of the ephemeral reward task, we proposed that the immediacy of the reward for choice of either alternative, together with the delay of the second reward for choice of the optimal alternative, may have been responsible for the difficulty of the task. We tested that hypothesis by forcing the rat to choose 20 s before obtaining the first reward and thereby shortening the relative delay to the second reward and reducing the rats'
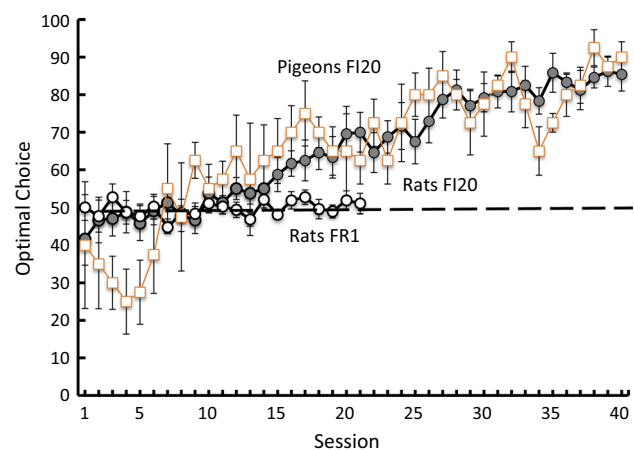
Fig. 2 Experiment 2: optimal choice by rats with a fixed-interval 20-s response requirement (*filled circles*). For comparison purposes, the mean data from Experiment 1 in which only a single peck was required for initial choice are also presented (*open circles*). The data from Zentall et al. (2016) for the FI20s condition are also presented in the figure (*open squares*). Error bars = ±1 SEM

impulsivity. It is somewhat paradoxical that increasing the delay to reinforcement, a manipulation that typically results in slower learning (see, e.g., Renner 1964) appears to make it easier for rats to learn to choose optimally.

One issue that needs to be addressed is the relatively short intertrial interval (5 s) used in the present study. It could be argued (as one reviewer did) that the short intertrial interval essentially removed the penalty for making a suboptimal choice because choice of the suboptimal alternative ended the current trial but started the next trial with a minimal delay. The introduction of the fixed-interval 20-s schedule at the start of each trial, however, increased the penalty for making the suboptimal choice because it effectively delayed reinforcement on the next trial by 25 s.

Comparing the procedures used with the different species suggests that learning this task was not directly related to the duration of the intertrial interval because the capuchin monkeys had 15 min between trials and they did not acquire the task, whereas the parrots had only 90 s between trials and they learned it. More relevant to the present research, Zentall et al. (2016) used a 90-s intertrial interval and for their control group (without the fixed-interval 20-s schedule as a commitment response but with a 20-s delay prior to the choice response) failed to find any indication of learning to choose the ephemeral alternative by pigeons. Thus, pigeons did not learn to choice optimally even with a delay of 110 s between trials.

It should be noted that the earlier research with fish, primates, and parrots training involved only 100 trials, whereas in the pigeon and rat research much more training was provided. First, in general, the additional training did not help either the pigeons or the rats when the choice involved a single response but it did help when there was a considerable delay between the initial choice and the first reinforcement. Second, when the initial choice involved a fixed-interval 20-s schedule, neither the rats nor the pigeons acquired the optimal choice anywhere near as quickly as the fish or the parrots.

General discussion

The fact that the delay to reinforcement provided by the 20-s delay facilitated acquisition of optimal choice may suggest the basis for species differences in learning this task. Given that the delay to reinforcement decreases impulsivity in rats and also in pigeons, perhaps animals that are inherently impulsive (as defined by delay discounting measures) will have difficulty acquiring this task. If this assumption is correct, how can it account for the relative ease with which cleaner wrasse and grey parrots acquire this task in the absence of the delay to reinforcement?

In the case of cleaner wrasse, one can speculate that in their natural environment they need to be relatively cautious because they feed by swimming into the mouth of predatory fish (see Bshary and Grutter 2005; Bshary et al. 2008). Impulsive wrasse are likely to have a shorter life. If the cleaner wrasse learn to be (or are naturally) cautious, it may be easier for them to associate choice of the optimal alternative with the second reward. That is, they would be predicted to show less discounting with delays and have shallower delay discounting functions. Furthermore, this ability appears to be an ability that they acquire with experience (or maturation), because young cleaner wrasse do not acquire the task as readily as adults (Salwiczek et al. 2012).

But what about the grey parrots (Pepperberg and Hartsfield 2014)? Although it is possible that they too are a not very impulsive species, the three parrots that were used in the Pepperberg and Hartsfield (2014) experiment had had extensive prior training. One parrot had been exposed to "continuing studies on comparative cognition and interspecies communication" (p. 299), while the other two had received considerable training on referential communication. It is possible that this training had the effect of reducing their natural impulsivity. In fact, one of those parrots was found to show great impulse control when given a choice between an immediate desirable reward and a delayed (by as much as 15 min) more desirable reward (Koepke et al. 2015).

On the other hand, most primates would be expected to be quite impulsive because they live in a social environment and any delay in obtaining food would likely result in losing it to other members of their troop. Thus, Salwiczek et al. (2012) may have been correct that the species differences that they found can be attributed to species differences in foraging behavior, but the mechanism involved in those differences is likely to be differences in impulsivity. In fact, when the rewards were not visible and were thus delayed (either by hiding them under a cup or by having to use a "joy stick" to click on the chosen stimulus), monkeys were more successful.

It would be instructive to test this impulsivity hypothesis further by assessing the cleaner wrasse's delay discounting functions. If reduced impulsivity is responsible for the ease with which they acquire the ephemeral reward task, they should show slower rates of delay discounting than other species that have been tested.

References

Ainslie G (1975) Specious reward: a behavioral theory of impulsiveness and impulse control. Psychol Bull 82:463–496

Bitterman ME (1965) Phyletic differences in learning. Am Psychol 20:396–410

Boysen ST, Berntson GG, Hannan MB, Cacioppo JT (1996) Quantity based interference and symbolic representations in chimpanzees (*Pan troglodytes*). J Exp Psychol Anim Behav Proc 22:76–86

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

Bshary R, Grutter AS, Willener AST, Leimar O (2008) Pairs of cooperating cleaner fish provide better service quality than singletons. Nature 455:964–967

Doyle JR (2013) Survey of time preferences, delay discounting models. Judgm Decis Mak 8:116–135

Koepke AE, Gray SL, Pepperberg IM (2015) Delayed gratification: a grey parrot (*Psittacus erithacus*) will wait for a better reward. J Comp Psychol 129:339–346

Pepperberg IM, Hartsfield LA (2014) Can Grey parrots (*Psittacus erithacus*) succeed on a "complex" foraging task failed by nonhuman primates (*Pan troglodytes, Pongo abelii, Sapajus apella*) but solved by wrasse fish (*Labroides dimidiatus*)? J Comp Psychol 128:298–306

Prétôt L, Bshary R, Brosnan SF (2016a) Comparing species decisions in a dichotomous choice task: adjusting task parameters improves performance in monkeys. Anim Cognit. doi:10.1007/s10071-016-0981-6

Prétôt L, Bshary R, Brosnan SF (2016b) Factors influencing the different performance of fish and primates on a dichotomous choice task. Anim Behav 119:189–199

Rachlin H, Green L (1972) Commitment, choice and self-control. J Exp Anal Behav 17:15–22

Rayburn-Reeves RM, Stagner JP, Kirk CR, Zentall TR (2013) Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): qualitative differences in behavioral flexibility. J Comp Psychol 127:202–211

Renner KE (1964) Delay of reinforcement: a historical review. Psychol Bull 61:341–361

Salwiczek LH, Prétôt L, Demarta L, Proctor D, Essler J, Pinto AI, Wismer S, Stoinski T, Brosnan SF, Bshary R (2012) Adult cleaner wrasse outperform capuchin monkeys, chimpanzees, and orang-utans in a complex foraging task derived from cleaner-client reef fish cooperation. PLoS ONE 7:e49068. doi:10.1371/journal.pone.0049068

Smith AP, Zentall TR (2016) Suboptimal choice in pigeons: choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. J Exp Psychol Anim Behav Proc 42:212–220

Zentall TR, Case JP, Luong J (2016) Pigeon's paradoxical preference for the suboptimal alternative in a complex foraging task. J Comp Psychol 130:138–144

Zentall TR, Case JP, Berry JR (2016) Early commitment facilitates optimal choice by pigeons. Psychol Bull Rev. doi:10.3758/s13423-016-1173-8