



The Ephemeral Reward Task: Pigeons and Rats Fail to Learn Unless Discouraged from Impulsive Choice

Thomas R. Zentall*, Jacob P. Case, Jonathon R. Berry, and Jasmine Luong

Department of Psychology, University of Kentucky

*Corresponding author (Email: zentall@uky.edu)

Citation – Zentall, T. R., Case, J. P., Berry, J. R., & Luong, J. (2018). The ephemeral reward task: Pigeons and rats fail to learn unless discouraged from impulsive choice. *Animal Behavior and Cognition*, 5(1), 169-180.
<https://doi.org/10.26451/abc.05.01.12.2018>

Abstract – The failure of certain species to learn a particular task while others learn it easily can help identify the learning mechanisms involved. In the *ephemeral reward task*, animals are given a choice between two distinctive stimuli, A and B, each containing an identical bit of food. If they choose A they get the food on A and the trial is over. If they choose B they get the food on B and they are allowed to get the food on A before the trial is over. Thus, it is optimal to choose B. Although cleaner fish (wrasse) and parrots acquire the optimal response easily, several primate species do not. Furthermore, pigeons and rats also appear to be unable learn to choose optimally. The failure of primates, pigeons, and rats to learn this task and the ease with which cleaner fish and parrots learn it raises important questions about the learning mechanisms involved in those differences. To account for these paradoxical findings, we proposed that certain species may have difficulty with this task because they tend to respond impulsively to the initial choice which has similar immediate outcomes and they do not associate the choice and reinforcement with the second reinforcement. To test this hypothesis, we temporally separated the initial choice from the first reinforcement by imposing a 20-s delay between the choice and its outcome. Under these conditions both pigeons and rats gradually acquired the optimal choice response. We suggest that impulsive choice may make it difficult to acquire certain tasks and imposing a delay between choice and outcome may decrease impulsivity and allow for closer to optimal task performance.

Keywords – Ephemeral choice task, Suboptimal choice, Prior commitment, Delay of reinforcement

When animals fail to acquire what appears to be a relatively simple task, it raises questions that can lead to a better understanding of principles of learning that may apply more generally to other tasks and to other species. Such a task is one we refer to as the *ephemeral reward task*.

Psychologists have been studying the acquisition of relatively simple tasks by animals for over 100 years; from their escape from puzzle boxes (Thorndike, 1898) to simple conditioned reflexes (Pavlov, 1927). More recently the ability to show improvement with the learning in a serial reversal of a simple discrimination has been used as a proxy for the assessment of animal intelligence (Bitterman, 1965). In many cases intuition about animal intelligence has been confirmed with animals deemed to be more like humans doing better on these measures than those that are less like us. Thus, apes tend to show more improvement over reversals than monkeys, monkeys more improvement than rats, rats more improvement than pigeons and so on (Bitterman, 1965, 1975). But as we will see, this is not always the case.

The Paradoxical Ephemeral Reward Task

An apparent exception to this general *rule of thumb* is the performance by different species on the *ephemeral reward task*. With this task, animals are given a choice between two distinctive stimuli (e.g., a red plate and a yellow plate) each containing an identical bit of food. If, for example, the animal chooses the food on the red plate, the yellow plate is removed and the trial is over. But if the animal chooses the food on the yellow plate it can also have the food on the red plate (Bshary & Grutter 2002; Salwiczek et al., 2012). With this task, the food maximizing or optimal solution is for the subject always to choose the food on the red plate because it results in two bits of food per trial, whereas if it chooses the food on the yellow plate it receives only one.

Cleaner Fish and Primates

Bshary and Grutter (2002) found that bluestreak cleaner wrasse (*Labroides dimidiatus*), a species of fish that cleans the mouth and gills of large sometimes predatory fish, are generally able to learn to make the optimal choice (to regularly choose the dish that lets them have both pieces of food) in under 100 trials. Surprisingly, however, several primate species including capuchin monkeys (*Cebus apella*), orangutans (*Pongo spp.*), and several chimpanzees (*Pan troglodytes*) when trained on this task, were not able to learn to choose optimally within 100 trials (Salwiczek et al., 2012). Salwiczek et al. attributed the optimal performance by fish to the natural foraging behavior of this reef dwelling species. As it turns out, these fish have a symbiotic relation with large fish, feeding on the client fish's parasites and mucus. When a client fish visits the reef, it is an *ephemeral* resource that should be serviced immediately because the client may leave for the territory of other cleaner fish, whereas reef-dwelling client fish are relatively *permanent* and can be serviced later. Thus, the wrasse has learned to go first for the ephemeral client (the optimal alternative) rather than the permanent resident (the suboptimal response). Learning is presumed to be involved because juvenile wrasse also have great difficulty with this task. Salwiczek et al. argued that primates, however, live in a quite different environment with ecological constraints that do not encourage optimal choice with this task because they often encounter foods unpredictably and opportunistically.

Parrots

This hypothesis is challenged by the finding that grey parrots also find this task relatively easy (Pepperberg & Hartsfield, 2014) and the natural feeding environment of the parrot is more like that of the primates than the fish. To account for their finding, Pepperberg and Hartsfield suggest that the fish and the parrots choose between the two plates with their mouth, whereas the primates choose with their hands. Given that primates have two hands, they may have a tendency to reach for both rewards at the same time, a strategy that would not be permitted with this task, so it may lead to confusion and disrupted learning. The fish and parrots, on the other hand, would naturally have to choose between the two options with their mouth. This distinction has received some support because when the task was modified for the primates by requiring them to make their choices on a computer screen, using a joystick (more like a single mouth), they showed learning of the optimal choice (Prétôt, Bshary, & Brosnan, 2016a).

Pigeons

To test the hypothesis that animals that have only a single means of choosing (i.e., with their mouth) could learn this task, whereas those that have the ability to choose both alternatives simultaneously (one with each hand) would have great difficulty, we tested pigeons with this task using two different procedures (Zentall, Case, & Luong, 2016). In the first experiment, we tested the pigeons using the manual presentation of the alternatives (as was done earlier with fish, parrots, monkeys, and apes), each on a uniquely colored disk, yellow or blue, on each of which was placed a single dried pea. Similar to the task used with other animals, if the pigeon chose the pea on one of the colors (e.g., the

yellow one) the other disk (i.e., the blue disk) was removed and the trial was over. But if the pigeons chose the pea on the blue disk, it was free to eat the pea on the yellow disk as well (the optimal color was counterbalanced). Surprisingly, not only were the pigeons unable to learn to choose the optimal (ephemeral) alternative that allowed them to have both peas but they showed a significant tendency to choose the suboptimal (permanent) alternative; the one that provided them with a single pea (see Figure 1). Apparently, choice with the mouth or beak was not the only distinguishing characteristic of those species that were able to easily acquire this task.

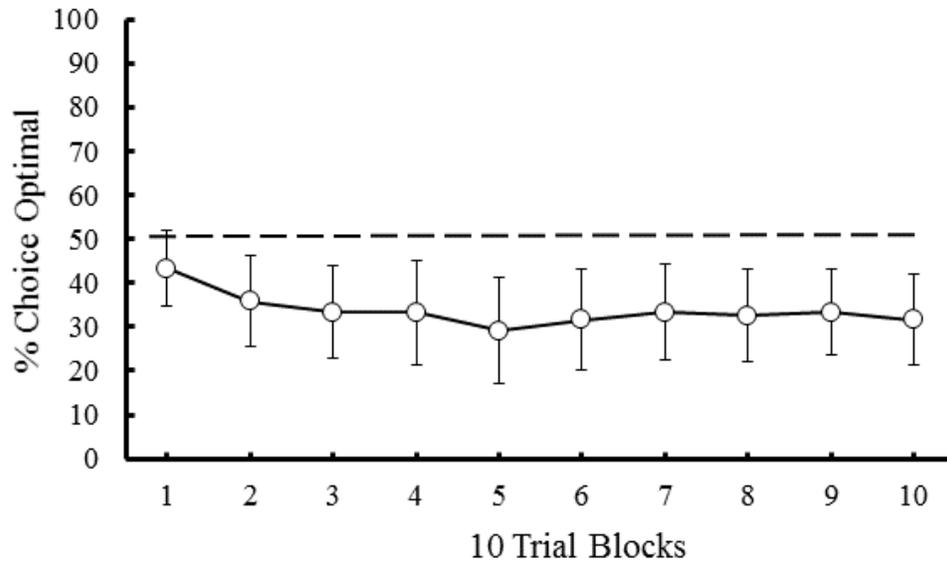


Figure 1. Pigeons' performance on the manually presented ephemeral reward task (error bars = \pm SEM; after Zentall et al., 2016, Experiment 1).

To test the generality of the suboptimal choice, in Experiment 2 we replicated the phenomenon with pigeons in an automated (operant) chamber in which choice of one color provided them with access to a grain feeder for 2.0 s and the trial was over, whereas choice of the other color provided them with access to the grain feeder for 2.0 s and they could then peck the other color to obtain a second 2.0 s access to the grain feeder. Once again we found a similar significant tendency to choose the suboptimal alternative, even after 400 trials of training (see Figure 2).

It appears that in both experiments the pigeons were not associating the second reinforcer with their initial choice. But why did they show a significant preference for suboptimal alternative? Earlier research with other animals referred to non-learners but did not indicate if there was a suboptimal preference. We considered two non-independent hypotheses. First, if at the start of training the pigeons sampled the two alternatives randomly from trial to trial, they would receive twice as many reinforcements directly associated with a peck to the suboptimal alternative. Consider two trials on which they initially chose each of the alternatives once. If they chose the optimal alternative they would receive one reinforcement associated with the optimal alternative and one reinforcement associated with the suboptimal alternative. But if they chose the suboptimal alternative they would receive only the reinforcement associated with the suboptimal alternative. Thus, considering the outcome of the two trials, they would have received two reinforcements associated with the suboptimal alternative, but only one reinforcement associated with the optimal alternative. Second, all trials ended with a response and reinforcement following a peck to the suboptimal alternative. Thus, any tendency to better remember the stimulus that preceded the last reinforcement obtained on a trial (a recency effect) would favor the suboptimal alternative (see Figure 3A).

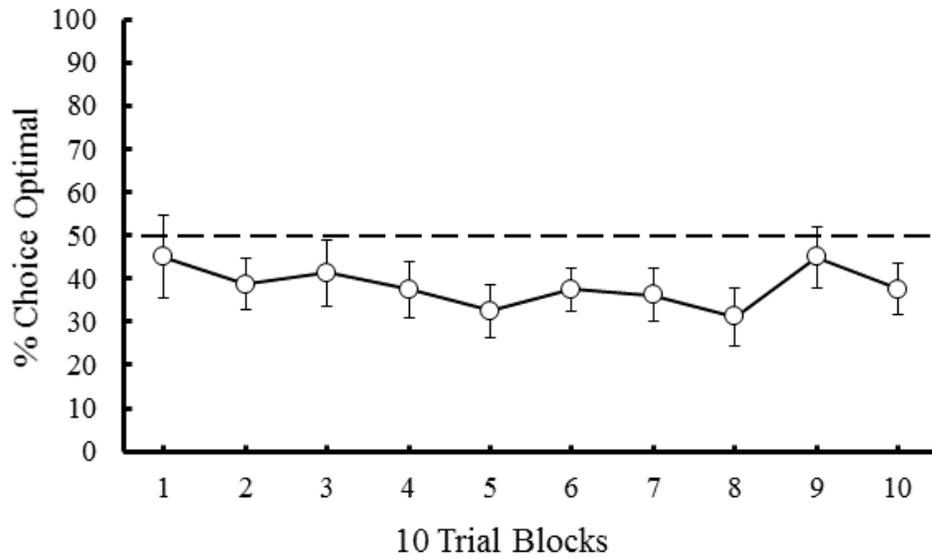


Figure 2. Pigeons' performance on the automated (operant) chamber ephemeral reward task (error bars = \pm SEM; after Zentall et al., 2016, Experiment 2).

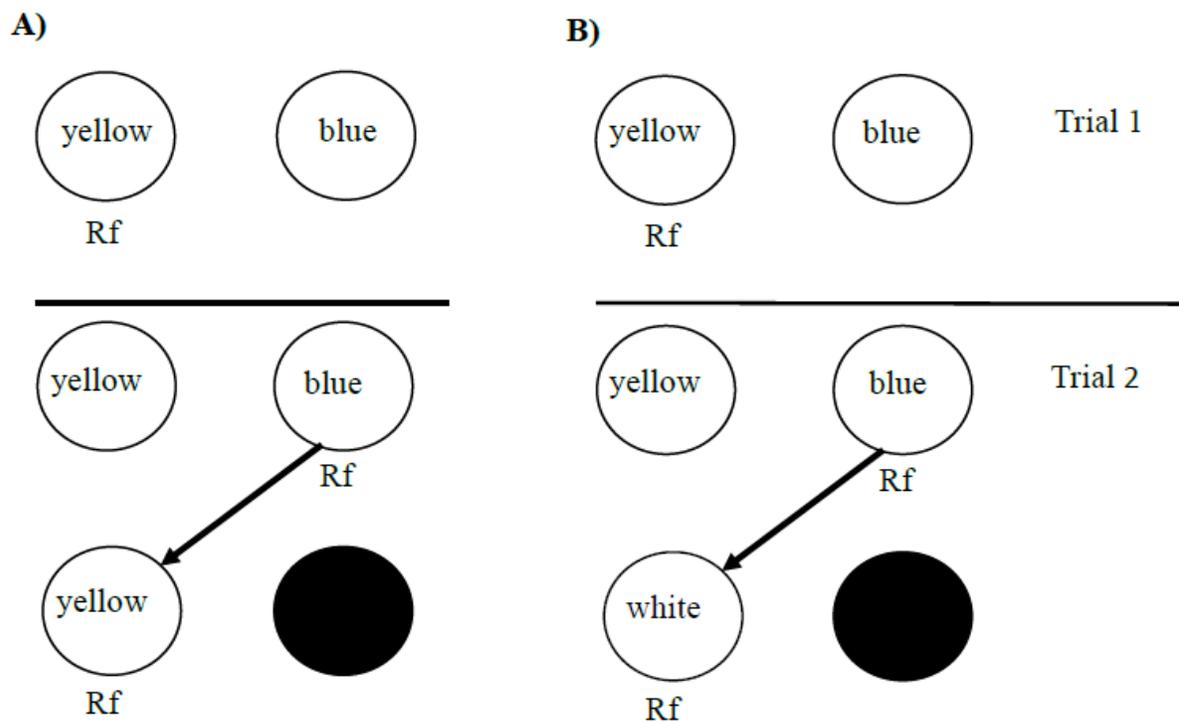


Figure 3. Design of Zentall, Case, & Luang (2016), Experiment 3: (A) Procedure for control group in which there are twice as many reinforcements associated with the suboptimal choice color and all trials end with a response to the suboptimal choice color. (B) Procedure for the experimental group in which an equal number of reinforcements are associated with the optimal and suboptimal choice color.

To test these hypotheses proposed to account for the preference for the suboptimal alternative, in Experiment 3, we arranged the contingencies in the operant box such that following initial choice of the optimal alternative and during reinforcement, the color of the suboptimal alternative changed to white and a peck to the white stimulus was reinforced (see Figure 3B). Thus, given one choice to each of the optimal and suboptimal alternatives, there should be no inherent bias because there would be one reinforcement associated with yellow, one with blue, and one with white.

Preferences by this group were compared to those of a new control group which received the same task as described in Experiment 2, with no change in color after choice of the optimal alternative. Once again, the control group showed a significant preference for the suboptimal alternative. Although the experimental group did not acquire a preference for the optimal alternative, it did choose the optimal alternative significantly more than the control group (see Figure 4). That is, the experimental group now chose the suboptimal and optimal alternatives about equally.

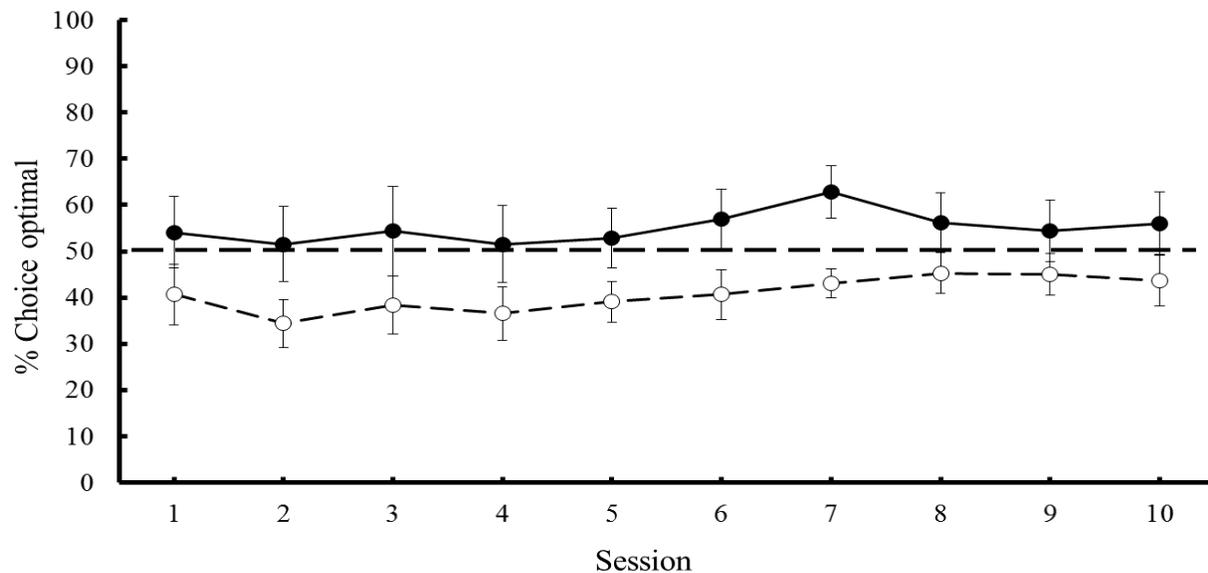


Figure 4. Pigeons' performance on the automated (operant) chamber ephemeral reward task. The control group (white circles) is a replication of Zentall et al., (2016, Experiment 2). For the experimental group (black circles) choice of the optimal alternative provided reinforcement and changed the color of the other alternative to white (error bars = \pm SEM; after Zentall et al., 2016, Experiment 3).

Rats

Given that primates and pigeons appeared to have difficulty with this task, whereas fish and parrots found the task relatively easy, we proceeded to ask if rats, a species used in much comparative psychological research, would be able to master this task (Zentall, Case, & Berry, 2017b, Experiment 1). For the rat experiment, we used retractable levers and to make the task a visual discrimination rather than a spatial discrimination, and we activated a light over one of levers (the optimal lever for half of the rats, the suboptimal lever for the rest). Although the rats too failed to learn to choose the optimal lever in over 800 trials, unlike the pigeons, they chose the two options about equally (see Figure 5).

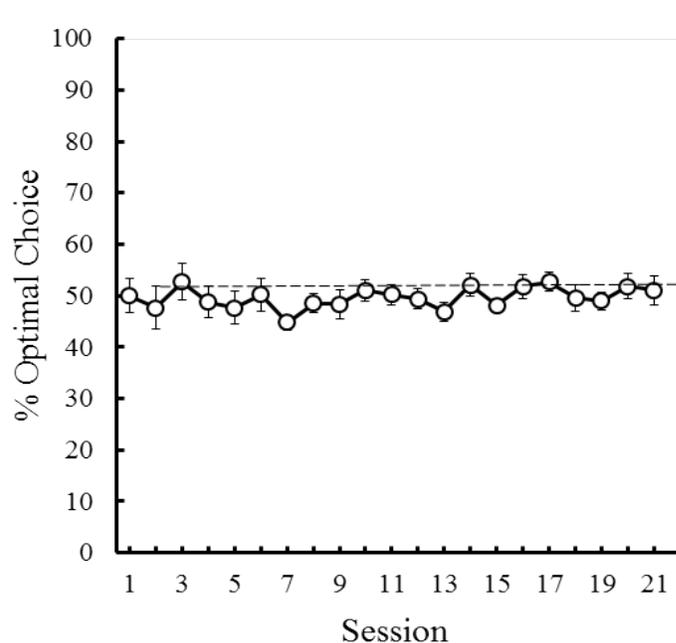


Figure 5. Rats' performance on the automated ephemeral reward task (error bars = \pm SEM; after Zentall et al., 2017b, Exp. 1).

Delay of Reinforcement in the Ephemeral Reward Task

For some reason, certain species (e.g., monkeys, apes, pigeons and rats) appear to be unable to easily associate the second reinforcement with choice of the first stimulus. That is, it appears that some animals treat the task as a choice between two immediately present reinforcers and fail to consider the relation between the second reinforcement and the original choice. It occurred to us that this effect may be indirectly related to the phenomenon of delay discounting (Ainslie, 1975). In delay discounting subjects are given a choice between a small immediate reinforcement (e.g., 1 pellet) and a larger but delayed reinforcement (e.g., 4 pellets). Although it is typically optimal to choose the larger later reinforcer, subjects often choose the smaller sooner one (Green & Meyerson, 1995). The ephemeral choice task is somewhat different because initially the two alternatives provide equal reinforcement and, given the optimal choice, the second reinforcement is somewhat delayed. That delay is generally quite short; on the order of 1 s with the manual presentation (Zentall et al., 2016, Exp. 1), and about 2.5 s in the operant chamber version of the task (Zentall et al., 2016, Exp. 2). Nevertheless, it appears to be sufficiently long to reduce the likelihood that the reinforcement that follows the response to the second stimulus on a trial becomes associated with the initial choice of the optimal alternative.

Impulsivity

A review of the delay discounting literature provides a possible procedural variation on delay discounting that may not only explain why certain species have so much trouble acquiring this task but also may identify conditions under which they can acquire it. Rachlin and Green (1972) describe a procedure in which pigeons that normally choose the suboptimal smaller-sooner reinforcer will show better "self-control" and prefer the larger-later reinforcer. To accomplish this, Rachlin and Green required the pigeons choose whether to make their choice of the smaller-sooner or larger-later after t s or to receive the larger later (without a choice) after t s. That is, the pigeons were allowed to make a commitment to the larger-later reinforcer so as not to present themselves with the opportunity to impulsively choose the smaller-sooner when it would become available. By avoiding the smaller-sooner reinforcer they could ensure that they would receive the larger-later reinforcer. When the prior commitment time was sufficiently long, most of the pigeons made the commitment for the larger-later reinforcer.

Pigeons

In the ephemeral reward task, the problem is that both alternatives appear to provide equal immediate amounts of food and the optimal alternative provides additional food but only after a delay. Would pigeons be more likely to integrate the two reinforcements if the time between choice and the first reinforcement was longer? In our next experiment (Zentall, Case, & Berry, 2017a) we inserted a delay between the initial choice and the first reinforcer (i.e., the pigeons had to make a prior commitment). In this experiment we again gave pigeons a choice between two alternatives, one optimal that allowed the pigeon to peck the second stimulus for a second reinforcer; the other suboptimal which provided only one reinforcer. For the experimental group, the initial choice turned off the other alternative and started a fixed-interval 20-s schedule, such that the first peck to the chosen stimulus after 20 s provided the first reinforcer. If the optimal alternative had been chosen, a single peck to the remaining stimulus provided a second reinforcer, if the suboptimal alternative had been chosen it also started a fixed-interval 20 s schedule followed by reinforcement but then the trial was over. To ensure that any change in the pigeon's choice did not result from the lengthening of the trial by 20 s, a control group was included for which all trials began with an orienting signal, a stimulus that had to be pecked on a fixed interval 20 s schedule before the standard choice was presented requiring a single peck. Furthermore, to ensure that the orienting signal itself could not account for any group differences in choice, an orienting signal was added to the experimental group but the required response to that signal (to produce the choice between the two alternatives) was a single peck. The design of this experiment appears in Figure 6.

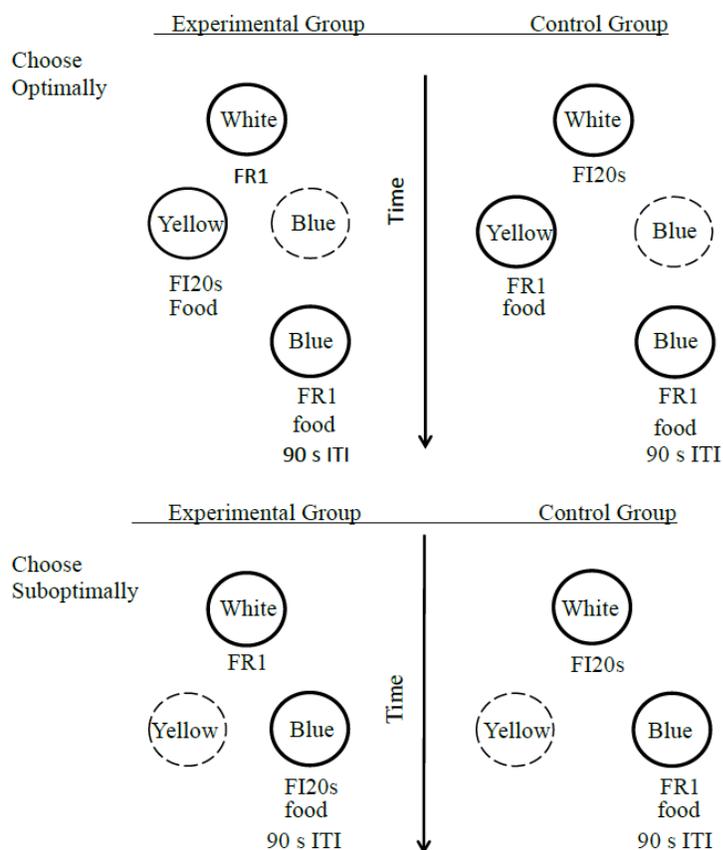


Figure 6. Design of the Zentall et al. (2017a) experiment: Events that occurred following an optimal choice (top) and following a suboptimal choice (bottom). Prior Commitment group on the left, No Prior Commitment group on the right. Optimal and suboptimal colors (yellow and blue) were counterbalanced over pigeons, and sides (left and right) were counterbalanced over trials. 90-s intertrial interval (ITI) separated the trials.

The results of this experiment were quite clear. On the one hand, the control group showed the characteristic significant preference for the suboptimal alternative, although with continued training (400 trials) choice of the suboptimal alternative by the control group approached indifference. On the other hand, the experimental group showed clear evidence of learning to make the optimal choice (see Figure 7). It too began by choosing suboptimally but within 70 trials it quickly reached indifference between the optimal and suboptimal alternatives and by the end of training (400 trials) the experimental group was choosing the optimal alternative 90% of the time. Apparently making a prior commitment eventually allowed the pigeon to integrate the two reinforcers that occurred when the optimal alternative was chosen.

The mechanism responsible for learning the delayed differential consequences of choosing optimally may be as simple as Weber's law, which states that the discrimination between two stimuli – in this case temporal duration – depends on the ratio of the difference between them to the absolute value of the stimuli. Thus, the difference between immediate reinforcement and a 1-2.5 s delay to the second reinforcement may be quite discriminable and difficult to associate with the initial choice, however, the difference between the 20 s fixed interval and the same 20 s plus 1-2.5 s may not be as discriminable. For this reason, at the time of choice, the two reinforcers associated with the optimal choice may be more closely represented as two reinforcements. Another way of looking at this effect is, when the choice involves a single peck for reinforcement, the immediacy of the first reinforcement may elicit an impulsive choice but delaying the first reinforcement may decrease the likelihood of an impulsive choice.

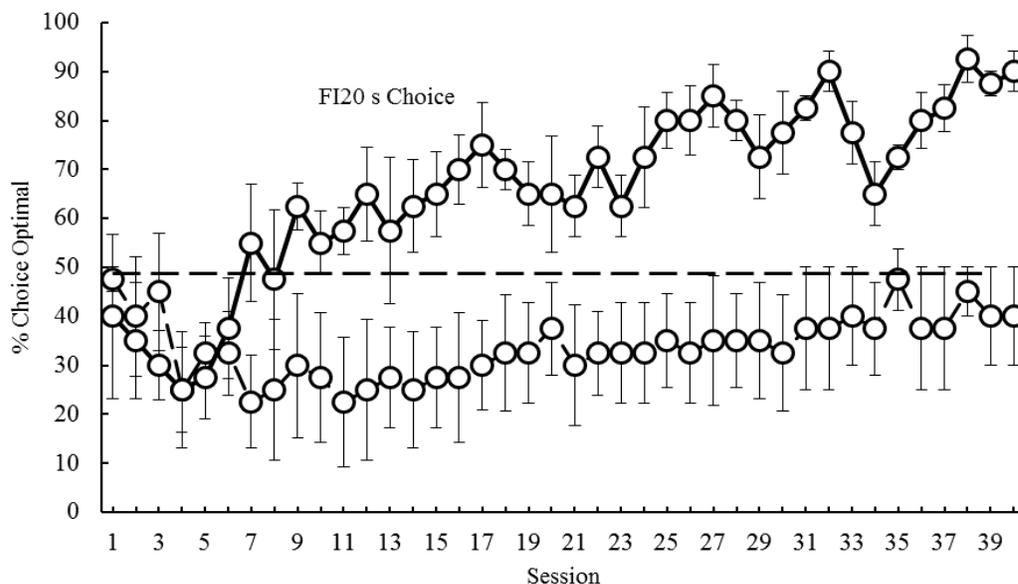


Figure 7. Percentage optimal choice for pigeons that had to complete a fixed interval 20-s schedule (FI 20s Choice) to obtain initial reinforcement and pigeons that had to make a single peck to obtain initial reinforcement (error bars = \pm SEM; after Zentall et al., 2017a).

Rats

If pigeons show more optimal choice when reinforcement is delayed, would rats benefit from the separation of their choice and the first reinforcement as well? In a follow-up ephemeral reward experiment we required rats, like the pigeons, to complete a fixed interval 20 s schedule to receive the first reinforcement but required only a single response to obtain the second (Zentall et al., 2017b, Experiment 2). The results were very similar to the results with pigeons (see Figure 8). It appears that the insertion of a delay between choice and reinforcement allowed the rats to become less impulsive and

choose more carefully. The findings with both pigeons and rats are somewhat counterintuitive as one more typically thinks of delay of reinforcement as not being conducive to optimal learning. For example, in a simple simultaneous discrimination, delaying reinforcement following choice generally leads to slower learning (Capaldi, 1978). In the ephemeral reward task, however, delaying reinforcement actually allows animals to learn to choose optimally.

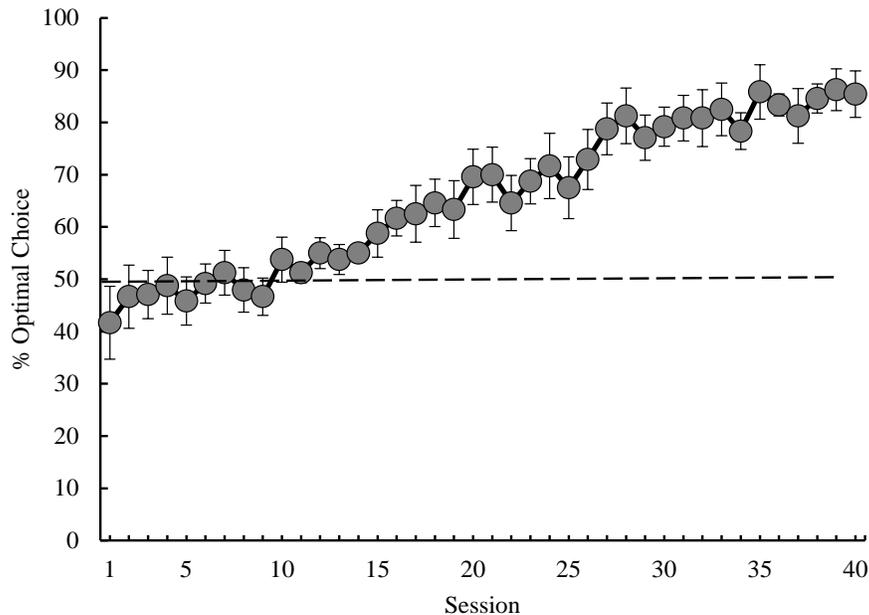


Figure 8. Percentage optimal choice for rats that had to complete a fixed interval 20 s schedule (FI 20s) to obtain initial reinforcement (error bars = \pm SEM; After Zentall et al., 2017b). Compare with the data presented in Figure 5.

Cleaner Fish

Although we have shown that pigeons and rats can learn to choose optimally when reinforcement following choice is delayed, other species, such as wrasse and parrots appear to choose optimally much more quickly when trained on this task and without the need for inclusion of the delay of reinforcement. It may be, however, that wrasse, fish that swim into the mouth of predatory fish, have learned to make choices carefully, as impulsive choices may have serious consequences because clients can punish the cleaners by terminating the cleaning interactions (Gingins, Werminghausen, Johnstone, Grutter, & Bshary, 2013) or even eating them. So the cleaners must often show ‘self-control’ by feeding against their preference, compared with other similar non-cleaner species (Gingins & Bshary, 2014). The fact that juvenile wrasse do not perform this task as well as adults suggests either that this cautious behavior must be learned or that it develops with maturation. It may be that animals, like wrasse, that do not choose impulsively are more likely to integrate the first and second reinforcement.

Parrots

But what about parrots? Or perhaps one should ask what about the parrots in the Pepperberg and Hartfield (2014) study because those parrots had had extensive prior training on a variety of tasks. One parrot had been exposed to “continuing studies on comparative cognition and interspecies

communication” (p. 299) whereas 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 the parrots in the Pepperberg and Hartfield study 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, Gray, & Pepperberg, 2015). It remains to be seen whether parrots that have not had the long history of training with other tasks would also show the same optimal choice with this ephemeral reward task.

Primates

In spite of their presumed superior intelligence and their ability to show considerable self-control in other contexts (Beran, 2015), with this task, primates do not readily learn to choose optimally when trained on the original ephemeral reward task. However, there is evidence that they can learn to choose optimally under certain conditions and those conditions are quite informative. Prétôt, Bshary, and Brosnan (2016a) adapted the ephemeral reward task by presenting the stimuli on a computer monitor and requiring monkeys to respond by moving a cursor to the selected stimulus using a joystick, receiving a reward at a pellet dispenser. Although the authors suggest that this procedure was more ecologically relevant to the primates, it is not obvious how this would be true but it did serve to separate the response from the reinforcement. Not only did the monkey’s hand not touch the stimulus directly but the reinforcer was not visible at the time of choice, and it appeared at a different location.

In another experiment, Prétôt, Bshary, and Brosnan (2016b) found that monkeys could learn to choose optimally if each reinforcer was placed under a distinctively colored cup. That is, the food was not visible at the time of choice. Once again, optimal choice could be trained by not allowing the monkeys to have direct access to the food and presumably reducing impulsive choice. In another experiment, the authors tested the monkeys with the original task with visible food but instead of the plates being distinctively colored, the food was distinctively colored pink or black. Again they found better acquisition of the optimal choice. The authors suggested that coloring the food made its physical properties the focus of attention, rather than the color of the plate on which the food was presented. It is also possible that the unusual color of the food may have reduced impulsive choice by the monkeys.

The separation of choice from reinforcement also may be relevant to the acquisition of other tasks. In an often cited study by Boysen, Berntson, Hannan, and Cacioppo (1996) a chimpanzee was trained on a *reverse contingency task*. The chimpanzee was offered a choice between two plates, both of which had candy but one always had more pieces than the other. Because of the reverse contingency, the chimpanzee always received the candy on the *unchosen* plate. As smart as this chimpanzee was, she was unable to consistently choose the plate with the fewer candies. As it happened, this chimpanzee also had been trained in the use of Arabic numerals to symbolically represent the number of objects in a set (Boysen & Berntson, 1989). That is, she had learned the association between Arabic numerals and the number of objects that each represented. Interestingly, when the task was changed such that the choice was between two plates, each one with an Arabic numeral, the chimpanzee quickly learned to choose the one with the smaller number on it in order to receive the number of candies represented by the unchosen Arabic numeral. Once again, it may be that removing the impulse to choose the larger number of candies by removing the candies from view, allowed the chimpanzee make the more optimal choice.

Summary

The ephemeral reward task is a relatively novel task that provides somewhat surprising results: wrasse and parrots acquire the optimal response quite easily, whereas primates, pigeons and rats do not. Whenever one encounters a task which appears that it should be relatively easy to acquire but rats, pigeons, and even primates have great difficulty learning it, the reason for the difficulty should be of interest to psychologists interested in learning. Furthermore, the fact that cleaner fish and parrots learn this task quite easily confirms that there is nothing inherently difficult about this task. Biologists may not

be surprised by these species differences and would likely attribute them to evolutionarily selected, genetic differences, together with the degree to which the species natural environment is compatible with the task and the laboratory environment. Psychologists, however, interested in the mechanisms of learning, would likely consider the pigeons', rats', and primates' difficulty in acquiring this task to be examples of "meaningful failures" because they seem surprising given what is known about the learning abilities of these species.

The ecological account proposed by (Salwiczek et al., 2012), that wrasse need to learn to service ephemeral visitors to the reef before servicing resident clients, does not appear to account for the parrots' ability to acquire the optimal response (Pepperberg & Hartsfield, 2014) an environment more similar to that of the primates than of the fish. Furthermore, the hypothesis that fish and parrots choose with their mouth whereas primates choose with their hands does not appear to easily distinguish species that can acquire the optimal response from those that cannot because pigeons, which choose with their beaks, do not easily acquire the optimal response, and they are not any better than rats. We proposed that impulsive choice may be responsible for the failure to integrate the first and second reinforcer when the optimal choice is made. Using the commitment model proposed by Rachlin and Green (1972), we forced pigeons and rats to choose 20 s before the first reinforcement was presented and found that under those conditions both species learned to choose optimally. It would be instructive to determine if other species, such as monkeys and apes, species that have difficulty choosing optimally with the ephemeral reward task as originally trained (Salwiczek et al., 2012) would also benefit from the temporal separation of choice from reinforcement.

We believe that when animals have great difficulty learning tasks that should be relatively easy, it may be that impulsive choice is involved. And paradoxically, although delay of reinforcement is typically thought to retard task acquisition, in certain cases, the separation of the *choice* response from reinforcement may actually facilitate learning. The strong attraction to immediate reinforcement over more optimal delayed reinforcement is what accounts for the strong preference for smaller-sooner over larger-later reinforcement in delay-discounting research (Ainslie, 1975). It also appears to explain the strong suboptimal attraction to the appearance of conditioned reinforcers when delay to reinforcement is controlled (see also, Zentall, Andrews, & Case, 2017).

References

- Ainslie, G., (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82, 463–496.
- Beran, M. J. (2015). The comparative science of "self-control": What are we talking about? *Frontiers in Psychology*, 6, 51.
- Bitterman, M. E. (1965). Phyletic differences in learning. *American Psychologist*, 20, 396–410.
- Bitterman, M. E. (1975). A comparative analysis of learning. *Science*, 188, 699–709.
- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, 103, 23–31.
- Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1996). Quantity based interference and symbolic representations in chimpanzees (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 76–86.
- Bshary, R., & Grutter, A. S. (2002). Experimental evidence that partner choice is a driving force in the payoff distribution among cooperators or mutualists: The cleaner fish case. *Ecology Letters*, 5, 130–136.
- Capaldi, E. J. (1978). Effects of schedule and delay of reinforcement on acquisition speed. *Animal Learning & Behavior*, 6, 330–334.
- Gingins, S., & Bshary, R. (2014). Pairs of cleaner fish prolong interaction duration with client reef fish by increasing service quality. *Behavioural Ecology*, 26, 350–358.
- Gingins, S., Werminghausen, J., Johnstone, R. A., Grutter, A. S., & Bshary, R. (2013). Power and temptation cause shifts between exploitation and cooperation in a cleaner wrasse mutualism. *Proceedings of the Royal Society B*, 280, 20130553.
- Green, L., & Myerson, J. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, 64, 263–276.

- Koepke, A. E., Gray, S. L., & Pepperberg, I. M. (2015). Delayed gratification: A grey parrot (*Psittacus erithacus*) will wait for a better reward. *Journal of Comparative Psychology*, *129*, 339–346.
- Pavlov, I. (1927/1960). *Conditioned reflexes*. New York: Dover.
- Pepperberg, I. M., & Hartsfield, L. A. (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*)? *Journal of Comparative Psychology*, *128*, 298–306.
- Prétôt, L., Bshary, R., & Brosnan, S. F. (2016a). Comparing species decisions in a dichotomous choice task: Adjusting task parameters improves performance in monkeys. *Animal Cognition*, *19*, 819–834.
- Prétôt, L., Bshary, R., & Brosnan, S. F. (2016b). Factors influencing the different performance of fish and primates on a dichotomous choice task. *Animal Behaviour*, *119*, 189–199.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, *17*, 15–22.
- Salwiczek, L. H., Prétôt, L., Demarta, L., Proctor, D., Essler, J., Pinto, A. I., 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.
- Thorndike, E. L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *Psychological Review Monograph Supplement*, *2*, 1–109.
- Zentall, T. R., Andrews, D. M., & Case, J. P. (2017). Prior commitment: Its effect on suboptimal choice in a gambling-like task. *Behavioural Processes*, *145*, 1-9.
- Zentall, T. R., Case, J. P., & Berry, J. R. (2017a). Early commitment facilitates optimal choice by pigeons. *Psychonomic Bulletin & Review*, *24*, 957–963.
- Zentall, T. R., Case, J. P., & Berry, J. R. (2017b). Rats' acquisition of the ephemeral reward task. *Animal Cognition*, *20*, 419–425.
- Zentall, T. R., Case, J. P., & Luong, J. (2016). Pigeon's paradoxical preference for the suboptimal alternative in a complex foraging task. *Journal of Comparative Psychology*, *130*, 138–144.