Enhancing "self-control": The paradoxical effect of delay of reinforcement

Thomas R. Zentall [1]

Abstract

Delay of reinforcement is generally thought to be inversely correlated with speed of acquisition. However, in the case of simultaneous discrimination learning, in which choice results in immediate reinforcement, delay of reinforcement can improve acquisition. For example, in the ephemeral reward task, animals are given a choice between two alternatives, A and B. Choice of A provides reinforcement, and the trial is over. Choice of B provides reinforcement and access to alternative A (thus, two reinforcements). Many animals appear unable to learn to choose B consistently, but inserting a 20-s delay between choice and outcome has been shown to facilitate optimal choice. Similarly, pigeons given a choice between a signal for one pellet and a signal for two pellets (each occurring without a delay) have difficulty learning to choose the two-pellet alternative, unless the reinforcement is delayed. In a version of object permanence, food is placed in one of two containers, and the pigeon must choose the container with the food. Pigeons have difficulty reliably choosing the correct container unless a brief delay is inserted between baiting and choice. Finally, pigeons have been shown to prefer a suboptimal alternative (a 20% chance of getting a cue for reinforcement) over an optimal alternative (a 100% chance of getting a cue for 50% reinforcement). However, if pigeons are forced to wait 20 s following their choice to receive the cues, no preference for the suboptimal alternative is found. Thus, impulsive choice may be reduced by delaying the consequence of that choice.

Keywords Delay of reinforcement · Ephemeral reward · Object permanence · Magnitude of reinforcement discrimination · Suboptimal choice

In his research comparing the predictions of optimal foraging theory to laboratory research on schedules of reinforcement, Stephen Lea (1979) showed the strong preference that animals have for reinforcement that comes with shorter delay over longer delay, even when the density of reinforcement favors the longer delay reinforcement. Delay of reinforcement has long been considered a primary determinant of the effectiveness of a reinforcer (e.g., Catania, 1979; Kimble, 1961). Historically, response acquisition has been viewed as being negatively correlated with the interval between a response and reinforcement (Thorndike, 1911), although it also has been suggested that delay of reinforcement affects performance, but has little effect on learning (Warden & Haas, 1927; see also Watson, 1917). Hull (1952) proposed that different findings on the effect of delay of reinforcement on learning could be explained in terms of the degree to which conditioned reinforcers might bridge the delay between the response and the reinforcer (see also Spence, 1947). With regard to discrete trial discrimination learning, studies often have found that delay of reinforcement retards learning (Carlson & Wielkiewicz, 1972; Culbertson, 1970; Keesey, 1964).

Although the earlier research was conducted with rats in the context of discrete-trial procedures, similar results have been found with pigeons using operant procedures (Baum, 1973; Fantino, 1969; Mazur, 1987). In the context of operant conditioning, it appears that the effect of reinforcement delay depends on what the animal is doing during the delay (e.g., pecking at a lit key or not pecking at a lit key; see Lattal, 2010, for a review).

Delay discounting and the effect of a prior commitment

A clear effect of delay of reinforcement can be seen in the delay discounting effect found in many species, including

✉ Thomas R. Zentall
zentall@uky.edu

[1] Department of Psychology, University of Kentucky, Lexington, KY 40506-0044, USA

pigeons, rats, and humans. With this task, animals have a choice between two alternatives—one that provides a small amount of food after a short delay, the other that provides a larger amount of food, but after a longer delay. For example, when pigeons were given a choice between eight pellets immediately or 16 pellets after a short delay, they became indifferent between the two alternatives when the 16-pellet alternative was delayed by as little as 10 s (Oliveira, Green, & Myerson, 2014). The relation between delay of reinforcement and magnitude of reinforcement is well described by the hyperbolic discounting function in which V is the present value of a future reward, A is the amount of the reward, D is the delay of the reward, and k is the slope of the discounting function:

$$V = \frac{A}{1 + kD}. \tag{1}$$

An interesting characteristic of the hyperbolic delay discounting function can be seen in Fig. 1. When the delay to the smaller, sooner reward is short, it may have greater value than the larger, later reward; however, when the delay to the smaller, sooner reward is longer, the value of the larger, later reward may be greater.

This relation is more intuitive if viewed as a function of Weber's Law. Let us say that at Choice Time 1, one pellet, if delayed by 0.5 s, is preferred to four pellets, if delayed by 5 s. The ratio of the sooner to the later would be 1 to 10, but if the choice is made at Choice Time 2, 10 s earlier, the delay to sooner would be 10.5 s and the delay to later would be 15 s, a ratio of 10.5 to 15, or close to 2 to 3. Thus, according to this theory, the values of the two delays should be more similar,
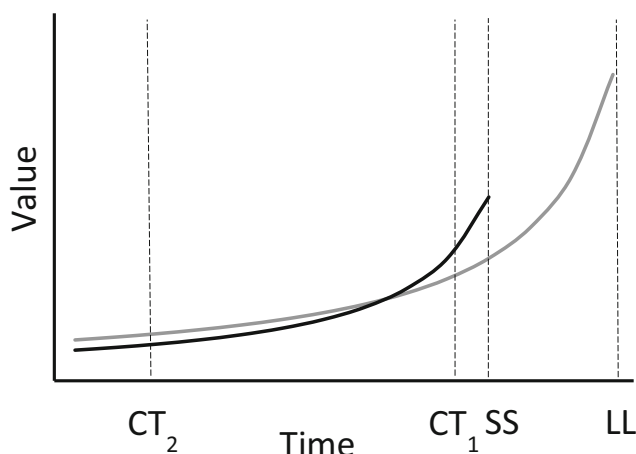

Fig. 1 Hyperbolic delay discounting functions representing the loss of value for each outcome as a function of the delay between choice time (CT) and either the smaller sooner reward (SS) or the larger later reward (LL). CT_1 = choice at time 1; CT_2 = choice at time 2. *Note.* At CT_1 the value of SS is greater than the value of LL, but at CS_2 the value of LL is greater than the value of SS

and magnitude of reinforcement should play a larger role in choice. Rachlin and Green (1972) showed that getting pigeons to make a prior "commitment" got them to switch from a preference for the smaller, sooner to that of the larger, later.

The positive effect of delayed reinforcement

Delay discounting is often associated with impulsivity (Odum, 2011) or the lack of self-control because, in general, a preference for the larger, later would result in a higher rate of reinforcement. It is easy to speculate, however, how impulsivity might have been evolutionarily selected. After all, nature is generally not able to "promise" a larger, later reinforcement. In nature, there is completion from other conspecifics, especially in animals that tend to forage socially. In nature, for many species the adage *he who hesitates is lost* is likely to apply.

On the other hand, in our Western culture, we tend to value self-control because important rewards often require self-control (e.g., getting a college degree, withholding aggression, avoiding large credit card debt) and, at least in our modern environment, future rewards can be made more predictable. However, our hunter-gatherer ancestors may have needed to be somewhat impulsive as they foraged for food, especially when they encountered escaping prey.

With the delay discounting procedure, the effect of making a prior commitment (delaying reinforcement) follows from the nature of the hyperbolic discounting function. If the choice is made earlier, the value of the smaller, sooner is predicted to cross over the value of the larger, later, and such a crossover has been found by (Rachlin & Green, 1972). Although in the case of prior commitment, the added delay likely devalues both alternatives; because the delay function is hyperbolic, it devalues the smaller reinforcer faster than the larger one.

The principle of prior commitment might be applicable to other designs in which animals are given a choice between two different magnitudes or probabilities of reinforcement, but they are not choosing optimally. In what follows, I will describe the results of several experiments in which we have found that pigeons (and sometimes rats) have difficulty learning to choose the larger reinforcer when an objectively smaller reinforcer is available. The tasks are sometimes quite different, but what most of them have in common is a choice in which both alternatives are associated with reinforcement, but one is clearly a better outcome than the other.

The ephemeral reward task

First reported by Bshary and Grutter (2002), the ephemeral reward task involves presentation to a subject of two alternatives (plates), each of which contains a small bit of food. If the subject chooses Alternative A, it gets the food from that plate

and the trial is over. If the subject chooses Alternative B, it gets the food from that plate, and it also can eat the food on the other plate, Alternative A. Perhaps, not surprisingly, the subjects, in this case wrasse (cleaner fish), learn to choose Alternative B quickly (within 100 trials). What is surprising is that several species of primates do not learn to choose Alternative B within 100 trials (Salwiczek et al., 2012). The authors suggest that the cleaner fish have a natural tendency to acquire such a task because they live on reefs and they clean the mouths of larger fish, some of which also live on the reef, but others who merely visit the reef. The authors propose that the cleaner fish must learn to swim out and service the visitors because they are transitory and will quickly leave (the visitors correspond to Alternative B), whereas those that live on the reef will remain and can be serviced later (residents correspond to Alternative A). The primates have had no such experience, and thus they find the ephemeral reward task difficult.

The authors propose that the cleaner fish generalize from servicing large fish on or near the reef where they live to eating from plates provided to them in laboratory fish tanks. Such generalization seems quite remarkable. It must have seemed so as well to Pepperberg and Hartsfield (2014), who repeated the experiment with African grey parrots, whose natural ecology is quite similar to that of primates and is quite different from that of the wrasse. Yet, like the wrasse, the parrots acquired the ephemeral reward task quite easily.

Pepperberg and Hartsfield (2014) noted that both the fish and the parrots made their choice with the mouth or beak, whereas the primates chose with their hand. It is not clear why that would make a difference, but with two hands, perhaps there is some tendency for the primates to attempt to choose both alternatives, one with each hand.

We tested the one-beak versus two-hand hypothesis by conducting a similar experiment with pigeons with different-colored cues (Zentall, Case, & Luong, 2016). Surprisingly, not only did the pigeons fail to acquire the optimal two-reward choice, but they actually showed a significant preference for the suboptimal one-reward alternative.

Careful examination of the task identified a potential artifact. With this task, all trials involved a response to Alternative A. Choice of A ended the trial and choice of B ended with a response to A. Thus, initially, there would have been twice as many reinforcements associated with choice of the suboptimal Alternative A as with the optimal Alternative B. In a follow-up experiment, we replicated the result in an automated operant box in which pecking at either of two colored lights resulted in reinforcement from the feeder below.

To test the artifact hypothesis, we reran the experiment, but when the pigeon chose the optimal alternative, during the time of reinforcement we replaced the color of the optimal alternative with a third color, C, and a response to C provided a second reinforcement. Thus, choice of the optimal alternative no longer ended with reinforcement associated with the suboptimal

choice, A. This group showed a significant reduction in choice of the suboptimal alternative, but these pigeons still failed to acquire a significant preference for the optimal alternative.

The inability to maximize reinforcement with the ephemeral reward task appears to be unrelated to generally held notions of animal intelligence (Bitterman, 1975) and not directly attributable to the natural ecology of the species tested. In spite of its apparent simplicity, although wrasse and parrots easily acquired the task, primates and pigeons did not. To test the generality of this failure, Zentall, Case, and Berry (2017b) repeated the original experiment with rats as subjects and showed that they too did not show a preference for the optimal alternative.

The failure of several species to acquire the optimal choice response with the ephemeral reward task brought to mind the results of delay discounting research, which also show consistent suboptimal choice under a variety of conditions. Rachlin and Green (1972) found that suboptimal choice can be reduced by using a prior commitment procedure, in which choice of the smaller, sooner is reduced by making the pigeons choose at an earlier time—that is, by eliminating the immediacy of the outcome of the suboptimal alternative.

To test this hypothesis, much like with the commitment procedure, Zentall, Case, and Berry (2017a) imposed a 20-s delay (using a fixed-interval 20-s schedule) between the pigeons' choice and the first reinforcement. If the optimal alternative was chosen, following reinforcement, the other alternative appeared, and a single peck provided a second reinforcement. With this commitment procedure, the pigeons developed a strong preference for the optimal alternative (see Fig. 2). To test the generality of this finding, Zentall et al. (2017b) tested the procedure with rats and found a similar result. Now the rats too developed a preference for the optimal alternative. Thus, imposing a delay between choice and reinforcement facilitated optimal choice with this task.

One might interpret the results of the original ephemeral reward task as being an example of delay discounting because in the ephemeral reward task, the second reinforcement is somewhat delayed. To do so, however, one would have to argue that although the immediate outcome following choice is the same regardless of the choice, the short delay between the first reinforcement and the second reinforcement (about 1 s; see Zentall et al., 2016, Experiment 1) is sufficient to act like a larger (i.e., extra) later reinforcer. The delay discounting interpretation may have more difficulty accounting for the results of the one versus two pellet experiment presented in the following section.

Discrimination of large versus small reward

Several years ago, we attempted to train pigeons on a simultaneous discrimination in which choice of a blue light resulted
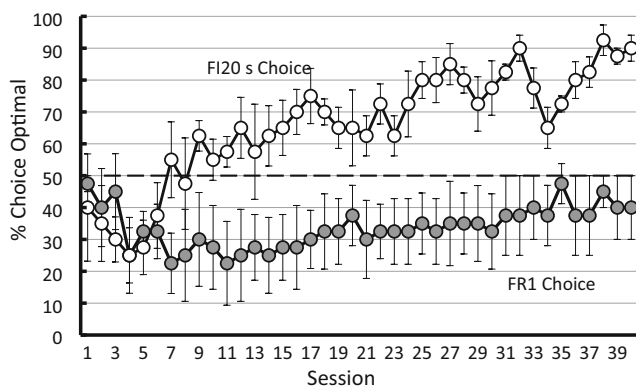
Fig. 2 Ephemeral reward task: Pigeons have a choice between two alternatives. If they choose suboptimally, Alternative A, they get rewarded and the trial is over. If they choose optimally, Alternative B, they get rewarded and they can also have the reward associated with A. Thus, A is worth 1 reward and B is worth 2 rewards. In the FR1 choice condition, a single peck is required to make their initial choice, followed immediately by reinforcement. In the FI20s choice condition, the first peck after 20 s following their initial choice is followed by reinforcement. After Zentall, Case and Berry (2017a)

in feeder access of 1.5 s, and choice of a simultaneously presented red light resulted in feeder access of 3.0 s. Surprisingly, we found that the pigeons had great difficulty learning the discrimination.

Recently, House, Peng, and Zentall (2020) returned to this magnitude-of-reinforcement discrimination. We hypothesized that the problem may have been that the pigeons were having trouble discriminating between the durations of reinforcement associated with each color because the immediate effect of choice of each color was exactly the same (access to the feeder). Perhaps the discrimination would be easier if the discrimination was between one and two pellets of food because the pigeons could see the difference in the magnitude of reinforcement.

In light of the results of the ephemeral reward task, however, we decided to include a group of pigeons for which the pellet outcomes were delayed, by requiring 10 pecks to either color. The results were quite striking. The control group for which a single peck was required to produce the outcome associated with each color showed little sign of learning to choose the color associated with the two-pellet alternative, whereas the experimental group for which 10 pecks were required learned the discrimination (see Fig. 3). We hypothesized that the immediate reinforcement following choice made the two outcomes difficult to discriminate. It is also possible, however, that requiring 10 pecks allowed for better processing of the stimuli (see, e.g., Elsmore, 1971; Roberts, 1972) or, alternatively, resulted in contrast between the greater effort and reinforcement (a so-called justification of effort effect; Clement, Feltus, Kaiser, & Zentall, 2000; Friedrich, Clement, & Zentall, 2005; Friedrich & Zentall, 2004). Paradoxically, by delaying both outcomes, the pigeons learned to obtain a greater amount of food.
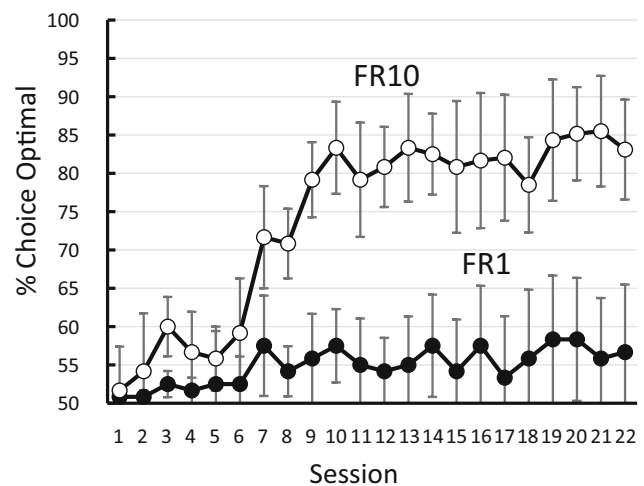

Fig. 3 The one-pellet versus two-pellet discrimination task: When a single peck was required, and choice of one alternative provided pigeons immediately with one pellet, whereas choice of the other alternative provided pigeons immediately with two pellets, they showed little evidence of optimal choice. When 10 pecks were required to either alternative, however, the pigeons learned to choose the optimal two-pellet alternative. After House, Peng, and Zentall (2020)

Object permanence

Various forms of object permanence have been used to track the cognitive development of young children (Piaget, 1963). Object permanence is assumed to involve the understanding that when an object is placed into a container or behind an occluder, it continues to exist and can be found. In the simplest form of object permanence, visible displacement, in the presence of the subject, an object may be placed in one of two containers and the subject is free to recover the object. In the more difficult form of object permanence, invisible displacement, after the object is placed in one of the containers, the container with the object inside is moved, thus invisibly displacing the object. One way this has been tested with children is by placing a container at each end of a beam that can be rotated, and after placement of the object, the beam is rotated (Bai & Bertenthal, 1992). In the easier form of invisible displacement, the beam is rotated 90°, so the orientation of the beam is now different. In the harder version, the beam is rotated 180°, so the container with the object and the empty container exchange places (Barth & Call, 2006). Research with dogs (Miller, Gipson, Vaughan, Rayburn-Reeves, & Zentall, 2009) has found that they have no problem with the visible displacement, and they do very well with the 90° invisible displacement, but they do not appear to be able to follow the 180° rotation.

We have recently conducted a similar experiment with pigeons and found them to have great difficulty with the simplest, visible displacement form of the task, even after many sessions of training (Zentall & Raley, 2019). Observation of the pigeons suggested that they readily learned to associate the

sight and sound of grain being placed into one of the containers. They would flap their wings and became highly agitated, but they chose the container with the food at low levels of accuracy above chance. Could the relative immediacy of reinforcement following baiting cause the pigeons to choose impulsively? If so, could we get a better measure of object permanence involving visible displacement if we imposed a brief delay between baiting the container and choice?

In a follow-up experiment with new birds, conducted the same way as the original experiment but with a 5-s delay between baiting and test, we found that the pigeons could learn which container was baited. But learning which container has the food does not demonstrate object permanence, because the pigeons could have learned to use the sight of the experimenter's hand or the location of the sound of the grain falling into the container as a cue to choose that container. That is, the original test (the first few trials) is the only appropriate measure of object permanence because object permanence should occur without training.

Following acquisition of the visible displacement form of the task, we transferred the pigeons to the 90° invisible displacement, again with a 5-s delay between baiting and test. Importantly, the pigeons were highly accurate on the immediate transfer test. Finally, we transferred the pigeons to the 180° invisible displacement, and, surprisingly, they transferred at a high level of accuracy. Thus, inserting a 5-s delay between baiting the container and testing the birds for object permanence improved the performance of pigeons on the various forms of the object permanence task.

This finding is particularly interesting in that it is quite different from the examples previously discussed. That is, in the earlier research, it was a choice between good and better, such that there were no "incorrect" choices. Instead, the object permanence task is more like a typical simultaneous discrimination, with one stimulus associated with reinforcement and the other stimulus associated with the absence of reinforcement, yet the pigeons were not able to learn the discrimination without the inserted delay. This finding is consistent with the idea that the immediacy of reinforcement is responsible for the failure to choose accurately on this task.

The commitment procedure developed in the context of delay discounting was the motivation for the introduction of the delay in the various examples provided in the preceding sections. It is not likely, however, that the crossover in the assumed hyperbolic delay functions responsible for the success of the commitment procedure is involved in the effect of delay on acquisition of the one-pellet versus two-pellet task, the object permanence task, or even the ephemeral reward task. Instead, I propose that when choice leads to immediate reinforcement, it often leads to impulsive choice, and the introduction of a delay between choice and reinforcement leads to better self control.

In the next section, I will describe one more example of the effect of an added delay on the reduction in suboptimal choice.

The task has to do with a form of suboptimal choice related to unskilled gambling behavior, such as buying lottery tickets or playing roulette or slot machines.

Suboptimal choice in a gambling task

For several years, we have been studying a task in which pigeons and rats show suboptimal choice, but in this task, it not just that the pigeons do not learn to make the optimal choice. Instead, the pigeons show a strong preference for the suboptimal choice (Stagner & Zentall, 2010; see also Mazur, 1989; Spetch, Belke, Barnet, Dunn, & Pierce, 1990). In the version of this task most similar to unskilled gambling by humans, pigeons are given a choice between two alternatives. One alternative, 20% of the time, provides a cue that signals they will get 10 pellets of food, but 80% of the time they get a signal that no food will be coming. The optimal alternative provides a cue that signals that they will always get three pellets of food. Thus, their choice is between an average of two pellets of food and a sure three pellets of food (Zentall & Stagner, 2011). Surprisingly, the pigeons showed a strong preference for the two-pellet alternative.

Pigeons also show a similar suboptimal preference when the outcomes involve different probabilities of reinforcement (Stagner & Zentall, 2010). For example, they prefer an alternative that 20% of the time gives them a signal that they will always be fed, to an alternative that 100% of the time gives them a signal that they will get food 50% of the time. The research suggests that it is not the probability or magnitude of reinforcement associated with initial choice that determines the preference, but instead the probability or magnitude of reinforcement associated with the signals for reinforcement that follow. Interestingly, the signal for the absence of reinforcement that occurs on 80% of the choices of the suboptimal alternative fails to inhibit choice of that alternative.

Consistent with this hypothesis, tests of the value of the signal for the absence of reinforcement using a combined cue test failed to show that it functions as a conditioned inhibitor (Laude, Stagner, & Zentall, 2014). In the combined cue test, the presumed inhibitory stimulus is presented in combination with a known excitatory stimulus, and the reduction in responding, relative to the excitatory stimulus by itself, is taken as a measure of inhibition (Rescorla, 1969).

Further support for the hypothesis that the probability of reinforcement associated with the signals for reinforcement that follow choice is responsible for suboptimal choice comes from a design in which 50% of the time one alternative provides a signal for reinforcement, whereas 100% of the time the other alternative provides a signal for reinforcement (Smith & Zentall, 2016). Consistent with the signal value hypothesis, as both signals for reinforcement are perfect predictors of

reinforcement, the pigeons are indifferent between the two alternatives.

In a follow-up experiment that extended training for 75 sessions, we found that under similar conditions pigeons gradually developed a significant preference for the suboptimal alternative (Case & Zentall, 2018). This finding led us to propose that in addition to the value of the signals for reinforcement, pigeons' choice of the suboptimal alternative is also affected by the contrast between the expected value of reinforcement associated with the initial choice and the value of signal for reinforcement that followed. Curiously, the pigeon should expect 50% reinforcement for choice of the suboptimal alternative, but when it occurs, the appearance of the cue for reinforcement signals 100% reinforcement. Hence, there is positive contrast. On the other hand, given choice of the optimal alternative, the pigeon would expect 100% reinforcement, and the appearance of the signal for 100% reinforcement involves no contrast. Thus, even when the optimal alternative involves no uncertainty (100% reinforcement), pigeons develop a preference for the suboptimal alternative that provides reinforcement only half of the time.

In the suboptimal choice (gambling) task, there is already a delay between the initial choice alternative and reinforcement, so impulsive choice would not appear to be an issue. In most of the procedures used (e.g., Stagner & Zentall, 2010), however, the signal for reinforcement associated with the suboptimal alternative appears immediately following the choice response (typically one peck). Thus, one could view the suboptimal choice as sometimes providing an immediate conditioned reinforcer.

If inserting a delay between initial choice and reinforcement facilitates the acquisition of optimal choice using the several procedures outlined in the preceding sections, could it also affect suboptimal choice when applied to the delay between
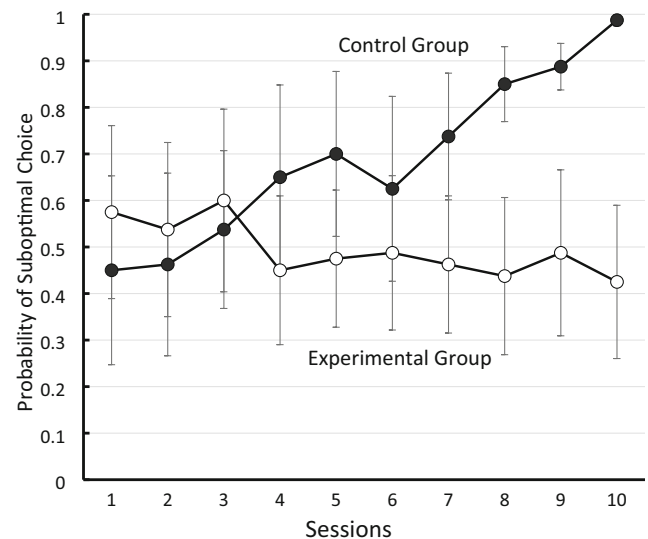

Fig. 5 Suboptimal choice (gambling) task. Pigeons in the experimental group were required to complete a fixed-interval 20-s schedule following their choice to obtain reinforcement. Control pigeons made a single peck to obtain reinforcement. After Zentall, Andrews, and Case (2017)

choice and the conditioned reinforcer in this gambling-like task? Zentall, Andrews, and Case (2017) tested this hypothesis using a design in which choice of the suboptimal alternative was followed by signaled reinforcement 25% of the time, whereas choice of the optimal alternative was followed by unsignaled reinforcement 75% reinforcement (see Fig. 4). For pigeons in the experiment group, choice of either alternative initiated a fixed interval 20-s schedule, at the end of which the stimulus signaling reinforcement (or its absence) appeared. For pigeons in the control group (with trial duration equated), choice of either alternative led immediately to the scheduled signaling stimulus. Pigeons in the control group showed the
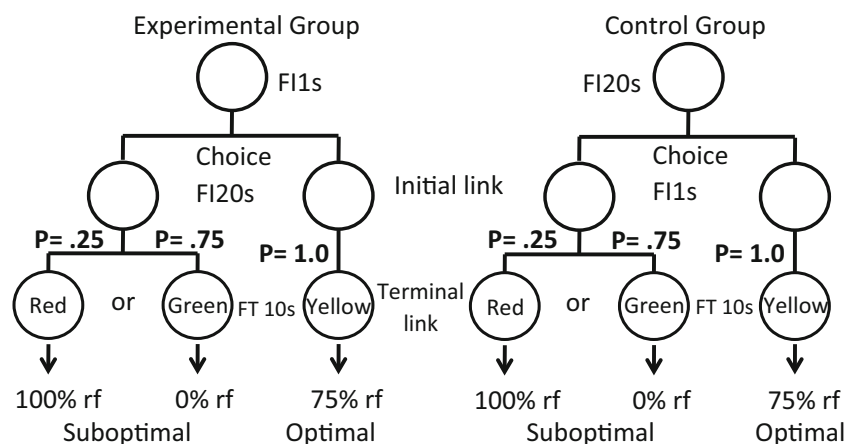

Fig. 4 Design of the suboptimal choice (gambling) task. Pigeons in the experimental group made their choice by completing a fixed-interval 20-s schedule. Pigeons in the control group made their choice by completing a brief fixed-interval 1-s schedule. To attempt to equate the trial duration for both groups, pigeons in the control group started each trial by completing a fixed-interval 20-s schedule to the center key, whereas pigeons in the experimental group started each trial by completing a fixed-interval 1-s schedule to the center key. After Zentall, Andrews, and Case (2017)

typical strong preference for the suboptimal alternative, whereas those in the experimental group were relatively indifferent between the two alternatives (see Fig. 5). Although the delay did not eliminate suboptimal choice by the pigeons in the experimental group, it did result in a substantial reduction in suboptimal choice (see also McDevitt, Spetch, & Dunn, 1997).

Conclusions

The history of research delay of reinforcement suggest that delay typically leads to a weaker association between a stimulus and reinforcement that follows. Research described here suggests that in learning that involves a simultaneous discrimination, under a variety of conditions, adding a delay between choice and the outcome of that choice (reinforcement or a conditioned reinforcer) can discourage animals from choosing a suboptimal alternative.

The prior commitment procedure developed by Rachlin and Green (1972) provided the impetus for a number of experiments exploring the effect of the insertion of a delay between a choice response and reinforcement. In the ephemeral reward task, the failure of rats and pigeons to learn to choose the alternative that provided them with two reinforcements rather than one was reminiscent of delay discounting and suggested that the immediacy of reinforcement may have been a factor. Inserting a delay between choice and the first reinforcement led to optimal choice by both species.

In a related task, pigeons had difficulty learning to choose a stimulus that provided them with two pellets of food rather than one, when reinforcement following choice was immediate. But the pigeons learned readily when the choice required 10 pecks rather than one to make their choice.

In a somewhat different task, pigeons were not able to show object permanence in the simplest visible displacement version of the task, and, surprisingly, they showed only minimal ability to learn by trial and error. Once a delay was inserted between baiting and choice, however, they not only learned to choose correctly when food was visibly displaced but they transferred that learning to a 90° invisible displacement and then to the more difficult 180° invisible displacement.

Finally, in a quite different task, pigeons were found to show a strong preference for an alternative that infrequently signaled a high probability reinforcer over an alternative that always signaled a more frequent but lower probability reinforcer. In this research, there was a clear delay to reinforcement following choice, but there was no delay between the choice and the conditioned reinforcer that followed. When the choice was followed by a delay prior to the appearance of the conditioned reinforcer, however, substantially less suboptimal choice was found.

It is proposed that in a variety of discrimination tasks, difficulty in learning to make the optimal response may be constrained by the immediacy of reinforcement, which may lead to impulsive choice. Although adding a delay between the choice response and reinforcement may appear counterintuitive, it may facilitate learning and improve the performance of simultaneous discriminations in several contexts. Importantly, procedures that decrease the likelihood of impulsive choice, by definition, lead to what in humans would be considered better self-control.

References

Bai, D. L., & Bertenthal, B. I. (1992). Locomotor status and the development of spatial search skills. *Child Development*, *63*(1), 215–226. doi:https://doi.org/10.2307/1130914

Barth, J., & Call, J. (2006). Tracking the displacement of objects: A series of tasks with great apes (*Pan troglodytes, Pan paniscus, Gorilla gorilla*, and *Pongo pygmaeus*) and young children (*Homo sapiens*). *Journal of Experimental Psychology: Animal Behavior Processes, 32*(3), 239–252. doi:10.1037/0097-7403.32.3.239

Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior, 20*(1), 137–153. doi: https://doi.org/10.1901/jeab.1973.20-137

Bitterman, M. E. (1975). The comparative analysis of learning. *Science, 188*(4189), 699–709. doi:https://doi.org/10.1126/science.188.4189.699

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*(1), 130–136. doi:https://doi.org/10.1046/j.1461-0248.2002.00295.x

Carlson, J. G., & Wielkiewicz, R. M. (1972). Delay of reinforcement in instrumental discrimination learning of rats. *Journal of Comparative and Physiological Psychology, 81*(2), 365–370. doi:10.1037/h0033531

Case, J. P., & Zentall, T. R. (2018). Suboptimal choice in pigeons: Does the predictive value of the conditioned reinforcer alone determine choice? *Behavioural Processes, 157,* 320–326. doi: https://doi.org/10.1016/j.beproc.2018.07.018

Catania, A. C. (1979). *Learning.* Englewood Cliffs, NJ: Prentice-Hall.

Clement, T. S., Feltus, J., Kaiser, D. H., & Zentall, T. R. (2000). "Work ethic" in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review, 7*(1), 100–106. doi:https://doi.org/10.3758/BF03210727

Culbertson, J. L. (1970). Effects of brief reinforcement delays on acquisition and extinction of brightness discrimination is rats. *Journal of Comparative and Physiological Psychology, 70,* 317–325. doi: 10.1037/h0028736

Elsmore, T. F. (1971). Effects of response effort on discrimination performance. *The Psychological Record, 21,* 17–24.

Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior, 12,* 723–730. doi: https://doi.org/10.1901/jeab.1969.12-723

Friedrich, A. M., Clement, T. S., & Zentall, T. R. (2005). Discriminative stimuli that follow the absence of reinforcement are preferred by pigeons over those that follow reinforcement. *Learning & Behavior, 33*(3), 337–342. doi:https://doi.org/10.3758/BF03192862

Friedrich, A. M., & Zentall, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behavioural*

Processes, 67(3), 405–415. doi:https://doi.org/10.1016/j.beproc.2004.07.001

House, D., Peng, D., & Zentall, T. R. (2020). *Pigeons can learn a difficult discrimination by delaying reinforcement following choice.* Manuscript submitted for publication.

Hull, C. L. (1952). *A behavior system: An introduction to behavior theory concerning the individual organism.* New Haven, CT: Yale University Press.

Keesey, R. E. (1964). Intracranial reward delay and the acquisition of a brightness discrimination. *Science, 143*(3607), 700–701. doi:https://doi.org/10.1126/science.143.3607.702

Kimble, G. A. (1961). *Hilgard and Marquis' conditioning and learning.* New York, NY: Appleton-Century-Crofts.

Lattal, K. A. (2010). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior, 93*(1), 129–139. doi:https://doi.org/10.1901/jeab.2010.93-129

Laude, J. R., Stagner, J. P., & Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 40*(1), 12–21. doi:10.1037/xan0000010

Lea, S. E. G. (1979). Foraging and reinforcement schedules in the pigeon: Optimal and non-optimal aspects of choice. *Animal Behaviour, 27*(3), 875–886. doi:https://doi.org/10.1016/0003-3472(79)90025-3

McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the Experimental Analysis of Behavior, 68*(3), 317–327. doi: https://doi.org/10.1901/jeab.1997.68-317

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analysis of behavior: Vol 5. Effects of delay and of intervening events on reinforcement value* (pp. 55–73). Hillsdale, NJ: Erlbaum.

Mazur, J. E. (1989). Theories of probabilistic reinforcement. *Journal of the Experimental Analysis of Behavior, 51*(1), 87–99. doi:https://doi.org/10.1901/jeab.1989.51-87

Miller, H. C., Gipson, C. D., Vaughan, A., Rayburn-Reeves, R., & Zentall, T. R. (2009). Object permanence in dogs: Invisible displacement in a rotation task. *Psychonomic Bulletin & Review, 16*(1), 150–155. doi:10.3758/PBR.16.1.150

Odum, A. L. (2011). Delay discounting: I'm a k, you're a k. *Journal of the Experimental Analysis of Behavior, 96*(3), 427–439. doi:https://doi.org/10.1901/jeab.2011.96-423

Oliveira, L., Green, L., & Myerson, J. (2014). Pigeons' delay discounting functions using a concurrent-chains procedure. *Journal of the Experimental Analysis of Behavior, 102*(2), 151–161. doi: https://doi.org/10.1002/jeab.97

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*(3), 298–306. doi:https://doi.org/10.1037/a0036205

Piaget, J. (1963). *The psychology of intelligence.* Totowa, NJ: Littlefield Adams.

Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior, 17*(1), 15–22. doi: 10.1901/jeab.1972.17-15

Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin, 72*(2), 77–94. doi:https://doi.org/10.1037/h0027760

Roberts, W. S. (1972). Short-term memory in the pigeon: Effects of repetition and spacing. *Journal of Experimental Psychology, 94*(1), 74–83. doi: https://doi.org/10.1037/h0032796

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 orangutans in a complex foraging task derived from cleaner-client reef fish cooperation. *PLOS One, 7,* e49068. doi: https://doi.org/10.1371/journal.pone.0049068

Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is based primarily on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Behavior Processes, 42*(2), 212–220. doi: https://doi.org/10.1037/xan0000092

Spence, K. W. (1947). The role of secondary reinforcement in delayed reward learning. *Psychological Review, 54*(1), 1–8. doi: https://doi.org/10.1037/h0056533

Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior, 53*(2), 219–234. doi: https://doi.org/10.1901/jeab.1990.53-219

Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review, 17,* 412–416. doi: https://doi.org/10.3758/PBR.17.3.412

Thorndike, E. L. (1911). *Animal intelligence.* New York, NY: Macmillan.

Warden, C. J., & Haas, E. L. (1927). The effect of short intervals of delay in feeding upon speed of maze learning. *Journal of Comparative Psychology, 7*(2), 107–116. doi: https://doi.org/10.1037/h0074089

Watson, J. B. (1917). The effect of delayed feeding upon learning. *Psychobiology, 1*(1), 51–60. doi:https://doi.org/10.1037/h0074422

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. doi: https://doi.org/10.1016/j.beproc.2017.09.008

Zentall, T. R., Case, J. P., & Berry, J. R. (2017a). Early commitment facilitates optimal choice by pigeons. *Psychonomic Bulletin & Review, 24*(3), 957–963. doi:https://doi.org/10.3758/s13423-016-1173-8

Zentall, T. R., Case, J. P., & Berry, J. R. (2017b). Rats' acquisition of the ephemeral reward task. *Animal Cognition, 20,* 419–425. doi:https://doi.org/10.1007/s10071-016-1065-3

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*(2), 138–144. doi:https://doi.org/10.1037/com0000026

Zentall, T. R., & Raley, O. L. (2019). Object permanence in the pigeon: Insertion of a delay prior to choice facilitates visible- and invisible-displacement accuracy. *Journal of Comparative Psychology, 133*(1), 132–139. doi:https://doi.org/10.1037/com0000134

Zentall, T. R., & Stagner, J. P. (2011). Maladaptive choice behavior by pigeons: An animal analog of gambling (sub-optimal human decision making behavior). *Proceedings of the Royal Society B: Biological Sciences, 278*(1709), 1203–1208. doi:https://doi.org/10.1098/rspb.2010.1607