

COMPARATIVE COGNITION & BEHAVIOR REVIEWS

What Suboptimal Choice Tells Us About the Control of Behavior

Thomas R. Zentall

University of Kentucky

When animals make decisions that are suboptimal, it helps us to identify the processes that have evolved to produce this behavior. In an earlier article, I discussed three examples of suboptimal choice or bias (Zentall, 2016): (a) *sunk cost*, the tendency to continue on a losing project because of the amount already invested; (b) *unskilled gambling*, in which the loss is greater than the return; and (c) *justification of effort*, the bias to prefer conditioned stimuli that in training required more effort to obtain. Here I discuss three additional examples of suboptimal choice that we have studied in animals: (a) *when less is better*, in which animals prefer one piece of food (one preferred item) over two pieces of food (one preferred item plus one less preferred item); (b) suboptimal choice on the *ephemeral choice task*, in which animals prefer one piece of food now over two pieces of the same food, one now but the second briefly delayed; and (c) suboptimal choice in the *midsession reversal task*, errors of anticipation and perseveration. Each of these examples may help to identify the relative limits on behavioral flexibility found when animals are exposed to conditions that may be different from those that they would normally encounter in their natural environment. They also may help us to understand the origins of similar behavior when it occurs in humans.

Keywords: *suboptimal choice, less is better, ephemeral reward, midsession reversal*

Those of us who study the behavior of animals assume that they have evolved to maximize their success (e.g., at finding food), and much of learning theory (Skinner, 1938; Thorndike, 1911) is based on this premise. Animals select those responses that lead to the increased probability of reinforcement over those that do not. When animals' behavior is consistent with this theory, it strengthens our belief in the validity of the theory. However, when animals show a preference for alternatives that result in less food over those that result in more food, it is important to try to understand why they do.

Kacelnik (2006) suggested that rationality in decision making can be defined in different ways. When defined by philosophers and psychologists, it has been judged in terms of the reasoning or thought processes that accompany the decisions. When defined by economists, it does not require thought processes but refers to behavior that is internally consistent and is compatible with expected utility maximization. When defined by biologists, it is broader and goes beyond the organism to allow for inclusive fitness (including benefit to one's kin).

Sometimes, what appears to be an irrational choice may reflect a change in state. An animal's preference

for one kind of food over another may reverse if it has been sated on the preferred food, or an animal that has a choice between eating and being with conspecifics may choose the latter because being close to others may enhance feeding rate or may offer safety from predation (Kacelnik, 2006). Alternatively, the condition that the animal is in may cause it to choose less food over more food. For example, an animal may choose a low probability but possibly larger amount of food over a frequent but smaller amount of food but one that will not allow it to survive through the night (Stephens, 1981; see also Houston, McNamara, & Steer, 2007).

When animals prefer an alternative that provides them with less food over one that provides them with more (i.e., they choose suboptimally), on one hand, it may cause us to question the processes that underlie that behavior. In an earlier article in this journal (Zentall, 2016), I described a task in which pigeon showed a strong preference for one alternative that on 20% of the trials provided them with a signal for reinforcement and on 80% of the trials provided them with a signal for the absence of reinforcement, over a second alternative that always provided them with a signal for 50% reinforcement (Stagner & Zentall, 2010). With this procedure, not only do pigeons quickly show a preference for the 20% signaled over the 50% unsignaled reinforcement but they show no evidence that they learn to correct that preference with extensive training. Furthermore, that preference is not simply controlled by the uncertainty of the reinforcement associated with the higher probability of reinforcement alternative, because even when the alternatives are between 50% signaled reinforcement and 100% reinforcement, pigeons do not show a preference for the optimal alternative (McDevitt, Dunn, Spetch, & Ludvig, 2016; Smith & Zentall, 2016). In addition, a similar pattern of suboptimal choice can be shown when reinforcement magnitude is manipulated. For example, pigeons prefer a 20% chance of obtaining a signal for 10 pellets of food over a 100% chance of obtaining a signal for three pellets of food (Zentall & Stagner, 2011).

When animals choose suboptimally, it may tell us something about the natural environment in which the animals have evolved (Fortes, Pinto, Machado, & Vasconcelos, 2018). Several mechanisms may be

responsible for this suboptimal choice. First, in nature, when an animal approaches a stimulus that signals the presence of food, it is likely that the probability of reinforcement will increase. Not so in this choice task in which choice frequency has no effect on the probability of reinforcement. Second, in nature, when an animal encounters a signal for the absence of food, that signal can generally be ignored, because the animal will simply reject it and look elsewhere for food (Fortes et al., 2018; Vasconcelos, Machado, & Pandeirada, 2018). That is, in nature there is no need to remain in its presence, so it does not acquire inhibitory value, whereas the animal must remain in its presence in the laboratory choice experiment.

Although the predictive value of the conditioned reinforcer that follows choice of each alternative, independent of its probability of occurrence, appears to predict choice (Smith & Zentall, 2016), evidence suggests that there may be a third factor (Case & Zentall, 2018; McDevitt et al., 2016). Case and Zentall (2018) found that when pigeons are given a choice between 50% signaled reinforcement and 100% reinforcement, they initially show indifference between the two alternatives; however, with continued training they show a significant preference for the suboptimal alternative (see also Kendall, 1974). Case and Zentall suggested that the preference for the suboptimal alternative may result from positive contrast between the expected value of reinforcement following choice of the suboptimal alternative and the value of the conditioned reinforcer that follows on half of the trials. Positive contrast would not be expected between choice of the optimal alternative and the conditioned reinforcer that follows, because the expected value of reinforcement is consistent with the value of reinforcement that follows. A similar mechanism was suggested by McDevitt et al. (2016), who proposed that the conditioned reinforcement that followed choice of the suboptimal alternative represented “good news,” whereas the conditioned reinforcement that followed choice of the optimal alternative was not newsworthy. Although identifying the predispositions responsible for suboptimal choice with this procedure will likely require further research, the inability of the pigeons to learn to choose optimally suggests that there are conditions under which pigeons do not appear to have the flexibility to overcome these predispositions.

In the earlier article (Zentall, 2016), I identified two other cases in which pigeons fail to choose optimally. The first was research on the *sunk cost effect* in which pigeons prefer to complete pecking on one reinforcement

Author Note: Thomas R. Zentall, Department of Psychology, University of Kentucky, Lexington, KY 40506-0044.

Correspondence concerning this article should be addressed to Thomas R. Zentall at zentall@uky.edu.

schedule over changing to another reinforcement schedule, even though changing to the other schedule would have reduced the time and effort (amount of pecking) to reinforcement. For example, pigeons first learned to peck 30 times for food when the color was green and 10 times for food when the color was red. They then learned that after pecking green a variable number of times, they would be given a choice between completing the pecks to green and switching to peck the red 10 times. Surprisingly, the pigeons preferred to return to pecking green, even when returning to green required as many as 25 more pecks (Pattison, Zentall, & Watanabe, 2012; see also Magalhães & White, 2014; Navarro & Fantino, 2005).

The second additional line of research described in the Zentall (2016) article actually involved a bias rather than a suboptimality. Pigeons were trained to peck a light to receive a choice between two colors. On some trials, a single peck was required and the choice was between, for example, red and yellow and choice of red was reinforced. On other trials, 20 pecks were required and the choice was between, for example, green and blue and choice of green was reinforced. On probe trials, pigeons were given a choice between red and green, the two colors both associated with reinforcement. Surprisingly, the pigeons showed a preference for green, the color that during training they had to work harder to obtain. When a similar effect has been found in humans (e.g., Aronson & Mills, 1959), it has been referred to as the *justification of effort* effect; however, we prefer to interpret this preference as a *contrast* effect. That is, the positive contrast between 20 pecks and green was greater than the positive contrast between one peck and red.

In the present article I examine three additional phenomena, each of which demonstrates a behavior that is suboptimal. The first is commonly referred to as the less is better effect; the second is the failure to learn to choose optimally on a task in which choice of one alternative provides two reinforcements, whereas the other provides only one (the ephemeral reward task); and the third is the failure to choose optimally on the midsession reversal task.

The Less Is Better Effect

Economists have traditionally held that when humans are given sufficient information, they generally make rational choices (Persky, 1995). This is the basis of rational choice theory (Becker, 1976). However, Tversky and Kahneman (1974) challenged this notion by showing

that humans tend to use various affective heuristics in making decisions and those heuristics can be shown to lead to suboptimal decisions. Such an example is the less is better effect (sometimes referred to as the less is more effect), demonstrated in several experiments by Hsee (1998). In one example, Hsee asked subjects to estimate the value of a set of 24 dishes, all in good condition, or to estimate the value of a set of 40 dishes, but only 31 were in good condition. Surprisingly, the set of 24 dishes was valued higher than the set of 40 dishes. Apparently, the nine dishes of poor quality depreciated the value of the 31 good-quality dishes. The average quality of the set, as a whole, apparently overshadowed the objective judgment of the value of the set. But this effect may be unique to humans, who may be sensitive to the aesthetics of the two sets of dishes.

In another study, subjects were asked to imagine that a friend had given them a \$55 wool coat from a store where coats cost between \$50 and \$500, or alternatively a \$45 wool scarf from a store where scarves cost between \$5 and \$50 (Hsee, 1998). The subjects said that they would be happier with the scarf than with the coat because the purchase of the scarf would reflect greater generosity than the purchase of the coat. The scarf was at the high end of the range, whereas the coat was at the low end of the range. This finding suggests that if gift givers want their gift recipients to perceive them as generous, it would be better for them to give a high-value item from a low-value product category (e.g., a \$45 scarf) than a low-value item from a high-value product category (e.g., a \$55 coat).

Would animals show the same bias if food of different quality was used rather than dishes or clothing? According to optimal foraging theory (Stephens & Krebs, 1986), other factors being equal (e.g., the possibility of predation), nature should select against any tendency to prefer an alternative that provides less food. Kralik, Xu, Knight, Khan, and Levine (2012) tested this hypothesis. They found that monkeys readily would eat grapes and sliced cucumbers, but when offered a choice between them, they preferred the grapes. When the monkeys were offered a choice between a grape by itself or a grape and a slice of cucumber, however, they generally showed a strong preference for the grape alone.

A similar effect was found by Beran, Ratliff, and Evans (2009) for two of four chimpanzees when given a choice between a slice of banana and a similar slice of banana plus a slice of apple. Similarly, chimpanzees were indifferent between a preferred pellet and a similar pellet plus either a less preferred piece of carrot or a

less preferred piece of apple (Sanchez-Amaro, Pereto, & Call, 2016). And when Beran, Evans, and Ratliff (2009) manipulated the quantity rather than the quality of the combined option, four chimpanzees preferred a 20 g slice of banana over the same 20 g slice of banana plus an additional 5 g slice of banana.

Dogs, too, have been found to show a less is better effect (Pattison & Zentall, 2014). Several dogs were found to eat a slice of carrot or a slice of cheese, but when given a choice, they preferred the cheese. However, when given a choice between the cheese and a combination of the cheese and the carrot, these dogs preferred the cheese alone (see Figure 1).

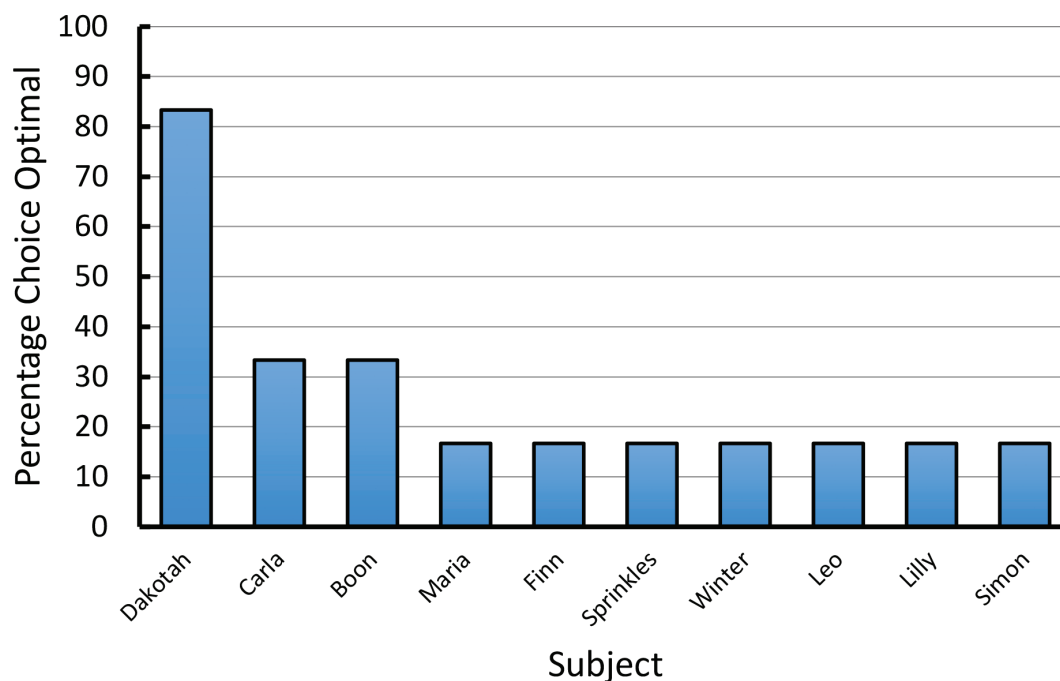
When a similar experiment was conducted with pigeons, the results were less clear (Zentall, Laude, Case, & Daniels, 2014, Experiment 1). Although all of the pigeons would readily eat milo seeds and dried peas (both a part of their normal laboratory diet), they strongly preferred the peas. In Experiment 1, the pigeons were kept at a typical level of food restriction (they were fed once a day, and when tested they had not eaten in about 24 hours) and were given a choice between one pea and a pea together with a milo seed. Unlike the monkeys and the dogs, however, the pigeons preferred the milo

plus the pea over the pea alone. That is, they chose optimally, two bits of food over one.

Although one might be tempted to conclude that the difference between the pigeons and the dogs and primates reflected species differences in their susceptibility to the less is better effect, we first looked for differences in the experimental conditions. Both the monkeys and dogs were minimally deprived of food (for both species, the food presented represented special treats). For the pigeons, however, the experimental procedures followed a longer period without food. Perhaps the species differences had to do with the subjects' relative level of motivation. Thus, in a follow-up experiment (Zentall et al., 2014, Experiment 2), the level of food deprivation was manipulated. Once again, the pigeons that were about 24 hours food deprived failed to show the less is better effect. However, pigeons that were food deprived for only 4 hr showed a reliable less is better effect (Figure 2).

These results paint a more nuanced picture of the less is better effect. Based on the results with pigeons, the average *quality* of the food offered affects choice when the animals have been deprived of food for a relatively short time, but when food deprivation is greater,

Figure 1. Percentage preference for the optimal alternative (a piece of cheese plus a piece of carrot) over the suboptimal alternative (a piece of cheese alone) for each dog in the study (Pattison & Zentall, 2014).



optimal choice based on the *quantity* of food in each alternative appears to control choice. Of interest, in the experiment with dogs, one of the dogs failed to show the less is better effect (see Figure 1), and of all of the dogs, that dog had been an adult rescue originally picked up as a stray. Although none of the dogs had experienced extreme deprivation prior to the experiment, it is likely that the dog that failed to show the less is better effect had been food insecure for some time in its life prior to being put up for adoption. Thus, in addition to the animals' current level of food deprivation, prior experience with high levels of food deprivation may lessen the likelihood of finding the less is better effect (Zentall et al., 2014, Experiment 2). It would be interesting to know if humans, too, would choose optimally if the stakes were higher.

The Ephemeral Reward Task

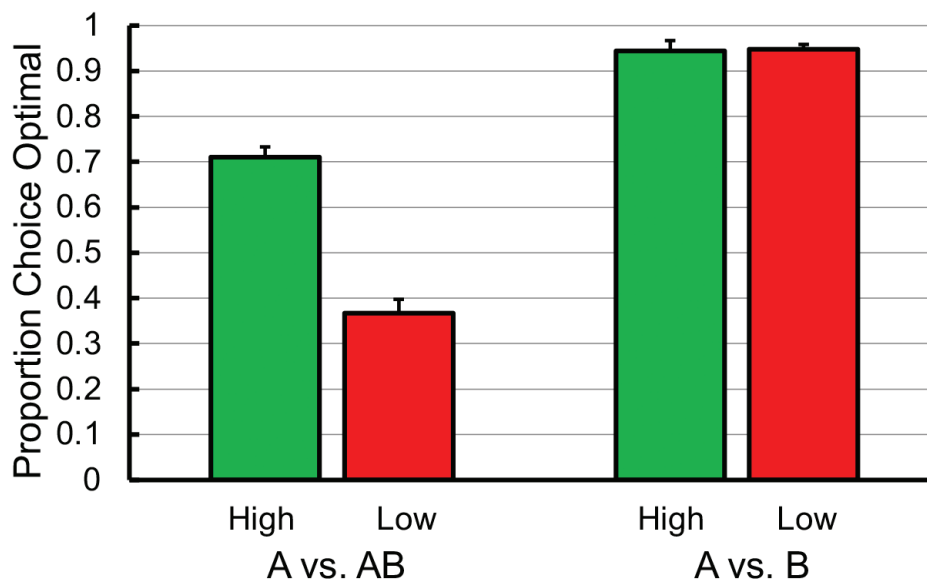
In the ephemeral reward task, animals are given a choice between two alternatives, each of which is associated with a similar piece of food. The contingency is such, however, that if alternative A is selected, reinforcement follows and the trial is over, but if alternative B is selected, reinforcement follows and the subject can be reinforced for responding to alternative A. Thus, choice of alternative B results in twice as much food as alternative A. This task was originally studied with wrasse

(cleaner fish) as a model of their foraging behavior on a reef, their natural habitat (Bshary & Grutter, 2002). Wrasse obtain food by cleaning the mouths of larger fish. Some of the fish that they service live on the reef and are relatively permanent, whereas other fish that they service are visitors to the reef and they must be serviced quickly, before they swim away. The ephemeral reward task is assumed to be analogous to foraging on the reef: Choice to service a visiting fish (analogous to choice of alternative B) means that the wrasse can also service the resident fish (alternative A), whereas first servicing the fish on the reef (alternative A) may lead to the loss of being able to service the visitor (alternative B).

The hypothesis that this task is a model of the wrasses' reef-foraging behavior was tested by comparing the performance of this task by wrasse to that of several primate species (Salwiczek et al., 2012). Consistent with the model, the wrasse learned to choose alternative B readily, whereas monkeys and orangutans did not learn to choose alternative B reliably within 100 trials, and only two of four chimpanzees learned to do so.

Although differences in species-typical foraging behavior may explain the unexpected differences in species ability to readily acquire the optimal choice of B over A, other hypotheses have been proposed. Pepperberg and Hartsfield (2014) proposed that if the rapid acquisition of this task was determined by the species' natural foraging behavior, parrots, which live in an

Figure 2. Proportion of optimal choice plotted for the high food restricted (High) and low food restricted (Low) groups. A = the more preferred grain; B = the less preferred grain; AB = both grains. Error bars = ± 1 SEM (Zentall et al., 2014).



environment similar to primates and eat similar food (e.g., fruit, nuts, and berries), also should have difficulty acquiring this task. However, Pepperberg and Hartsfield found that the parrots acquired the ephemeral reward task at about the same rate as the wrasse. Pepperberg and Hartsfield suggested that the difference in learning between primates and fish or parrots might be attributable to way each species chooses the alternatives. Fish and parrots choose with their mouth (or beak), whereas primates choose with their hands. Why choosing with the hand would make the task difficult is not clear, but the presence of two bits of food and having two hands with which to choose may result in conflict (e.g., trying to choose both A and B at the same time) that would not be present with a single possible choice response.

The results with parrots led us to test this theory further with pigeons and also with rats. According to the theory proposed by Pepperberg and Hartsfield, pigeons should quickly learn to choose optimally with this task, whereas the rats should not.

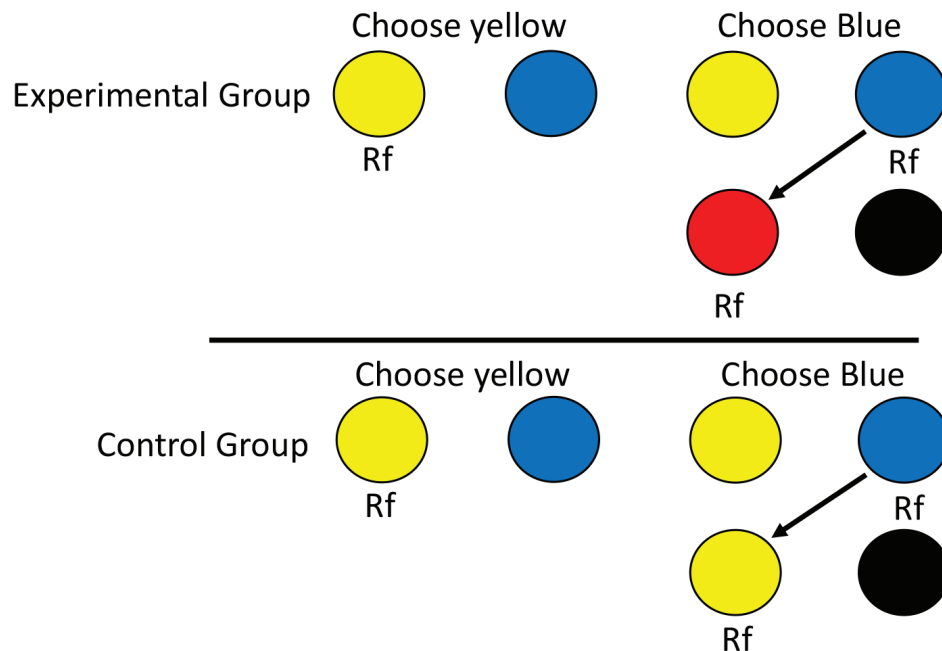
The results with pigeons were quite surprising (Zentall, Case, & Luong, 2016). Not only were the pigeons not able to learn to reliably choose alternative B (the one that gave them two pieces of food), but they consistently showed a significant preference for

alternative A (the one that gave them only one). The preference for alternative A was particularly unexpected because such a bias indicates that there was some learning about the task contingencies but they were different from the ones intended by the design of the task.

The reason that the pigeons chose suboptimally may be related to the differential frequency of reinforcement associated with the two alternatives. If one assumes that initially the pigeons chose randomly between the two alternatives, they would have had more experience with alternative A. In fact, they would have experienced the reinforcement associated with alternative A on every trial, whereas they would have experienced the reinforcement associated with alternative B only if they had chosen B first. Furthermore, every trial ended with reinforcement associated with alternative A, so even when B was chosen, the last alternative experienced was always A (see the design of the control group in Figure 3).

To test this hypothesis, we used an operant version of the task in which alternatives A and B were represented by colors projected on response keys (Zentall et al., 2016, Experiment 3). Once again, we arranged the contingencies such that the choice of A provided reinforcement but ended the trial, whereas the choice of B provided reinforcement and allowed a second reinforcement for

Figure 3. Design of Zentall et al. (2016, Experiment 3). When pigeons chose yellow, reinforcement (Rf) was provided and the trial was over. When pigeons chose blue, reinforcement was provided and they could peck the other color to receive a second reinforcement. For the control group, the other color remained yellow. For the experimental group, the yellow color changed to red.



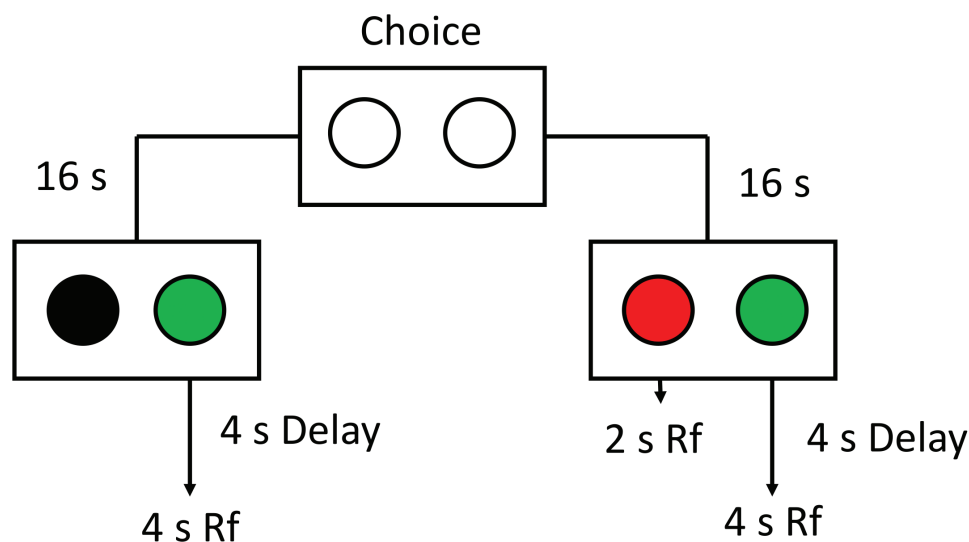
responding to A. To test the hypothesis that the preference for alternative A resulted from the fact that there were more reinforced responses to alternative A than reinforced responses to alternative B, for the experimental group, choice of alternative B provided reinforcement; while the pigeons were eating, the color of alternative B changed to a different color (C), and a peck to C provided the second reinforcement (see the design of this experiment in Figure 3). Thus, for the experimental group, initial random choice would still provide one reinforcer for the choice of A and two reinforcers for the choice of B, but it would also equate reinforcement that followed a response to each of the three colors. The results indicated that relative to a control group that replicated the original significant preference for the choice of alternative A, the experimental group chose alternative B significantly more. However, the experimental group did not actually show a preference for alternative B, as did the fish and parrots; rather, they were indifferent about the two alternatives.

Given that the pigeons did not show the optimal performance shown by the parrots and fish, when we trained rats on the original task, it was not surprising that the rats failed to acquire the task either (Zentall, Case, & Berry, 2017b). Although the rats did not learn to choose alternative B reliably, neither did they show the significant preference for alternative A shown by the pigeons. Instead, they showed indifference about the two alternatives.

The puzzle remained to understand why some species appear to be unable to learn to choose alternative B, the option that provides them with two pieces of food over alternative A, which provides them with only one. The puzzle is reminiscent of research on delay discounting, the choice between an immediate small amount of food and a delayed larger amount of food (Ainslie, 1974). In delayed discounting, it should be noted that the preference that most animals show for the suboptimal smaller-sooner alternative is another example of suboptimal choice that has been extensively studied (see Estle, Green, Myerson, & Holt, 2007; Mazur, 1997). In the ephemeral reward task, although the immediate consequences of choosing option A or B is exactly the same, the delayed consequence of choosing alternative B is additional food.

Delay discounting is often viewed somewhat negatively as a sign of the lack of self-control; however, in nature it may be a functional choice. In natural environments the “promise” of a larger later reward is often accompanied by a reduction in the probability of reinforcement, because any reinforcer that is delayed is less certain due to both intraspecies and extraspecies competition. Furthermore, for most species, delay implies travel time, and travel may increase the chances of encountering a predator or a more dominant conspecific. Thus, what is generally considered an impulsive response to a smaller-sooner reward, under natural conditions may actually be quite adaptive. In the ephemeral reward task,

Figure 4. Commitment procedure used by Rachlin and Green (1972). Pigeons could choose between a later choice between smaller-sooner (2-s immediate reinforcement [Rf]) and larger-later (4-s delayed Rf) or only larger later (4-s delayed Rf).



however, the delay between the first reinforcement and the second reinforcement for choice of the B alternative is relatively short (perhaps only 1 s in the case of the manual presentation of the two alternatives and only about two seconds in the case of the operant analog). Nevertheless, in the ephemeral reward task, presentation of the second reinforcement does not appear to become associated with choice of alternative B. From the subject's perspective, it is as if the choice of either alternative results in a common outcome, immediate reinforcement, and a second reinforcer sometimes appears mysteriously and independently of the alternative chosen.

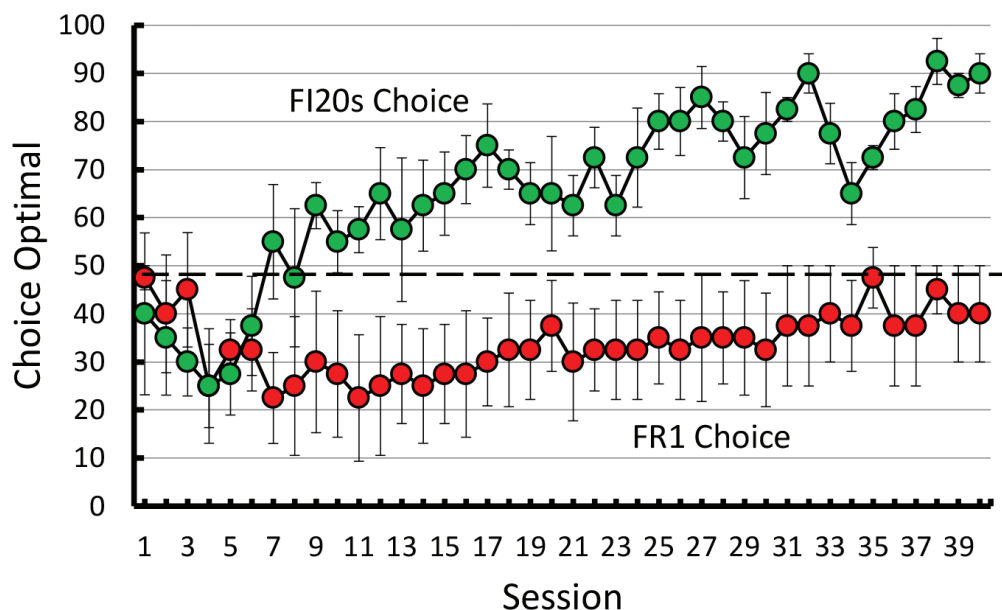
One way to encourage the association of the second reinforcer with choice of alternative B may be to use a procedure developed by Rachlin and Green (1972) to mitigate the effects of delay discounting. Rachlin and Green found that pigeons that preferred a smaller immediate reinforcer over a larger later reinforcer would choose optimally if the pigeons would make a "commitment" to the larger-later reinforcer at a time prior to the actual choice. They gave pigeons a choice between (a) having a choice 8 s later between the smaller-sooner reinforcer and larger-later reinforcer and (b) being forced to have access to only the larger-later alternative (see Figure 4). Of interest, although the larger-later option was available independent of their initial choice, the pigeons preferred

the option that did not give them a second choice. Apparently, by choosing not to be able to select the smaller-sooner reward later, they could avoid the "temptation" to choose the smaller-sooner alternative.

This prior commitment procedure has implications for humans who are trying to avoid impulsive decision making (see Laibson, 1997). For example, to discourage oneself from smoking, at a time when the urge to smoke is not great, one can commit to refrain from smoking at a later time by not having cigarettes available and possibly making them difficult to obtain (e.g., going on a camping trip). Similarly, the pigeons chose not to have the smaller-sooner alternative available even after the short 8-s wait.

In the suboptimal choice task studied by Spetch, Belke, Barnett, Dunn, and Pierce (1990) and others, in which pigeons were given a choice between 50% signaled reinforcement and 100% reinforcement, McDevitt, Spetch, and Dunn (1997) examined the effect of inserting a delay between choice and the signals that followed. They found that when a dark gap occurred following choice but before the onset of the stimuli that signaled reinforcement or its absence, pigeons tended to choose optimally. More recently, Zentall, Andrews, and Case (2017) extended the duration of the chosen stimulus for 20 s following choice and found a similar preference for

Figure 5. Percentage optimal choice for pigeons that had to complete a fixed-interval 20-s schedule (FI20s Choice) to obtain initial reinforcement (green) and pigeons that had to make a single peck to obtain initial reinforcement (red; error bars = ± 1 SEM). FR1 = response requirement to S1 (Zentall et al., 2017a).



the optimal alternative. Delaying the onset of the signals for reinforcement (or its absence) following choice can be thought of as making a commitment, some time prior to the appearance of the signals for reinforcement (or its absence).

If we apply something like the commitment procedure to the ephemeral reward task, it would suggest that inserting a delay between the initial choice and the first reinforcer may encourage the pigeons choose optimally. Specifically, it would encourage the pigeons to associate the second reinforcer with the initial choice of alternative B. Using the operant procedure, we (Zentall, Case, & Berry, 2017a) gave the pigeons a choice between the A and B alternatives. Once a pigeon made its choice, the other stimulus turned off and a fixed-interval 20-s schedule was in effect (the first response after 20 s provided reinforcement). If alternative A had been selected, reinforcement was provided and the trial was over. If alternative B had been selected, following reinforcement, alternative A was presented again and only a single peck was required to provide a second reinforcement. Once again, the control group that did not have the 20-s delay between choice and reinforcement showed a significant preference for alternative A. The experimental group, however, learned to choose optimally (see Figure 5).

Given the success that we found with pigeons learning to choose optimally with a delay inserted between choice and the first reinforcement (Zentall et al., 2017a) we tried a similar procedure with rats and found that they too learned to choose optimally when we inserted a 20-s delay (fixed-interval 20-s schedule) between their first choice and the first reinforcer (Zentall et al., 2017b). Thus, the insertion of a delay between choice and the first reward appears to have some generality.

But why is it that the wrasse and parrots choose optimally with this task without the delay inserted between choice and reinforcement? In the case of the wrasse, it is likely that impulsive choice would not be adaptive. Impulsively swimming into the mouth of a large fish could have unfortunate consequences. Thus, in general, cautiously approaching a potential reinforcer may be prudent. Furthermore, the wrasse appear to use a tactile dancing behavior as a signal to avoid being eaten (Grutter, 2004). It is also quite possible that the client fish gives the wrasse an indication that it is safe to begin cleaning.

In the case of parrots, the three parrots that were used by Pepperberg and Hartsfield (2014) had received extensive training for several years and may have learned, generally, to avoid choosing impulsively. One parrot had been involved in several studies on

comparative cognition and interspecies communication, and the other two had received considerable training on referential communication. Extensive prior training may reduce impulsivity in general and enable animals to associate the two reinforcers that follow choice of the B alternative.

Recent research with monkeys has found that they, too, learned to choose optimally when the foods that they were choosing were distinctively colored (one pink, the other black; Prétôt, Bshary, & Brosnan, 2016b, Experiment 2). Why distinctive coloring would facilitate acquisition is not obvious, but the unusual color of the food may have made the monkeys choose more carefully (i.e., it may have reduced the monkeys' tendency to choose impulsively).

In another experiment, when the rewards were hidden under distinctive cups that the monkeys had to lift or point to, they also chose optimally (Prétôt, Bshary, & Brosnan, 2016b, Experiment 3). That is, the rewards were not immediately visible, a change that may have reduced impulsivity.

In another experiment with monkeys, a computer version of the task was used in which the monkeys had to move a cursor to the chosen stimulus to receive the reward at a different location (Prétôt, Bshary, & Brosnan, 2016a). It may be that this spatial and temporal delay was sufficient to reduce the monkeys' impulsivity.

Impulsivity may not be the only factor that distinguishes the conditions under which animals choose suboptimally in the ephemeral reward task. It is likely that further research will be needed to disentangle the variables that contribute to this phenomenon.

The Midsession Reversal Task

The midsession reversal task is a variation of the serial reversal task in which a simple simultaneous discrimination is acquired and is then repeatedly reversed (see, e.g., Mackintosh, McGonigle, Holgate, & Vanderver, 1968). The goal of the serial reversal task is to determine how much an organism can benefit from the experience of successive reversals (i.e., how much it can *learn to learn*). Ideally, such a task may lead to the development of a win–stay/lose–shift strategy in which a reversal results in only a single error.

The midsession reversal task is a multisession task in which there is an additional cue to the reversal because all sessions start with the same correct stimulus (S1) and incorrect stimulus (S2), and typically the reversal occurs midway through the session (often after

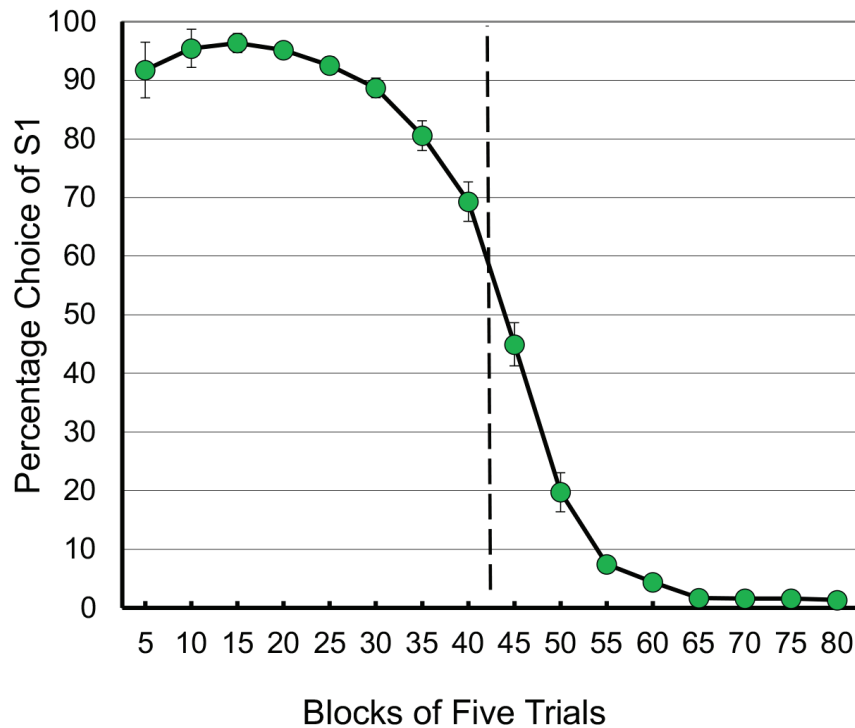


Figure 6. Mean percentage choice of the first correct stimulus (S1) as a function of trial number for Sessions 31 to 50 (the last 20 sessions of training). S1 was correct for the first 40 trials of each session. Incorrect stimulus was correct for the last 40 trials of each session. Error bars = ± 1 SEM (Rayburn-Reeves et al., 2011).

40 trials of an 80-trial session). That is, from Trial 1 to Trial 40, S1 is correct and S2 is incorrect, whereas from Trial 41 to Trial 80, S2 is correct and S1 is incorrect. Surprisingly, even after considerable training (e.g., 50 sessions), pigeons make many more errors than what would be optimal (Rayburn-Reeves, Molet, & Zentall, 2011). Specifically, pigeons continue to make anticipatory errors by choosing S2 as the reversal approaches, as well as making perseverative errors by choosing S1 after the reversal has occurred (see Figure 6). In fact, the function displayed in Figure 6 appears much like a psychophysical timing function that one might see following training on a temporal discrimination, when testing with stimulus durations between the two training durations (e.g., Stubbs, 1976). That is, it appears that the pigeons are attempting to time the point in the session when the reversal will take place. This hypothesis has been confirmed by research in which, following training, pigeons were tested with longer and shorter intertrial intervals, thus causing the time from the start of the session to the reversal to either increase or decrease, respectively (McMillan & Roberts, 2012; Smith, Beckmann, & Zentall, 2017). When the intertrial interval is increased, the pigeons begin to reverse sooner in the

session (after fewer trials); when the intertrial interval is reduced, the pigeons begin to reverse later in the session.

It is surprising that the pigeons would attempt to time from the start of the session to the reversal, because that time would be somewhat variable, as it would depend on the rate with which the pigeons proceed through the trials. More important, the pigeons have a more reliable cue for the reversal available—the choice and consequence of the choice from the preceding trial(s). Specifically, if their previous choice was correct, it generally provides a cue to choose it again, whereas if it was not correct, it should serve as a cue to choose the other stimulus (i.e., such a strategy would result in the development of a win–stay/lose–shift strategy). Surprisingly, when the point in the session in which the reversal occurs is made unpredictable (i.e., it varies randomly from session to session), the pigeons' accuracy is not any better (Rayburn-Reeves et al., 2011). In fact, when by chance the reversal comes early in the session, pigeons tend to make many more perseverative errors, and when it comes late in the session, they tend to make many more anticipatory errors. Curiously, when the reversal is unpredictable, pigeons tend to be most accurate when the reversal occurs at the middle of the session. It is as if,

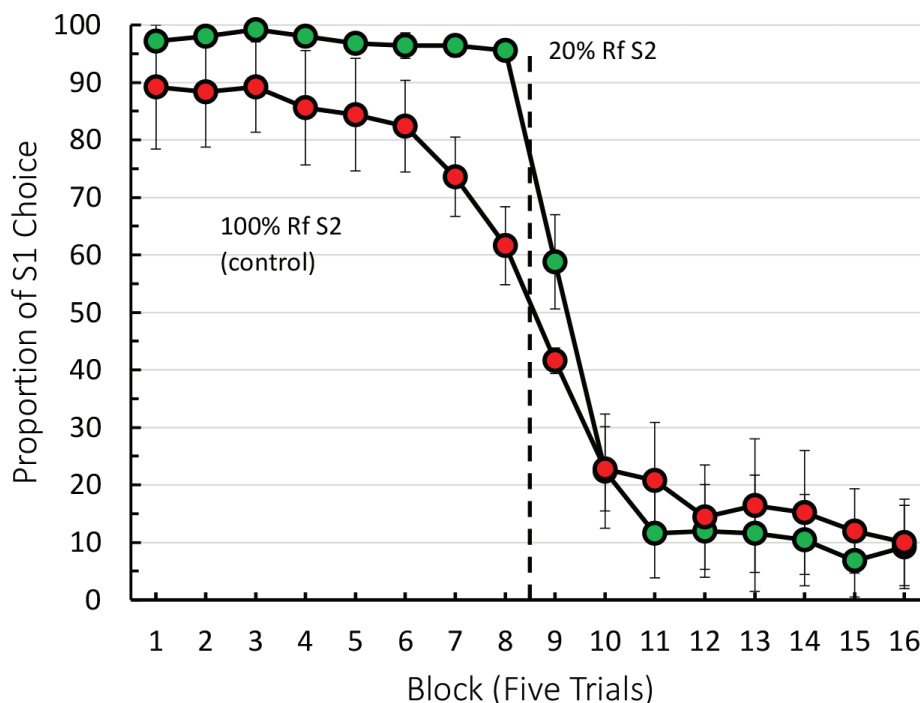
during training, although the reversal is unpredictable, the pigeons average the time into the session at which the reversal occurs.

It is important to note that not all animals choose suboptimally with this task. When monkeys were trained on this task, they showed virtually no anticipatory errors, but they did show some perseverative errors (Rayburn-Reeves, James, & Beran, 2017). Surprisingly, rats, but not pigeons, trained on a spatial midsession reversal task (e.g., left is correct for the first 40 trials, right is correct for the next 40 trials) show near-optimal accuracy (Rayburn-Reeves, Stagner, Kirk, & Zentall, 2013). Pigeons do show near-optimal accuracy with a spatial midsession reversal task when the intertrial interval is very short (1.5 s); however, their accuracy can be attributed to the repetitive response pattern involving the location of the stimulus and the feeder (Laude, Stagner, Rayburn-Reeves, & Zentall, 2014). Near-optimal accuracy by the pigeons appears to result from the short intertrial interval and the spatial nature of the task, because when the task with the short intertrial intervals involved visual stimuli such that the spatial location of the correct stimulus could not be anticipated, the typical numbers of both anticipatory and perseverative errors were found (Laude et al., 2014).

The question remains why pigeons persist in attempting to estimate the time from the start of the session to the midpoint of the session when the feedback from the preceding trial(s) would be a far more effective cue. One hypothesis is that they have difficulty remembering both the stimulus last chosen and the outcome of that choice, both of which would be needed to decide what to do on the next trial. The fact that reducing the duration of the intertrial interval to as short as 1.5 s did not improve accuracy on the visual discrimination form of the task suggests, however, that memory for the stimulus chosen and the resulting outcome is not likely the only problem.

A more direct test of the forgetting hypothesis was conducted by Smith et al. (2017). They inserted cues during the intertrial interval that could remind the pigeon which alternative it had selected (if the stimulus chosen had been red, the intertrial interval was lit by a house light in the ceiling; if the stimulus chosen had been green, the intertrial interval was lit by a house light at the top of the response panel) and the outcome of that choice (if the response had been correct, the feeder light remained on throughout the intertrial interval). Although providing appropriate feedback during the intertrial interval significantly improved task accuracy,

Figure 7. Mean percentage choice of the first correct stimulus (S1) as a function of trial number for Sessions 41 to 50. S1 was correct for the first 40 trials of each session. Incorrect stimulus (S2) was correct for the last 40 trials of each session. For the experimental group, correct S2 responses were reinforced 20% of the time. Rf = reinforcement. Error bars = ± 1 SEM (unpublished research; Zentall et al., 2019).



the pigeons continued to make both anticipatory and perseverative errors.

The results of a recent experiment by Santos, Soares, Vasconcelos, and Machado (2017) may provide a clue as to why pigeons do not perform more accurately with this task. Santos et al. compared pigeons' accuracy on the standard midsession reversal with a procedure in which correct choices of S1 were reinforced 100% of the time but correct choices of S2 were reinforced only 20% of the time. One would have expected this procedure to reduce anticipatory errors but increase perseverative errors (i.e., bias the pigeons to choose S1 both before and after the reversal). Not surprising, this procedure virtually eliminated anticipatory errors, but unexpectedly it did not increase perseverative errors. Thus, this manipulation resulted in a net increase in midsession reversal accuracy. In unpublished research, Zentall, Andrews, Case, and Peng (2019) have since replicated this effect (see Figure 7). Paradoxically, reducing the overall probability of reinforcement by 40% (80% after the reversal) resulted in an increase in overall accuracy. How is this phenomenon to be explained?

With the correct choice of S2 reinforced only 20% of the time, choice of S2 provides unreliable (ambiguous) feedback because the feedback from making a "correct" S2 response is often the same as feedback from making an incorrect S2 response. However, the feedback from making either a correct or incorrect S1 response remains a reliable cue. Thus, it appears that in the procedure in which there is 20% reinforcement of correct S2 responses, the unreliability of feedback from the S2 response encourages the pigeons to use the feedback from S1 responses as the primary basis for responding to S2. Because nonreinforcement of the choice of S1 is the only reliable cue to the possibility of reinforcement for the choice of S2, anticipatory errors are eliminated. Furthermore, because, in the absence of anticipatory errors, nonreinforcement of choice of S1 is now the only reliable event, it could serve as a better cue to choose S2. Thus, there is not a significant increase in perseverative errors, and midsession reversal accuracy actually improves.

It appears that in the case of the standard midsession reversal, why the pigeons use the time from the start of the trial to the reversal rather than the feedback from their choices is not because of the *absence* of reliable cues but presumably because of the symmetry or excess of reliable cues. Pigeons appear to have a bias to attend to cues for reinforcement rather than cues for the absence of reinforcement, and the cues for reinforcement appear to compete with each other. As the pigeon approaches

the midpoint of the session, the current cues may indicate that S1 is the correct response; however, anticipation of responding to S2 competes with the correct S1 response. Similarly, after the reversal, memory of the correct S1 response competes with the current feedback from the correct S2 response. Paradoxically, the task is made easier by devaluing the feedback from responding to S2 and encouraging the pigeons to rely primarily on feedback from the incorrect choice of S1.

If this hypothesis is correct, there may be other ways to shift the pigeons' attention from reinforcement associated with choice of S2 to nonreinforcement associated with choice of S1. For example, the symmetry between choice of S1 and S2 could be altered by making the response requirement to S2 greater (e.g., FR10) than the response requirement to S1 (FR1). Such an experiment is currently in progress.

Conclusions

When animals choose suboptimally, it may tell us something about the processes that underlie the behavior. Biologists would say that we have taken behavior that has evolved to have adaptive value for animals in their natural environment and placed it in an unnatural context in which it is no longer adaptive. Their point may be correct, and it encourages us to ask if there is a heuristic that in the natural environment might make this behavior adaptive.

The less is better effect tells us that in choosing between options that are not homogeneous, an organism may use a heuristic that averages the quality of the two options because under certain natural conditions such averaging may be more efficient than trying to quantify food availability. It may be that the quality of fruit on a tree, for example, is easier to estimate than the quantity of fruit on a tree, and the averaging heuristic may be used even when the quantities are easy to discriminate, as they should be in the less is better experiments.

The ephemeral reward task presents the animal with an unusual problem that may be related to the less is better task. Although the second, delayed reward is nominally the same as the first, delayed rewards typically have less value to animals than immediate rewards. If the animal averages the two rewards it may make them less valued than the single immediate reward (Zentall et al., 2016). Even in cases in which there is indifference between the two alternatives, the immediacy of a reward may make it difficult to associate that choice with a later reward. For this reason, delaying the initial reward may

help the animal to integrate the two rewards and demonstrate a significant degree of optimal choice.

Explanation of the pigeon's suboptimal choice on the midsession reversal task poses an even more perplexing problem. Choice by the pigeon on the typical task provides the pigeon with a great deal of information. Errors let the pigeon know that the reversal either has not yet occurred or has already occurred. Apparently, however, the pigeon has difficulty remembering both what it chose on the preceding trial and the outcome of that trial (Smith et al., 2017). If this is the case, although the time since the start of the session may appear to be a less reliable cue, it may involve less interference than accurate memory of the events on the preceding trial. Using time from the start of the session may turn out to be a heuristic that provides a less ambiguous cue for the pigeon. Paradoxically, greatly reducing the feedback for choice of the stimulus that will provide reinforcement during the last half of the session (S2) encourages the pigeon to better use the feedback from choice of S1, the stimulus that provides reinforcement during the first half of the session. If the pigeon attends to the consequence of choosing S1 and ignores the consequence of choosing S2, it appears to be able to use that information to determine the occurrence of the reversal and to stop using the less efficient timing heuristic.

Although biologists may object to exposing animals to tasks that are not typically encountered in nature, psychologist may use such tasks as diagnostics to explore the tendencies that animals have evolved to deal with the natural world. If environments change slowly enough, the genetic predispositions that appear to be responsible for the kinds of suboptimal choice described in the present review could be selected against and new predispositions emerge. However, in a rapidly changing environment, if these animals are to avoid extinction, they should be able to modify those naturally occurring tendencies. That is, flexibility in being able to maximize reinforcement when exposed to artificial laboratory procedures may provide a measure of a species' ability to adapt to the possibility of naturally occurring rapid environmental change.

The common theme that runs through the three examples of suboptimal choice described in the present article is that evolved heuristics that function reasonably well in nature may be less efficient under certain laboratory conditions. Is it possible that evolved heuristics can account for other examples of suboptimal choice by animals?

Zentall (2016) identified three examples of suboptimal or biased choice by animals. The first is *sunk cost*,

the tendency to continue with a schedule of reinforcement in which the animal is already engaged rather than switching to an alternative that would provide reinforcement sooner. The bias to continue a task that is already started may be suboptimal under certain laboratory conditions (Macaskill & Hackenberg, 2012; Magalhães, & White, 2014; Navarro & Fantino, 2005; Pattison et al., 2012), but in nature, switching between tasks may incur added costs including travel time, the possibility that expected resource will no longer be there, and the possibility of predation. Thus, in nature a preference for the relative known over a potentially better but relative unknown would be a useful heuristic.

The second example of suboptimal choice described in the earlier article was referred to as the *justification of effort effect* (Clement, Feltus, Kaiser, & Zentall, 2000). To demonstrate this bias, pigeons are trained to peck a white key, and on some trials, after a single peck, the pigeon can choose between red light and a yellow light, and a peck to red is reinforced. On other trials, 20 pecks are required, and then the pigeon can choose between a green light and a blue light, and a peck to green is reinforced. When pigeons are then given a choice between the two colors that were correct—red and green (with no prior pecking required)—they tend to prefer the green key. Clement et al. attributed this effect to contrast between the state of the pigeon just prior to the appearance of the red and green lights and the conditioned reinforcing value of the correct colored lights. They proposed that there should be greater positive contrast upon the appearance of the green light. The evolved heuristic that could account for this effect may be that if rewards that follow greater effort are given additional value (compared to rewards that follow less effort), it may provide the animals with added incentive to continue foraging, which should have additional survival value.

The third example of suboptimal choice described in the earlier article was thought to be analogous to human unskilled gambling (e.g., lottery tickets and slot machines). Pigeons could choose an option that 20% of the time provided them with a reliable cue for reinforcement and 80% of the time a reliable cue for the absence of reinforcement. Alternatively, they could choose an option that always provided them with a cue for 50% reinforcement. Under these conditions, the pigeons showed a strong preference for the 20% reinforcement alternative (Stagner & Zentall, 2010). Later research found that pigeons actually preferred 50% signaled reinforcement over 100% reinforcement (Case & Zentall, 2018). The fact

that the preference for the suboptimal alternative occurs only when reinforcement for choice of that alternative is signaled suggests that it is the value of the signal rather than its probability that determines the preference (20% signaled reinforcement over 50% unsignaled reinforcement; Stagner & Zentall, 2010). However, the fact that 50% signaled reinforcement was preferred over 100% reinforcement (Case & Zentall, 2018) suggests that there may also be contrast between the expected value of reinforcement (50% expected) and the obtained value of reinforcement (100% obtained) given choice of the suboptimal alternative, whereas there would be little contrast involving the optimal alternative (100% reinforcement expected and 100% reinforcement obtained). The hypothesis that contrast between what is expected and what is signaled to occur is essentially the same as what McDevitt et al. (2016) referred to as the “signal for good news” that occurs upon the appearance of the conditioned reinforcer following choice of the suboptimal alternative.

In all six of the examples of suboptimal or biased choice by pigeons presented in the present article, together with those presented in Zentall (2016), the suboptimal behavior can be explained in terms of evolved heuristics that work reasonably well in nature but sometimes fail under laboratory conditions. The relative stability of suboptimal choice in the case of both the signaled lower probability of reinforcement experiments and the midsession reversal experiments suggests that it may be very difficult for some animals to overcome those heuristics. Furthermore, it suggests that under certain conditions, some animals may not have the flexibility to easily modify their behavior if sudden changes in their environment require it.

References

- Ainslie, G. W. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior*, *21*, 485–489. doi:10.1901/jeab.1974.21-485
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, *59*, 177–181. doi:10.1037/h0047195
- Becker, G. (1976). *The economic approach to human behavior*. Chicago: The University of Chicago Press.
- Beran, M. J., Evans, T. A., & Ratliff, C. L. (2009). Perception of food amounts by chimpanzees (*Pan troglodytes*): The role of magnitude, contiguity, and wholeness. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*, 516–524. doi:10.1037/a0015488
- Beran, M. J., Ratliff, C. L., & Evans, T. A. (2009). Natural choice in chimpanzees (*Pan troglodytes*): Perceptual and temporal effects on selective value. *Learning and Motivation*, *40*, 186–196. doi:10.1016/j.lmot.2008.11.002
- Boysen, S. T., Berntson, G. G., Hannan, M. B., & Cacioppo, J. T. (1997). Quantity-based interference and symbolic representation in chimpanzees (*Pan-troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 76–86. doi:10.1037/0097-7403.22.1.76
- 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. doi:10.1046/j.1461-0248.2002.00295.x
- 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:10.1016/j.beproc.2018.07.018
- 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*, 100–106. doi:10.3758/BF03210727
- Estle, S. J., Green, L., Myerson, J., & Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science*, *18*, 58–63. doi:10.1111/j.1467-9280.2007.01849.x
- Fortes, I., Pinto, C., Machado, A., & Vasconcelos, M. (2018). The paradoxical effect of low reward probabilities in suboptimal choice. *Journal of Experimental Psychology: Animal Learning and Cognition*, *44*, 180–193. doi:10.1037/xan0000165

- Grutter, A. S. (2004). Cleaner fish use tactile dancing behavior as a preconflict management strategy. *Current Biology*, *14*, 10180–1083. doi:10.1016/j.cub.2004.05.048
- Houston, A. I., McNamara, J. M., & Steer, M. D. (2007). Violations of transitivity under fitness maximization. *Biology Letters*, *3*, 365–367. doi:10.1098/rsbl.2007.0111
- Hsee, C. K. (1998). Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making*, *11*, 107–121. doi:10.1002/(SICI)1099-0771(199806)11:2<107::AID-BDM292>3.0.CO;2-Y
- Kacelnik, A. (2006). Meanings of rationality. In S. Hurley & M. Nudds (Eds.), *Rational animals?* (pp. 87–106). Oxford, England: Oxford University Press. doi:10.1093/acprof:oso/9780198528272.003.0002
- Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, *21*, 463–473. doi:10.1901/jeab.1974.21-463
- Kralik, J. D., Xu, E. R., Knight, E. J., Khan, S. A., & Levine, J. W. (2012). When less is more: Evolutionary origins of the affect heuristic. *PLoS ONE*, *7*, e46240. doi:10.1371/journal.pone.0046240
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, *112*, 443–477. doi:10.1162/003355397555253
- Laude, J. R., Stagner, J. P., Rayburn-Reeves, R. M., & Zentall, T. R. (2014). Midsession reversals with pigeons: Visual versus spatial discriminations and the intertrial interval. *Learning & Behavior*, *42*, 40–46. doi:10.3758/s13420-013-0122-x
- Macaskill, A., & Hackenberg, T. D. (2012). The sunk cost effect with pigeons: Some determinants of decisions about persistence. *Journal of the Experimental Analysis of Behavior*, *97*, 85–100. doi:10.1901/jeab.2012.97-85
- Mackintosh, N. J., McGonigle, B., Holgate, V., & Vanderver, V. (1968). Factors underlying improvement in serial reversal learning. *Canadian Journal of Psychology*, *22*, 85–95. doi:10.1037/h0082753
- Magalhães, P., & White, K. G. (2014). The effect of a prior investment on choice: The sunk cost effect. *Journal of Experimental Psychology: Animal Learning and Cognition*, *40*, 22–37. doi:10.1037/xan0000007
- Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement. *Animal Learning & Behavior*, *25*, 131–147. doi:10.3758/BF03199051
- McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, *105*, 23–40. doi:10.1002/jeab.192
- McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the Experimental Analysis of Behavior*, *6*, 317–327. doi:10.1901/jeab.1997.68-317
- McMillan, N., & Roberts, W. A. (2012). Pigeons make errors as a result of interval timing in a visual, but not a visual-spatial, midsession reversal task. *Journal of Experimental Psychology: Animal Behavior Processes*, *38*, 440–445. doi:10.1037/a0030192
- Navarro, A. D., & Fantino, E. (2005). The sunk cost effect in pigeons and humans. *Journal of the Experimental Analysis of Behavior*, *83*, 1–13. doi:10.1901/jeab.2005.21-04
- Pattison, K. F., & Zentall, T. R. (2014). Suboptimal choice by dogs: When less is better than more. *Animal Cognition*, *17*, 1019–1022. doi:10.1007/s10071-014-0735-2

- Pattison, K. F., Zentall, T. R., & Watanabe, S. (2012). Sunk cost: Pigeons (*Columba livia*) too show bias to complete a task rather than shift to another. *Journal of Comparative Psychology*, *126*, 1–9. doi:10.1037/a0023826
- 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. doi:10.1037/a0036205
- Persky, J. (1995). The ethology of homo economicus. *Journal of Economic Perspectives*, *9*, 221–231. doi:10.1257/jep.9.2.221
- 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. doi:10.1007/s10071-016-0981-6
- 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. doi:10.1016/j.anbehav.2016.06.023
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, *17*, 15–22. doi:10.1901/jeab.1972.17-15
- Rayburn-Reeves, R. M., James, B. J., & Beran, M. J. (2017). Within-session reversal learning in rhesus macaques (*Macaca mulatta*). *Animal Cognition*, *20*, 975–983. doi:10.1007/s10071-017-1117-3
- Rayburn-Reeves, R. M., Molet, M., & Zentall, T. R. (2011). Simultaneous discrimination reversal learning in pigeons and humans: Anticipatory and perseverative errors. *Learning & Behavior*, *39*, 125–137. doi:10.3758/s13420-010-0011-5
- Rayburn-Reeves, R. M., Stagner, J. P., Kirk, C. R., & Zentall, T. R. (2013). Reversal learning in rats (*Rattus norvegicus*) and pigeons (*Columba livia*): Qualitative differences in behavioral flexibility. *Journal of Comparative Psychology*, *127*, 202–211. doi:10.1037/a0026311
- 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:10.1371/journal.pone.0049068
- Sanchez-Amaro, A., Pereto, M., & Call, J. (2016). Differences in between-reinforcer value modulate the selective-value effect in great apes (*Pan troglodytes*, *P. paniscus*, *Gorilla gorilla*, *Pongo abelii*). *Journal of Comparative Psychology*, *130*, 1–12. doi:10.1037/com0000014
- Santos, C. dos, Soares, C., Vasconcelos, M., & Machado, A. (2017, June). *Reinforcement bias performance in the midsession reversal task*. Poster presented at the meeting of the Society of the Quantitative Analysis of Behavior, Denver, CO.
- Simonson, I., & Tversky, A. (1992). Choice in context—tradeoff contrast and extremeness aversion. *Journal of Marketing Research*, *29*, 281–295.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.
- Smith, A. P., Beckmann, J. S., & Zentall, T. R. (2017). Mechanisms of midsession reversal accuracy: memory for preceding events and timing. *Journal of Experimental Psychology: Animal Learning and Cognition*, *43*, 62–71. doi:10.1037/xan0000124
- Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate. *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 212–220. doi:10.1037/xan0000092

- Spetch, M., Belke, T., Barnet, R., Dunn, R., & Pierce, W. (1990). Suboptimal choice in a percentage reinforcement procedure: Effects of signal condition and terminal-link length. *Journal of the Experimental Analysis of Behavior*, *53*, 219–234. doi: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:10.3758/PBR.17.3.412
- Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences. *Animal Behaviour*, *29*, 628–629. doi:10.1016/S0003-3472(81)80128-5
- Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Stubbs, D. A. (1976). Response bias and the discrimination of stimulus duration. *Journal of the Experimental Analysis of Behavior*, *11*, 223–238. doi:10.1901/jeab.1968.11-223
- Thorndike, E. L. (1911). *Animal intelligence: Experimental studies*. New York, NY: Macmillan. doi:10.5962/bhl.title.55072
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. doi:10.1126/science.185.4157.1124
- Vasconcelos, M., Machado, A., & Pandeirada, J. N. S. (2018). Ultimate explanations and suboptimal choice. *Behavioural Processes*, *152*, 63–72. doi:10.1016/j.beproc.2018.03.023
- Zentall, T. R. (2016). When humans and other animals behave irrationally. *Comparative Cognition & Behavior Reviews*, *11*, 25–48. doi:10.3819/ccbr.2016.110002
- 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:10.1016/j.beproc.2017.09.008
- Zentall, T. R., Andrews, D. M., Case, J. P., & Peng, D. N. (2019). Less information results in better midsession reversal accuracy by pigeons. Manuscript submitted for publication.
- Zentall, T. R., Case, J. P., & Berry, J. R. (2017a). Early commitment facilitates optimal choice by pigeons. *Psychonomic Bulletin & Review*, *24*, 957–963. doi: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: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*, 138–144. doi:10.1037/com0000026
- Zentall, T. R., Laude, J. R., Case, J. P., & Daniels, C. W. (2014). Less means more for pigeons but not always. *Psychonomic Bulletin & Review*, *21*, 1623–1628. doi:10.3758/s13423-014-0626-1
- Zentall, T. R., & Stagner, J. P. (2011). Suboptimal choice by pigeons: Failure to support the Allais paradox. *Learning and Motivation*, *42*, 245–254. doi:10.1016/j.lmot.2011.03.002