COMPARATIVE COGNITION & BEHAVIOR REVIEWS

Meliorating the Suboptimal-Choice Argument

Marilia Pinheiro de Carvalho
Federal University of Pará

Cristina Santos and Catarina Soares
University of Minho

Armando Machado
University of Aveiro

In relatively simple choice tasks, some animals seem to behave irrationally by making suboptimal choices. Zentall (2019) suggests that these animals may choose according to a variety of heuristics that are adaptive in their natural environments but maladaptive in the contrived laboratory settings. We argue that Zentall's specific heuristics range from the reasonable and testable to the conceptually confused and even inconsistent. To be useful, heuristics must be clearly defined, delimited, and coordinated with known behavioral processes.

Keywords: *heuristic, less-is-better effect, ephemeral reward task, midsession reversal task, optimality, nonhumans*

Zentall's (2019) target article, "What suboptimal choice tells us about the control of behavior," is in three parts. The first part reviews a set of studies that have yielded surprising findings: In relatively simple choice tasks, animals seem to behave irrationally by making suboptimal choices. The second part introduces a set of hypotheses to account for the surprising findings: Animals may behave according to a variety of heuristics that are adaptive in their natural environments but maladaptive in the contrived laboratory settings. The third part explains what suboptimal choice in fact tells us about the control of behavior.

In this commentary we argue that Part 1 is timely, interesting, and important; that Part 2, potentially the article's greatest contribution, includes imaginative, testable hypotheses alongside conceptually confused and even inconsistent hypotheses; and that Part 3 may be too vague to be useful. We conclude with some general remarks on the nature of the problems brought to our attention by the target article.

Part 1. Suboptimal Choice as a Subset of Surprising Research Findings

Like many other scientific research fields, the field of learning has always had its share of puzzling, seemingly irrational phenomena. A well-known case discovered in the 1960s is instinctive drift:

> A pig was conditioned to pick up large wooden coins and deposit them in a large "piggy bank." The coins were placed several feet from the bank and the pig required to carry them to the bank and deposit them, usually four or five coins for one reinforcement.... Pigs condition very rapidly, they have no trouble taking ratios, they have ravenous appetites (naturally), and in many ways are among the most tractable animals we have worked with. However, this particular problem-behavior developed in pig after pig, usually after a period of weeks or months, getting worse every day. At first the pig would eagerly pick up one dollar, carry it to the bank, run back, get another; carry it rapidly and neatly, and so on, until the ratio was complete. Thereafter, over a period of weeks the behavior would become slower and slower. He might run over

eagerly for each dollar, but on the way back, instead of carrying the dollar and depositing it simply and cleanly, he would repeatedly drop it, root it, drop it again, root it along the way, pick it up, toss it up in the air, drop it, root it some more, and so on.... The behavior persisted and gained in strength in spite of a severely increased drive—he finally went through the ratios so slowly that he did not get enough to eat in the course of a day.... This problem behavior developed repeatedly in successive pigs. (Breland & Breland, 1961, p. 683).

What surprises us in this case is that behavior appropriate to the reinforcement contingencies was initially learned but then replaced by behavior inappropriate to the contingencies. The food promptly earned at first was significantly delayed later on.

Another case, of a different kind, concerns avoidance and extinction. Common sense predicts that, having learned to avoid electric shock by jumping to the other side of a shuttle box when a light turns off, a dog will continue to jump when the shock is discontinued. In contrast, early versions of two-factor theory (e.g., Mowrer, 1947) predicted that the jumps would cease because the light-off stimulus would extinguish its association with the shock. The results are generally closer to the commonsense prediction (e.g., Solomon & Wynne, 1954). In this case, the experimental findings—the resistance to extinction following avoidance training—surprise us because they are inconsistent with a theory. (The one-trial, long-delay taste-aversion learning is another case of behavior that the theory predicted should not be acquired.)

The cases described by Zentall in the target article are similar to the preceding cases in the sense that they also surprise us because they are at odds with our basic conceptions or theories of behavior and learning. The novelty of Zentall's cases is that they all involve explicit choice. In the less-is-better and the ephemeral

Author Note: Marilia Pinheiro de Carvalho, Federal University of Pará, Núcleo de Teoria e Pesquisa do Comportamento, Campus Universitário do Guamá, Rua Augusto Corrêa, 01, CEP: 66075-110, Belém, Pará.

Correspondence concerning this article should be addressed to Marilia Pinheiro de Carvalho at marilia.pinheiro.carvalho@gmail.com

reward tasks, the animal chooses the option that yields less food, and in the midsession reversal (MSR) task the animal chooses the incorrect option despite the presence of seemingly easy-to-use local cues to the correct option.

When research findings violate basic conceptions and theories—when they surprise us—they typically set the agenda for subsequent empirical and theoretical research. Empirically, researchers attempt to assess the reliability and generality of these findings, including the conditions under which they occur. Theoretically, researchers revise their conceptions and theories to reconcile them with the surprising findings. Hence, as Zentall claims, the surprising findings of suboptimal, and seemingly irrational, choice may indeed help us better understand the learning processes. Their description is both interesting and important.

Part 2. Hypothetical Processes of Suboptimal Choice

The second part of the target article presents Zentall's specific hypotheses to "save the appearances" in each of the suboptimal choice categories. They are arguably the author's most important contribution. We consider them in context, following the same order as in the target article.

The Less-Is-Better Effect

In the seminal work by Silberberg, Widholm, Bresler, Fujita, and Anderson (1998), long-tailed macaques, rhesus monkeys, and a chimpanzee were indifferent between a mixture option (a slice of a preferred food plus a slice of a less preferred food) and a single-item option (a slice of the preferred food). To explain this surprising result, the authors proposed that the primates valued the preferred item of each option but not the less preferred item, so the two options had similar values. However, a decade later, Beran, Ratliff, and Evans (2009, Experiments 2–3) showed that chimpanzees preferred the single-item option to the mixture option, the "less-is-better" effect.

Subsequent research showed that the effect cannot be due to a bias for smaller amounts of food, because when the animals choose between two slices of the preferred food and one slice of it, or between two slices of the less preferred food and one slice of it, they prefer the two slices. The effect is also not because the more preferred and less preferred items have positive and negative value, respectively, such that the mixture would have a smaller net value than the single-item option, for

subjects readily eat the less preferred item (Kralik, Xu, Knight, Khan, & Levine, 2012); eating the less preferred item seems inconsistent with the negative value attributed to it. The effect also rules out the hypothesis that the less preferred item has no value, as Silberberg et al. (1998) suggested, because in this case the subject would be indifferent rather than prefer the single-item option.

The current best hypothesis seems to be that both items have positive value, one greater than the other, and that the animals choose between the largest value (single-item option) and the average of the two values (mixture option). Because the average of two different numbers is always less than the larger of the two numbers, the animals prefer the single-item option (e.g., Chase & George, 2018; Kralik et al., 2012; Zentall, 2019). Zentall claims that this averaging rule or heuristic applies when animals choose between different quality foods.

The hypothesis is ingenious, but it is challenged by a few facts. First, virtually identical experimental conditions have yielded two sets of results: preference for the single-item option and indifference between the two options; some studies have even produced both sets of results (e.g., Beran, Ratliff, & Evans, 2009; Sánchez-Amaro, Peretó, & Call, 2016; Silberberg et al., 1998, Experiment 4; Zentall, Laude, Case, & Daniels, 2014). We do not know what causes one or the other set. Second, not all subjects in the experiments that confirmed the less-is-better effect showed the effect (e.g., Beran, Ratliff, & Evans, 2009, Experiments 1 and 3; Chase & George, 2018, Experiment 1; Pattison & Zentall, 2014; Sánchez-Amaro et al., 2016). How then do we explain the opposite choice, or the source of the individual differences? Third, when the mixture option included two highly preferred foods—one slightly more preferred than the other—and the single-item option included one of these two, the animals preferred the mixture option, but the degree of preference varied with the content of the single-item option (Sánchez-Amaro et al., 2016, Experiment 2). These data suggest that if the mixture option included two qualitatively distinct but equally preferred items—for example, A and B, with Value(A) = Value(B)—and the single-item option included one of these items (e.g., B), then the animals would not be indifferent, as the averaging hypothesis predicts, but prefer the mixture. Zentall's averaging hypothesis remains to be more stringently tested.

Moreover, given a choice between a mixture option containing 20 g of the preferred food placed side by side with 5 g of the same preferred food, and a single-item option containing 20 g of the preferred food, chimpanzees were either indifferent or preferred the single, smaller option. But when the two items composing the mixture were presented one on top of the other, the chimpanzees preferred the larger option (Beran, Evans, & Ratliff, 2009, Experiment 2). Zentall mentions the first result, consistent with the averaging hypothesis, but not the second, inconsistent with it.

Beran, Evans, and Ratliff (2009) suggested that *fragmented food* may be aversive, hence the preference for the single-item option over the larger, fragmented option. When the researchers placed one item atop the other, they may have disguised the fragmentation and thereby reinstated preference for the larger option. According to this hypothesis, then, avoidance of the mixture in the prototypical choice task stems from avoidance of an option that offers fragmented, nonuniform items. Although preferring the single-item option still incurs the cost of favoring the lesser amount of food, avoiding fragmented/nonuniform items might protect the animal from previously-interfered-with and discarded food. If this proves to be the case, the less-is-better effect, suboptimal as it may be from the point of view of caloric intake, may be optimal from the point of view of protecting the animal from dangerous food. This new, unanticipated function would remove the suboptimal tag from the effect.

Future research needs to contrast these hypotheses as well as determine how quality and quantity combine to determine reinforcement value. Zentall's hypothesis that deprivation level is important may help us understand the conditions in which the gastronomic value of food trumps its nutritional value.

The Ephemeral Reward Task

In one variation of the task, pigeons choose between two keys, red (R) and green (G). A peck at R yields a bit of food and ends the trial; a peck at G yields a bit of food but then, because R remains present, a peck at R yields another bit of food and ends the trial. Exclusive preference for G yields twice as many rewards as exclusive preference for R; surprisingly, the pigeons slightly prefer R.

Zentall advances the differential reinforcement hypothesis, although we prefer to phrase it in terms of associative strength: Whenever G is chosen, both G and R associate with the reward, G with the first reward and R with the second; when R is chosen, only R associates with the reward. As trial effects accumulate, R gains more associative strength than G, hence the preference for R.

But if some animals behave suboptimally and prefer R (e.g., pigeons, rats), others behave optimally and prefer

G (e.g., cleaner fish and parrots). To explain the species differences, Zentall brings in the concept of impulsivity. The animals that prefer R act impulsively, and for that reason the second reward of option G is not associated with G. The animals that prefer G are less impulsive and for that reason associate G with both rewards. However, the impulsive animals can learn to choose optimally if we reduce their impulsivity by inserting a delay between choice and food (i.e., if R is chosen, food follows 20 s later and the trial ends; if G is chosen, food follows 20 s later, but because R remains present a single peck at it delivers food *without delay*). Zentall argues that inserting the delay may "encourage the pigeons to associate the second reinforcer with the initial choice of alternative B" (p. 9).
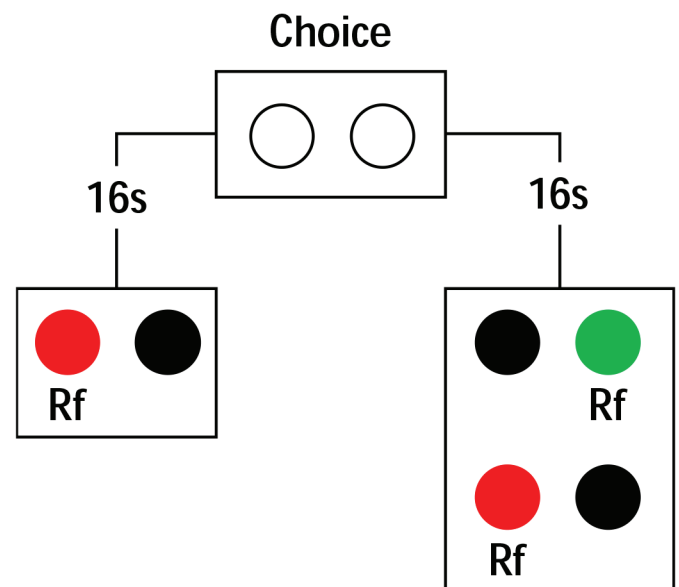
The hypothesis is confusing, perhaps because it is insufficiently elaborated. First, it does not explain how impulsivity prevents associations when present (pigeons) and permits them when absent or reduced (cleaner fish). Whether in its colloquial sense of rapid, unplanned choices with little regard to long-term consequences or in its technical senses of preference for a smaller-sooner (SS) reward over a larger-later (LL) reward, or a steep delay-discounting gradient, in none of these senses is the relation between impulsivity and association clear. Second, all else equal, increasing the interval between two events makes their association harder to learn. Yet, without justification, the hypothesis assumes the opposite trend: that increasing the delay between choice of G and the second food facilitates their association. The concepts of impulsivity and association seem to be defined and coordinated in a rather Pickwickian way.[1]

The experimental test of the hypothesis showed that, with the delay inserted, pigeons and rats preferred G over R; their choices became optimal. Zentall argues that the delay reduced impulsivity in the same way it does in the commitment-response variation of the self-control procedure (see Figure 4 of the target article): Given a choice between an SS reward and an LL reward, pigeons prefer SS when consumption is imminent but seem to prefer LL when consumption is delayed (i.e., away from "temptation"). The evidence for the latter is that, well in advance of consumption, pigeons prefer an option that commits them to the LL reward to an option that preserves the choice between the SS and LL rewards. In Figure 4 of the target article, pigeons prefer the left key of the initial link. However, when the pigeons choose the right key, and 16 s later choose between the SS and LL rewards, they typically choose the SS reward. The SS reward is less valued than LL far from consumption but more valued than LL close to consumption. Hence, to extend the commitment response interpretation to the ephemeral reward task, one would need to show that option G is both (a) less valued than option R when consumption is imminent and (b) more valued than option R when consumption is delayed. To that end, the ephemeral reward task could be inserted in a response commitment procedure (see Figure 1).

An interpretation of the ephemeral reward task that does not rely on the concept of impulsivity starts with Zentall's differential reinforcement hypothesis just mentioned and then adds to it the idea that pigeons and rats may fail to associate G with the second reward because of interference by the intervening events, including the first reward and pecking the R key. At this light, the 20-s delay is effective because it increases the exposure to G, strengthens its memory trace, and reduces the interference effects; the association of G with the second reward is facilitated. To test this hypothesis, one could increase the duration of R following the first reward and see if, as this duration increased, preference for G decreased.

---

[1] To make matters more confusing, Zentall advances in the Discussion section yet another hypothesis: Option G includes two rewards, delivered d1 and d2 seconds after choice; option R includes only one reward delivered d1 seconds after choice. If the value of an option equals the average of the values of its rewards, and if the value of a reward decays with delay, then G will have less value than R. The hypothesis assumes that, regardless of impulsivity, the animal learned to associate the choice of G with the second reward!

Figure 1. Commitment response procedure for the ephemeral reward task.

The Midsession Reversal Task

Given two stimuli (S1 and S2), choices of S1 are reinforced during the first 40 trials and choices of S2 during the last 40 trials. At the steady state, pigeons show anticipatory and perseverative errors, which are consistent with timing the moment of reversal from the beginning of the session. Why the pigeons resort to timing rather than use local cues (e.g., win-stay/lose-shift strategy), is the surprising outcome.

Zentall's first account of these results was that the pigeons have difficulty remembering their previous choices and outcome, both when the task requires a visual (i.e., S1 = red key, S2 = green key) or a spatial (i.e., S1 = left key, S2 = right key) discrimination with a relatively long inter-trial interval ITI (ITI; e.g., greater than 1.5 s; Laude, Stagner, Rayburn-Reeves, & Zentall, 2014; Rayburn-Reeves, Laude, & Zentall, 2013; Rayburn-Reeves, Molet, & Zentall, 2011). Given the alleged difficulty remembering the trial events, the chosen stimulus, and the received outcome, pigeons use the time from the beginning of the session as a cue to switch from S1 to S2.

The hypothesis of memory limitations is more plausible for visual than spatial discriminations, for if Williams (1972) found that pigeons do not learn a win-stay/lose-shift strategy when a color discrimination is required, Shimp (1976) found that pigeons learn this strategy with a spatial discrimination, even when the ITI lasts 6 s. Moreover, focusing on the pigeons' memory difficulties may well explain the relatively small number of anticipatory and perseverative errors around the reversal moment, but only at the cost of failing to explain the large number of correct choices on the remaining trials! That is, we should not forget the obvious fact that on the vast majority of trials, pigeons perform correctly.

Furthermore, timing requires memory. To wit, timing models such as the scalar expectancy theory (Gibbon, 1977, 1991) or the learning-to-time model (Machado, 1997; Machado, Malheiro, & Erlhagen, 2009), for example, include memory components—memory bins that store reinforcement times in scalar expectancy theory, and variable-strength associative links connecting behavioral states to operant responses, in the learning-to-time model. Any account that substitutes timing for memory to explain the MSR task results may shift the focus from local cues (responses and their outcomes) to a global cue (time), but it cannot do without the learning of response-outcome relations.

Zentall's second account argues that pigeons use timing due to the symmetry and excessive number of reliable cues in the MSR task. The cues for reinforcement, in particular, may compete with one another in complex ways, both before and after the reversal:

> Why the pigeons use the time from the start of the trial [sic] to the reversal rather than use 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.... 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. (Zentall, 2019, p. 12)

Timing in the MSR task is no longer justified by memory limitations but by cue competition. However, Zentall's cue-competition account needs to be further elaborated, preferably under the form of a dynamic model to accommodate the following findings. In the standard MSR task, when the ITI doubles, pigeons start to choose S2 around the first quarter of the session, which corresponds approximately to the time of the original reversal (McMillan & Roberts, 2012; Smith, Beckmann, & Zentall, 2017). This result supports the idea that time from the beginning of the session is indeed the main cue for switching from S1 to S2, and it is compatible with the cue-competition hypothesis. But when the ITI's duration is halved, pigeons switch shortly after the reversal, not at the end of the session, which would be the time corresponding to the original reversal (McMillan & Roberts, 2012; unpublished data from our lab also confirms this effect). This result suggests that time is not the only operative cue, that pigeons remain sensitive to the outcome of their choices (e.g., extinction following several S1 choices), remember the response-outcome pairing of the previous trial, and adapt their behavior to the current contingencies when time no longer predicts reinforcement. In any case, how time and local cues combine to cue behavior remains unclear.

In addition, according to Zentall's cue-competition hypothesis, when there are fewer cues available and S1 is the only reliable source of information (i.e., S1 = 100% and S2 = 20% reinforcement), pigeons adhere to the feedback provided by S1. They do not incur anticipatory errors and reduce perseverative errors significantly. Although Zentall does not elaborate the case, the cue-competition hypothesis seems to predict that when S2 is the only reliable source of information (S1 = 20%, S2 = 100% reinforcement), by virtue of the asymmetry and reduction of the number of reliable cues, overall performance should also improve. The data suggest

otherwise, though. Pigeons make a large number of anticipatory errors and overall performance worsens (Santos, Soares, Vasconcelos, & Machado, in press).

Unfortunately, Zentall did not extend his account to the case of S1 = 20% and S2 = 100% reinforcement. Reasoning by analogy yields the following: When the only reliable source of information is S2, pigeons adhere to the feedback of S2, that is, they continue to peck S2 after reinforcement and switch to S1 after nonreinforcement. But feedback from S2 can be obtained only by pecking S2. Hence, Zentall's account seems to predict anticipation errors. However, the account does not predict the proportion of these errors, nor does it explain why errors remain relatively rare during the very first trials.

An alternative explanation is that the difference in reinforcement probabilities between S1 (20%) and S2 (100%) biases the pigeons' time estimate of the reversal moment. But in Zentall's cue-competition account, pigeons would not need to time when there is only one reliable cue. More generally, it remains to be explained why making S1 or S2 the only reliable sources of information has such asymmetrical effects on performance (Santos et al., in press). In the absence of a quantitative dynamic model, cue competition remains so vague that its empirical test is virtually impossible.

Part 3. Beyond the Heuristic Value of Heuristics

After analyzing the various cases of suboptimal choice, Zentall concludes that "suboptimal behavior can be explained in terms of evolved heuristics that work reasonably well in nature but sometimes fail under laboratory conditions" (p. 14). This conclusion may be premature, because the evidence for the specific heuristics is conspicuously missing. With respect to the less-is-better effect, Zentall does not show that, in their natural settings, animals choose based on the average quality of each option and not on the quantity of each option or, more likely, on both quality and quantity. With respect to the ephemeral reward task, no specific, well-defined heuristic was proposed. With respect to the MSR task, Zentall also does not show that in natural settings animals rely on timing when local cues are ambiguous, excessive, or symmetric.

Suboptimal choice behavior seems to demand heuristics. But some cases of apparent suboptimal choice may turn out to be cases of optimal choice when we unravel a hitherto unknown biological function; for example, aversion to fragmented food may prevent the animal from eating discarded food. Other cases of suboptimal behavior may simply denounce the shortcomings of our current theories, for example, at identifying the functional choice set or the operative constraints in a learning situation. Zentall himself provides an example when he discusses self-control choices that seem irrational until we consider the animal's time horizon—the high-probability amount did not yield enough nutrients to keep the animal alive for the night. Once the researchers unraveled this critical variable, the suboptimal tag was abandoned, and with it the reason for the heuristic.

Reading the various accounts of suboptimal choice that Zentall proposes—all revealing their author's ingenuity and creativity—one cannot help but think that without conceptual clarity regarding definition and use, and without either a plausibility argument or a modicum of supporting evidence, heuristics remain "just so stories." But even when we define heuristics clearly and provide either a strong plausibility argument or empirical evidence supporting them, we still need to coordinate heuristics with known behavioral processes. Take the timing heuristic inspired by the MSR task results. Discriminative stimuli interact in different ways: They may have additive or subtractive effects on behavior, for example. They may even cancel each other, as when one stimulus cues the animal to move forward and another stimulus cues the animal to move backward. In this case, a third stimulus may determine the animal's behavior (e.g., a displacement activity). Stimulus interactions are behavioral processes, or part of behavioral processes, but the operation of a behavioral process is not a heuristic. And if we say that the heuristic consists of using a global cue (e.g., time) when local cues (responses and their outcomes) are excessive or unreliable, we still need to identify the basic mechanisms that implement the heuristic—that detect when local cues are *excessive* or *unreliable*, for example.

To conclude, what indeed does suboptimal choice tell us about the control of behavior? Like any other set of surprising findings, suboptimal choice tells us that we do not understand the phenomenon of choice—our theories of choice are either wrong or incomplete. Furthermore, each specific instance of suboptimal choice tells us that we need to conduct more studies to characterize the phenomenon and identify its boundary conditions. For example, we do not know whether the less-is-better effect extends to mixtures of three or more differently valued items, or whether the ephemeral reward task effect occurs when the two options remain spatially different (e.g., a concurrent chain schedule with two white keys as initial links, one terminal link consisting

of a 2-s red light followed by food, and the other terminal link consisting of a 2-s green light followed by food and then a 2-s red light also followed by food).

We also need to fathom new hypotheses about the behavioral processes of suboptimal choice, clarify, and test them. To that end, knowledge of an animal's phylogenetic and ontogenetic histories and of how the animal adapts to its natural environment is likely to be crucial. But this knowledge may be as crucial as it is hard to obtain; most of our statements about phylogenetic history or reinforcement history qualify more as ad hoc speculation than knowledge. As for heuristics, to move them from the domain of private science where "anything goes" to the domain of public science where conceptual coherence and empirical sensitivity rule, we need to define them clearly, identify the conditions that activate them, and coordinate them with currently known behavioral processes. Otherwise, heuristics will multiply and range widely yet may move us little further forward.

References

Beran, M., Evans, T., & Ratliff, C. (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., Ratliff, C., & Evans, T. (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

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist, 16*, 681–684. doi:10.1037/h0040090

Chase, R., & George, D. (2018). More evidence that less is better: Sub-optimal choice in dogs. *Learning & Behavior, 46*, 462–471. doi:10.3758/s13420-018-0326-1

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279–325. doi:10.1037/0033-295X.84.3.279

Gibbon, J. (1991). Origins of scalar timing. *Learning and Motivation, 22*, 3–38. doi:10.1016/0023-9690(91)90015-Z

Kralik, J., Xu, E., Knight, E., Khan, S., & Levine, W. (2012). When less is more: Evolutionary origins of the affect heuristic. *PLoS ONE, 7*, e46240. doi:10.1371/journal.pone.0046240

Laude, J. R., Stagner, J. P., Rayburn-Reeves, R., & 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

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 104*, 241–265. doi:10.1037/0033-295X.104.2.241

Machado, A., Malheiro, M., & Erlhagen, W. (2009). Learning to time: A perspective. *Journal of the Experimental Analysis of Behavior*, 92, 423–458. doi:10.1901/jeab.2009.92-423

McMillan, N., & Roberts, W. (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

Mowrer, O. H. (1947). On the dual nature of learning—A reinterpretation of "conditioning" and "problem solving." *Harvard Educational Review, 17*, 102–148.

Pattison, K., & 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

Rayburn-Reeves, R. M., Laude, J. R., & Zentall, T. R. (2013). Pigeons show near-optimal with-stay/lose-shift performance on a simultaneous-discrimination, midsession reversal task with short intertrial intervals. *Behavioural Processes, 92*, 65–70. doi:10.1016/j.beproc.2012.10.011

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

Sánchez-Amaro, A., Peretó, 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., Soares, C., Vasconcelos, M., & Machado, A. (in press). The effect of reinforcement probability on time discrimination in the midsession reversal task. *Journal of the Experimental Analysis of Behavior*.

Shimp, C. (1976). Short-term memory in the pigeon: The previously reinforced response. *Journal of the Experimental Analysis of Behavior, 26*, 487–493. doi:10.1901/jeab.1976.26-487

Silberberg, A., Widholm, J., Bresler, D., Fujita, K., & Anderson, J. (1998). Natural choice in nonhuman primates. *Journal of Experimental Psychology: Animal Behavior Processes, 24*, 215–228. doi:10.1037/0097-7403.24.2.215

Smith, A., Beckmann, J., & 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

Solomon, R. L., & Wynne, L. C. (1954). Traumatic avoidance learning: The principles of anxiety conservation and partial irreversibility. *Psychological Review, 61*, 353–385. doi:10.1037/h0054540

Williams, B. (1972). Probability learning as a function of momentary reinforcement probability. Journal of the Experimental Analysis of Behavior, 17, 363–368. doi:10.1901/jeab.1972.17-363

Zentall, T. R. (2019). What suboptimal choice tells us about the control of behavior. Comparative Cognition & Behavior Reviews, 14, 1–17. doi:10.3819/CCBR.2019.140001

Zentall, T. R., Laude, J., Case, J., & Daniels, C. (2014). Less means more for pigeons but not always. Psychonomic Bulletin & Review, 21, 1623–1628. doi:10.3758/s13423-014-0626-1