

Original citation:

McDevitt, Margaret A. , Dunn, Roger M., Spetch, Marcia L. and Ludvig, Elliot Andrew. (2016) When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, 105 (1). pp. 23-40.

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/74461>

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

"This is the peer reviewed version of the following article: McDevitt, Margaret A. , Dunn, Roger M., Spetch, Marcia L. and Ludvig, Elliot Andrew. (2016) When good news leads to bad choices. *Journal of the Experimental Analysis of Behavior*, 105 (1). pp. 23-40., which has been published in final form at <http://dx.doi.org/10.1002/jeab.192> . This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#)."

A note on versions:

The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher's version. Please see the 'permanent WRAP URL' above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

When Good News Leads to Bad Choices

Margaret A. McDevitt

McDaniel College

Roger M. Dunn

San Diego State University

Marcia L. Spetch

University of Alberta

Elliot A. Ludvig

University of Warwick

Author Note

Support was provided by the Natural Sciences and Engineering Research Council of Canada to MLS.

Correspondence concerning this article should be addressed to: Margaret A. McDevitt, Department of Psychology, McDaniel College, 2 College Hill, Westminster, MD 21157. E-mail: mmcdevit@mcdaniel.edu

Abstract

Pigeons and other animals sometimes deviate from optimal choice behavior when given informative signals for delayed outcomes. For example, when pigeons are given a choice between an alternative that always leads to food after a delay and an alternative that leads to food only half of the time after a delay, preference changes dramatically depending on whether the stimuli during the delays are correlated with (signal) the outcomes or not. With signaled outcomes, pigeons show a much greater preference for the suboptimal alternative than with unsignaled outcomes. Key variables and research findings related to this phenomenon are reviewed, including the effects of durations of the choice and delay periods, probability of reinforcement, and gaps in the signal. We interpret the available evidence as reflecting a preference induced by signals for good news in a context of uncertainty. Other explanations are briefly summarized and compared.

Keywords: Suboptimal choice, preference, signals, context, conditioned reinforcement, probabilistic, pigeons

When Good News Leads to Bad Choices

Outcomes are rarely certain in nature. Chasing prey or foraging for seeds does not always lead to food. The shortest line at the grocery store is not always the quickest. Submitting a manuscript may or may not result in publication. Our lives are filled with choices that involve uncertainty, or probabilistic outcomes. The role of this uncertainty in the behavior of humans and other animals has been the target for much comparative research.

Despite the importance of making good choices, human and nonhuman animals do not always choose optimally (see Ariely, 2008; Fantino, 2012; Kahneman, 2011; Zentall, 2015). One striking example of such suboptimal behavior comes from pigeons making decisions for delayed food rewards (e.g., Dunn & Spetch, 1990; Gipson, Alessandri, Miller, & Zentall, 2009; Kendall, 1974).

Consider a hungry pigeon given repeated choices between two options. One option always leads to food after a 40-s wait. The other option sometimes leads to food after 40 s and sometimes leads to no food after 40 s. The better choice is obvious – choosing the option that leads to food 100% of the time doubles the overall food intake compared to choosing the 50% option. This optimal choice would be expected not only from theories of rational choice but also from biological perspectives such as optimal foraging theory (Pyke, Pulliam, & Charnov, 1977); the importance of food procurement for survival should favor behavioral mechanisms that maximize energy intake while minimizing time, energy expenditure, and exposure to danger.

Indeed, under some circumstances, pigeons display strong preference for the more certain, and hence more optimal, option consistent with these expectations. Just as humans sometimes make irrational choices, however, pigeons sometimes behave suboptimally, choosing the 50% option as often or even more often than the one that provides more frequent food (e.g.,

Dunn & Spetch, 1990; Gipson, Alessandri, Miller, & Zentall, 2009; Kendall, 1974, 1985; Spetch, Belke, Barnet, Dunn, & Pierce, 1990). Kendall was the first to show that, under certain conditions, pigeons will choose an alternative that provides food only 50% of the time over one that always provides food (Kendall, 1974, 1985). In his studies, the delay to food was the same on the 50% and 100% options. Thus, choice of the 50% option was clearly suboptimal in terms of acquiring food. This peculiar finding, that pigeons will sometimes choose the suboptimal option has been replicated many times and with several procedural variations in a growing body of literature we review here.

The critical variable that has emerged from this work is the presence of stimuli that predict, or signal, the upcoming outcome in an uncertain context. Figure 1 shows a typical signaled procedure (top panel) and an unsignaled procedure (bottom panel). Both procedures involve a choice between a 100% option and a 50% option, and both have the same delay periods. The only difference between the two procedures is that the red and green delay stimuli are perfectly correlated with the food and no-food outcomes in the signaled procedure and uncorrelated in the unsignaled procedure. Although the presence or absence of signals does not alter the food payoff (it remains a choice between 100% and 50% food), the pigeons' behavior is very sensitive to the change. For example, Dunn and Spetch (1990, Experiment 3) directly compared signalled and unsignalled conditions in an ABA design. Figure 2 shows that pigeons frequently chose the 50% option when the outcomes were signalled, but switched to preference for the 100% option when they were not signalled. Note that any choice of the 50% option is suboptimal in the sense that it lowers the overall rate of food obtained. This preference for signalled reinforcers also resembles some casino games (e.g., slot machines), with implications for understanding human gambling and other real-world examples of suboptimal behavior (Lalli

& Mauro, 1995; Laude, Beckmann, Daniels, & Zentall, 2014; Zentall & Stagner, 2011a). For an excellent review of the implications of suboptimal choice for gambling in humans, see Zentall and Laude (2013).

Our review focuses on the psychological mechanisms that explain *why* pigeons are sometimes drawn to suboptimal choices in the case of probabilistic outcomes. We summarize the body of literature that illustrates the function of predictive signals and other key variables in driving suboptimal choice. We then describe a framework for understanding the role of the signal for food (good news) and the signal for no food (bad news) and briefly review other approaches.

Factors That Affect Suboptimal Choice

Six variables have proved critical in the demonstration of suboptimal choice: (1) the contingencies between the delay stimuli and the outcomes, (2) the response requirement in the choice phase, (3) the delay to food reward, (4) the contiguity between the choice response and presentation of the signal for good news, (5) the probability of reinforcement, and (6) factors related to self-control.

1. Signal Contingencies

As noted above, Kendall (1974) showed that signals for reward can exert powerful effects on behavior that run counter to the overall rate of food delivery. Frequent choice of the suboptimal option occurred only when the delayed food on the 50% side was signalled by a stimulus that immediately followed choice of that option. The importance of the signal contingencies has now been demonstrated across a variety of probabilistic choice procedures (e.g., Dunn and Spetch, 1990; Spetch et al., 1990; Stagner, Laude, & Zentall, 2012; Zentall & Stagner, 2012).

In one recent procedure that generated extreme swings in preference, Stagner and Zentall (2010) compared 20% reinforcement with 50% reinforcement. When outcomes on both alternatives were not signalled, choices of the suboptimal 20% alternative were infrequent (about 15% of the choices). When outcomes were signalled on the 20% alternative but not on the 50% alternative, the 20% alternative was chosen on roughly 90% of the trials – an extreme preference for the suboptimal option. Another variant of this basic procedure pits a small certain reinforcer against a low probability of a larger reinforcer. As Zentall and Stagner (2011a) noted, varying both amount and probability provides a closer analogue to the human gambling experience. They compared one alternative that ended with a 100% probability of 3 pellets to another alternative that ended with a 20% probability of 10 pellets and 80% probability of no pellets. Their results are presented in Figure 3 and show that when the outcomes were not signalled during the delays, pigeons preferred the alternative that ended with 3 pellets. When the outcomes were signalled, they preferred the probabilistic alternative that ended with an average of 2 pellets.

The importance of signal contingencies is also highlighted in recent neuroscience studies with monkeys. Bromberg-Martin and Hikosaka (2009) gave monkeys a choice between two options that both led 50% of the time to a large water reward and 50% of the time to a small water reward, using visual saccades as the choice response. For both options, looking at one of the choice stimuli led to one of two visually distinct stimuli that remained on for a 2.5-s delay and was then followed by the presentation of the water reward directly into the monkey's mouth. The critical difference between the two options was that for one of them the delay stimuli signalled which reward would occur, whereas for the other option, the delay stimuli were uncorrelated with the reward. Both monkeys developed a strong preference for the option with signaled outcomes. Moreover, a recent follow-up study demonstrated that thirsty monkeys are

even willing to forego some water in order to receive this advance information about which reward was upcoming (Blanchard, Hayden, & Bromberg-Martin, 2015).

These studies also provided evidence for a potential neural mechanism for this preference for good news signals (Blanchard et al., 2015; Bromberg-Martin & Hikosaka, 2009). In the monkeys, midbrain dopamine neurons responded differently to the delay cues. At the onset of the informative cue that signalled the big reward, there was a burst of neuronal firing; no such burst was seen at the onset of any of the other cues. These results suggest that the informative signal for reward might affect preference through an amplified positive reward prediction error signal of the sort typically encoded by dopamine neurons (e.g., Schultz, Dayan, & Montague, 1997).

The role of the signal contingencies in suboptimal choice is also consistent with an extensive literature on similar comparisons between signalled and unsignalled contingencies in observing behavior (in which subjects respond in order to produce informative cues, e.g., McMillan, 1974, Wyckoff, 1952; see Dinsmoor, 1983 for a review) and related choice procedures (Alsop & Davison, 1986; Fantino & Moore, 1980; Green & Rachlin, 1977; Hursh & Fantino, 1974; Roper & Zentall, 1999).

2. Response Requirement in the Choice Phase

A critical feature of Kendall's (1974, 1985) results was that high levels of suboptimal choice occurred only when a single peck was required for the transition from the choice phase to the delay phase. Kendall (1974, Experiment 1) reported preference for the 50% alternative with a fixed-ratio (FR) 1 choice requirement. In a comparable set of exploratory conditions in a second experiment with a variable-interval (VI) 20-s requirement, suboptimal choice was lower and less consistent.

This dependence on the nature of the choice response was replicated in more extensive comparisons by Dunn and Spetch (1990, Experiment 3). They varied the schedule in the choice phase across four values as shown in Table 1. Preference for the suboptimal alternative was obtained with the FR 1 but shifted to the optimal alternative as the response requirement increased. Note that the opposite pattern occurs when the outcomes are not signalled: Spetch and Dunn (1987) reported that preference for the optimal (100%) alternative over an unsignalled 50% alternative decreased (but did not reverse) as the choice schedule was lengthened from VI 45 s to VI 180 s.

3. Duration of Delays to Food

In general, longer delays to food lead to increased suboptimal choice. For example, Spetch et al. (1990) manipulated the duration of the delay to food in three experiments. Suboptimal choice was greater with 30-s delays compared to 5-s delays; greater with 30-s delays compared to 10-s delays; and greater with 50-s delays compared to 10-s delays. Suboptimal choice did not differ in comparisons between longer delays (e.g., 30 s vs. 50 s). Spetch, Mondloch, Belke, and Dunn (1994, Experiment 3) also reported greater suboptimal choice with 30-s delays than with 5-s delays. The overall pattern in the data from the two studies (Spetch et al. 1990, 1994) is clear in Table 2. In sum, in procedures with a single response requirement in the choice phase and outcomes signalled during the delays, suboptimal choice systematically increased as the delays to food were increased from 5 to 30 s and did not vary with longer delays.

4. Immediacy of Signals for Good News

To test the role of the signal for food in reinforcing the choice response, McDevitt, Spetch, and Dunn (1997) used a “gap” procedure, developed earlier by Belke and Spetch (1994, Experiment 2). The rationale was simple: If a signal for food functions as a conditioned

reinforcer for the choice response that produces it, then inserting even a small, 5-s gap between the choice response and the presentation of the signal should substantially reduce or even eliminate the conditioned reinforcement.

McDevitt et al. (1997) tested this prediction by inserting a gap prior to the onset of one or more of the delay stimuli. The gap eliminated the first 5 s of a terminal-link stimulus, so that the response keys were darkened for 5 s following the choice response, after which the terminal-link stimulus was presented for the remainder of the delay. It is important to note that the delay to primary reinforcement remained unchanged when the gap was added in order to maintain the effect of primary reinforcement while degrading only the conditioned reinforcement. Figure 4 shows the conditions in their study. The NONE condition consisted of a standard baseline procedure in which pigeons chose between a 100% option and a signalled 50% option. The mean choice proportion for the 50% option during baseline was .54, indicating that the pigeons behaved suboptimally by choosing the 50% option just as often as the 100% option.

As shown in Figure 5, when the gap preceded only the 50% S- (signal for no food) or the 100% S+ (signal for food), there was little change in the overall choice proportions, indicating that these delay stimuli did not function as conditioned punishers or conditioned reinforcers, respectively. In contrast, placing a gap preceding only the 50% S+ resulted in a large shift away from the 50% option. Thus, despite the fact that the signals for food on the two alternatives provided identical information about the delay to food, only the 50% S+ appears to function as a strong conditioned reinforcer. When the gaps preceded both the 50% S+ and the 50% S- (in the All and 50% Both conditions), the level of suboptimal choice decreased. A similar procedure, with starlings, also showed that delaying the onset of the informative cues reduces suboptimal choice (Vasconcelos, Monteiro, & Kacelnik, 2015). Taken together, these findings provide

insight into the functions of the various stimuli in the suboptimal choice procedure and, of particular note here, illustrate that signals for food may function differently if there is a context of uncertainty. There are parallels in other recent explorations of the effects of uncertainty (e.g., Anselme, Robinson, & Berridge, 2013). Uncertainty enhances the value of the stimulus that resolves it.

5. Probability of reinforcement

Although the initial demonstrations of suboptimal preference were shown in a procedure in which a signalled 50% food alternative was pitted against a 100% food alternative (see Figure 1), larger effects have been found with other probabilities of reinforcement. For example, Dunn and Spetch (1990) conducted a pilot test with two birds in which the suboptimal alternative signalled 80% food and the optimal alternative provided 100% food, thus decreasing the difference in primary reinforcement between the two alternatives compared to the standard procedure. Both birds showed extreme preference for the suboptimal alternative (94% and 95% preference). In another example, Gipson et al. (2009, Experiment 2) reduced the probability of reinforcement on the optimal alternative from a signalled 100% to an unsignalled 75%, and found that pigeons preferred the suboptimal alternative (signalled 50% reinforcement) an average of 76% of the time. In those studies, the higher level of preference compared to the standard procedure is not surprising because the difference between the probabilities on the two alternatives was reduced.

Zentall, Laude, Stagner, and Smith (2015) compared 75% unsignalled reinforcement with 25% or 50% signalled reinforcement. The lower percentage of food was preferred in both comparisons with no difference in the level of preference. In another study noted above, Stagner and Zentall (2010) reduced the probability of food associated with both alternatives, pitting a

signalled 20% food alternative against an unsignalled 50% food alternative, and found preference for the suboptimal alternative was even higher (87% or above across three conditions). Using a similar design, Vasconcelos et al. (2015) reduced the probability of reinforcement on the unsignalled alternative in starlings down from 20% to 0% in 5% increments. They found that suboptimal choice started to decrease only when the probability of reinforcement was 5%.

At this point, it is clear that suboptimal preference occurs across a wide range of comparisons and is not directly tied to the absolute difference in probability of reinforcement on the two alternatives. A full parametric manipulation of the probability of the suboptimal and optimal alternatives remains to be done.

6. Factors Related to Self-Control

Dunn and Spetch (1990) observed that suboptimal choice with probabilistic alternatives shares features with self-control procedures in that suboptimal choice, like impulsivity, is characterized by preference for a smaller, more immediate reward over a larger, more delayed reward (e.g., Rachlin & Green, 1972). In the suboptimal procedures described above, the small immediate reward is a conditioned reinforcer, and the large delayed reward is the primary (food) reinforcer. As in the typical self-control procedures, suboptimal choice results when the smaller immediate reward begins to exert as much or more control than the larger delayed reward.

The same underlying mechanism may be responsible for both types of choice: namely, a subject's sensitivity to the immediacy of an outcome. Impulsivity has been described by the slope of a delay discounting function in which variations in delay are associated with the discounting of another reinforcement variable (such as amount or probability). The steepness of the discounting function corresponds with the degree of impulsivity, or lack of self-control,

shown by a subject. Preliminary research has shown that some of the manipulations that affect temporal discounting on self-control tasks also affect suboptimal choice with probabilistic alternatives. For example, Laude, Beckmann, et al. (2014) found that pigeons' suboptimal choice with probabilistic alternatives was correlated with the steepness of the slope on a delay discounting function obtained on a separate task. In other words, pigeons which were more impulsive as indicated by the discounting function showed a greater degree of suboptimal choice. Figure 6 compares the percentage of suboptimal choices for pigeons with steeper discounting functions (high impulsivity) and pigeons with shallower discounting functions (low impulsivity) in the Laude et al. results. Research with humans has shown that pathological gamblers have a much steeper delay discounting function than non-gamblers, and gamblers with a history of substance abuse have an even steeper function (Petry, 2001).

Research with humans also suggests that lower income levels (i.e., greater deprivation) may be associated with steeper delay discounting functions, indicating that impulsiveness might increase in conditions of economic deprivation (e.g., Green, Myerson, Lichtman, Rosen, & Fry, 1996; but see Oliveira, Calvert, Green, & Myerson, 2013). Although the evidence of the effects of deprivation on discounting in animals is mixed (e.g., Bradshaw & Szabadi, 1992; Eisenberger, Masterson, & Lowman, 1982; Oliveira et al., 2013), Laude, Pattison, and Zentall (2012) found that deprivation did affect suboptimal choice in pigeons. They presented pigeons with a choice between signalled 50% food (the suboptimal alternative) and unsignalled 75% food (the optimal alternative). Half of the pigeons were maintained at 75% of free-feeding body weight (high deprivation) and half were maintained at 90% of free-feeding body weight. Intuitively, one would expect the hungrier pigeons to more strongly prefer the alternative with the higher probability of food. Instead, at the end of training the birds maintained at the lower weights

showed, on average, an 84% preference for the suboptimal alternative, compared to a 32% preference for birds maintained at the higher weights. Thus, the choice behavior of the hungrier pigeons resulted in significantly fewer food reinforcers. Figure 7 shows the mean number of reinforcers per session for each group of pigeons in the Laude et al. study, illustrating the critical impact of suboptimal choice on food acquired.

Environmental enrichment is another variable that influences self-control and may influence suboptimal choice. Indeed, Pattison, Laude, and Zentall (2013) demonstrated an effect of environmental enrichment on suboptimal choice behavior. Pigeons given daily access to a large group cage (enriched condition) required significantly more sessions to develop suboptimal choice behavior compared to pigeons caged individually (isolated condition). Although the enriched environment did not prevent suboptimal preference, it did delay its development (but see Laude, Beckmann, et al., 2014).

Is Bad News Actually Bad?

Given the clear evidence of the role of the signals for food, it would be reasonable to assume that signals for no food would have symmetrical punishing effects on choices. Thus, signals for no food on the 50% alternative should counteract the reinforcement by signals for food. This prediction would also be made from conventional wisdom: The signals for no food are essentially “timeouts” and hence should punish the choice response on the 50% alternative.

Several lines of evidence, however, suggest that the timeout stimulus has very little impact on the choice response. As seen in condition S- of Figure 5, McDevitt et al. (1997) showed that disrupting the contiguity between the choice response and onset of the timeout stimulus by inserting a gap had very little effect on choice. The duration of the timeout also does not seem to matter. For example, Spetch et al. (1994) held all other features of the choice

procedure constant while manipulating the duration of the timeout (5 or 55 s) and found no effect. Pisklak, McDevitt, Dunn, and Spetch (in press) reported a comparison with bad news on both alternatives. The good news outcome was food after 10 s on both alternatives. The bad news on the optimal alternative was food after 40 s; the bad news on the suboptimal alternative was no food after 40 s. The suboptimal alternative was preferred when the outcomes were signaled. Preference was sensitive to manipulations of the duration of delay to food but not to the duration of delay to no food.

Stagner, Laude, and Zentall (2011) hypothesized that perhaps the S- stimulus was ineffective because the pigeons avoided it, but their investigation found that pigeons still preferred the 50% alternative even when the signal for the timeout was a diffuse stimulus that could not be avoided by turning away. Laude, Stagner, and Zentall (2014) tested whether the timeout stimulus on the suboptimal alternative acquired inhibitory properties. Although the stimulus initially exerted some inhibition, the inhibitory effects decreased over training while preference for the suboptimal alternative increased.

Finally, the literature on observing includes several investigations of the role of informative signals. In the typical observing procedure, schedules of reinforcement and extinction alternate without correlated stimuli. Observing responses produce correlated stimuli, good news if the reinforcement schedule is in effect or bad news if it is not. The signal for good news maintains the observing response but the signal for bad news does not (Dinsmoor, 1983; Silberberg & Fantino, 2010). Nor is there direct evidence that the signal for bad news is punishing (Dinsmoor, Browne, & Lawrence, 1972; cf: Killeen, Wald, & Cheney, 1980).

Thus, data from a variety of procedures converge to support the notion that signals for bad news are not punishing and appear to have little or no direct influence on suboptimal choice.

These data also provide evidence against an intuitive account of the signalling effect, namely that the information per se is driving suboptimal choice. Although the signals for good news are valuable and provide conditioned reinforcement for responding, signals for bad news, although just as informative, do not appear to have either positive or negative value. The evolutionary benefit of having signals for good news function as conditioned reinforcers is intuitive, as organisms are likely to benefit from devoting attentional and behavioral resources to stimuli correlated with food. In contrast, the apparent absence of negative value associated with signals for bad news is not as intuitive. One possibility is that these signals function differently because signals for food require organisms to engage in preparatory behaviors, whereas signals for the absence of food do not. In any case, an explanation based solely on a traditional information approach cannot account for the effect of the signals in suboptimal choice procedures. An alternative account based on the temporal information provided by signals for reinforcement, however, may represent a more useful approach (e.g., Shahan & Cunningham, 2015).

Is Suboptimal Choice for 50% versus 100% Simply Indifference?

Although choice between signalled and unsignalled options can clearly lead to extreme suboptimal preference (e.g., Gipson et al., 2009), suboptimal choice for signalled 50% over 100% food is typically only slightly above chance, often with high inter-individual variability. Therefore, Zentall et al. (2015) suggested that the individual differences might be a procedural artifact and reflect side biases that occur when birds are indifferent between options. In support of this hypothesis, they reported results which suggest that pigeons are sensitive to the value of a conditioned reinforcer but not to its frequency. Specifically, they found equal suboptimal choice for pigeons that chose between signalled 25% and unsignalled 75% reinforcement as compared to pigeons that chose between signalled 50% and unsignalled 75% reinforcement. In addition,

Stagner et al. (2012) found indifference between 20% signalled and 50% signalled reinforcement. A recent study by Pisklak et al. (in press) also found that the level of preference for signalled 25% over 100% reinforcement was similar to that for signalled 50% over 100% reinforcement. This pattern of results led Zentall et al. (2015) to suggest that pigeons should also be expected to be indifferent between signalled 50% and 100% reinforcement because each of these alternatives provide signals for food that differ only in frequency.

For reasons discussed in greater detail below, we argue that the signal for food on the 50% option does not have the same value as the signal on the 100% option, and that simple biases do not account for suboptimal choice. There is evidence for preference rather than indifference. First, contingency reversals were implemented in the original studies by Kendall (1974, 1985), with most birds reversing their preferences when the sides associated with the 50% and 100% reinforcement were changed. This contingency reversal procedure was also used by Belke and Spetch (1994) to determine whether three birds that showed strong preferences for the suboptimal option were showing side biases. As shown in Figure 8, these birds all showed a reversal of their suboptimal preference.

Second, the results of the gap study by McDevitt et al. (1997) provide compelling evidence against the notion that choice between 50% and 100% reinforcement represents “schedule indifference ... masked by spatial preference” (Zentall et al., 2015). If pigeons were indeed indifferent between the 50% and 100% options because both provide food signals of equal value, then choice should be equally affected by inserting a gap prior to onset of the food signal on either alternative. As seen in Figure 5, however, suboptimal choice was selectively affected by inserting a gap prior to the food signal on the 50% alternative. Specifically, suboptimal choice declined when a gap was inserted after all choices (condition ALL), after any

choice of the 50% option (condition BOTH), or only prior to the food signal on the 50% option. In contrast, inserting a gap prior to the food signal on the 100% option had no substantial effect.

Together, these results strongly suggest that the signal for food on the 50% alternative and the signal for food on the 100% alternative are not equivalent reinforcers that differ only in frequency. The suggestion that pigeons are indifferent to the frequency of conditioned reinforcers is also inconsistent with several studies that have explicitly shown frequency of conditioned reinforcement to be an important determinant of choice behavior in other procedures (McDevitt & Williams, 2010; Shahan, Podlesnik, & Jimenez-Gomez, 2006; Williams & Dunn, 1991). Why then, might Stagner et al. (2012) have observed indifference, on average, when pigeons chose between a signalled 20% reinforcement alternative and a signalled 50% reinforcement alternative? One potential explanation lies in the fact that altering the probability of reinforcement also changes the value of the conditioned reinforcer. The 20% alternative provides less overall primary reinforcement than the 50% alternative, and that leaner context serves to enhance the value of the conditioned reinforcer that signals food. Thus, primary reinforcement is reduced at the same time that the value of the more immediate conditioned reinforcement is enhanced, resulting in roughly equivalent value between the two alternatives.

Signals for Good News: The SiGN Hypothesis

Dunn and Spetch (1990) proposed an explanatory framework based on the assumption that choices depend on two sources of reinforcement: (1) the delayed primary reinforcement (e.g., food) and (2) the reinforcement provided by signals that follow a choice response. Their model was inspired by Fantino's delay reduction theory (DRT, Fantino, 1969; Fantino, Preston & Dunn, 1993). Fantino's theory assumes that an event reinforces a choice to the extent that it signals a reduction in the average waiting time to food. Dunn and Spetch (1990) proposed that

suboptimal choice could be accommodated within a delay reduction framework by assuming that the choice response serves as the reference point for assessing a signal for a delay to food. Specifically, they proposed that any event that signals a reduction in the delay beyond that already signalled by the choice response will function as a conditioned reinforcer. Our hypothesis also assumes that bad news plays no role other than to create a context of uncertainty. We will refer to this explanation as the Signals for Good News (SiGN) hypothesis. Here we will examine how the SiGN hypothesis holds up to the wealth of evidence that has been developed in recent years, particularly through the work of Zentall and his colleagues.

A critical feature of the SiGN hypothesis is that the occurrence of the choice response itself establishes a new context for comparison; it resets the local “reference point” (Kahneman, 2011) against which any changes in state are compared. Research with humans has shown that the value associated with a particular outcome is influenced by the current reference point. For example, news of a raise in pay would be rewarding to an employee who is expecting a possible pay cut; the same news would be less rewarding if the employee already knows the raise is coming. In these cases, the context sets a reference point against which the new state (a raise) is evaluated. Similarly, for the pigeon choosing between the 50% and 100% options, the value of a delay stimulus will depend on the local context provided by the choice itself. A signal predicting food at the end of a trial would be good news in the context of a gamble with a 50% chance of getting nothing. In contrast, a signal predicting food at the end of a trial would have less value if the food was already perfectly predicted by the choice response, as is the case on the 100% option.

It follows from this analysis that the reinforcing effects of the delay stimuli should depend critically on the contingencies between these stimuli and the outcome at the end of the

trial. If the food and no-food outcomes on the 50% alternative are not differentially signalled (as in the "unsignalled" condition in Figure 1), then choice of that alternative is followed by a redundant stimulus that signals no change in the possibility of food or no food. In this case, onset of the delay stimulus does not signal good news because there is no improvement in local context. Thus, the SiGN hypothesis is consistent with the numerous studies that have demonstrated that suboptimal choice occurs only in procedures with signalled outcomes (e.g., Kendall, 1974; Spetch et al., 1990; Stagner et al., 2012; Zentall & Stagner, 2012).

The SiGN hypothesis puts the emphasis on the local context created by the transition between the choice response and the presentation of the signals for food. When the option that always leads to food (100%) is chosen, the choice peck itself and onset of the food signal predict the same delayed reward so the signal provides no improvement in the local context. Therefore, the food signal on the 100% alternative should not function to reinforce the choice. In contrast, choice of the 50% alternative may or may not lead to food. Therefore, onset of the food signal provides a substantial improvement in the local context and should function as an effective conditioned reinforcer. The SiGN hypothesis assumes that suboptimal choice in signalled procedures is driven by this imbalance in the conditioned reinforcement on the two alternatives. The results from the "gap" procedure (McDevitt et al., 1997) highlight this difference in the function of the two signals for food; suboptimal preference was shown to be driven largely by an immediate conditioned reinforcer on the 50% option.

The SiGN hypothesis accounts for the shift away from suboptimal choice when the response requirement is changed, as in procedures with VI choice schedules (e.g., Dunn & Spetch, 1990; Kendall, 1974, 1985). In a VI schedule, responses are only occasionally effective in producing a transition from the choice phase to the delay phase. In addition, the duration of

the initial link is longer, introducing greater temporal uncertainty as to when the delay phase will begin. Both these factors mean that the food signal on the 100% option is no longer redundant and its onset will signal an improvement in the prediction of food, with the degree of improvement increasing as a function of the VI duration. Therefore, when VI schedules are used in the choice phase, signals for food on both the 100% and the 50% alternatives should function as conditioned reinforcers. The less predictive the choice responses are, the stronger both food signals would be as reinforcers. As the disparity in conditioned reinforcement is reduced on the two alternatives, choice should be increasingly determined by the only other difference, the relative rates of food (primary reinforcement).

Another key variable for the SiGN hypothesis is the duration of the delay to food following the choice response. It is important to remember that, according to the SiGN hypothesis, the primary reinforcement provided by the food and the conditioned reinforcement provided by signals for good news operate jointly to influence choice behavior. If the delays to food are very short, primary reinforcement should be a strong determinant of choice. As delays to food increase, the effect of primary reinforcement should diminish as illustrated by numerous studies of temporal discounting (e.g., McDevitt & Williams, 2001).

The interplay between delayed primary reinforcement (food) and immediate conditioned reinforcement can also account for the large individual differences often seen in suboptimal choice procedures. Specifically, if birds differ in their individual rates of delay discounting, then the relative control by the food and by the signal for good news could be expected to differ across individuals. Indeed, Laude, Beckmann, et al. (2014) reported that suboptimal preference was correlated with the steepness of individual rates of delay discounting functions.

At present, the SiGN hypothesis is a conceptual description of the factors that are assumed to operate jointly to produce suboptimal choice. A formalization of this hypothesis in a quantified model based on reinforcement learning is currently underway (Ludvig, Spetch, Dunn, & McDevitt, 2015). Briefly, in the computational model, animals learn to predict upcoming reinforcers through an error-correction mechanism, similar to the Rescorla-Wagner learning rule (Rescorla & Wagner, 1972; Sutton & Barto, 1998). In addition, whenever the prediction changes upward due to the introduction of a new cue (e.g., at the start of the terminal link), a good news bonus is generated. This good news bonus directly sums with the reinforcer strength, increasing preference for that option. The exact form of the relationship between the good news bonus and the probability and delay of reinforcement, however, is as yet unclear, and significant work is still necessary to explore the interplay of these factors in the model and the consequent implications for behavior.

Relationship to Other Models of Suboptimal Choice

Hyperbolic-Decay Model

Our SiGN hypothesis shares some features in common with Mazur's Hyperbolic-Decay Model (HDM), which was originally developed to describe how the value of a reinforcer decreases as reinforcer delay increases (Mazur, 1984, 1989, 2001). HDM has also successfully described preference for variable reinforcement schedules over fixed reinforcement schedules, as well as the basic finding related to suboptimal preference for signalled reinforcement alternatives. According to HDM, conditioned reinforcement is a hyperbolic function of the delay (D) to primary reinforcement associated with an alternative. Simply put, as the delay to a reinforcer increases, the value of that alternative decreases. The most important aspect of HDM for the present discussion concerns which time variables are included in the estimation of D.

HDM includes only time spent in the presence of stimuli that are correlated with food. Time spent in the presence of stimuli not correlated with food (i.e., signals for timeout) are not included, and therefore do not affect the value of an alternative.

Similar to the SiGN hypothesis, HDM emphasizes that conditioned reinforcement is a primary determinant of behavior. However, HDM assumes that conditioned reinforcement is the only contributor to value of an alternative (Mazur, 1989, 2001), whereas the SiGN hypothesis assumes that choice behavior is a result of the influence of both primary and conditioned reinforcement.

Both the SiGN hypothesis and HDM assume that time spent in a signalled timeout does not detract from the value of an alternative. In fact, HDM assumes no role at all for a signalled timeout in the value of an alternative. In this way, HDM can account for suboptimal choice, because time spent in the presence of all delay stimuli is included in the estimation of D (and thus decreases the value of the alternative), except when timeouts are signalled by unique stimuli (Mazur, 1991). The time spent in a signalled timeout has no role in HDM, but according to the SiGN hypothesis, the possibility of a timeout alters the local context, allowing the $S+$ to provide good news.

Like the SiGN hypothesis, HDM predicts that the value of an alternative (and preference for that alternative) will decrease as the delay to reinforcement increases. In the standard choice between a 100% alternative and a signalled 50% alternative, HDM can never predict preference for the 50% alternative over the 100% when the delays on the two alternatives are equal (Mazur, 1995). In contrast, the SiGN hypothesis allows for the possibility that preference may be observed for the suboptimal alternative.

Contrast

Zentall and colleagues have proposed another account of preference in suboptimal choice procedures (Gipson et al., 2009; Zentall & Stagner, 2011a). Similar to the SiGN hypothesis, they suggest that the signal for food is enhanced in value because it is contrasted with the value of the choice stimulus. Thus, when the signal for food occurs in the context of an alternative that sometimes ends in food and sometimes ends in no food, the value of that signal is enhanced by the contrast between the conditions in the choice phase and the delay phase. The contrast view provides no negative role for the signal for no food, so that the enhanced value of the signal for food is not counteracted by decreased value of the signal for no food. Contrast is absent when the probabilistic outcomes are not signalled, and therefore preference is expected to shift to the alternative with the richer reinforcement schedule.

The contrast account and HDM make similar predictions, but as noted above, HDM cannot account for preference (i.e., greater than .50) for a low probability food alternative over a 100% food alternative, whereas the contrast account can make such a prediction (Gipson et al., 2009; Mazur, 1995). Contrast and the SiGN hypothesis make very similar predictions, which is not surprising given the similarity in their analyses. Attempts to distinguish between the contrast account and DRT (Fantino, 1969), which initially inspired the SiGN hypothesis, have been complicated by the difference in the procedures used by the two approaches (Singer, Berry, & Zentall, 2007; Singer & Zentall, 2011). The SiGN hypothesis differs from DRT in that it focuses on local context effects, a focus that is shared with the contrast model. Comparisons between SiGN and the contrast account are limited for now because the contrast account has not yet been extended to manipulations of the choice phase and the delays to reinforcement. As noted by Gipson et al. (2009), contrast is the least quantitative of the three approaches. Future work

formalizing the SiGN and contrast hypotheses will be important for distinguishing these similar conceptual frameworks and providing clear, testable predictions.

Disengagement

Another hypothesis comes from a recent reinforcement-learning model of observing behavior (Beierholm & Dayan, 2010). According to this model, suboptimal choice would occur because animals occasionally disengage from (i.e., stop paying attention to) lower-valued conditioned stimuli during the delay to food. Disengagement is formalized in the model as a reduced likelihood of learning from the reinforcer when it happens. As a result, animals underestimate the value of those conditioned stimuli because they fail to associate any food that eventually arrives with the predictive cue. Less valuable conditioned stimuli lead to higher degrees of disengagement, and thus greater underestimation of the value of alternatives that lead to those conditioned stimuli.

This model provides a possible explanation for situations where animals choose between signalled and unsignalled alternatives (e.g., Blanchard et al., 2015; Bromberg-Martin & Hikosaka, 2009; Roper & Zentall, 1999; Stagner & Zentall, 2010). In the signalled case, one cue is perfectly predictive of reward, and the other cue is perfectly predictive of non-reward. The animals do not disengage from the cue predictive of reward, and thus correctly learn its value. In the unsignalled case, neither cue is perfectly predictive of reward, and thus the animals occasionally disengage during the delay and underestimate the value of those cues. As a result, animals show a preference for the signalled, suboptimal option.

The disengagement model, however, fails for situations where the choice is between a 50% alternative and a 100% alternative in which both sides are signalled. In this situation, the model predicts no disengagement on the 100% alternative and thus predicts a large preference

for the 100% alternative, yet animals sometimes prefer the 50% alternative (Dunn & Spetch, 1990; Kendall, 1974; Spetch et al., 1994).

Implications & Future Directions

Suboptimal choice as evidenced in the procedures summarized above provides a challenge for models of choice that minimize or eliminate the role of conditioned reinforcement (Davison & Baum, 2006; Shahan, 2010, 2013; see Bell & McDevitt, 2014). Other lines of evidence make similar arguments for the concept of conditioned reinforcement as incorporated in the SiGN hypothesis (e.g., Bell & Williams, 2013; Williams, 1994).

Despite extensive investigation across variations of the suboptimal choice procedure, several variables are still largely unexplored. For example, the importance of contiguity between the signal and the outcome has not yet been evaluated in this paradigm –specifically, would inserting a gap between the offset of the signal and the presentation of food have the same effects as inserting a gap between the choice response and the signal? The contingency between the good news signal and food has also not been parametrically manipulated; the suboptimal alternative has almost always provided one stimulus that is perfectly predictive of food and another that is perfectly predictive of no food (but see Zentall & Stagner, 2011b). Would the reinforcing effectiveness of the good news stimulus be disrupted more by reducing its correlation with food or by adding some food outcomes after the other stimulus? As quantitative models become available, it will also be important to systematically explore the probabilities of food on each alternative. For example, if the optimal alternative provides 100% reinforcement, how would choice be affected if the percentage of food signalled on the suboptimal alternative varied between 10 and 90%? Similarly, how would choice be affected if the suboptimal alternative was held constant at 50%, but the optimal alternative is varied from 100% to 60%? A full

understanding of the phenomena will require further parametric investigations of the limits of suboptimal choice.

There is a wealth of evidence for suboptimal choice in pigeons, but few investigations with other species. In a series of experiments with rats using an adjusting-delay procedure, Mazur (2005, 2007) failed to find any influence of signals on probabilistic choice. One contributing factor may be differences in the delays used with the two species, and another factor may be differences in the way the two species discount primary reinforcement (Mazur & Bondi, 2009). Recent work by Beckmann and Chow (2015) suggests that species-specific differences in sign vs. goal tracking may provide a further complication in comparisons of choice tasks across species. Monkeys, in contrast to rats, do display a similar form of suboptimal choice (Bromberg-Martin & Hikosaka, 2009; Blanchard et al., 2015) as do starlings (Vasconcelos et al., 2015). Further work evaluating suboptimal choice with other species may provide a basis for speculation about the evolutionary pressures that underlie suboptimal choice.

At present, there appear to be only a few investigations of suboptimal choice in humans using procedures derived from the pigeon work. Lalli, Mauro, and Mace (Experiment 1, 2000) developed a 100% vs. 50% procedure for two children with mild developmental delays. In conditions with 30-s delays to food, the children preferred the 100% alternative when outcomes were not signalled but preferred the 50% alternative when the outcomes were signalled. Preference for the 50% alternative was not obtained with shorter (10 s) delays to food. In a second experiment with another child with mild developmental delay, Lalli et al. inserted a 10-s gap before the signals for the outcomes (much like Belke & Spetch, 1994, and McDevitt et al., 1997). In an ABAB sequence, the child preferred the 50% alternative when there was no gap, and preferred the 100% alternative when there was a gap.

More recently, Molet et al. (2012) developed a human version of the suboptimal choice task by embedding the contingencies in a video game. The goal of the game was to kill generals on enemy space ships, and hence the reward was number of generals killed. Participants chose between two planetary systems. Each choice was followed by entry into one of two planets distinguished by different colors (the delay stimuli). For one planetary system (signalled low probability), the planets signalled how many generals would be on the ship that could be destroyed; in this system, one planet occurred 20% of the time and had 10 generals aboard, and the other planet occurred 80% of the time and had no generals aboard. Therefore this planetary system paid off with an average of 2 generals killed. For the other planetary system (unsignalled high probability), there were again two planets with one appearing 20% of the time and one occurring 80% of the time. In this case, however, there were three generals aboard the ships in both planets; thus the average reward was three generals killed. There was a difference in suboptimal choice depending on gambling habits: most participants who did not frequently gamble showed low levels of suboptimal choice, whereas those participants who gambled regularly chose the low probability planet, on average, around half the time. The results of this interesting experiment suggest that humans who gamble frequently may be particularly drawn to stimuli that signal big wins in the context of uncertainty, much like pigeons. Interestingly, there was large individual variability among the humans who gambled frequently, as there is for pigeons.

Making good choices is crucial to survival in a complex and uncertain world. Here, we have reviewed evidence that pigeons, starlings, and sometimes monkeys and humans, exhibit significant suboptimal choices, giving up food, water, or money for advanced signals about upcoming rewards. These suboptimal choices reflect a desire for good news and may result from

simple, fundamental learning mechanisms that interact with the reward contingencies of the world in surprising and interesting ways.

References

- Alsop, B., & Davison, M. (1986). Preference for multiple versus mixed schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 45(1), 33-45. doi:10.1901/jeab.1986.45-33
- Anselme, P., Robinson, M. F., & Berridge, K. C. (2013). Reward uncertainty enhances incentive salience attribution as sign-tracking. *Behavioural Brain Research*, 23853-61. doi:10.1016/j.bbr.2012.10.006
- Ariely, D. (2008). *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York, NY: HarperCollins
- Beckmann, J. S., & Chow, J. J. (2015). Isolating the incentive salience of reward-associated stimuli: Value, choice, and persistence. *Learning & Memory*, 22(2), 116-127. doi:10.1101/lm.037382.114
- Beierholm, U. R., & Dayan, P. (2010). Pavlovian-instrumental interaction in 'observing behavior'. *PLoS Comput Biol*, 6(9), e1000903.
- Belke, T. W., & Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: Preference for unreliable reinforcement. *Journal of the Experimental Analysis of Behavior*, 62(3), 353. doi:10.1901/jeab.1994.62-353
- Bell, M. C., & McDevitt, M. A. (2014). Conditioned Reinforcement. In *The Wiley-Blackwell Handbook of Operant and Classical Conditioning* (eds F. K. McSweeney and E. S. Murphy), John Wiley & Sons, Ltd, Oxford, UK. doi: 10.1002/9781118468135.ch10
- Bell, M. C., & Williams, B. A. (2013). Conditioned reinforcement in chain schedules when time to reinforcement is held constant. *Journal of the Experimental Analysis of Behavior*, 99(2), 179-188. doi:10.1002/jeab.10

- Blanchard, T. C., Hayden, B. Y., & Bromberg-Martin, E. S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron*, *85*(3), 602-614.
- Bradshaw, C. M., & Szabadi, E. (1992). Choice between delayed reinforcers in a discrete-trials schedule: The effect of deprivation level. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology*, *44B*(1), 1-16.
- Bromberg-Martin, E. S., & Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, *63*(1), 119-126.
doi:10.1016/j.neuron.2009.06.009
- Davison, M., & Baum, W. M. (2006). Do conditional reinforcers count? *Journal of the Experimental Analysis of Behavior*, *86*, 269–283. doi:10.1901/jeab.2006.56-05
- Dinsmoor, J. A. (1983). Observing and conditioned reinforcement. *Behavioral and Brain Sciences*, *6*, 693-728. doi:10.1017/S0140525X00017969
- Dinsmoor, J. A., Browne, M. P., & Lawrence, C. E. (1972). A test of the negative discriminative stimulus as a reinforcer of observing. *Journal Of The Experimental Analysis of Behavior*, *18*(1), 79-85. doi:10.1901/jeab.1972.18-79
- Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes: Conditioned reinforcement effects. *Journal of the Experimental Analysis of Behavior*, *53*(2), 201–218.
doi:10.1901/jeab.1990.53-201
- Eisenberger, R., Masterson, F. A., & Lowman, K. (1982). Effects of previous delay of reward, generalized effort, and deprivation on impulsiveness. *Learning and Motivation*, *13*(3), 378–389.
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of*

Behavior, 12(5), 723–730. doi:10.1901/jeab.1969.12-723

Fantino, E. (2012). Optimal and non-optimal behavior across species. *Comparative Cognition & Behavior Reviews*, 7, 44-54. Retrieved from <http://comparative-cognition-and-behavior-reviews.org/doi:10.3819/ccbr.2012.70003>

Fantino, E., & Moore, J. (1980). Uncertainty reduction, conditioned reinforcement, and observing. *Journal of the Experimental Analysis of Behavior*, 33(1), 3-13. doi:10.1901/jeab.1980.33-3

Fantino, E., Preston, R. A., & Dunn, R. (1993). Delay reduction: Current status. *Journal of the Experimental Analysis of Behavior*, 60(1), 159-169. doi:10.1901/jeab.1993.60-159

Gipson, C. D., Alessandri, J. J. D., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons. *Learning & Behavior*, 37(4), 289–98. doi:10.3758/LB.37.4.289

Green, L., Myerson, J., Lichtman, D., Rosen, S., & Fry, A. (1996). Temporal discounting in choice between delayed rewards: The role of age and income. *Psychology and Aging*, 11(1), 79-84. doi:10.1037/0882-7974.11.1.79

Green, L., & Rachlin, H. (1977). Pigeons' preferences for stimulus information: Effects of amount of information. *Journal of the Experimental Analysis of Behavior*, 27(2), 255-263. doi:10.1901/jeab.1977.27-255

Hursh, S. R., & Fantino, E. (1974). An appraisal of preference for multiple versus mixed schedules. *Journal of the Experimental Analysis of Behavior*, 22(1), 31. doi:10.1901/jeab.1974.22-31

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus, & Giroux.

Kendall, S. B. (1974). Preference for intermittent reinforcement. *Journal of the Experimental Analysis of Behavior*, 21(3), 463–473. doi:10.1901/jeab.1974.21-463

Kendall, S. B. (1985). A further study of choice and percentage reinforcement. *Behavioural*

Processes, 10, 399-413. doi:10.1016/0376-6357(85)90040-3

- Killeen, P., Wald, B., & Cheney, C. D. (1980). Observing behavior and information. *The Psychological Record*, 30(2), 181-190.
- Lalli, J. S., & Mauro, B. C. (1995). The paradox of preference for unreliable reinforcement: The role of context and conditioned reinforcement. *Journal of Applied Behavior Analysis*, 28(3), 389-394. doi:10.1901/jaba.1995.28-389
- Lalli, J. S., Mauro, B. C., & Mace, F. C. (2000). Preference for unreliable reinforcement in children with mental retardation: The role of conditioned reinforcement. *Journal of Applied Behavior Analysis*, 33(4), 533-544. doi:10.1901/jaba.2000.33-533
- Laude, J. R., Beckmann, J. S., Daniels, C. W., & Zentall, T. R. (2014). Impulsivity affects suboptimal gambling-like choice by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 40(1), 2-11. doi:10.1037/xan0000001
- Laude, J. R., Pattison, K. F., & Zentall, T. R. (2012). Hungry pigeons make suboptimal choices, less hungry pigeons do not. *Psychonomic Bulletin & Review*, 19, 884-891. doi:10.3758/s13423-012-0282-2
- Laude, J. R., Stagner, J. P., & Zentall, T. R. (2014). Suboptimal choice by pigeons may result from the diminishing effect of nonreinforcement. *Journal of Experimental Psychology: Animal Learning and Cognition*, 40(1), 12-21. doi:10.1037/xan0000010
- Ludvig, E. A., Spetch, M. L., Dunn, R. M., & McDevitt, M. A. (2015). When good news leads to bad choices: A reinforcement-learning model of information-driven suboptimal choice. *RLDM: The Multi-disciplinary Conference on Reinforcement Learning and Decision Making*. Edmonton, Canada
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal*

of Experimental Psychology: Animal Behavior Processes, 10(4), 426-436.

doi:10.1037/0097-7403.10.4.426

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

Mazur, J. E., (1991), Choice with probabilistic reinforcement: Effects of delay and conditioned reinforcers. *Journal of The Experimental Analysis of Behavior*, 55(1), 63-77.

Mazur, J. E. (1995). Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *Journal of The Experimental Analysis of Behavior*, 63(2), 139-150.

doi: 10.1901/jeab.1995.63-139

Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, 108(1), 96-112. doi:10.1037/0033-295X.108.1.96

Mazur, J. E. (2005). Effects of reinforcer probability, delay, and response requirements on the choices of rats and pigeons: Possible species differences. *Journal of the Experimental Analysis of Behavior*, 83(3), 263-279. doi:10.1901/jeab.2005.69-04

Mazur, J. E. (2007). Species differences between rats and pigeons in choices with probabilistic and delayed reinforcers. *Behavioural Processes*, 75(2), 220-224.

doi:10.1016/j.beproc.2007.02.004

Mazur, J. E., & Biondi, D. R. (2009). Delay-amount tradeoffs in choices by pigeons and rats: Hyperbolic versus exponential discounting. *Journal of the Experimental Analysis of Behavior*, 91(2), 197-211. doi:10.1901/jeab.2009.91-197

McDevitt, M. A., Spetch, M. L., & Dunn, R. (1997). Contiguity and conditioned reinforcement in probabilistic choice. *Journal of the Experimental Analysis of Behavior*, 68(3), 317-327.

doi:10.1901/jeab.1997.68-317

- McDevitt, M. A., & Williams, B. A. (2001). Effects of signalled versus unsignalled delay of reinforcement on choice. *Journal of the Experimental Analysis of Behavior*, *75*(2), 165-182. doi:10.1901/jeab.2001.75-165
- McDevitt, M. A., & Williams, B. A. (2010). Dual effects on choice of conditioned reinforcement frequency and conditioned reinforcement value. *Journal of the Experimental Analysis of Behavior*, *93*(2), 147-155. doi:10.1901/jeab.2010.93-147
- McMillan, J. C. (1974). Average uncertainty as a determinant of observing behavior. *Journal of the Experimental Analysis of Behavior*, *22*(2), 401-408. doi: 10.1901/jeab.1974.22-401
- Molet, M., Miller, H. C., Laude, J. R., Kirk, C., Manning, B., & Zentall, T. R. (2012). Decision making by humans in a behavioral task: Do humans, like pigeons, show suboptimal choice?. *Learning & Behavior*, *40*(4), 439-447. doi:10.3758/s13420-012-0065-7
- Oliveira, L., Calvert, A. L., Green, L., & Myerson, J. (2013). Level of deprivation does not affect degree of discounting in pigeons. *Learning & Behavior*, *41*(2), 148-158. doi:10.3758/s13420-012-0092-4
- Pattison, K. F., Laude, J. R., & Zentall, T. R. (2013). Environmental enrichment affects suboptimal, risky, gambling-like choice by pigeons. *Animal Cognition*, *16*(3), 429-434. doi: 10.1007/s10071-012-0583-x
- Petry, N. M. (2001). Pathological gamblers, with and without substance abuse disorders, discount delayed rewards at high rates. *Journal of Abnormal Psychology*, *110*(3), 482-487. doi:10.1037/0021-843X.110.3.482
- Pisklak, J. M., McDevitt, M. A. Dunn, R. M., & Spetch M. L. (in press) When good pigeons make bad decisions: Choice with probabilistic delays and outcome. *Journal of the Experimental Analysis of Behavior*.

- Pyke, G.H., Pulliam, H. R., & Charnov, E. L. (1977). Optimal foraging: A selective review of theory and tests. *The Quarterly Review of Biology*, 52(2) 137-154.
<http://www.jstor.org/stable/2824020>
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17(1), 15-22. doi: 10.1901/jeab.1972.17-15
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.
- Roper, K. L., & Zentall, T. R. (1999). Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure. *Learning and Motivation*, 30, 201-220. doi:10.1006/lmot.1999.1030
- Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.
- Shahan, T. A. (2010). Conditioned reinforcement and response strength. *Journal of the Experimental Analysis of Behavior*, 93(2), 269–289. doi:10.1901/jeab.2010.93-269
- Shahan, T. A. (2013). Attention and conditioned reinforcement. In G. J. Madden, W. V. Dube, T. D. Hackenberg, G. P. Hanley, K. A. Lattal (Eds.), *APA Handbook of Behavior Analysis, Vol. 1: Methods and Principles* (pp. 387- 410). Washington, DC: American Psychological Association. doi:10.1037/13937-017
- Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered. *Journal of the Experimental Analysis of Behavior*, 103, 405–418.
doi:10.1002/jeab.142
- Shahan, T. A., Podlesnik, C. A., & Jimenez-Gomez, C. (2006). Matching and conditioned

reinforcement rate. *Journal of the Experimental Analysis of Behavior*, 85, 167–180.

doi:10.1901/jeab.2006.34-05

Silberberg, A., & Fantino, E. (2010). Observing responses: Maintained by good news only?

Behavioural Processes, 85(1), 80-82. doi:10.1016/j.beproc.2010.06.002

Singer, R. A., Berry, L. M., & Zentall, T. R. (2007). Preference for a stimulus that follows a

relatively aversive event: Contrast or delay reduction?. *Journal of the Experimental*

Analysis of Behavior, 87(2), 275-285. doi:10.1901/jeab.2007.39-06

Singer, R. A., & Zentall, T. R. (2011). Preference for the outcome that follows a relatively aversive

event: Contrast or delay reduction?. *Learning and Motivation*, 42(3), 255-271.

doi:10.1016/j.lmot.2011.06.001

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(2), 219–234. doi:

10.1901/jeab.1990.53-219.

Spetch, M. L., & Dunn, R. (1987). Choice between reliable and unreliable outcomes: Mixed

percentage-reinforcement in concurrent chains. *Journal of the Experimental Analysis of*

Behavior, 47(1), 57-72. doi:10.1901/jeab.1987.47-57

Spetch, M., Mondloch, M., Belke, T., & Dunn, R. (1994). Determinants of pigeons' choice between

certain and probabilistic outcomes. *Animal Learning & Behavior*, 22(3), 239–251. doi:

10.3758/BF03209832.

Stagner, J. P., Laude, J. R., & Zentall, T. R. (2011). Sub-optimal choice in pigeons does not depend

on avoidance of the stimulus associated with the absence of reinforcement. *Learning and*

Motivation, 42(4), 282–287. doi:10.1016/j.lmot.2011.09.001

- Stagner, J. P., Laude, J. R., & Zentall, T. R. (2012). Pigeons prefer discriminative stimuli independently of the overall probability of reinforcement and of the number of presentations of the conditioned reinforcer. *Journal of Experimental Psychology. Animal Behavior Processes*, 38(4), 446–452. doi:10.1037/a0030321
- Stagner, J. P., & Zentall, T. R. (2010). Suboptimal choice behavior by pigeons. *Psychonomic Bulletin & Review*, 17(3), 412-416. doi:10.3758/PBR.17.3.412
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. Cambridge, MA: MIT Press.
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific Reports*, 5, 13874.
- Williams, B. A. (1994). Conditioned reinforcement: Neglected or outmoded explanatory construct?. *Psychonomic Bulletin & Review*, 1(4), 457-475. doi:10.3758/BF03210950
- Williams, B. A. & Dunn, R. (1991). Preference for conditioned reinforcement. *Journal of the Experimental Analysis of Behavior*, 55(1), 37-46. doi:10.1901/jeab.1991.55-37
- Wyckoff Jr, L. B. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review*, 59(6), 431-442. doi:10.1037/h0053932
- Zentall, T. R. (2015). When animals misbehave: analogs of human biases and suboptimal choice. *Behavioural Processes*, 112, 3-13. doi:10.1016/j.beproc.2014.08.001
- Zentall, T. R., & Laude, J. R. (2013). Do pigeons gamble? I wouldn't bet against it. *Current Directions in Psychological Science*, 22(4), 271–277. doi: 10.1177/0963721413480173
- Zentall, T. R., Laude, J. R., Stagner J. P. & Smith, A. P. (2015) Suboptimal choice by pigeons: Evidence that the value of the conditioned reinforcer rather than its frequency determines choice. *Psychological Record*, 65:223- 229. doi 10.1007/s40732-015-0119-2

Zentall, T. R., & Stagner, J. (2011a). Maladaptive choice behavior by pigeons: An animal analogue and possible mechanism for gambling (sub-optimal human decision-making behavior).

Proceedings of the Royal Society B: Biological Sciences, 278, 1203-1208.

doi:10.1098/rspb.2010.1607

Zentall, T. R., & Stagner, J. (2011b). Sub-optimal choice by pigeons: Failure to support the Allais paradox. *Learning and Motivation*, 42, 245-254. doi: 10.1016/j.lmot.2011.03.002

Zentall, T. R., & Stagner, J. P. (2012). Do pigeons prefer information in the absence of differential reinforcement? *Learning & Behavior*, 40(4), 465–475. doi:10.3758/s13420-012-0067-5

Table 1. Choice proportions for the suboptimal alternative as a function of initial-link schedule in signaled concurrent chains with fixed time 50-s schedules in the terminal links.

Initial-Link Schedule	Choice Proportion
FR 1	.60
VI 10	.41
VI 30	.19
VI 80	.18

Notes: All data shown are from Experiment 3 of Dunn and Spetch, 1990, and are averaged across four pigeons and in some cases across replications.

Table 2. Choice proportions for the suboptimal alternative as a function of terminal-link (TL) duration in signaled concurrent chains with fixed-ratio 1 schedules in the initial link in procedures that manipulated TL duration.

TL Duration (in seconds)	Spetch et al. 1990, Exp 1	Spetch et al. 1990, Exp 2	Spetch et al. 1990, Exp 3	Spetch et al. 1994, Exp 3	Mean
5	.15	.25		.11	.17
10	.12		.30		.21
30	.62	.56		.60	.59
50	.58		.57		.58
90	.67				.67

Notes: All data shown are averaged across birds and in some cases across replications. The data are from 4, 8 and 4 birds, respectively, in Spetch et al., 1990, and 8 birds in Spetch et al., 1994. The data for Spetch et al., 1994 are no longer available so the choice proportions were derived from Figure 5 of that study. The table excludes conditions in which the intertrial interval was increased to 40 s.

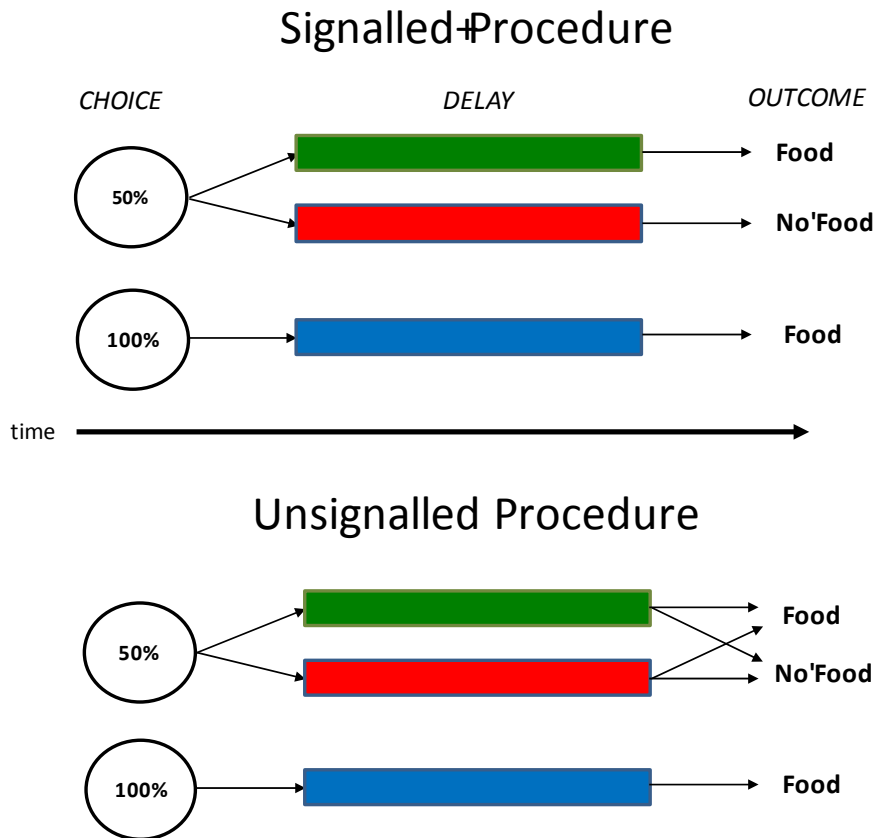


Figure 1. Schematic of the typical procedures used to study choice between 100% and 50% reinforcement in pigeons. Choice of the 100% alternative always leads to a delay period followed by food, whereas choice of the 50% alternative leads to the same delay but is followed by food only half of the time. Choice of the 50% alternative is considered suboptimal because it yields less food. The signalled (top) and unsignalled (bottom) procedures differ only in the correlation between the delay stimuli and the outcome on the 50% alternative. In the unsignalled procedure, the delay stimuli on the 50% alternative provide no information about whether the trial will end in food or no food (each outcome occurs 50% of the time following each stimulus).

In the signalled procedure, one stimulus signals that the delay will end in food and a visually different stimulus signals that the delay will end in no food.

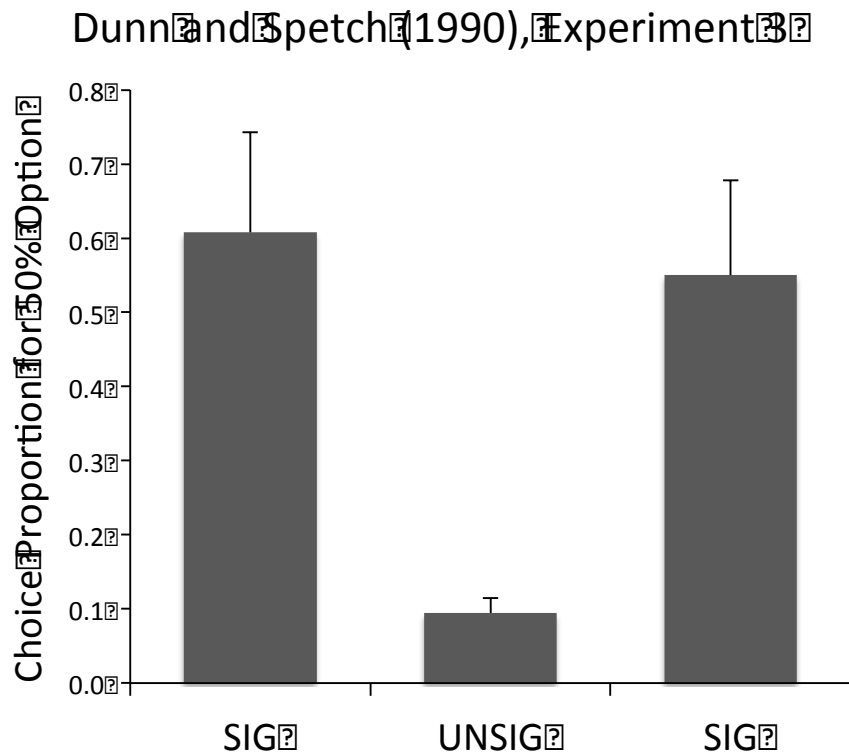


Figure 2. Choice proportions for the 50% (suboptimal) option averaged across the 5 birds in Experiment 3 of Dunn and Spetch (1990). All birds were tested in signalled (Sig) and unsignalled (Unsig) procedures in the order shown, and each condition was in effect until a stability criterion was met. Data are from the last nine sessions per condition. Error bars are standard error of the mean. Figure is adapted from “Choice with Uncertain Outcomes: Conditioned Reinforcement Effects,” by R. Dunn and M. L. Spetch, 1990, *Journal of the Experimental Analysis of Behavior*, 53, p. 212.

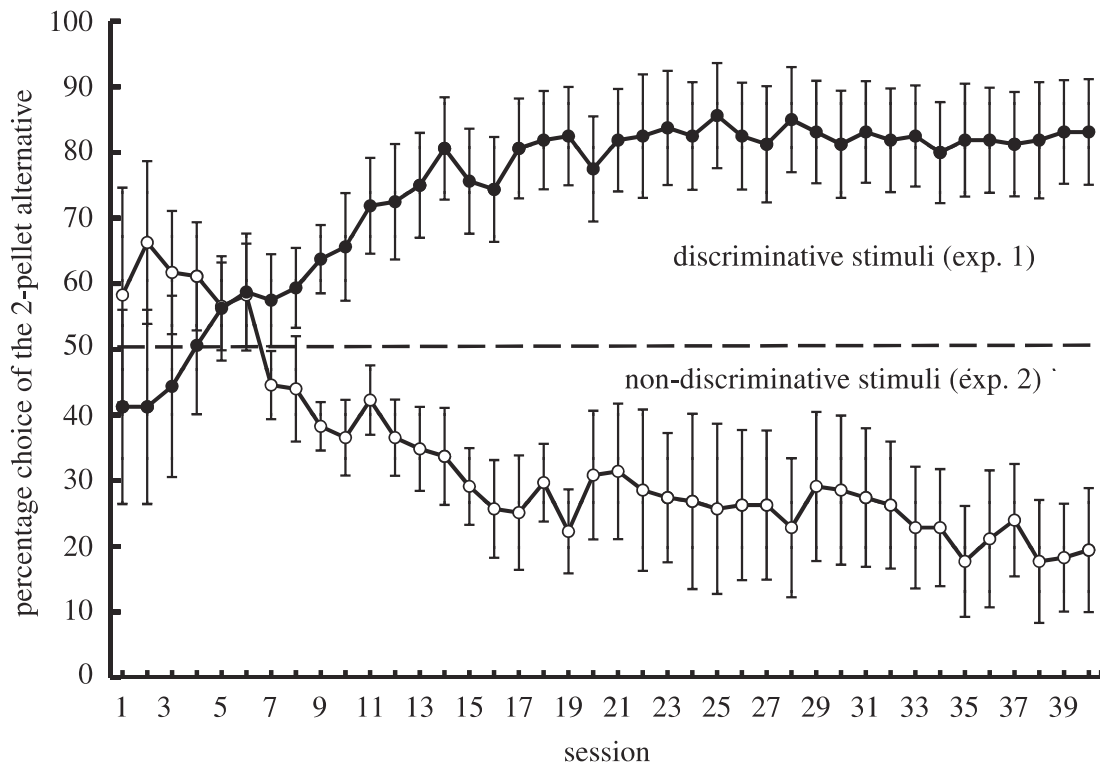


Figure 3. Percentage choice of the suboptimal (2-pellet) alternative by pigeons across sessions in a signalled procedure (discriminative stimuli) and an unsignalled procedure (non-discriminative stimuli). The suboptimal alternative provided 10 pellets with a probability of .2 and 0 pellets with a probability of .8. The optimal alternative always provided 3 pellets. Figure is reproduced from “Maladaptive Choice Behavior by Pigeons: An Animal Analogue and Possible Mechanism for Gambling (Sub-Optimal Human Decision-Making Behavior),” by T. R. Zentall, and J. Stagner, 2011, *Proceedings of the Royal Society B: Biological Sciences*, 278, p. 1205.

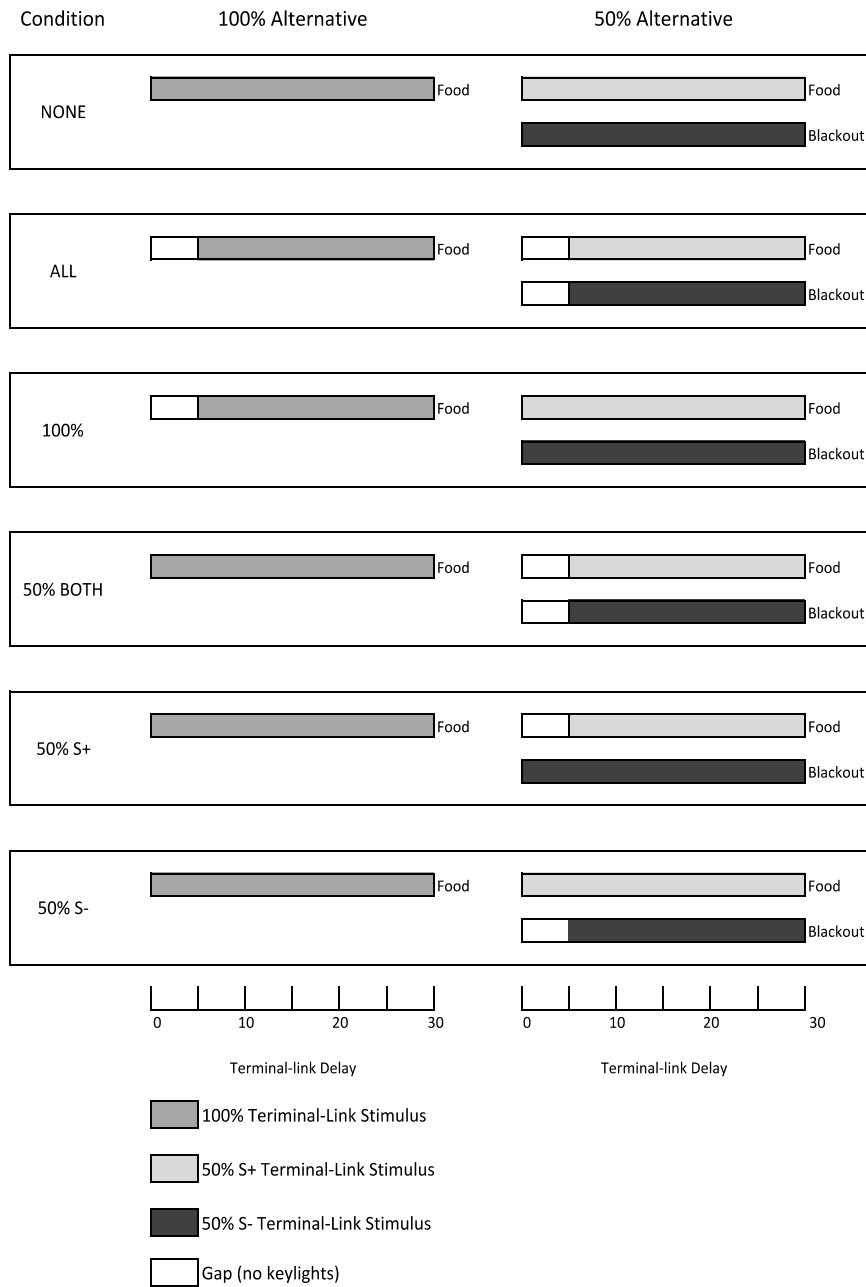


Figure 4. Schematic of the six conditions of the gap procedure tested by McDevitt et al (1997).

In all conditions pigeons chose between a 100% alternative, which always provided food after a

30-s delay, and a signalled 50% alternative, which provided food after the delay on only half of the trials. The NONE was the baseline in which the delay stimuli immediately followed the choice. The other conditions had a 5-s gap (indicated by white bar) inserted between the choice and onset of the one or more of the delay stimuli. Figure is reconstructed from “Contiguity and Conditioned Reinforcement in Probabilistic Choice,” by M. A. McDevitt, M. L. Spetch, and R. Dunn, 1997, *Journal of the Experimental Analysis of Behavior*, 68, p. 322.

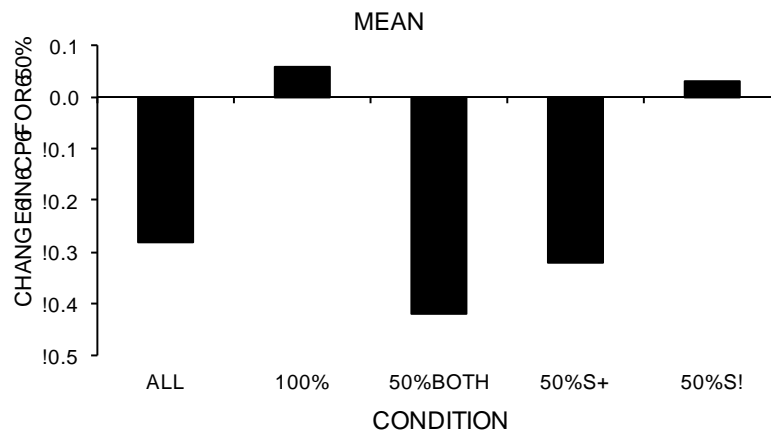


Figure 5. Change in choice proportions for the 50% (suboptimal) alternative in conditions in which a 5-s gap preceded one or more delay stimuli, relative to the NONE condition in which there were no gaps. Suboptimal choice decreased substantially in all conditions that included a gap before the food signal on the 50% alternative (conditions ALL, BOTH and S+), whereas there was little change when the gap was inserted before only the 100% delay stimulus or the signal for no food (S-) on the 50% alternative. Figure is reconstructed from “Contiguity and Conditioned Reinforcement in Probabilistic Choice,” by M. A. McDevitt, M. L. Spetch, and R. Dunn, 1997, *Journal of the Experimental Analysis of Behavior*, 68, p. 324.

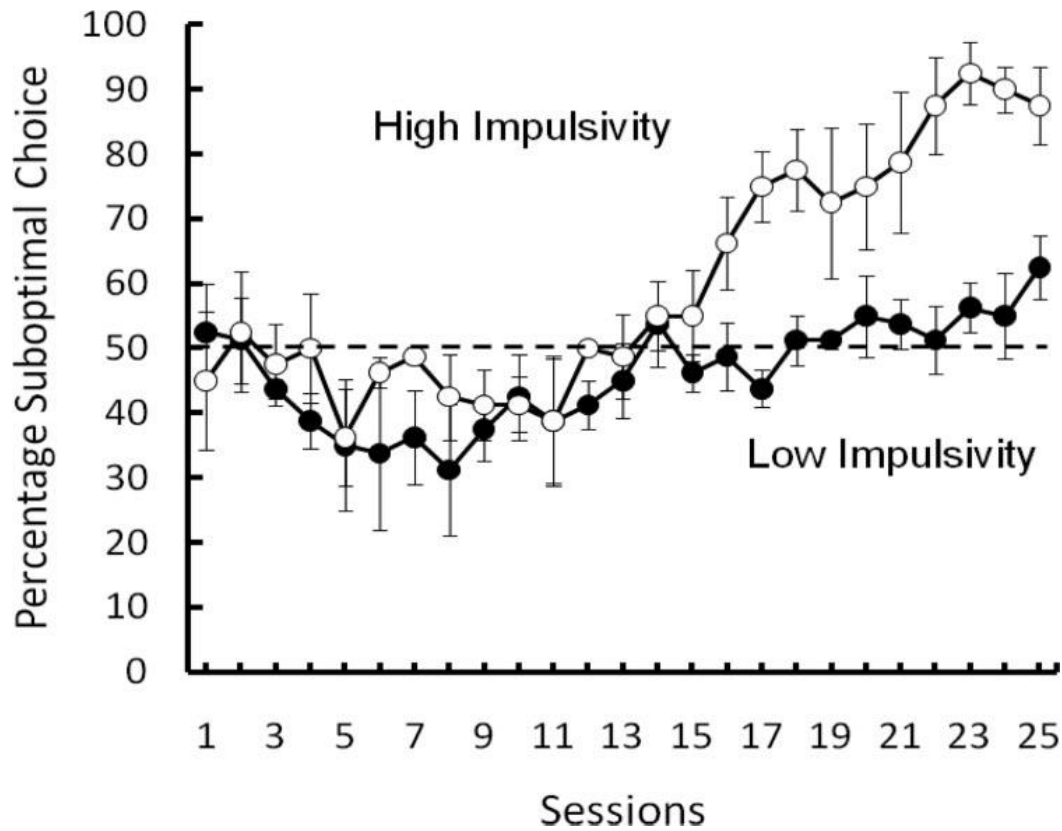


Figure 6. Mean percentage of suboptimal choice across sessions for pigeons that scored high on impulsivity in a separate delay discounting task (open circles) and for pigeons that scored lower on impulsivity in the discounting task (filled circles) in Laude, Beckmann et al., 2014. Figure reproduced from “Impulsivity Affects Suboptimal Gambling-like Choice by Pigeons,” by J. R. Laude, J. S. Beckmann, C. W. Daniels, and T. R. Zentall, 2014, *Journal of Experimental Psychology: Animal Behavior Processes*, 40, p. 8.

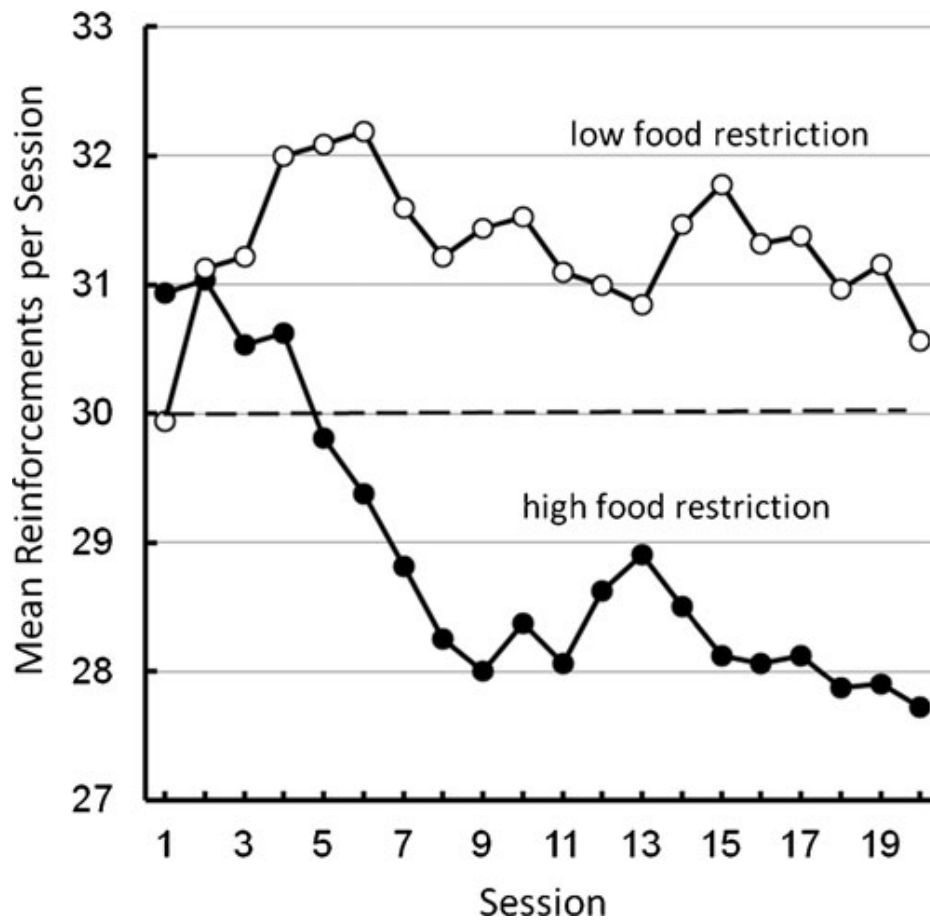


Figure 7. Mean number of reinforcers obtained in each session of a suboptimal choice procedure for pigeons maintained under high food restriction (75% of free feeding weight) or low food restriction (90% of free feeding weight) in Laude et al., 2012. The choice task involved signalled 50% versus unsignalled 75% reinforcement. Figure reproduced from “Hungry Pigeons Make Suboptimal Choices, Less Hungry Pigeons Do Not,” by J. R. Laude, K. F. Pattison, and T. R. Zentall, 2012. *Psychonomic Bulletin & Review*, 19, p. 888.

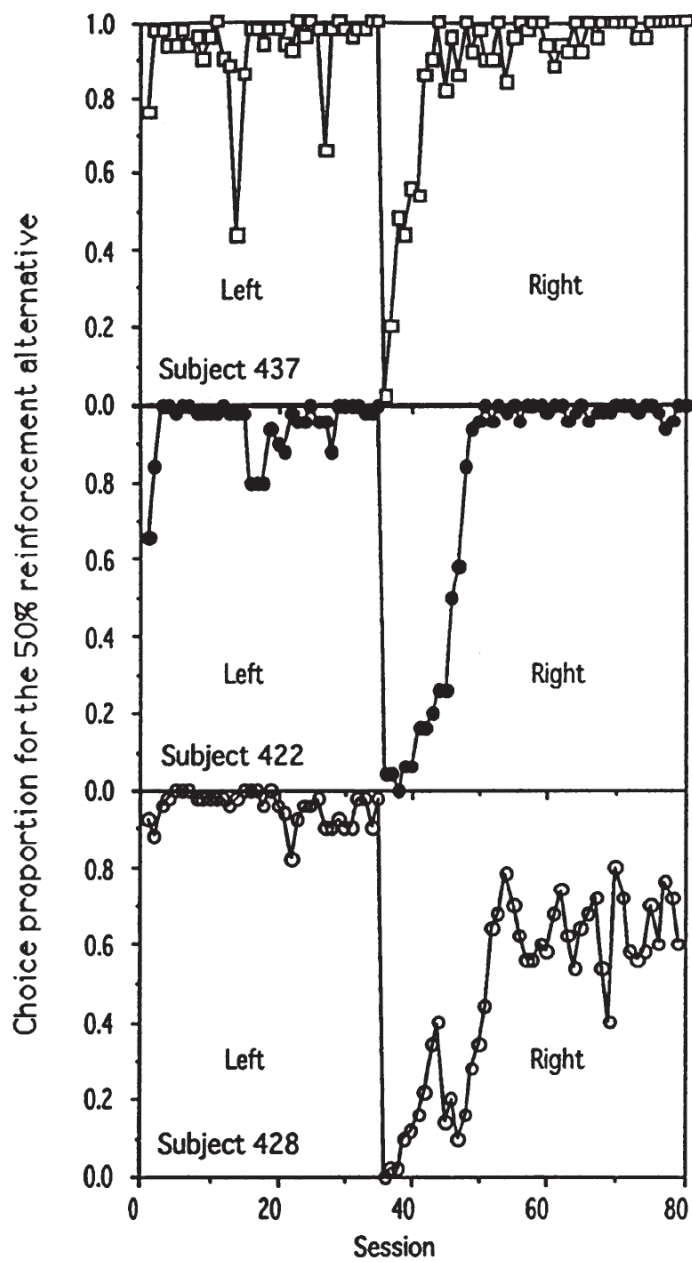


Figure 8. Proportion of choices for the suboptimal (signalled 50%) alternative over the 100% alternative for three pigeons that were given contingency reversals in Belke and Spetch, 1994.

For the first block of sessions choice of the left key led to 50% reinforcement and during the next block of sessions the contingencies were switched so that the right key led to 50% reinforcement.

Figure reproduced from "Choice Between Reliable and Unreliable Reinforcement Alternatives

Revisited: Preference for Unreliable Reinforcement,” by T. W. Belke and M. L. Spetch, 1994,
Journal of the Experimental Analysis of Behavior, 62, p. 361.