Operant Analyses of Choice Involving Probabilistic Reinforcers

BY

JEFFREY M. PISKLAK

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology
University of Alberta

Abstract

In the natural world, for both human and non-human animals alike, the consequences of choices can be delayed, infrequent, and often highly uncertain (probabilistic). These factors, and their resulting interactions, can have diverse effects that, at face value, seem to control the allocation of behaviour in counterintuitive ways. This dissertation investigates two distinct paradigms of choice that examine the role of probabilistic reinforcement on choice behaviour. More specifically, in Chapter 2, the conditionally reinforcing function of stimuli in a concurrent-chains pigeon (*Columba liva*) "suboptimal" choice paradigm were manipulated. Findings revealed that, contrary to the predictions of some theories of suboptimal choice, the frequency of reinforcement, and not just its signal value, was a necessary factor in determining how pigeons allocated their responding across choice alternatives. Chapter 3 assessed the impact of increased initial-link ratio requirements on a similar pigeon task to that of Chapter 2. As predicted, increasing the response requirement attenuated the effect of suboptimal preference. Chapter 4 evaluated the role of probabilistic outcomes in the context of the extreme-outcome effect in both pigeons and humans (*Homo sapiens*). Findings revealed that, like humans, pigeons will exhibit greater preference of a probabilistic ("risky") alternative for high-value choices relative to low-value choices. However, this finding was

found to be limited only to cases where a zero-value outcome (extinction) is a possible extreme within the choice context. When zero was removed from this context, only humans, and not pigeons, continued to display the extreme-outcome effect.

Preface

The research conducted within this thesis forms part of two international research collaborations: The work reported within Chapter 2 and Chapter 3 involved Dr. Margaret A. McDevitt at McDaniel College; Dr. Roger M. Dunn at San Diego State University, and Dr. Marcia L. Spetch at the University of Alberta. The work reported within Chapter 4 involved Dr. Christopher R. Madan at the University of Nottingham; Dr. Elliot A. Ludvig at the University of Warwick, and Dr. Marcia L. Spetch at the University of Alberta.

All the technical apparatuses referred to within Chapter 3 and Chapter 4 were designed, according to the specifications of Dr. M. Spetch and Dr. D. Grant, by the University of Alberta's Psychology Shop. All data analyses are my original work, as well as the literature review of Chapter 1 and conclusions of Chapter 5. Chapter 2's data collection was conducted by McDaniel College undergraduate students for course credit under the supervision of Dr. M. McDevitt. Data collection for the remaining chapters was conducted by me with the aid of laboratory research associates whom I supervised. Concept formation, manuscript composition, and editing for chapters 2-4 was a collaborative process involving each chapter's listed authors.

All procedures within Chapter 3 and Chapter 4 involving animal research procedures were approved by the University of Alberta Biological Sciences Animal Care and Use Committee, Project Name "COGNITIVE PROCESSES IN PIGEONS", No. AUP00002018. The procedures of Chapter 4 involving human research were approved by University of Alberta Research Ethics Board 2, Project Name "DECISION MAKING AND GAMBLING," No. Pro00058367.

Chapter 2 of this thesis has been published as Pisklak, J. M., McDevitt, M. A., Dunn, R., & Spetch, M. (2018). Frequency and value both matter in the suboptimal choice procedure. *Journal of the Experimental Analysis of Behavior*, *111*(1), 1-11. doi:10.1002/jeab.490.

Chapter 3 is an original work yet to be published and is currently undergoing peer review. The working title and author list for this work is as follows: Pisklak, J. M., McDevitt, M. A., Dunn, R., & Spetch, M. (2018). Suboptimal Choice and Initial Link Requirement.

Chapter 4 of this thesis is currently awaiting publication release as: Pisklak,

J. M., Madan, C. R., Ludvig, E. A & Spetch, M. L. (in press). The power of nothing: Risk preference in pigeons, but not people, is driven primarily by avoidance of zero outcomes. *Journal of Experimental Psychology: Animal Learning and Cognition.* Experiment 1 and 2 of this research were completed as part of unpublished work for my Master's thesis in Psychology at the University of Alberta. Experiments 3 and 4 were completed as part of my Doctoral research.

For my parents, who were far too patient with me.

"You see, I went on with this research just the way it led me. That is the only way I ever heard of true research going. I asked a question, devised some method of obtaining an answer, and got a fresh question. Was this possible or that possible? You cannot imagine what this means to an investigator, what an intellectual passion grows upon him! You cannot imagine the strange, colourless delight of these intellectual desires! The thing before you is no longer an animal, a fellow-creature, but a problem! Sympathetic pain,—all I know of it I remember as a thing I used to suffer from years ago. I wanted—it was the one thing I wanted—to find out the extreme limit of plasticity in a living shape."

"But," said I, "the thing is an abomination—"

-H.G. Wells, *The Island of Doctor Moreau*

Acknowledgements

Contents

List of Figures

List of Tables

Nomenclature

DRT Delay Reduction Theory:

$$\frac{B_1}{B_2} = \left(\frac{R_1}{R_2}\right)\left(\frac{T-t_1}{T-t_2}\right), T > t_1, T > t_2$$

extinction [Operant] The procedure of withholding a reinforcer.

FE Forced-Exposure Trials

FI Fixed-Interval Schedule: One response is reinforced after a fixed amount of time has elapsed.

FR Fixed-Ratio Schedule: A specific number of responses is required to obtain reinforcement.

HDM Hyperbolic Decay Model:

$$V = \sum_{j=1}^{n} p_j \left(\frac{A}{1+KD_j}\right)$$

HVA Hyperbolic Value-Added Model:

$$\frac{B_1}{B_2} = b \left(\frac{r_{i1}}{r_{i2}}\right)^{a_j} \left(\frac{V_{t1}-a_t V_i}{V_{t2}-a_t V_i}\right), V_{t1} > a_i V_i, V_{t2} > a_i V_i$$

ITI Inter-Trial Interval

RRM Reinforcement Rate Model

SCM Sequential Choice Model

Sig-Both Chapter 2 experimental condition containing a choice between a 20% signalled suboptimal alternative versus a 80% signalled optimal alternative.

Sig-Sub Chapter 2 control condition containing a choice between a 20% signalled suboptimal alternative versus a 80% unsignalled optimal alternative.

SiGN Signals for Good News Hypothesis

signalled When a stimulus or event is maximally correlated ($r = 1$) with the delivery or absence of a terminal reinforcer

unsignalled When a stimulus or event is partially correlated ($0 < r < 1$) with the delivery or absence of a terminal reinforcer

VI Variable-Interval Schedule: One response is reinforced after a variable amount of time has elapsed.

VR Variable-Ratio Schedule: A variable number of responses is required to obtain reinforcement.

Chapter 1   An Introduction to Probabilistic Choice

For many, the concept of "choice" is a complex one that can range from profound philosophical issues of free-will – on both political and metaphysical level – to normative claims of what should be done, to purely descriptive claims of what is in fact done. But even more so, choice often implies a single explicit determination of some kind on the part of the chooser. In the context of behavioural research, however, choice is more encompassing than this. As one researcher succinctly phrased it, "all behaviour is choice" (Baum, 2010, p. 161). Within any given context there are a range of behaviours that could be performed, and some will be more probable than others. For a pigeon recently deprived of food, key pecking for edibles becomes eminently more probable than other, often unmeasured, background activities such as grooming, cooing, scratching, and the like. For a human confined to a waiting room, lamentably devoid of reading and listening materials, conversation with other unlucky individuals might come to dominate their observed behaviour (Conger & Killeen, 1974). Examples aside, this trade-off between behaviours of varying probability has been most elegantly captured by Herrnstein's quantitative law of effect (Equation 1.1):

$$B = \frac{kR}{R + R_e} \qquad\qquad (1.1)$$

Here, B is equal to the obtained rate of a specific target behaviour, k is equal to the asymptotic rate of that behaviour, R is equal to that behaviour's rate of reinforcement, while R_e is equal to the rate of reinforcement for other potential background activities. The equation states that the rate of any given response will be a hyperbolic function of the rate of reinforcement of that response and other extraneous responses (Herrnstein, 1970). Equation 1.1's effect is depicted in Figure 1.1.

Herrnstein's equation has met with an impressive amount of success empirically, describing behaviour across a range of species and reinforcers (see de Villiers, 1977, for a review), but moreover, it describes a fundamental truth that all observed behaviour can be understood as a choice, but this is not choice as it is routinely conceived. There is no assumed "intentionality" or "act of will" as a philosopher might say. From this perspective, this is choice as the allocation of behaviour over time (Baum, 2010). The organism does not "choose" a behaviour to engage in, rather the various contingencies of
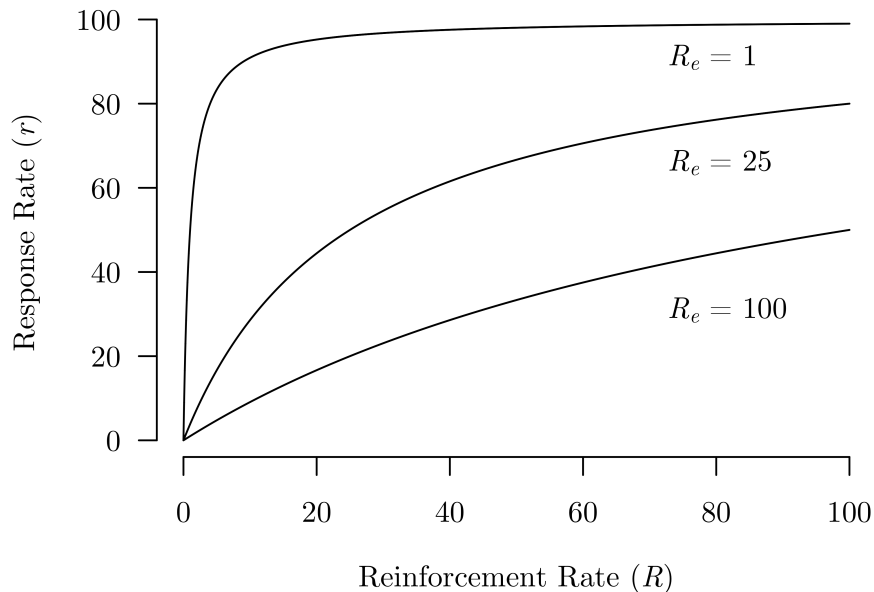
Figure 1.1: A plot of Herrnstein's quantitative law of effect showing the response rate of a target behaviour (B) as a function of its reinforcement rate (R) and the reinforcement rate of competing background activities (R_e).

reinforcement (and punishment) in the organism's environment *selects* the behaviour it engages in (Skinner, 1981), just as environments naturally select which species thrive and which species go extinct.

The focus of this dissertation centres around issues of choice and how it is influenced by the occurrence of probabilistic reinforcement. The objective of this chapter is to introduce necessary background elements of probabilistic choice that inform the three studies discussed later in Chapters 2-4. A heavy emphasis will be placed on the foundational role of animal research to illuminate fundamental principles of behaviour. While many will find this a disagreeable point, in most, if not all, of the cases described below, the basic truths that have emerged from operant research involving, typically small, laboratory animals such as pigeons and rats have been shown to extend readily to the human case. And even for those scenarios where it does not (e.g. Section 1.4's discussion of description-based choices), the study of animals is valuable

for their own sake (not just ours) and much can be gleaned from examining where the differences between distinct species lay. Indeed, the research to discussed later in Chapter 4 attests to the veracity of this.

1.1. **The Role of Matching**

In the operant tradition of psychological research, early studies detailing the effect of probabilistic outcomes on choice behaviour often centred around examinations of schedule induced behaviour. In their landmark title, *Schedules of Reinforcement*, Ferster and Skinner (1957) presented pigeons and rats with differing combinations of reinforcement schedules. One notable combination tested pigeons with two independent and concurrently available variable-interval (VI) 1-minute schedules. Interval schedules reinforce a behavioural response only when the response occurs *after* a given period of time. The variable nature of the schedules employed further meant that the reinforcement delivery operated around a psuedo-randomized average, and not a fixed, amount of time. Thus, whether any given response produced reinforcement (i.e., the rate of reinforcement) was a probabilistic function of time. Following steady-state performance, it was observed that this contingency led Ferster and Skinner's pigeons to allocate their responding between the two alternatives at rates approximately equivalent to one another, thus maximizing their obtained rate of reinforcement. It is noteworthy that had the pigeons developed an exclusive preference towards one alternative, they would have been limiting their food intake needlessly because, following the reinforcement of one alternative on an interval schedule, there is a period of time in which further responding will not yield reinforcement. Thus, from the standpoint of a pigeon maximizing reinforcement, it is optimal to switch towards the adjacent alternative where a greater amount of time will have elapsed since the last reinforcer delivery. This move makes the probability of any given response being reinforced necessarily higher.

Later work by other researchers would reveal that when the rates of reinforcement on each alternative's VI schedule are altered, predictable changes in the rate of behaviour are also observed such that the relative rate of responding will *match* the relative rates of reinforcement according to the following

Equation (Herrnstein, 1961):

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2} \tag{1.2}$$

B_1 and B_2 equal the predicted rate, time, or total amount of responding on each alternative, and R_1 and R_2 are the programmed rates of reinforcement provided by each alternative. Additionally, Equation 1.2 can be straightforwardly expanded to encompass n alternatives and sources of reinforcement:

$$\frac{B_1}{B_1 + B_2 + B_3 + \cdots + B_n} = \frac{R_1}{R_1 + R_2 + R_3 + \cdots + R_n} \tag{1.3}$$

The experiment described above consisted of outcomes that were determined as a probabilistic function of time (i.e., a VI schedule). In contrast to the VI case, a scenario can be set up whereby it is disadvantageous, in terms of maximizing reinforcement, to switch between alternatives but the delivery of reinforcement is still probabilistic. Consider a scenario where a pigeon is confronted with concurrently available variable-ratio (VR) 50 and VR 70 schedules (Herrnstein & Loveland, 1975). Ratio schedules make reinforcement delivery contingent on a set amount of responses, as opposed to a set amount of time. This means that ratio schedules have a feedback relation wherein more responding produces more reinforcement. Additionally, in terms of maximizing reinforcement, switching between alternatives limits reinforcer delivery needlessly. It is more advantageous to respond only to the alternative containing the lower response requirement. It is perhaps unsurprising then that ratio schedules (fixed or variable) typically select for exclusive preference of the schedule with the highest rate of reinforcer delivery per unit of behaviour (i.e., the VR 50 in the above scenario). In such cases Equation 1.2 provides a poor account of choice behaviour and a generalized logarithmic form (Baum, 1974) of the equation is often used to quantify the deviation obtained. Equation 1.4 shows the equation detailing the generalized matching relation:

$$\log \frac{B_1}{B_2} = a \cdot \log \frac{R_1}{R_2} + \log c \tag{1.4}$$

The two specific types of deviation from matching are bias, c, and sensitivity, a.

When $\log c = 0$ and $a = 1$, equation Equation 1.4 makes the same behavioural predictions as Equation 1.2. Choices between ratio schedules typically still follow the predictions of matching though, as might be expected, it occurs with a systematic bias towards the higher frequency alternative. Interestingly, evidence suggests that the bias observed is not straightforwardly controlled by the smaller of the two response requirements but is in fact – by virtue of the feedback relation – controlled by the ratio of mean response requirements in the two schedules (Beautrais & Davison, 1977). Thus, the choice bias observed for any two ratio schedules will actually be unique to that pair.

A wealth of literature has provided unambiguous support of behaviour matching to rates of reinforcement across a range of responses, species, and reinforcer types (e.g. Baum & Rachlin, 1969; Brownstein & Pliskoff, 1968; Catania, 1963; Conger & Killeen, 1974; Mcsweeney, 1975; Schroeder & Holland, 1969; Silberberg & Fantino, 1970). Research has also sought to assess the generality of matching to other characteristics of the reinforcing outcome, specifically its magnitude. Magnitude can be quantified in terms of reinforcement duration (e.g., hopper access time) or absolute amounts of reinforcement (e.g., number of pellets delivered). When these independent variables are isolated from reinforcement rate the findings vary. Some studies find matching and some find undermatching (de Villiers, 1977), though the consensus seems to be that undermatching is the more common result (M. C. Davison & McCarthy, 1988). Undermatching describes a situation where, contrary to the prediction of matching, less responding is seen to one alternative as the reinforcement ratio between the alternatives increases (i.e., $a < 1$). Thus, as the value of M_1/M_2 increases, where M refers to the magnitude of a given alternative, less responding is seen on B_1 relative to B_2 than would be expected by pure matching. The prevalence of undermatching seen towards reinforcer magnitude suggests that its application to matching might be best considered as part of the total reinforcement received (rate $\times$ magnitude) or, even better, as a general "value" measure that is determined on the basis of numerous parameters of reinforcement, e.g., rate $\times$ magnitude $\times$ immediacy $\times$ any other measurable aspect (Killeen, 1972; Rachlin, 1971). The relation between rate and magnitude can be straightforwardly expressed in the generalized logarith-

mic form as:

$$\log \frac{B_1}{B_2} = a_R \cdot \log \frac{R_1}{R_2} + a_M \cdot \log \frac{M_1}{M_2} + \log c \qquad (1.5)$$

1.2. **The Role of Delay**

The research outlined in Section 1.1 focused heavily on how responding is distributed under conditions of immediate reinforcement. Once a schedule requirement has been satisfied, the reinforcer is immediately presented. However, the consequences of our choices are rarely so prompt as this. In many cases the choices we make have only distant outcomes which we may or may not opt to subject ourselves to. Some early research examining the influence of delay between a choice response and a reinforcer found that, when expressed as the reciprocal of delay, called "immediacy", the equations of Section 1.1 demonstrated matching to immediacy ratios (Chung & Herrnstein, 1967). However, attempts at replicating this have not met with success (Neuringer, 1969; Shimp, 1969a; Williams & Fantino, 1978). Furthermore, findings emerged that defied straightforward explanation in these terms (e.g., Duncan & Fantino, 1970; Herrnstein, 1964). For instance, in a commonly used choice procedure called a concurrent-chains procedure, pigeons could be presented simultaneously with two, left and right, pecking keys each operating on VI 1s schedules called *initial links*. If the schedule on the left initial link is satisfied, then its key changes colour and a VI 15s schedule, called a *terminal link*, goes into effect and terminates with food delivery. If the right initial link is satisfied, the right key changes colour and a FI 15s schedule goes into effect and terminates in food delivery. Using these parameters, strong preference for the left initial link is obtained despite the fact that its rate of reinforcer delivery is equivalent to that of the left. Moreover, even when rates of delivery on the right key are substantially higher than the left (e.g., FI 4), a consistent bias to the variable schedule will still be seen (Herrnstein, 1964), a finding seemingly at odds with the heavy control of reinforcement rate described in Section 1.1 and basic notions of "rational" choice.

The choice scenario just described, in which a probabilistic ("risky") delay is presented against a fixed ("safe") delay, represents a cornerstone of choice research. Both initial links result in an equivalently sized reinforcer at the end

of the terminal link, but the option containing the VI terminal link sometimes provides food after both short and long delays. Powerful evidence has since demonstrated that the reason for this is likely due to a hyperbolic discounting of reinforcers (Rachlin, 2000).

As the delay between a response and a reinforcer increases, the effectiveness (i.e., value) of that reinforcer diminishes according to a hyperbolic function described by Equation 1.6 (Mazur, 1984, 1986),

$$V = \sum_{j=1}^{n} p_j \left(\frac{A}{1 + KD_j} \right) \tag{1.6}$$

where V represents the value of a delayed reinforcer, p_j is the probability of a given delay to reinforcement D_j, A is the undiscounted value of the reinforcer (i.e., when $D = 0$), and K is a free parameter specifying the rate at which the value declines as a function of delay. This function, called the *hyperbolic decay model* (HDM), is depicted in Figure 1.2 and shows how the value of a reinforcer declines with ever increasing delay at three separate values of K, with $A = 1$.

As shown in Figure 1.2, most of the reinforcer's value is lost at short delays. This straightforward relation between reinforcer effectiveness and delay explains the preference seen for variable schedules over fixed. Consider the case described above comparing the FI 15s and VI 15s terminal links. Assuming $A = 1$ and $K = 1$, the fixed alternative has a value of $V = 1/(1+1\times15) = 0.06$. For simplicity we shall assume the variable alternative has only two possible delays, 1s and 29s (i.e., mean delay of 15s), that occur with equal likelihood. This creates a value at each j delay that sums to an overall value: $V_1 + V_2 = 0.27 = V$. Here $V_1 = 0.5\times(1/(1+1\times1)) = 0.25$ and $V_2 = 0.5\times(1/(1+1\times29)) = 0.02$. Since the value of the variable schedule's outcome (0.27) is higher on average than the fixed's (0.06), it is predicted to be the preferred outcome. It can be seen from this simple example that the preference for the VI schedule is due to its occasional short delay, which is disproportionately reinforcing – a conclusion others have reached independently of Mazur's work (M. C. Davison, 1972; Duncan & Fantino, 1970; Sherman & Thomas, 1968).

One of the central insights of Equation 1.6 is that it nicely explains why humans and other animals will often act against their own interest in "self-
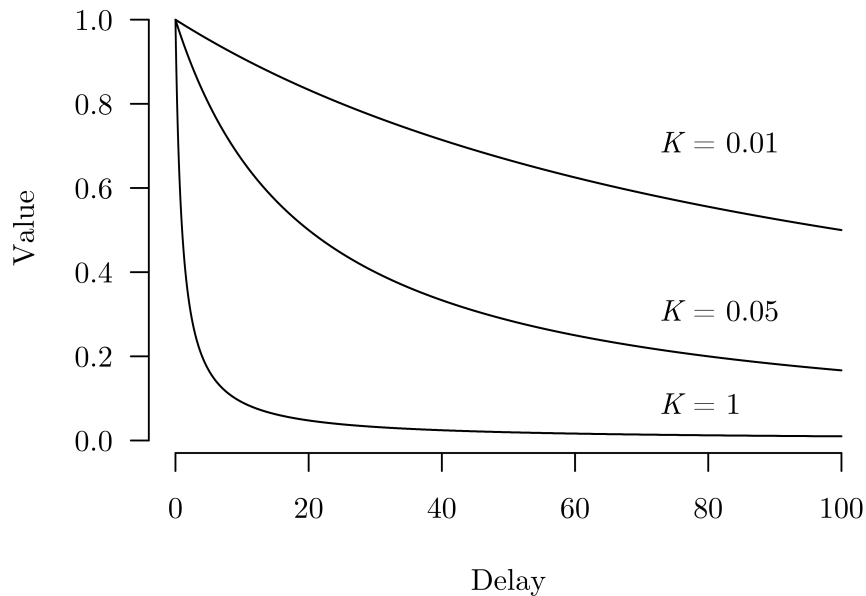
Figure 1.2: A plot of the HDM showing the effect of delay, D, on reinforcer value, V, at three levels of K.

control" tasks (Rachlin, 2000). Self-control tasks present a choice between a "larger-later" and a "smaller-sooner" option. The larger-later option provides more food overall for its selection, but at the cost of waiting (i.e., long terminal links). The smaller-sooner option, by contrast, provides less food overall but offers it up relatively quickly (short terminal links). A common finding, particularly with laboratory animals (Bateson & Kacelnik, 1996; Mazur & Romano, 1992; Mcdiarmid & Rilling, 1965), but with humans as well (Rosati, Stevens, Hare, & Hauser, 2007), is that, when the choice made in the initial-links is discrete (i.e., requiring a single response), the smaller-sooner option is preferred even in spite of its lower magnitude of reinforcement. Equation 1.6's explanation for this finding is the same as it was for the variable vs. fixed schedule preference described above: the long delays devalue the reinforcer too substantially relative to the value provided by the smaller-sooner option. Interestingly, if animals are provided an earlier option that allows them to *commit* to the larger-later option, they will take it provided the commitment

is made far enough in advance of the main choice. If the commitment option is provided only shortly before the presentation of the smaller-sooner vs. larger-later choice, they will opt to not commit and select the smaller-sooner option (Ainslie, 1974), a finding predicted by Equation 1.6.

Humans have frequently been shown to exhibit similar hyperbolic discounting functions as that observed in other laboratory animals. In many cases they are tested with hypothetical reinforcers (e.g., Friedel, DeHart, Madden, & Odum, 2014; Green, Fry, & Myerson, 1994; Rachlin, Raineri, & Cross, 1991; Odum, Baumann, & Rimington, 2006) but real reinforcers have been applied as well with little observed difference between the two (Johnson & Bickel, 2002; Lagorio & Madden, 2005; Madden, Begotka, Raiff, & Kastern, 2003). However; given the innumerable differences that exist between human and non-human animal research, care is still probably warranted in comparing them (Madden et al., 2004).

An interesting facet of Equation 1.6 is that it not only accounts for probabilistic delays to a guaranteed outcome, but it can be used to account for probabilistic outcomes themselves (i.e., a chance of being reinforced or not). Rachlin, Logue, Gibbon, and Frankel (1986), first showed that there is no functional difference between the probabilistic delivery of a reinforcer and delaying a reinforcer. That is to say, when a reinforcing outcome is made less likely to occur for a given response, its reinforcing effect is diminished in an analogous fashion to the diminishment caused by simply delaying that reinforcer. Mazur (1989) found this effect to be best described by Equation 1.6 when tested against other models of probabilistic reinforcement. Critically, however, for Equation 1.6 to apply there needs to be a measured duration and a probability associated with that duration. In animal studies, this often takes the form of time spent in the presence of stimuli correlated with a given alternative (i.e., conditional reinforcers). For instance, consider two trials in which a pigeon selects an initial link leading to a red terminal link lasting for 5s. On trial 1 the terminal link ends with no reinforcer delivered, whereas on trial 2, reinforcement is provided. In this case, the delay to reinforcement is $5 + 5 = 10$s, which produces a value of $V = 1/(1 + 1 \times 10) = 0.09$. Contrast this with a separate alternative that guarantees reinforcement on each trial. In this case the guarantee only has a delay to reinforcement of 5s and thus a corresponding higher value, $V = 1/(1 + 1 \times 5) = 0.17$. Here the guaranteed alternative is

almost twice as reinforcing as the probabilistic alternative (time in the initial links are not considered for simplicity).

1.3.   **The Role of Conditional Reinforcement**

A critical finding of Mazur's (1989) work on probabilistic reinforcement was that, when the inter-trial interval (i.e., the time between the outcome and the initial-link) was varied, no systematic effects on choice were observed. Thus, the delay (D) value of Equation 1.6 was best interpreted as time spent in the presence of stimuli signalling reinforcement, and not the total time to reinforcement (i.e., trial duration). Given this and other findings (see Experiment 3), Mazur concluded that Equation 1.6's HDM V parameter is best understood as the value of conditional reinforcers, and not simply the value of delayed reinforcers.

A conditional reinforcer is an otherwise neutral stimulus or event that has acquired the ability to reinforce a behaviour due to its relationship with other, typically unconditional, reinforcers. Classically, this "relationship" has been interpreted in terms of the pairing of stimuli (i.e., temporal contiguity), though it has long been known that basic pairing is neither necessary or sufficient to account for the breadth of findings on conditional reinforcement research (e.g., Schoenfeld, Antonitis, & Bersh, 1950; Stubbs, 1971). In terms of conditional reinforcement's application to studies of choice, the most widely accepted conception of it among researchers is *delay reduction theory* (DRT; Fantino, 1969), which posits that the conditional reinforcing value of a stimulus is determined by the degree to which it is correlated with a reduction in time to unconditional reinforcement (for a review of conditional reinforcement, see Williams, 1994). Equation 1.7 presents a ratio form of DRT that incorporates the basic matching processes of Section 1.1 and is based on the formulas of Squires and Fantino (1971):

$$\frac{B_1}{B_2} = \left(\frac{R_1}{R_2}\right)\left(\frac{T - t_1}{T - t_2}\right), T > t_1, T > t_2 \tag{1.7}$$

R_1 and R_2 correspond to the rates of reinforcement provided by both alternatives across their respective initial and terminal link. T is the mean time to unconditional reinforcement from the onset of the initial links (see Squires

& Fantino, 1971, for formal calculations of T), t_1 and t_2 correspond to the mean duration of the terminal links respectively. The expression $T\text{–}t$ inside the equation represents the core premise of delay reduction and determines how conditionally reinforcing a terminal link will be. As an illustration, consider an experiment by Fantino (1969) where one choice alternative offers a white VI 90s initial link followed by a green VI 30s terminal link and then food, while another alternative provides a white VI 30s initial link followed by a red VI 90s terminal link and then food. Note that in this scenario, both alternatives require the same overall amount of time and responding to obtain food; however, according to DRT, the green terminal-link signals a larger reduction in delay (97.5–30 = 67.5) than the red terminal-link (97.5–90 = 7.5) and should be preferred by a ratio of 9:1 based on effects of conditional reinforcement alone, which is indeed what was observed by Fantino (1969).

Both the hyperbolic decay and delay reduction interpretation of conditional reinforcement have differing relative merits in terms of what they can and cannot explain (Mazur, 1997); however, a critical advantage of DRT is that its mathematical formulation is specifically tailored to concurrent chain paradigms which form the basis for many contemporary analyses of conditional reinforcement and choice studies generally. For instance, in the Fantino (1969) experiment just described, the HDM has no explicit rule for dealing with the differing long initial links. Are they to be treated as functionally equivalent to their terminal-link counterpart, thereby predicting indifference? Should their duration be ignored and value only be determined on the basis of the terminal links? This would produce predictions consistent with Fantino but seems incorrect given the observed effects specific to initial-link manipulation (Fantino, 1977). Also, how does Equation 1.6's "value" translate to explicit quantitative predictions of preference?

Mazur addressed these various issues by introducing the concept of "value addition" into a model of concurrent chains called the *hyperbolic value-added model* or HVA (Mazur, 2001). The model is shown in Equation 1.8 and works off of a basic delay-reduction like mechanism. In it, Equation 1.6 is used to determine the value of each alternative's initial link (V_i) and the value of each alternative's terminal link (V_t). The two values are then subtracted ($V_t\text{–}V_i$) to determine the value added by the terminal link stimulus (i.e., the putative conditional reinforcer). Like DRT, the choice preference is then determined

by the ratio of value added by each alternative's terminal link. In essence, the hyperbolic valued-added model is identical to DRT, but with strict delay reduction $(T–t)$ being replaced by value addition $(V_t–V_i)$. Additionally, the model includes a sensitivity parameter, a_t, for the value addition component of the model and a separate parameter, a_j that controls the sensitivity to initial link reinforcement rates, r_{i1} and r_{i2}. The parameter b reflects systematic bias and is used to quantify departures from the model's formal predictions.

$$\frac{B_1}{B_2} = b \left(\frac{r_{i1}}{r_{i2}}\right)^{a_j} \left(\frac{V_{t1} - a_t V_i}{V_{t2} - a_t V_i}\right), \quad V_{t1} > a_i V_i, \quad V_{t2} > a_i V_i \qquad (1.8)$$

The incorporation of value addition into the general framework of conditional reinforcement, as it is understood from the HDM's perspective, is not a trivial one. It neatly allows for Equation 1.6 to be interpreted into a much wider set of scenarios, which is valuable given the equation's ability to account for both probabilistic delays and probabilistic outcomes.

1.4. The Description-Experiance Gap

The discussion of probabilistic outcomes in Sections 1.1, 1.2, and 1.3 centred around circumstances in which the probabilities are learned (i.e., effect a change in behaviour) through repeated exposure to outcomes contingent on a particular choice response. These might broadly be called "experience-based" choices. However, with humans, a hypothetical scenario can be contrived in which alternatives are presented verbally with the probabilities specified in advance of any outcome (of which there are often none given). For instance, Table 1 shows examples of two separate choices (A and B) Kahneman and Tversky (1979) posed to their participants.

When choice A was posed, most participants opted for the "safe" guarantee of money, meaning they were "risk-averse." However, when the question was re-framed in terms of losses (choice B) most participants opted for the "risky" (i.e., probabilistic) option, meaning they were "risk-seeking." A noteworthy point about the values used is that the preferences observed did not correspond to the expected value. For instance, in the gain scenario (A), the expected value for the risky option is $0.8 \times 4000 = \$3200$; whereas, it is simply \$3000 for the safe option. Thus, the "rational" preference to have is one

Table 1.1

Two choice examples used by Kahneman and Tversky (1979)

Choice		Decision Frame	Expected Value
A	80% chance to win $4000 (20% chance to win $0)	Risky Gain	$3,200
	- or -		
	100% chance to win $3000	Safe Gain*	$3,000
B	80% chance to lose $4000 (20% chance to lose $0)	Risky Loss*	-$3,200
	- or -		
	100% chance to lose $3000	Safe Loss	-$3,000

Note. * shows the more preferred alternative. Original values used by Kahneman and Tversky were presented in Israeli pounds.

that favours the risky option in this case; though this is not what people were generally found to prefer. Even when the expected value of the two options was equated preferences still corresponded to risk-aversion for gains and risk-seeking for losses (see problems 7 and 8 of Kahneman & Tversky, 1979).

Kahneman and Tversky dubbed their finding of risk-aversion for gains and risk-seeking for losses the *reflection effect*. The reflection effect represents a core finding of "description-based" choices; that is to say, choices which are made on the basis of described probabilities and not experienced outcomes. One might assume that these description-based choices are, in some way, broadly reflective of how a person will behave in a choice scenario with unknown outcomes and probabilities that have to be learned through experience. However, findings have emerged showing that the preferences obtained using described choices do not generalize to that of experience-based choices (Barron & Erev, 2003; Hau, Pleskac, & Hertwig, 2010; Hertwig, Barron, Weber, & Erev, 2004; Ungemach, Chater, & Stewart, 2009), and that experienced-based choices actually produce a reversal of the reflection effect (Ludvig & Spetch, 2011; Ludvig, Madan, & Spetch, 2014; Madan, Ludvig, & Spetch, 2017). Specifically, with experienced-based choices, greater risk-seeking is seen for gains than losses, even when the

expected value of the choices are equated.

One proposed explanation, dubbed the *extreme-outcome effect*, that has received substantial support from a sizable body of work (for a review see Madan, Ludvig, & Spetch, 2019) is that the most extreme outcomes inside the choice context are overweighted (i.e., disproportionately reinforcing or punishing relative to their counterparts). Using Table 1.1's values as an example, the most extreme values are +$4000 and –$4000. When the outcomes and the probabilities of all four of Table 1.1's alternatives are part of the same overall context in which choices occur, and have to be learned through repeated experience, the two extreme outcomes will sway preference such that the risky-gain will tend to be preferred when choice A is prompted and the safe-loss will tend to be preferred if choice B is prompted. This effect of extreme-outcomes may not be a uniquely human phenomenon either as it has been observed, in at least one case, to occur in pigeons on a foraging task for food rewards (Ludvig, Madan, Pisklak, & Spetch, 2014, however, see Chapter 4).

1.5. **Objective of the Current Work**

The current work builds off of, and is informed by, much of the past work highlighted in the sections above, detailing three separate studies that all relate to how probabilistic outcomes influence choice in pigeons (*Columba livia*) and humans (*Homo sapiens*). This is done with the goal of furthering knowledge about how probabilistic reinforcers function on certain aspects of choice behaviour. In particular, Chapter 2 and Chapter 3 centre around a paradigm of probabilistic choice that generates seemingly paradoxical results in laboratory pigeons (e.g., Dunn & Spetch, 1990; Kendall, 1974; Pisklak, McDevitt, Dunn, & Spetch, 2015), as well as other species such as rats, humans, and starlings (Ojeda, Murphy, & Kacelnik, 2018; Molet et al., 2012; Vasconcelos, Monteiro, & Kacelnik, 2015). Generally, when a choice is to be made between differing rates of the same reinforcer, a sufficiently food-deprived pigeon will form a preference for the higher of the two rates, with preferences scaling according to principles of matching (M. C. Davison & McCarthy, 1988). This fact notwithstanding, the contingencies inside the pigeon's environment can be arranged such that a pigeon will learn (be conditioned) to prefer the lower of the two rates. This is a finding that, at face value, seems to be at odds with

basic notions of expected utility (Bernoulli, 1954), optimal foraging (Stephens & Krebs, 1986), molecular maximization (Shimp, 1969b), molar maximization (Rachlin, Battalio, Kagel, & Green, 1981), and melioration (Herrnstein, 1997). In order to better understand the variables controlling this behaviour, Chapter 2 assessed the role of the signal-value frequency, while Chapter 3 explored the effect of increased response requirements. By contrast, Chapter 4 is directed at a different aspect of choice with probabilistic outcomes. In particular, it evaluates predictions of the extreme-outcome effect across two species, humans and pigeons, in four separate experiments. This work questions some assumptions that the two species might be operating according to a similar underlying mechanism of choice behaviour (Ludvig, Madan, Pisklak, & Spetch, 2014). Thus, Chapter 4 provides a cautionary tale about generalizing between evolutionarily disparate species in comparative analyses of behaviour. Chapter 5 summarizes the key findings of the previous three chapters and relates their significance to a more general understanding of how choice operates with probabilistic reinforcers.

Chapter 2   Frequency and Value Both Matter in the Suboptimal Choice Procedure

Pisklak, J. M., McDevitt, M. A., Dunn, R., & Spetch, M. (2018). Frequency and value both matter in the suboptimal choice procedure [Journal Article]. *Journal of the Experimental Analysis of Behavior, 111*(1), 1-11. doi:10.1002/jeab.490

2.1. **Abstract**

Pigeons chose between two options on a concurrent-chains task with a single response requirement in the initial link. The suboptimal option ended with food 20% of the time whereas the optimal option ended with food 80% of the time. During a Sig-Both condition, terminal-link stimuli on both options signalled whether or not food would occur. During a Sig-Sub condition, terminal-link stimuli on the suboptimal option provided differential signals, but stimuli on the optimal option did not differentially signal the food and no food outcomes. Initial-link choices revealed a clear preference for the optimal option in the Sig-Both condition, but preference shifted toward suboptimality in the Sig-Sub condition. These findings show that pigeon suboptimal choice is not singularly driven by signal value, as has been suggested, but also by reinforcer frequency.

2.2. **Introduction**

Under certain circumstances, pigeons, humans, and other animals will deviate from optimal choice behaviour (e.g., Dunn & Spetch, 1990; Molet et al., 2012; Ferster & Skinner, 1957; Vasconcelos et al., 2015). In one such procedure, developed by Kendall (1974), pigeons sometimes chose an alternative that provided food only half of the time over an alternative that always provided food (for reviews see McDevitt, Dunn, Spetch, & Ludvig, 2016; Zentall, 2016). These *suboptimal choices* are most likely to occur when food is delayed and differential signals for the outcomes occupy that delay. In a more recent demonstration of suboptimal choice (Stagner, Laude, & Zentall, 2012, phase 3), pigeons chose between two alternatives that sometimes provided food after a 10-s delay. In their "discrimination-one" group, choosing the optimal alternative led to food half of the time, but the outcome was not predictable (i.e., the stimuli present during the delay did not signal whether or not food would occur at the end of the delay). Choosing the other, suboptimal, alternative led to food only 20% of the time, but the delivery of food was predictable (i.e., after the choice was made, differential stimuli signalled whether that trial would end in food or not). Pigeons in this group showed strong preference for the suboptimal alternative, even though it provided less food overall.

A variation of this procedure, however, showed that pigeons can be sensitive to the overall food contingencies and can demonstrate an optimal pattern of preference. Pigeons in a second group ("nondiscrimination") in the Stagner et al. (2012) study were presented with the same task described above but with one important change: the outcome of the trial was no longer signalled on the suboptimal alternative. Even though the overall probabilities of food associated with the alternatives were identical for the two groups, the birds in the nondiscrimination group showed a strong preference for the optimal alternative, opposite to the preference shown in the discrimination-one group (see Stagner et al., 2012, Figure 2). Other studies using both between-subject designs (e.g., Smith, Beckmann, & Zentall, 2017) and within-subject designs (e.g., Dunn & Spetch, 1990; Spetch, Belke, Barnet, Dunn, & Pierce, 1990) have also demonstrated similarly divergent results in signalled and unsignalled conditions (for reviews see McDevitt et al., 2016; Zentall, 2016).

The key finding, that preference switches from an optimal pattern to a suboptimal pattern when outcomes on only the suboptimal alternative are signalled, has been central to attempts to find a theoretical explanation of the phenomenon. One prominent account, referred to here as the *signal-value hypothesis*, has been proposed by Zentall and his colleagues based on observations that the probability of food, and thus the frequency of signals for food, on a signalled alternative do not appear to affect the level of suboptimal preference (e.g., Zentall, 2016; Smith & Zentall, 2016). This finding is best illustrated by another condition in Stagner et al. which presented pigeons with a choice between a suboptimal alternative that provided food 20% of the time and an optimal alternative that provided food 50% of the time. In a condition called "discrimination both", both alternatives provided stimuli during the 10-s delays that differentially signalled whether or not food would be presented on that trial. In this condition, subjects were indifferent between the two alternatives even though they provided different rates of reinforcement, suggesting that "...the value of each alternative is determined by the predictability of reinforcement of the S^+ stimulus [the stimulus correlated with food] associated with that alternative and not by the frequency of its appearance or by the frequency of the S^+ stimulus that appeared on other trials" Stagner et al., p. 450. In addition, the value of the alternatives in this procedure does not appear to be influenced by the relative frequency of the delayed primary rein-

forcement. Other evidence that appears to be consistent with their proposal that signal value is all that matters was provided by Smith and Zentall (2016) who found indifference between two alternatives that differentially signalled food 50% and 100% of the time respectively. However, indifference has not been found in all studies (e.g., Belke & Spetch, 1994; Kendall, 1974; Gipson, Alessandri, Miller, & Zentall, 2009; Pisklak et al., 2015; Smith, Bailey, Chow, Beckmann, & Zentall, 2016).

Another recent model, based on considerations of foraging behaviour in the natural environment, similarly predicts that reinforcement frequency on a signalled alternative should have no effect on preference (Fortes, Vasconcelos, & Machado, 2016; Fortes, Machado, & Vasconcelos, 2017; Vasconcelos, Machado, & Pandeirada, 2018). One formalization of this functional model, called the *reinforcement rate model* (RRM) focuses on the rate of energy intake in the presence of informative (i.e., signalled) stimuli and non-informative (unsignalled) stimuli (Fortes et al., 2017). Animals are expected to disengage from stimuli that signal the absence of food and, as a result, the probability of receiving these stimuli does not affect choice. Consistent with these predictions, studies with starlings (Vasconcelos et al., 2015) and pigeons (Fortes et al., 2016) have found that preference for a signalled alternative over a more optimal unsignalled alternative remains strong despite decreases in reward probability on the signalled alternative, and this suboptimal preference decreases only at very low probabilities (e.g. < 0.1).

The evidence that pigeons and starlings are not sensitive to reinforcement frequency in the signalled suboptimal procedure appears to be at odds with large bodies of related work indicating that both primary and conditioned reinforcement frequencies do matter. For instance, many experiments have found that the relative frequency of primary reinforcement, by itself, is a powerful determinant of choice for concurrently available reinforcement schedules (see de Villiers, 1977). In addition, studies that have evaluated conditioned reinforcement using chain schedules have repeatedly demonstrated the importance of the frequency of conditioned reinforcers (see Kelleher & Gollub, 1962), a finding that extends to tests of choice using concurrently available chained schedules as well (Williams & Dunn, 1991). It is therefore of considerable interest that the relative frequencies of primary and conditioned reinforcement had no obvious effect in the suboptimal procedure reported by Stagner et al.

(2012) or subsequent studies (Fortes et al., 2016; Vasconcelos et al., 2015).

The present study was designed to test the proposal that relative frequency of reinforcement does not influence suboptimal choice by providing pigeons with a choice between two signalled alternatives that differ substantially in reward probability. According to both the signal-value hypothesis and the functional RRM, reinforcement frequency does not matter because the occurrence of signalled non-reinforced outcomes do not impact the value of the alternative. Consequently, a signalled alternative providing food with a low probability is equivalent to a signalled alternative providing food with a high probability because, for both, the signalled delays to no food are not counting against the value. Preference for the alternative providing the higher relative frequency of reinforcement would limit the conclusions of the signal-value hypothesis (Stagner et al., 2012) and RRM (e.g., Fortes et al., 2017). In the current study, pigeons were provided a suboptimal alternative ending with a 20% chance of food delivery and an optimal alternative ending with an 80% chance of food delivery. The 80% probability of food on the optimal alternative, in contrast to the 50% probability used by Stagner et al., was chosen to test the reasonable limits of predictions made by the signal-value hypothesis and RRM. If under conditions of a 60% difference in reinforcement frequency, pigeons are still indifferent when both alternatives are differentially signalled, it would be powerful evidence in favor of both those hypotheses. However, if frequency does matter, then pigeons should prefer the optimal alternative. Using a procedure similar to that employed by Stagner et al., both the suboptimal and optimal alternatives in a *Sig-Both* condition provided differentially signalled outcomes. In the *Sig-Sub* condition, which acted as a control, only the suboptimal alternative provided differentially signalled outcomes. If the frequency of reinforcement does not matter when all outcomes are signalled, pigeons should be indifferent between the alternatives in the Sig-Both condition, and they should show strong suboptimal choice in the Sig-Sub condition. To further aid correspondence between the present study and that of Stagner et al., unreinforced probe trials were included in which the food signals from each alternative were presented together following the Sig-Both condition. Stagner et al. included similar probe trials as a way to examine the "... relation between preference for the initial link alternative and preference for the S^+ ..." (p. 449).

2.3. **Method**

2.3.1.   *Subjects*   The subjects were ten adult ex-racing pigeons with prior experimental histories. They were maintained at approximately 85% of their free-feeding weights by grain obtained during experimental sessions and post-session feedings when necessary. The birds were housed in individual cages under a 12-hour light/dark cycle, with water and grit freely available. One pigeon (bird 5) died following the first phase of the experiment and was replaced by another (bird 822) for the second phase.

2.3.2.   *Apparatus*   Two operant chambers (approximately 360 mm wide, 320 mm long, and 350 mm high) were used. Three translucent response keys, 25 mm in diameter, were mounted on the front intelligence panel 260 mm above the floor and 72.5 mm apart. Each key required a force of approximately 0.15 N to operate and could be illuminated from the rear by standard IEE 28-V 12-stimulus projectors. A 28-V 1-W miniature lamp, located 87.5 mm above the center response key, provided general chamber illumination. This lamp was illuminated during the entire session, except during blackouts, during which it was turned off. Directly below the center key and 95 mm above the floor was an opening (57 mm high by 50 mm wide) that provided access to a solenoid-operated grain hopper. When the food hopper was activated, mixed grain was illuminated from above with white light by a 28-V 1-W miniature lamp. All blackouts and food outcomes lasted for 5 sec. A computer and a MED-PC interface, located in an adjacent room, controlled experimental events.

2.3.3.   *Procedure.* Pre-training. Prior to beginning the experimental procedure, each bird received pre-training for 2-3 sessions during which keypecks to individually illuminated response keys were reinforced with mixed grain according to a fixed-ratio (FR) schedule. The FR schedule was gradually increased from an FR 1 to at least an FR 20.

General Procedure. All pigeons were presented with a choice between a suboptimal alternative that ended with food 20% of the time and an optimal alternative that ended with food 80% of the time. Five birds started the experiment in the Sig-Both condition, and four birds started in the Sig-Sub condition. The first phase (Phase 1) consisted of 18 training sessions. Immediately following this phase, birds in the Sig-Both condition received 5 probe

sessions. In the second phase (Phase 2), the conditions were swapped between the two groups such that birds previously receiving the Sig-Both condition now received the Sig-Sub condition, and vice versa. Phase 2 consisted of 25 training sessions and was immediately followed by 5 probe sessions for birds in the Sig-Both condition. Sessions were conducted five days per week at approximately the same time each day. Undergraduate students conducted the experiment as part of a research course, and each pigeon was handled by up to two students.

Sig-Both Condition. Choice trials presented two initial-link stimuli – representing a suboptimal and an optimal alternative – both consisting of a hollow black circle stimulus on an illuminated side key. The side each alternative appeared on was randomly balanced as close as possible across subjects such that 5 birds received the suboptimal alternative on the left side and 4 received it on the right side. Bird 822 assumed all the same contingencies and assignments of Bird 5. The alternatives and their respective outcomes are diagramed in the top panel of Figure 2.1.
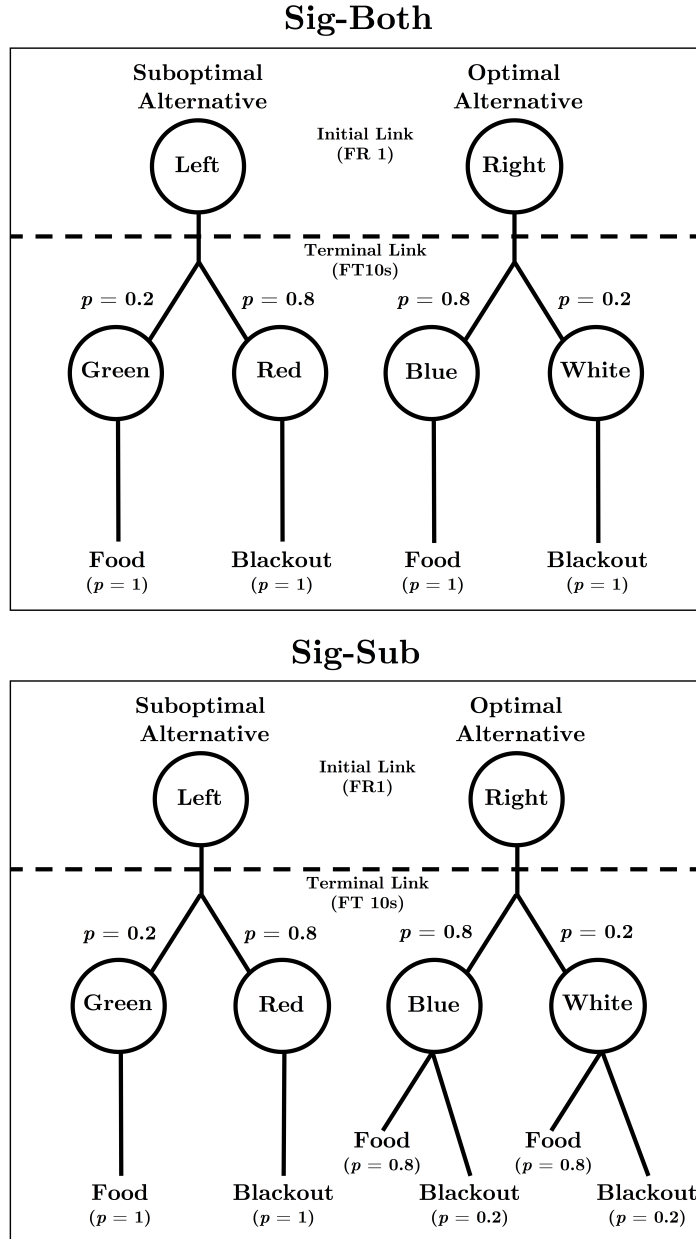
Sig-Both

Sig-Sub

Figure 2.1: A schematic displaying an example of the contingencies in effect for the Sig-Both condition (top panel) and the Sig-Sub condition (bottom panel).

A choice response consisted of a single peck to one of the side keys, which initiated the terminal link. At the start of the terminal link, the chosen key changed to one of two colours and the other key was darkened and inoperative. The terminal link lasted for 10 s, with the key colour signaling the outcome of the trial. For the suboptimal alternative, one colour (e.g., green) was presented on 20% of the trials, and always ended with food. The other colour (e.g., red) was presented on 80% of the trials, and always ended with a blackout. For the optimal alternative, one colour (e.g., blue) was presented on 80% of the trials and always ended with food. The other colour (e.g., white) was presented on 20% of the trials, and always ended with blackout. Thus, the suboptimal alternative provided food 20% of the time, the optimal alternative provided food 80% of the time, and both alternatives provided differential signals corresponding to the outcome of the trial. Terminal-link colour-side associations were the same for all birds.

Each training session began with 20 forced-exposure (FE) trials in which only one alternative was presented on each trial. The order of the presentation was semi-randomly determined, with the constraint that each alternative was presented a total of 10 times and a particular alternative could be presented a maximum of two times in a row. Following the forced-exposure trials, choice trials were presented for the remainder of the session. A 30-s intertrial interval separated each trial. Sessions terminated after 40 min. The mean (SD) number of choice trials per session in Phase 1 and Phase 2 for the Sig-Both condition was 46.50 (15.23) and 56.40 (11.24) respectively.

Probe Sessions. Probe sessions were the same as regular training sessions but with the addition of probe trials. Each probe trial consisted of the simultaneous presentation of the terminal-link stimulus that signalled food on the suboptimal alternative (e.g., green) and the terminal-link stimulus that signalled food on the optimal alternative (e.g., blue) for 10 s. These stimuli were presented on the same side as they had been as terminal-link stimuli during choice trials. Responses were recorded, but no outcome was delivered on probe trials. A probe trial was scheduled to occur immediately after the FE trials were completed (i.e., on trial 21), and then occurred after every fifth regular choice trial, until a maximum of five probe trials were presented. A total of five probe sessions were conducted.

Sig-Sub Condition. The Sig-Sub condition was the same as the Sig-Both condition, except that the terminal-link stimuli presented on the optimal alternative could precede either outcome type. That is, both colours (e.g., blue and white) terminated with food 80% of the time and blackout 20% of the time (see bottom panel of Figure 2.1). One of those colours (e.g., blue), was programmed to appear 80% of the time the optimal alternative was selected. The other colour (white) appeared on the remaining 20% of optimal alternative terminal links. The presentation order of the two colours was randomized. The mean (SD) number of choice trials per session in Phase 1 and Phase 2 of the Sig-Sub condition was 64.19 (8.30) and 58.36 (14.94) respectively. There were no probe sessions in the Sig-Sub condition.

Data Analysis. All statistical analyses were conducted using R 3.5.1 (R Core Team, 2018). Linear mixed-effects modelling, fit by maximum likelihood with subjects treated as a random effect, was conducted to examine the influence of condition, phase, and their interaction on the mean proportion of choices made to the suboptimal alternative during the last four sessions of each phase. This was calculated using the nlme R package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2018). An effect size in terms of r^2 is reported for each fixed-effect in the full model. Bayes Factors (BF_{10}) for each effect are also provided and tested against a null model that included only the intercept and an additive random effect of subject.

For probe trial assessment, the mean rate of responding to each probe stimulus was calculated across all probe sessions. Differences in the obtained rates were analyzed using a paired t-test. A JZS Bayes factor (BF_{10}) with a medium prior was computed to obtain the relative odds in favor of the alternative hypothesis against the null and was calculated using the BayesFactor R package (Morey & Rouder, 2015). An r^2 value is reported as effect size. For all tests, significant effects are defined as $p \leq .05$.

2.4. **Results**

Figure 2.2 displays the mean proportion of responding to the suboptimal alternative across the Sig-Both and Sig-Sub conditions. Choice of the suboptimal alternative was significantly below chance (i.e. the birds chose the optimal alternative more often) in the Sig-Both condition ($M = 0.12$,

95% CI $[0.03, 0.22]$). This is contrasted with the results of the Sig-Sub condition which showed considerably higher overall suboptimal responding ($M = 0.42, [0.15, 0.68]$). Results of the linear mixed-effect modelling revealed a significant main effect of condition, $\chi^2(1) = 6.89, p = .009, r^2 = .50, BF_{10} = 7.38$. Figure 2.3 displays the mean proportion of suboptimal choices as a function of session across both Phase 1 and 2. Visual inspection suggests a small increase in suboptimal preference for the Sig-Both condition and a small decrease in suboptimal preference for the Sig-Sub condition across the two phases. While no significant main effect of phase emerged, the obtained Bayes Factor does suggest at least some weak support for it, $\chi^2(1) = 0.02, p = .894, r^2 = .05, BF_{10} = 1.75$. However, there was no indication of a significant interaction between condition and phase, $\chi^2(1) = 0.64, p = .425, r^2 = .09, BF_{10} = 0.57$.
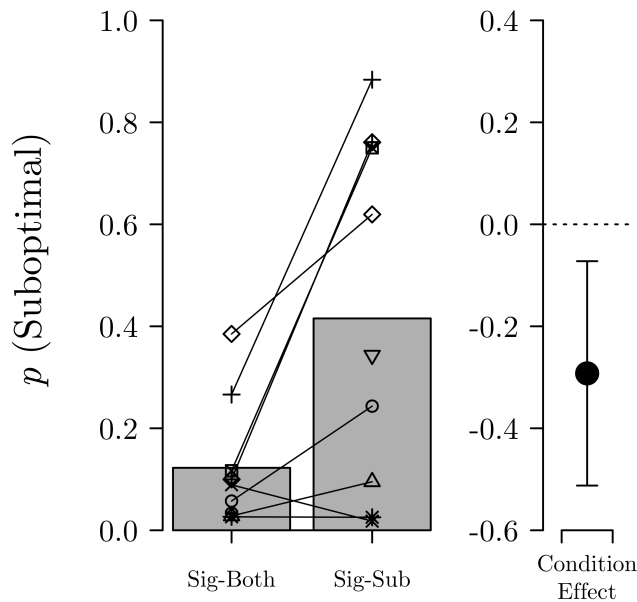
Figure 2.2: Bars on the left plot depict the mean proportion of suboptimal choices by condition with corresponding individual pigeon values plotted as distinct shapes. The right side depicts the approximate 95% confidence interval for a fixed-effect of condition. Data for birds 5 and 822 are included.
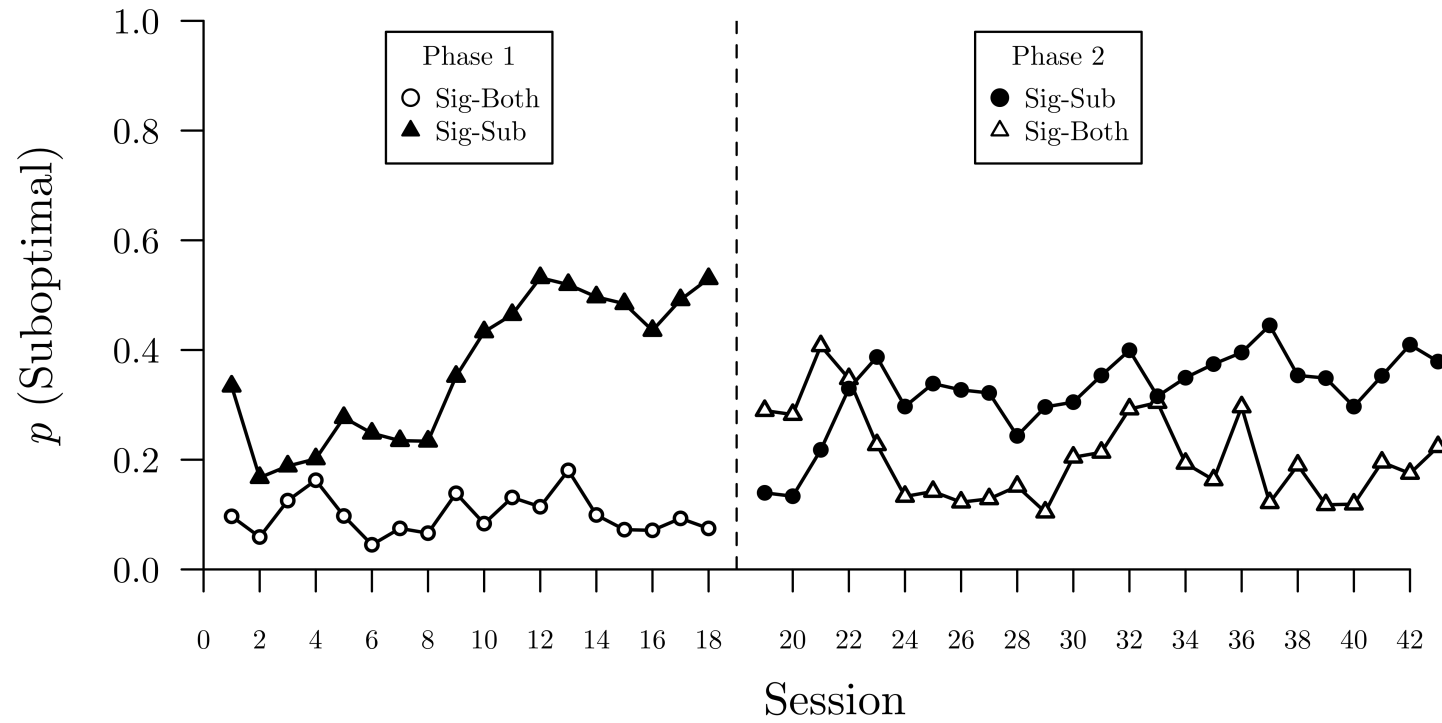
Figure 2.3: Mean proportion of suboptimal choices as a function of session across Phase 1 and 2. The Sig-Both conditions are indicated by empty markers and the Sig-Sub conditions are indicated by filled markers. Circles and triangles designate the two groups of subjects. The dashed vertical line indicates the point at which the Phase 1 changed to Phase 2. Data for birds 5 and 822 are included.

For probe trials, the mean rate of responding was significantly less on the terminal-link stimulus that signalled food on the suboptimal alternative ($M = 0.24[0.06, 0.42]$) than on the terminal-link stimulus that signalled food on the optimal alternative ($M = 1.10[0.58, 1.62]$), $t(8) = -3.34$, $p = .010$, $r^2 = .58$, $BF_{10} = 6.22$. Individual pigeon choice proportions and response rates for initial-link, terminal-link, and probe trial stimuli are listed in Table 2.1.

Table 2.1
Individual pigeon choice proportions and response rates

| | | Response Rates (r/s) | | | |
| | | FE Terminal Links | | Probe Trials | |
Pigeon No.	Prop. Suboptimal Initial Link Responses	20% S+	80% S+	20% S+	80% S+
Sig-Both					
4	0.27	2.64	2.43	0.01	2.05
6	0.12	0.80	1.11	0.12	0.97
34	0.03	0.68	0.87	0.24	0.26
42	0.09	2.64	1.00	0.06	0.45
48	0.39	1.74	1.57	0.68	0.68
714	0.03	1.66	1.40	0.00	1.05
799	0.10	1.39	2.27	0.46	0.95
822	0.03	1.91	2.98	0.16	2.26
2165	0.06	2.20	1.84	0.43	1.23
Mean (SD)	0.12 (0.12)	1.74 (0.71)	1.72 (0.72)	0.24 (0.23)	1.10 (0.67)
Sig-Sub					
4	0.88	3.00	2.54		
5	0.34	2.02	1.92		
6	0.75	0.62	1.13		
34	0.10	1.11	0.56		
42	0.02	2.06	0.37		
48	0.62	1.61	1.57		
714	0.03	1.88	0.77		
799	0.76	0.79	1.17		
2165	0.24	2.27	2.05		
Mean (SD)	0.42 (0.34)	1.71 (0.76)	1.34 (0.73)		

Note. FE = Forced Exposure. Initial and Terminal Link data are averaged over the last four sessions of each condition.

2.5. Discussion

The results from both the Sig-Both and the Sig-Sub conditions indicate that preference is not solely determined by signal-value but also by relative frequency of the reinforcers. In the Sig-Both condition, pigeons strongly preferred the optimal alternative. That is, when given a choice between two alternatives that both provide unambiguous food signals, the pigeons preferred the alternative that provided food, and signals for that food, at a higher frequency. When unambiguous food signals were provided on only the lower frequency (20%) alternative, as in the Sig-Sub condition, preference for the higher frequency alternative was attenuated significantly although preference did not become strongly suboptimal. This was the case irrespective of condition order. These results are inconsistent with predictions of both the signal-value hypothesis (e.g., Stagner et al., 2012) and the recent models inspired by optimal foraging theory (Fortes et al., 2017; Vasconcelos et al., 2018).

Probe trials with the S^+ terminal-link stimuli of the Sig-Both condition revealed strong preference for the stimulus correlated with the higher frequency of reinforcement. Although this preference contrasts with the indifference on probe trials seen by Stagner et al. (2012), the probe trial choices in both studies paralleled the initial-link choices. That is, pigeons in Stagner et al. showed indifference in both cases, whereas the present study showed preference for the optimal alternative in both cases. Interestingly, the probe trial results stand in contrast to the response rates on the terminal-link stimuli (see Table 2.1), which were almost identical for the suboptimal S^+ ($M = 1.74, [1.20, 2.28]$) and optimal S^+ ($M = 1.72, [1.16, 2.27]$) stimuli. One possible reason for the discrepancy between probe trial and terminal-link response rates is that when probes appear within a set of choice trials, the pigeons may generalize their response pattern from the choice phase of the regular training trials. We also question whether probe trials are informative in understanding the relation between initial- and terminal-link stimuli, as suggested by Stagner et al., because conditioned reinforcement is defined in terms of a stimulus' effect on the behaviour it follows. Therefore, the conditioned reinforcing function of the terminal links can only be determined on the basis of behaviour during the initial link of a concurrent-chains paradigm.

According to the signal-value hypothesis (e.g., Stagner et al., 2012), dif-

ferences in reinforcement rate should be of no consequence when outcomes are differentially signalled. Pigeons' choices should be predictable solely on the basis of the signal value and not the rate of primary or conditioned reinforcement. The results of the present study strongly suggest that is not the case. When outcomes were signalled on both alternatives, as in the Sig-Both condition, preference was decidedly optimal. In the Sig-Sub condition, when only the suboptimal alternative produced unambiguous signals, preference became more suboptimal, but did not show clear suboptimality as the signal-value hypothesis would predict.

These results stand in contrast to those obtained by (Stagner et al., 2012), who found indifference between signalled 20% and signalled 50% alternatives. The purpose of our study, however, was not replicate or falsify Stagner et al.'s results; rather, it was to test the extension of their results and temper their conclusions by demonstrating that the frequency of the signals for primary reinforcement can exert an influence in suboptimal choice paradigms. This is clearly shown by the strong preference for the optimal alternative in the Sig-Both condition because the alternatives were equal in signal value according to their model but differed in signal frequency. While a parametric analysis of different frequencies would be informative, the more salient point – which has hitherto been rejected by some and is the current focus of this paper – is that the frequency of the signals do matter.

In addition to the signal-value hypothesis, predictions can also be derived from the hyperbolic decay model (HDM), which also relies heavily on the predictive value of signaling stimuli and has been influential in understanding conditioned reinforcement (Mazur, 1997). The HDM assumes that the crucial determinant of the ability of a stimulus to function as a conditional reinforcer is whether it signals a relatively short delay to primary reinforcement. In the Sig-Both condition, the food signals on both alternatives were the same duration with only a small amount of additional delay accruing on the suboptimal alternative because of the multiple times through the initial-link period per food outcome. Thus, the HDM predicts approximately equal value with a slight bias towards optimality. In the Sig-Sub condition, the HDM predicts stronger preference for the suboptimal alternative because both terminal-link stimuli on the optimal alternative signal possible reinforcement and therefore the terminal links ending without food are incorporated into the time to rein-

forcement signalled by the stimuli. Accruing this delay devalues the optimal alternative. While neither prediction of the HDM appears strongly supported in the present study, its ordinal predictions are at least consistent with the present set of results. It is worth noting that the exact manner in which value translates into preference assessment on choice tasks remains an open question.

While most explanations of suboptimal choice have been framed in terms of operant principles of conditioned reinforcement, explanations can also be derived from ecological foraging theories (e.g., Vasconcelos et al., 2018). From this perceptive, so-called "suboptimal" behaviour arises from artificial constraints of the experimental environment. In a natural environment, encountering an extinction stimulus (i.e., a no-food signal) provides an opportunity to disengage from that stimulus and forage for food elsewhere. That opportunity is not afforded in most laboratory paradigms and, therefore, what seems suboptimal in the laboratory is not likely suboptimal in the real world. A recent attempt to formally quantify this functional approach, called the reinforcement rate model (RRM), focuses on the rate of energy intake in the presence of informative (i.e., signalled) stimuli and non-informative (unsignalled) stimuli (Fortes et al., 2017). While the RRM has met with success (Fortes et al., 2016, 2017; Vasconcelos et al., 2015), the present results provide a challenge to RRM, which predicts indifference in the Sig-Both condition because both alternatives are informative and thus functionally equivalent. Therefore, like the signal-value hypothesis, RRM predicts indifference in the Sig-Both condition and strong suboptimal preference in the Sig-Sub condition, neither of which occurred.

Another account of suboptimal choice is provided by the signals for good news (SiGN) hypothesis, which has theoretical roots in delay-reduction theory (Dunn & Spetch, 1990; Fantino, 1969). According to this hypothesis, preference is influenced by the primary reinforcement as well as the conditioned reinforcement provided by the stimuli presented during the delay to primary reinforcement. A central feature of the SiGN hypothesis is that the choice response functions as a reference point for the influence of terminal-link stimuli. When food and no-food outcomes occur probabilistically, the outcome of the choice response is uncertain. Terminal-link stimuli that signal a food outcome (i.e., good news) in the context of uncertainty create a reduction in overall delay to food relative to that signalled by the choice response alone.

Therefore, when the suboptimal alternative has differentially signalled terminal links, a conditioned reinforcer sometimes follows choice of that alternative. When the outcome of the choice response is certain (i.e., 100% reinforcement) or unsignalled, the terminal-link stimuli provide no reduction in the delay signalled by the choice response itself and therefore do not function as conditioned reinforcers. Thus, the uncertainty created by probabilistic outcomes gives conditioned reinforcement a critical role in the SiGN hypothesis since a higher frequency of immediate conditioned reinforcement on the suboptimal alternative can counter the higher relative frequency of delayed primary reinforcement on the optimal alternative.

Predictions based on the SiGN hypothesis are less specific than the signal-value hypothesis because SiGN postulates a combination of conditioned and primary reinforcement processes without a formal way to separate the strength of these sources of reinforcement. In the Sig-Sub condition, more conditioned reinforcement occurs on the suboptimal alternative because only it provides a signal for a reduction in the delay to food relative to the choice peck. By contrast, in the Sig-Both condition, the suboptimal and optimal alternatives each signal reductions. When the role of primary reinforcement is incorporated, predictions become less specific but the directionality of the predictions is still straightforward. Primary reinforcement will bias responding towards the alternative with the higher frequency. The extent of this bias will necessarily be further contingent on the duration of the terminal links as well as the magnitude of the primary reinforcement.

The SiGN hypothesis has been criticized for including primary reinforcement as an explanatory variable because the balance of the two sources of reinforcement (primary and conditioned) makes the hypothesis overly flexible in accounting for suboptimal choice data (Zentall, 2017). Nevertheless, the assumption that primary and conditioned reinforcement processes jointly determine the value of a choice alternative is consistent with neural evidence in monkeys (Blanchard, Hayden, & Bromberg-Martin, 2015), suggesting that predictive cues and primary reward are represented independently in the orbital frontal cortex and then combined into a single value scale in downstream areas. Moreover, the balance of the two processes is supported by multiple experiments that manipulate initial-link duration, terminal-link duration, and the immediacy of terminal-link stimuli (see Table 2 in McDevitt et al., 2016,

for a summary of results). For example, as the duration of the terminal links increases, the primary reinforcers on the two alternatives are necessarily devalued giving conditioned reinforcement greater control over choice responding and preference has been shown to follow accordingly (e.g., McDevitt, Pisklak, Spetch, & Dunn, 2018). Other work has shown that the conditioned reinforcement can also be selectively devalued, transferring control over to the alternative providing greater primary reinforcement (McDevitt, Spetch, & Dunn, 1997).

More recently, a temporal information-theoretic approach which frames conditioned reinforcement effects within a larger theory of Pavlovian conditioning has been proposed (Cunningham & Shahan, 2018). In suboptimal choice paradigms, conditioned reinforcement is predicted when a stimulus is "temporally informative", meaning it signals a larger reduction in time to primary reinforcement relative to the average time to primary reinforcement. In the language of information theory, a conditioned reinforcer provides *information* that *reduces uncertainty* about when primary reinforcement will occur. Thus, the information-theoretic approach is conceptually similar to both delay-reduction theory and the SiGN hypothesis, but is rooted in information theory and provides more formalized mathematical interpretations regarding suboptimal choice paradigms. Of critical note is its prediction with respect to the present Sig-Both condition. According to the information-theoretic approach, the amount of temporal information conveyed by an alternative (suboptimal or optimal) is determined by the average delay to food signalled by a terminal-link stimulus that ends in food, meaning that extinction signals are not treated as information by the model. Thus, in the case of the Sig-Both condition, the average time to food for the two alternatives is equivalent (see Equation 7 in Cunningham & Shahan, 2018), conveying the same amount of temporal information. Consequently, choice behaviour in the Sig-Both condition, when driven solely by the temporal information provided by the stimuli, should be distributed equally to the alternatives. However, like the SiGN hypothesis, the information-theoretic approach also considers the influence of primary reinforcement. Specifically, the tendency of organisms to match relative rates of primary reinforcement across alternatives also influences choice. If it is assumed that choice becomes exclusively determined by relative reinforcement rate when alternatives are equivalently signalled (as in the Sig-Both condition),

then basic matching provides an adequate approximation of the data. In the Sig-Sub condition, temporal informativeness favors the suboptimal alternative because only the suboptimal alternative is signalled. Assuming preference is weighted exclusively in favor of the temporal informativeness of the signals (i.e., not primary reinforcement rate) and no adjustment for hypersensitivity of suboptimal choice to relative signal value is made, a level of suboptimal choice slightly higher than indifference is predicted, which is consistent with the results of the present study.

As stated in the introduction, the proposition that responding in suboptimal choice procedures is due solely to the predictive value of those signals is counterintuitive. There is clearly adaptive significance to the influence of predictive signals, even to the point of generating suboptimality (Vasconcelos et al., 2018); however, it seems reasonable to assume that restrictions on this suboptimality exist. Our study demonstrates one such restriction. In the face of large differences of primary reinforcer frequency (i.e., 20% vs. 80%), the predictive value of signals appears insufficient to drive strong suboptimal choice. Relative frequency of reinforcement has long been shown to be a key determinant of preference across a variety of procedures (see de Villiers, 1977; Kelleher & Gollub, 1962, for summaries of the effects of primary and conditioned reinforcement respectively). The results presented here reaffirm the role of relative frequency of reinforcement in the suboptimal procedure.

2.6. **Author Note**

Jeffrey M. Pisklak, Department of Psychology, University of Alberta; Margaret A. McDevitt, Department of Psychology, McDaniel College; Roger M. Dunn, Department of Psychology, San Diego State University; Marcia L. Spetch, Department of Psychology, University of Alberta.

Chapter 3   Suboptimal Choice and Initial-Link Requirement

3.1. **Abstract**

Pigeons (n=14) were trained in a concurrent-chains suboptimal choice procedure that tested the effect of an increased ratio requirement in the initial links. Fixed-ratio 1 and 25 conditions were manipulated within subjects in a counter-balanced order. In all conditions, distinct terminal-link stimuli on a suboptimal alternative signalled either primary reinforcement (20% of the time) or extinction (80% of the time). On an optimal alternative, two distinct terminal-link stimuli each signalled a 50% chance of primary reinforcement. Preference for the suboptimal alternative was significantly attenuated, and in some individuals completely reversed, by the larger response requirement irrespective of condition order. This larger response requirement also generated a notable increase in between-subject variability. A measure of cumulative choice responding is introduced to mitigate the problems associated with traditional session averages. Basic ordinal predictions of some current theories of suboptimal choice are also considered in light of the results.

3.2. **Introduction**

In a now classic paradigm of choice behaviour, pigeons are given a choice between two fixed-ratio (FR) 1 initial links of concurrent chains. One initial link, on the "suboptimal" alternative, leads to delayed primary reinforcement probabilistically (e.g., a 50% chance of food), while a concurrently available "optimal" alternative provides a similarly delayed reinforcer on every trial (100% chance of food). Critically, during the delay, a terminal-link stimulus differentially signals the outcome to follow (i.e., the occurrence of reinforcement or extinction). Counterintuitively, under these conditions many pigeons will choose the suboptimal alternative as often or even more often than the optimal alternative (Dunn & Spetch, 1990; Kendall, 1974, 1985; McDevitt et al., 1997; Spetch, Mondloch, Belke, & Dunn, 1994), even though choosing the suboptimal alternative costs the pigeons in terms of overall rate of obtaining food in a session, or in the total number of food reinforcers obtained in an experiment (e.g., Pisklak et al., 2015). More contemporary variants of this procedure have found that, if the optimal alternative provides non-differential signals for 50% reinforcement, then preference for a signalled suboptimal alternative that provides a lower probability of reinforcement is even more extreme (e.g.,

Stagner et al., 2012; Stagner & Zentall, 2010).

One of the earliest explanations of this effect is that preference for the suboptimal alternative in these procedures is driven by a conditioned reinforcing function of the terminal-link stimuli. The degree to which a terminal-link stimulus functions as a conditioned reinforcer appears to depend on the degree to which it signals an improvement in the delay to primary reinforcement, relative to the delay signalled by the choice response. First presented by Dunn and Spetch (1990), this perspective was inspired by delay reduction theory (Fantino, 1977) - a prominent theory of conditioned reinforcement – and is now formally referred to as the signal for good news (SiGN) hypothesis (McDevitt et al., 2016). According to this view, the onset of the terminal-link stimulus is a conditioned reinforcer when it signals a reduction in wait time to food relative to the average wait time signalled by the initial-link response that produced it. The "good news" is the change in waiting time for food from the initial link to the terminal link that predicts food. When a single choice response is required in the initial links, the good news occurs on the suboptimal alternative when the terminal-link signal for food reduces the waiting time that was signalled by the choice response itself. According to SiGN, the signal for no food has no direct effect on choice; rather, its function is to provide a lean local context that enhances the "good news" signal. By contrast, when the optimal alternative provides a single terminal link with a guarantee of food, there is virtually no reduction in the delay to food signalled by the onset of the terminal link. Phrased differently, the terminal-link stimulus signals no meaningful improvement in wait time relative to the wait time signalled by the choice response itself. Similarly, when the optimal alternative's terminal link provides non-differential signals for chance levels of food, there is, again, no signalled reduction in waiting time moving from initial to terminal link. A hypothesis similar to the one originally proposed by Dunn and Spetch (1990) was put forward years later by Zentall (e.g., Stagner & Zentall, 2010), and neither have been formalized in purely quantitative terms. It is of historical note that the SiGN model was not posed as a general theory of choice; rather, it was formulated on the basic principles of delay reduction theory (DRT) to address conditioned reinforcement based on probabilistic primary reinforcement delivery. Without such modification, DRT is unable to account for the influence of differential signals in the suboptimal procedure established by Kendall (1974).

A different, yet conceptually similar, explanation to the SiGN model is found in the information-theoretic approach (Cunningham & Shahan, 2018). Under this framework, a terminal-link stimulus will function as a conditioned reinforcer if it conveys temporal information that reduces uncertainty about *when* the next primary reinforcer is likely to occur. "Information" here has a strict mathematical definition derived from information theory. Of critical note, in this approach temporal information is conveyed only by food-predictive signals. Signals for extinction are not considered a source of information, thus avoiding a difficulty in prior attempts at relating information theory and conditioned reinforcement (Shahan & Cunningham, 2015).

Another, often overlooked, framework is the hyperbolic value-added (HVA) model. This is a mathematical model that is based upon the logic of hyperbolic delay discounting and a delay reduction-like mechanism (Mazur, 2001). The model presumes that a reinforcer's value decreases according to a hyperbolic discounting function (e.g., Mazur, 1989), and that the terminal links of concurrent-chain schedules add additional value to that signalled by the initial-links. The difference between the value offered by the terminal links and the initial links is the "value added." The more value a terminal link adds, the more the alternative associated with it will be chosen.

An altogether different approach is a functional ecological hypothesis advanced on the basis of optimal foraging theory (Vasconcelos et al., 2018). From this perspective, selecting the suboptimal alternative is "suboptimal" only in the context of the laboratory paradigm described above. The natural-world analogue actually allows an animal to maximize its long-term rate of energy gain. This is because, if the animal selects the "suboptimal" alternative and receives a period of extinction, the animal is still able, in its natural environment, to retreat from that context and forage for food in other areas. The animal is not afforded this opportunity to retreat within the laboratory paradigm.

While there has been a recent wave of articles published on suboptimal choice (for reviews, see McDevitt et al., 2016; Vasconcelos et al., 2018; Zentall, 2016), most of the focus has been placed on the role of the terminal links in determining preference (e.g., Fortes et al., 2016; Fortes, Pinto, Machado, & Vasconcelos, 2018; McDevitt et al., 2018; Pisklak et al., 2015; Stagner et al., 2012; Stagner & Zentall, 2010). In most studies of suboptimal choice the initial-link schedule is a fixed-ratio (FR) 1. Only a small handful of early studies have

addressed the role of the initial links, and the common finding is that longer initial-link schedules tend to diminish suboptimal preference. For instance, in a seminal study, Kendall (1974) had pigeons choose between 50% and 100% reinforcement, both with FR 1 initial-link schedules (Experiment 1, N = 4) and with concurrent variable-interval (VI) 20 s schedules (Experiment 2, N = 4). When outcomes were correlated with the terminal-link stimuli (i.e., signalled), pigeons more often chose the 50% side in the condition with FR 1 schedules. Relative to Experiment 1, responding on the suboptimal alternative with correlated outcomes was notably diminished under the VI 20-s schedule used in Experiment 2. This basic finding was replicated in a later study comparing VI 20 and VI 60 schedules (Kendall, 1985). More systematically, Experiment 3 of Dunn and Spetch (1990) varied initial links within pigeons (n = 4) at values of FR 1, VI 10, VI 30, and VI 80. Despite large variability in the subjects' absolute preference, birds exhibited a notable decline of suboptimal choice as an increasing function of initial-link duration. More recently, Zentall, Andrews, and Case (2017) assessed the degree of suboptimal choice in two groups of pigeons. When the initial-link schedules were fixed-interval (FI) 1 s, the birds, overall, chose the suboptimal alternative more often than when they were FI 20 s.

In contrast to the studies described above in which the initial-link schedules have been manipulated simultaneously for both choice alternatives, Roper and Zentall (1999, Experiment 2) systematically increased the FR schedule for a signalled alternative while keeping the choice schedule at FR 1 for an unsignalled alternative. Each alternative in a pair provided the same overall rate of reinforcement, only differing in the whether or not the terminal-link stimuli signalled the reinforcers. The results showed, as expected, that choice of the signalled alternative decreased as the FR schedule on that alternative increased. Interestingly, the shift in preference varied depending on the overall probability of food. Preference shifted sooner, i.e., with relatively minimal increases in the FR schedule, when the probability of food on both alternatives was high (87.5%). When the probability of food was low (12.5%), preference was less affected by increases in the FR schedule, and subjects, on average, still preferred the signalled alternative even when its initial-link schedule was an FR 16. This finding is consistent with the idea that the influence of a signal for food is relative to its context (e.g., McDevitt et al., 2016; Stagner &

Zentall, 2010). When food is infrequent, the terminal-link stimulus associated with food should serve as a strong conditioned reinforcer because its presentation signals a dramatic increase in the probability of reinforcement. Thus, a signalled alternative is devalued less when its initial-link schedule is increased in a lean context, compared to the same manipulation in a rich context.

Overall, the results of manipulations of initial-link schedules are generally consistent with the SiGN model. When equivalent initial-link schedules are increased, terminal-link stimuli on both alternatives signal reductions in wait time and so the difference between the delay reduction signalled between the two alternatives decreases, leaving preference to be governed more by the remaining differences in primary reinforcement rates. Similarly, in the information theoretic approach, the relative difference in uncertainty reduction between the suboptimal and optimal alternatives shrinks slightly when initial-link durations are lengthened. Thus, more optimal responding is also predicted as initial-link duration increases, moving preference towards indifference. However, this only applies to instances where the terminal links on the optimal alternative are unsignalled (i.e., non-predictive), which is not the case in either the Kendall (1985) or Dunn and Spetch (1990) studies, or even in more recent work by Case and Zentall (2018). Under such conditions, the information theoretic approach, by itself, predicts only indifference (irrespective of initial-link length) and requires the inclusion of a matching relation that competes with temporal information - via a specific weighting function that formalizes the competition between conditioned and primary reinforcement (see Cunningham & Shahan, 2018) - to make predictions consistent with obtained findings. By contrast, the HVA model's principle of value addition makes ordinal predictions similar to SiGN without a strict need to appeal to an additional matching function on primary reinforcement. If one assumes that signalled periods of extinction and intertrial intervals (ITI) are not included in the calculation of value (see Mazur, 1997), and that each alternative's initial link has its own distinct value, then the amount of value added on each alternative by the terminal link grows as the initial-link durations increase. This growth makes the relative difference in value-addition between the choice alternatives shrink towards indifference, like the information-theoretic approach. In terms of an optimal foraging theory account, no formal analogue of initial-link duration/requirement seems to have been incorporated into the existing

functional model of suboptimal choice (see equations 2 and 3 of Vasconcelos et al., 2018). However, there does not appear to be, in principle, any reason it could not be incorporated as something equivalent to search time, for instance. A perhaps bigger problem is that, like the information theoretic approach, this specific optimal foraging account predicts indifference in situations where the optimal alternative provides a single informative terminal link, generating predictions at odds with much of the early suboptimal choice literature as well as some recent work which finds preference values higher than mere indifference (Case & Zentall, 2018; Dunn & Spetch, 1990; Kendall, 1974; Spetch et al., 1990). Moreover, unlike the information theoretic approach, it appears to have no additional means of accounting for competing primary reinforcement influences.

Related to notions of animal foraging is a theoretical model called the sequential choice model, or SCM (Kacelnik, Vasconcelos, Monteiro, & Aw, 2011). The SCM presupposes that the processes governing singular foraging situations determine how an animal will respond when faced with simultaneous foraging alternatives. While this model does not make explicit *a priori* predictions of choice on the basis of independent variables in suboptimal choice tasks like those discussed above, certain other predictions of the model are nevertheless interesting from the standpoint of understanding foraging behaviour. In particular, the SCM predicts that latencies observed during choice trials should be shorter than those observed during forced-exposure trials. The SCM also predicts that the latencies observed during sequential (forced-exposure) trials can be used to predict an animal's choices, with shorter latencies being the index of preference.

The present research was not designed to specifically evaluate the various models of suboptimal choice but rather to expand the empirical base on which the models can be evaluated or refined. One problem with the existing literature is that the conclusions about the role of initial links are based on only three studies with small samples and, in some cases, have concomitant manipulations that make a straightforward causal evaluation somewhat difficult. Further, these three studies all made use of the original procedure pioneered by Kendall (1974) that contains a single terminal link in the optimal alternative. Consequently, it is not clear if these conclusions would hold using more recent procedures that make the occurrence of primary reinforcement on the optimal

alternative completely unpredictable (50% chance) – a procedure that tends to generate very extreme suboptimal preference. The present study adapted such a procedure (used by Stagner & Zentall, 2010) which has been shown to produce high levels of suboptimal preference. To assess the impact of initial-link length on suboptimal choice, pigeons were tested, in a counter-balanced fashion, with FR 1 and FR 25 schedules in the initial links.

3.3.   **Method**

3.3.1.   *Subjects.* Subjects consisted of 14 retired adult racing pigeons (*Columba livia*), with extensive and varied learning histories, randomly selected from a University of Alberta colony room. Birds were housed inside $165 \times 69 \times 178$ cm flight cages containing approximately 7 birds per cage. The cages were located inside a temperature-controlled colony room, maintained at $20°$C, that ran on a 12-hour light-dark cycle with a 6:00 A.M. onset and 6:00 P.M offset. Free access to vitamin-enriched water and crushed oyster shell grit was provided inside each cage. All subjects were maintained at 80% of their free-feeding weight via post-experiment rations of Mazuri Gamebird food pellets (PMI Nutrition International LLC, Brentwood, MO). All procedures were approved by the University of Alberta Biological Sciences Animal Care and Use Committee, following the guidelines of the Canadian Council on Animal Care.

3.3.2.   *Apparatus and Stimuli.* Six custom built operant conditioning chambers for pigeons were used. Each chamber had an internal dimension of $29.2 \times 35.6 \times 38.1$ cm and was equipped with a 17 inch Viewsonic LCD monitor recessed into the back wall of the chamber. White noise was broadcast through each monitor's speakers to mask extraneous sounds from outside the experimental room. Sound pressure levels were equalized in each operant box at 65 dB using a A-weighting filter with a Brüel & Kjær Type 2239 Integrating Sound Level Meter. Each monitor was equipped with a Carrol Touch infrared touchscreen (Elo Touch Systems, Inc., Menlo Park, CA) to detect pecking by the pigeons. A white plastic barrier covered the lower third of the touchscreen to prevent incidental contact that might interfere with the collection of pecking responses (e.g., subject's feet and chest contacting the screen). On the adjoining right and left walls of each chamber was a 5.08×5.08 cm feeding port that

delivered food using a solenoid food hopper externally mounted to the chamber. Each food delivery was accompanied by illumination of the feeding port using a small incandescent light bulb housed within it. A Colbourne H20-94 photocell sensor was used to detect entry into each of the ports. Food delivery lasted for 3.5 s following entry of the food port. Only one hopper, left or right, was randomly selected to raise for each food delivery. Stimuli and responses were presented and logged using E-Prime® 2.0 Professional software.

Both initial- and terminal-link stimuli consisted of coloured circles, 25 mm in diameter, set against a grey background on each chamber's monitor. Circles appeared on either the left or right side of the screen 85 mm from the centre. Initial-link circles were filled white and had either a thin black vertical line or a thin black horizontal line within it. Terminal-link circles were filled either solid green, red, yellow, or blue. During periods where no initial- or terminal-link stimuli were present on the screen, empty black circles occupied their space as "inoperative keys."

3.3.3. *Procedure.*

Peck Training. All pigeons were first given a single 90 min session of an autoshaping procedure with an added FR 1 schedule of reinforcement. A circle, randomly-selected without-replacement from the initial- and terminal-link stimuli, was presented on the screen for 10 s or until a single response was detected. This was followed by the immediate presentation of food and then a 240 s ITI. Over the next three sessions, the ITI was shortened to 10 s and responses to the circles designated for terminal-link stimuli were gradually transitioned to a FI10s schedule of reinforcement. This transition occurred over the course of the first two of these sessions - progressing 1, 3, 6, and 10 s - with the third session consisting only of 10 s FI schedules. The reinforcement schedule for responses to the circles designated for initial-link stimuli remained at FR 1 across the three sessions. Each session had a 48 trial maximum and pigeons remained in the chamber for a full 90 min even upon completion of the 48 trials.

Testing Phase 1. Half the pigeons were assigned to an FR 1 condition and the other half assigned to an FR 25 condition. The procedure for both conditions is illustrated in (see Figure 3.1). For half the pigeons, the suboptimal initial-link stimulus was the vertical line and the optimal stimulus was

the horizontal line. For the remaining pigeons this was reversed. The side as-
signments for the initial links were balanced across pigeons, using fixed spatial
locations, as were the colours associated with their respective terminal links.
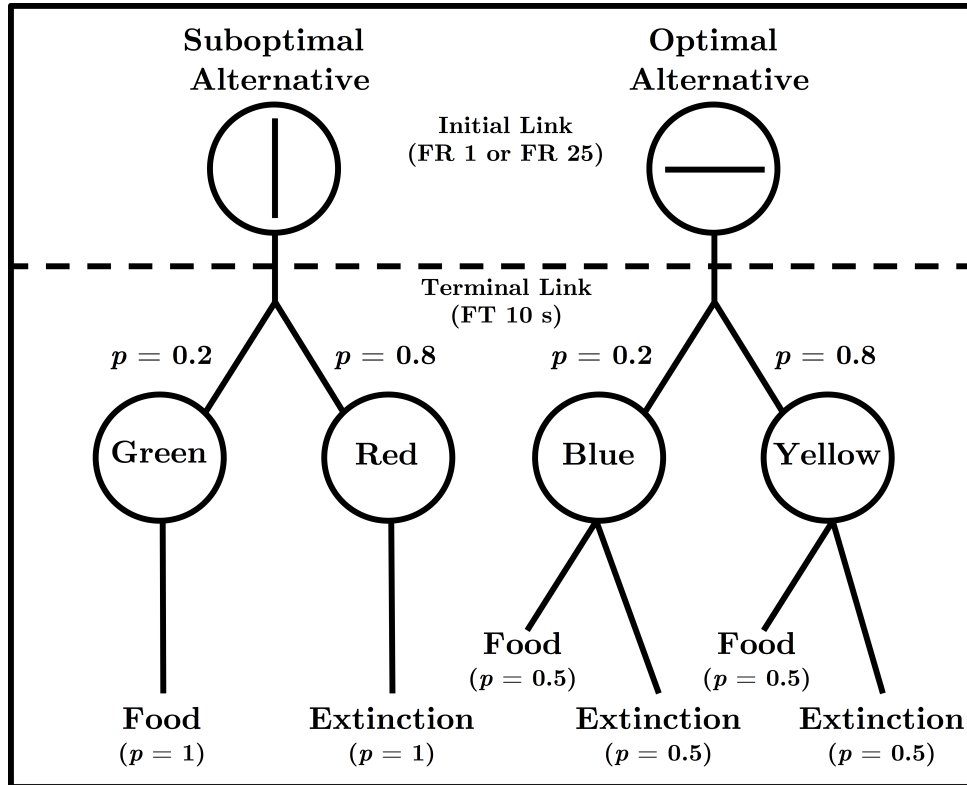
Figure 3.1: Schematic displaying an example of the contingencies in effect for
the FR1 and FR25 conditions

The initial-link stimuli remained on the screen until the bird completed the
FR requirement on one of them. For the FR1 schedule, this meant whichever
stimulus was pecked first, whereas for the FR 25 schedule, this meant whichever
one received 25 pecks first. If the FR was completed on the suboptimal alterna-
tive, the initial-link stimulus was replaced with one of two distinct terminal-link
stimuli appearing for a fixed 10 s duration. One terminal-link stimulus (e.g.,
green) was presented on 20% of the trials and was always followed by a food
presentation and a 10 s ITI. The other terminal-link stimulus (e.g. red), was
presented on 80% of the trials and was followed by only the ITI. Responding on
the optimal alternative was also followed by one of two distinct terminal-link

stimuli (e.g., yellow or blue) of fixed 10 s duration, on 20% and 80% of the trials, respectively. Regardless of which terminal-link stimulus was presented, 50% of the trials ended in food.

On forced-exposure trials, either the suboptimal or optimal initial-link stimulus was presented and a single peck or 25 pecks, depending on the condition, was required to enter the terminal link. During choice trials the initial links were presented concurrently. The first initial link to have its ratio requirement satisfied determined which set of terminal links was encountered. If, for instance, the suboptimal initial link was completed first, its terminal link was displayed and the corresponding optimal initial link would become inoperative. Peck counts to each alternative were reset each trial and each alternative was independent of the other. Each session lasted for 90 min and contained a maximum of 75 trials. There were 50 forced-exposure trials (25 suboptimal and 25 optimal) and 25 choice trials per session that a pigeon could potentially complete. The presentation order of the forced-exposure and choice trials were pooled together and randomized each session.

Testing Phase 1 lasted for 15 sessions. The FR25 condition included an additional session at the outset of the phase to shape responding up to the FR 25 requirement. In this session, the ratio requirement was increased by 1 after every two trials until FR 25 was reached on trial 49, at which point the ratio requirement remained at 25. For example, on trials 1 and 2, the ratio requirement on both alternatives was a FR 1; on trials 3 and 4 it was a FR 2; on trials 5 and 6 it was a FR 3, and so on until FR 25 was reached.

Testing Phase 2. This phase was conducted as per Phase 1 but the conditions were switched so that if a pigeon was on the FR1 condition they switched to FR25 and vice versa. Phase 2 lasted for 20 sessions. Pigeons switching to the FR25 condition, as before, received an initial shaping session that moved them from FR 1 to FR 25. Stimulus locations and colours remained as they had been in Phase 1.

3.3.4. *Data Analysis.* All statistical analyses were conducted using R 3.5.0 (R Core Team, 2018). Linear-mixed effects modeling (Pinheiro et al., 2018), fit by maximum likelihood, was used to assess the significance of main effects of condition (FR1 and FR25) and phase (1 and 2), with their corresponding interaction. Condition was nested within the random effect of

subject. Given the heterogenous spread of the variances obtained, a constant variance function structure was applied to all tested models to allow different spreads across condition and phase. A corresponding Bayes factor (BF_{10}) was calculated for each tested effect to obtain the relative odds in favor of the alternative hypothesis against a null model. Means are presented with corresponding 95% confidence intervals and collected from the last three sessions of each condition.

The latency to the first peck made during the initial links was compared between forced-exposure and choice trials using the same analysis as for the choice proportions. Simple effects analyses are reported using paired t-tests with a corresponding Hedges corrected effect size (g) and JZS Bayes factor (Morey & Rouder, 2015). These latencies were used to test predictions of the SCM. The calculations were based on the last three sessions of each phase, except that, as per Vasconcelos et al. (2015), choice predictions were determined on the basis of latencies obtained from the preceding 64 forced-exposure trials of each actual choice which, in some cases, included sessions that preceded the last three sessions of each phase. Latencies were normalized using a base-10 logarithmic transformation. For predictions of the FR1 and FR25 conditions, ordinary least-squares linear regression models were fit to the data and compared to an intercept-only (null) model. Corresponding R^2 values and Bayes Factors (BF_{10}) are provided for evaluation.

3.4. **Results**

Figure 3.2 plots each pigeon's proportion of suboptimal choices, along with the overall mean, as a function of session and phase. In the FR1 condition, pigeons showed a steep early trend towards exclusive suboptimal responding. This occurred irrespective of which phase the condition appeared in; though there was marginally more variability when this condition occurred in Phase 2. This corroborates the results of other laboratories running comparable conditions (Stagner et al., 2012; Stagner & Zentall, 2010).
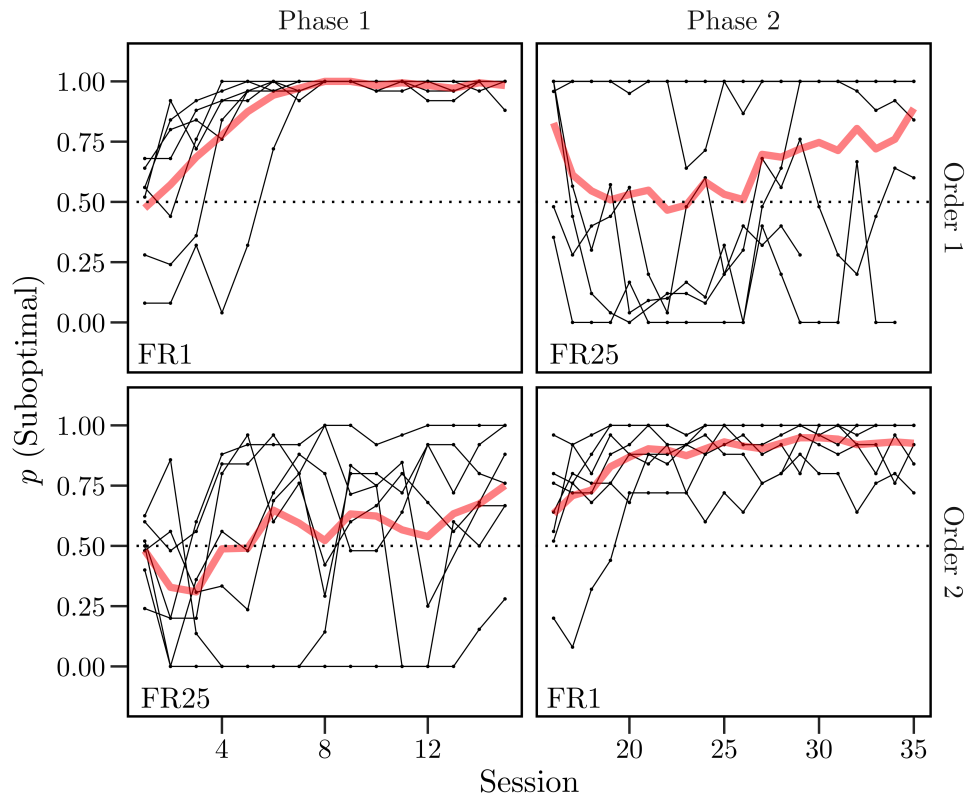
Figure 3.2: Plots showing proportion of suboptimal responding across sessions by phase and order. Order 1 received the FR1 condition first, order 2 received the FR25 condition first. Each black line depicts data from an individual pigeon. The transparent red line depicts the overall mean.

In the FR25 condition, no consistency between pigeons was seen and the data did not appear to follow a clear general trend. The session means suggest that there was some gradual tendency towards suboptimality in these conditions over time, but the variability across and within subjects makes this difficult to parse. In addition, responding towards alternatives on choice trials tended to be exclusive; that is to say, switching between alternatives almost never occurred after the initial choice response.

The high levels of variability observed in the FR25 condition may be an artifact of the calculation of session proportions - a scoring/plotting convention in pigeon suboptimal choice research. Each subject's proportion of suboptimal responses was calculated at the session level; however, what is not apparent in this measure is that the number of choice trials in each calculation is quite small for some birds in the FR25 condition. Table 3.1 shows the mean number of completed trials per session for each bird. In the FR25 condition, birds completed the initial links much more slowly. For some birds, the number of choice trials completed in a given session were less than two, which was deemed too few to permit a reasonable calculation of the choice proportion. This occurred on five separate sessions: once each for birds 75, 138, 304, 306, and 308. Figure 3.2 displays only session choice proportions that had two or more completed choice trials. Sessions with less than two choice trials are not included in the figure and adjacent sessions (with values ≥ 2) are connected via a straight line.

To mitigate these issues, a different plotting and scoring convention was applied. Figure 3.3 displays the cumulative record of choice responding for each pigeon by condition and phase across choice trials. Responses made to the suboptimal alternative were scored as a $+1$ and responses made to the optimal alternative were scored as a -1 and a cumulative sum across choice trials was calculated and plotted. As utilized here, steep positive slopes indicate preference for the suboptimal alternative, negative slopes indicate preference for the optimal alternative, and flat lines indicate indifference. In addition, because of the record's high, trial-by-trial resolution, subtle shifts in each pigeon's behaviour are easily seen and not subject to the loss of detail that session proportions and other averaging methods generate.
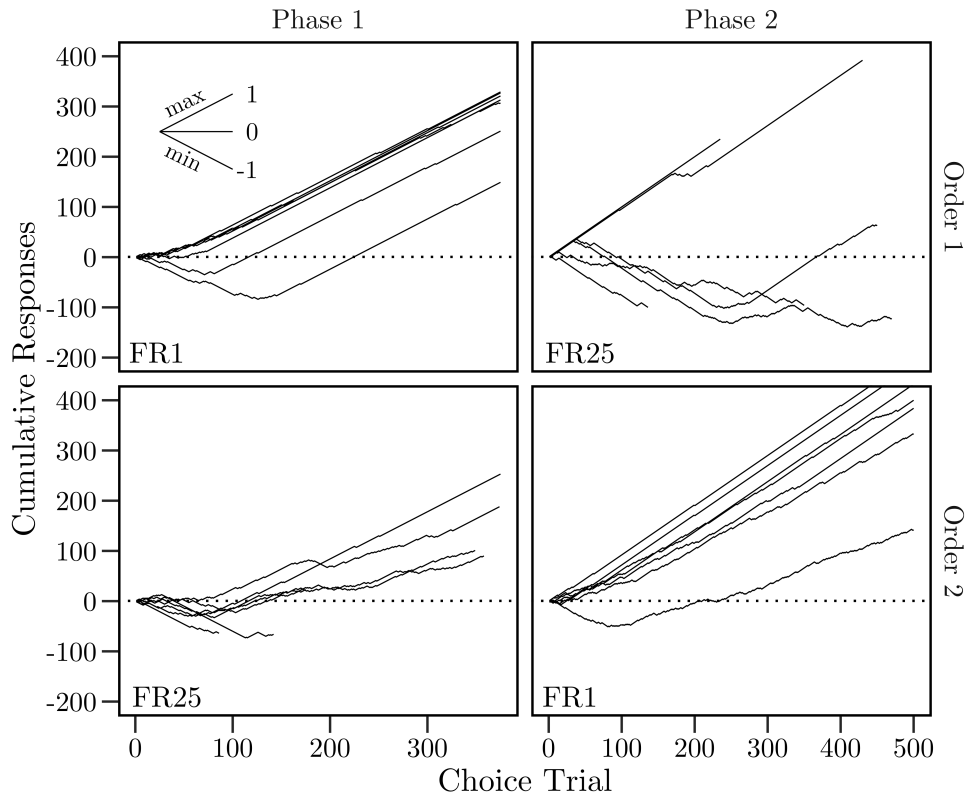
Figure 3.3: Pigeon choice data displayed as a cumulative record of responding. Choices made to the suboptimal alternative were scored as +1 and choices made to the optimal alternative were scored as −1. Each line depicts data from an individual pigeon. The legend in the top left panel displays the maximum and minimum slope the lines could take.
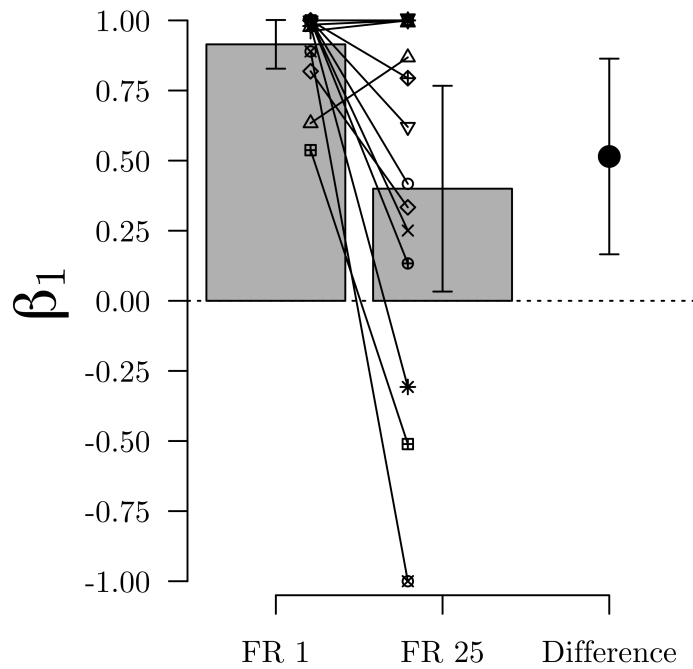
Figure 3.4: Bar plot depicting the mean slope (β_1) of cumulative responding
on choice trials across the last three sessions of the FR1 and FR25
conditions. Rightmost point displays the mean difference between the two
conditions. All error bars are 95% confidence intervals. Individual pigeon
values are plotted as distinct shapes.

Visual inspection of the cumulative records reveals that, in the FR25 condition, pigeons were either mostly suboptimal or mostly optimal in their direction of responding - with the former being more frequent. Only one pigeon showed a clear change of its direction of preference in this condition, suggesting that the overall gradual trend toward suboptimality shown by the mean choice proportion in the FR25 condition may be misleading. It is also worth noting that, while the direction of preference tended to remain consistent within individual birds, the slopes in the FR25 condition were shallower than their FR1 counterpart, indicating greater variability of preference. This is in contrast to the FR1 condition which generated, in many cases, maximally steep slopes after approximately 100 choice trials.

Figure 3.4 and Table 3.1 shows the obtained slope (β_1) of the cumulative

records, fit by ordinary least squares regression, for each bird during the last three sessions of each phase when learning effects are assumed to have subsided. In the FR1 condition the maximum possible number of choice trials, 75, were used to calculate each subject's slope whereas, in the FR25 condition, only a median of 70 choice trials (with an inter-quartile range of 24.5 to 75) could be used due to the lower rates of responding observed in this condition.

Table 3.1

Individual pigeon experiment values and results

Subject	Condition	Phase	Mean Trial/Session	C.P.	β	FE IL Median Latency (s)		FE IL Run Rate: (r/s)		FE Terminal Link: r/s			
						Sub	Opt	Sub	Opt	Sub 20%	Sub 80%	Opt 20%	Opt 80%
63	FR1	1	75	0.99	0.96	0.84	6.53			1.35	0	1.12	1.46
75			75	1	1	0.51	2.33			2.05	0.02	1.82	1.73
138			75	1	1	1.07	5.59			1.68	0	1.47	1.5
301			75	0.99	0.98	0.96	5.11			0.28	0.22	0.57	0.93
303			75	0.99	1	0.84	3.79			1.36	0.02	3.11	1.9
305			75	0.99	0.98	0.84	7.56			1.15	0.06	0.97	0.77
308			75	0.93	0.89	1.06	5.41			0.51	0.01	1.14	1.39
Mean			75.00	0.98	0.97	0.87	5.19			1.20	0.05	1.46	1.38
(SD)			(0.00)	(0.02)	(0.04)	(0.19)	(1.72)			(0.62)	(0.08)	(0.83)	(0.41)
49	FR25	1	70.87	0.71	0.42	1.69	7.33	3.19	2.39	2.15	0.08	2.5	3.01
51			74.73	0.88	0.87	1.41	6.96	3.14	2.28	2.91	0.06	1.95	3.01
98			29.8	0.6	0.33	11.87	17.64	0.59	0.65	0.57	0.18	1.1	0.95
135			69.4	0.83	0.62	5.35	8.89	1.47	1.4	1.24	0.09	1.46	1.4
304			35.07	0.68	0.13	5.83	11.22	1.52	1.35	2.37	1.03	2.58	2.24
306			16.47	0.22	-0.51	6.81	1.74	1.7	1.78	2.77	0.35	2.48	2.78
473			75	1	1	2.07	7.88	2.4	1.64	2.17	0.06	1.91	2.02
Mean			53.05	0.70	0.41	5.00	8.81	2.00	1.64	2.03	0.26	2.00	2.20
(SD)			(24.96)	(0.25)	(0.51)	(3.74)	(4.83)	(0.95)	(0.59)	(0.84)	(0.35)	(0.57)	(0.80)
49	FR1	2	75	1	1	0.33	5.84			2.41	0.07	2.75	2.52
51			75	0.87	0.63	0.56	2.12			2.94	0.03	1.83	2.8
98			75	0.87	0.82	1.23	2.1			0.26	0	0.62	0.4
135			75	1	1	1.31	5.61			1.29	0	1.36	1.46
304			75	1	1	1.37	3.38			1.94	0	1.88	1.71
306			75	0.76	0.54	1.15	1			2.37	0.01	2.42	2.35
473			75	1	1	0.66	6.46			2.05	0.03	1.78	1.87
Mean			75.00	0.93	0.86	0.94	3.79			1.89	0.02	1.81	1.87
(SD)			(0.00)	(0.10)	(0.20)	(0.42)	(2.17)			(0.88)	(0.03)	(0.69)	(0.80)
63	FR25	2	33.55	1	1	4.49	15.31	1.93	0.98	1.24	0.08	0.91	1.23
75			70.95	0.56	0.25	3.16	4.52	1.7	1.56	1.99	0.04	1.98	1.87
138			18	1	1	5.42	13.2	1.45	1.56	1.17	0.05	1.2	1.52
301			74.64	0.33	-0.31	12.32	13.85	1.02	1.03	0.29	0.09	0.91	1.4
303			67.85	0.88	0.79	4.44	6.44	1.99	1.45	1.11	0.04	2.12	1.86
305			63.75	1	1	4.26	14.49	1.59	1.09	0.99	0.02	0.94	0.67
308			20	0	-1	21.71	21.55	0.53	0.84	1.6	0.05	1.57	1.39
Mean			49.82	0.68	0.39	7.97	12.77	1.46	1.22	1.20	0.05	1.38	1.42
(SD)			(24.99)	(0.40)	(0.79)	(6.78)	(5.72)	(0.52)	(0.30)	(0.53)	(0.02)	(0.52)	(0.41)

Note. C.P. = Choice Proportion; IL = Initial Link; FE = Forced-Exposure Trials. Apart from the trial per session column, all values are from the last three sessions of each phase. All response rates reflect a mean local rate of responding. Run Rates are calculated as the rate of responding on the initial link, following the first response of a trial.

Across all birds, a significant main effect of condition (FR1 versus FR25) was obtained when tested against a null model containing the intercept, additive random effect of subject, and constant variance function structure, $\chi^2(1) = 9.88$, $p = 0.002$, $BF_{10} = 26.47$. No significant main effect of phase was obtained when added to the treatment only model, $\chi^2(1) = 2.31$, $p = 0.129$, $BF_{10} = 0.60$, nor was there any observed interaction between condition and phase, $\chi^2(1) = 0.09$, $p = 0.760$, $BF_{10} = 0.20$.

Analysis of mean log-latencies found a significant main effect of trial type (single vs. forced-exposure) $\chi^2(1) = 27.20$, $p < 0.001$, $BF10 > 150$ as well as an additional effect of condition, $\chi^2(1) = 33.35$, $p < 0.001$, $BF_{10} > 150$. A significant interaction between the two was also observed, $\chi^2(1) = 13.01$, $p < 0.001$, $BF_{10} = 89.17$. Simple effects analyses found that forced-exposure latencies in the FR1 condition, $M = 0.24$ 95% CI $[0.15, 0.33]$, were significantly longer than choice trial latencies, $M = -0.11$ $[-0.24, 0.01]$; $t(13) = 9.90$, $p < 0.001$, $g = 1.23$ 95% CI $[0.38, 2.08]$, $BF_{10} > 150$. This effect, while considerably smaller, held for the FR25 condition as well, $t(13) = 3.81$, $p = 0.002$, $g = 0.24$ $[-0.54, 1.02]$, $BF_{10} = 20.39$; with the forced-exposure trials obtaining a $M = 0.83$ $[0.65, 1.01]$ and the choice trials a $M = 0.69$ $[0.45, 0.92]$.

Figure 3.5 displays the observed proportion of suboptimal choices as a function of the SCM's predicted proportion of suboptimal choices. Linear regression showed moderate fits for both the FR1 ($R^2 = 0.57$, $BF_{10} = 100.94$) and FR25 conditions ($R^2 = 0.41$, $BF_{10} = 10.40$). However, the y-intercept obtained in the FR1 condition was noticeably higher than the SCM's predicted value of 0. Further inspection revealed a disproportionate amount of leverage caused by bird 306 in the FR1 condition. With that data point removed the FR1 condition's linear regression provided a markedly worse fit to the data ($R^2 = 0.03$, $BF_{10} = 0.33$).

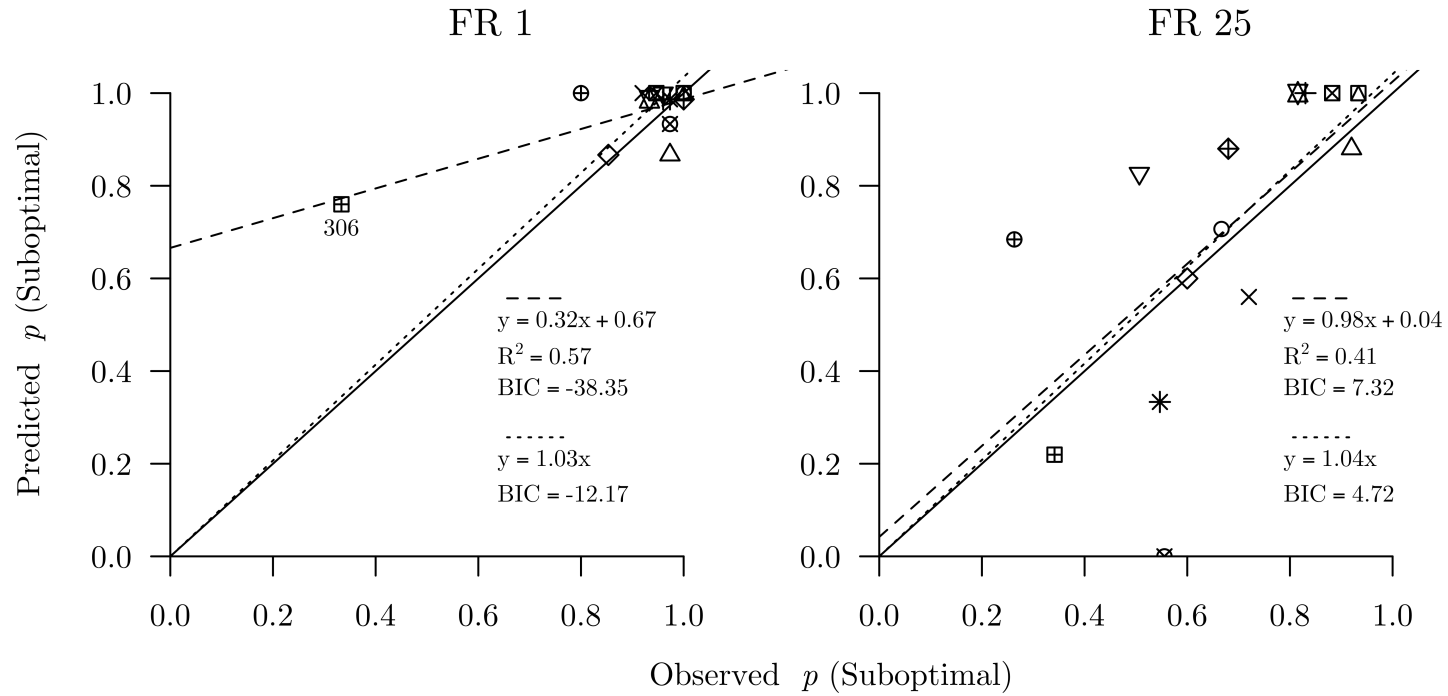Figure 3.5: Plots displaying the obtained proportion of suboptimal choices as a function of the SCM's predicted proportion. Left and right panels show the FR1 and FR25 conditions respectively. Each symbol corresponds to an individual subject. Dashed lines show the regression obtained with an unconstrained y-intercept and dotted lines show the y-intercept constrained at 0. Solid lines show a slope = 1 and intercept = 0.

3.5. **General Discussion**

The present results show that initial-link schedule is an important determinant of the level of suboptimal choice pigeons show when choosing between signalled and unsignalled outcomes. Although there was variability across birds in the magnitude of the effect, larger FR schedules resulted in significantly attenuated preference for the suboptimal alternative. The increased variability with longer initial-link schedules is also consistent with prior studies. For example, in two experiments, each using somewhat different procedures, Zentall et al. (2017) found greater inter-subject variability when the initial links were FI 20 s than when they were FI 1. In addition, the pigeons generally showed less suboptimal choice when the initial links were FI 20s. In the present procedure, when the response requirement on both initial-links was increased from FR 1 to FR 25, pigeons on average, still distributed their responding more frequently towards a suboptimal alternative. However, the increased initial-link requirement notably attenuated mean preference, and in some individuals completely shifted preference to the optimal alternative. This corroborates past findings using VI schedules with signalled optimal alternatives (Dunn & Spetch, 1990; Kendall, 1974, 1985) as well as the basic ordinal predictions of the SiGN model (McDevitt et al., 2016), the information theoretic approach (Cunningham & Shahan, 2018), and the HVA model (Mazur, 2001). Future research would do well to test a wider range of schedule values so that a more explicit quantitative assessment of these models' suboptimal choice predictions can occur.

Consistent with the SCM's predictions, mean log-latencies were found to be significantly shorter on choice than forced-exposure trials. However, explicit SCM choice predictions received less support. The observed latencies in the FR25 condition were largely in accordance with predictions of SCM; however, the FR1 condition obtained effects that did not as clearly fit the SCM predictions. The reasons for this may be measurement related. The infrared touchscreens yield a less precise measurement of pigeons' responses than that of mechanical key-switches. Perhaps with greater measurement precision of each response, the SCM would be better supported in the FR1 condition. Alternatively, given that most support for the SCM's choice predictions have come from research on European starlings (Freidin, Aw, & Kacelnik, 2009; Shapiro,

Siller, & Kacelnik, 2008; Vasconcelos, Monteiro, Aw, & Kacelnik, 2010; Vasconcelos et al., 2015), it may be that, for pigeons, the SCM's predictions in the FR 1 case simply do not hold.

The low response rates some birds exhibited in the FR25 condition reduced their overall exposure to the contingencies. Nevertheless, inspection of cumulative records shows that individual preferences proceeded in a fairly consistent direction for all but one bird in the FR25 condition. It is worth noting that cumulative records, despite their seemingly antiquated nature in psychology, provided an elegant and highly detailed way to view the progression of learning. Further, the clear differential effects observed with the initial-link latencies as well as response rates in the terminal links strongly suggest that the birds had learned the programmed contingencies despite some of them being slower to respond than others. It therefore seems unlikely that additional training would have substantially altered the observed preferences.

In the present study, an FR 25 initial-link schedule was selected to provide a straightforward extension of the FR 1 and because fixed schedules avoid the short delays occasionally provided by variable schedules. Although the low response rates of some birds in the FR25 condition may have been avoided by using a variable schedule, it also might have introduced variability due to other schedule effects. An open question remains concerning whether fixed and variable schedules differentially affect suboptimal choice. In the present study, suboptimal choice was reduced when the initial-link FR was lengthened in the same way that suboptimal choice was reduced by long VI schedules in past studies, such as Kendall (1974). Nevertheless, subtle differences between ratio and variable initial-link schedules may occur. For instance, the SiGN model emphasizes the role of the context set by the choice response and assumes organisms rely on all available cues to signal this context. With long FR schedules, the context may change over the initial link because each peck results in a slight reduction in time to food as the terminal link approaches. By comparison, the unpredictable nature of a variable-ratio (VR) schedule prevents individual responses from acquiring a precise discriminative function. Thus, it might be predicted that long VR schedules would produce even less suboptimal choice than equivalently long FR schedules. The HVA model, by contrast, appears to predict more suboptimal choice for variable than fixed schedules because the short interval that sometimes occur on variable sched-

ules provide higher value to the initial-links. Both models, however, would predict that the subtle differences between fixed and variable schedules would be small relative to the difference between short and long initial links.

Recently, much of the focus of suboptimal choice research has emphasized the signal value of the terminal links (e.g., Zentall, 2016), with little emphasis on the role of the preceding events. The results obtained here not only corroborate certain key findings and theories of suboptimal choice, but also emphasize the important role of the initial links.

3.6.   **Author Note**

Jeffrey M. Pisklak, Department of Psychology, University of Alberta; Margaret A. McDevitt, Department of Psychology, McDaniel College; Roger M. Dunn, Department of Psychology, San Diego State University; Marcia L. Spetch, Department of Psychology, University of Alberta.

Chapter 4   The Power of Nothing: Risk Preference in Pigeons, But Not People, Is Driven Primarily by Avoidance of Zero Outcomes

4.1. **Abstract**

Both human and non-human animals regularly need to make choices where the outcomes of their actions are unpredictable or probabilistic in some way. These are often termed "risky" choices. Faced with uncertain rewards, people (*Homo sapiens*) and pigeons (*Columba livia*) often show similar choice patterns. When the reward probabilities of risky choices are learned through experience, preferences in both species seem to be disproportionately influenced by the extreme (highest and lowest) outcomes in the decision context. Overweighting of these extremes increases preference for risky alternatives that lead to the highest outcome and decreases preference for risky alternatives that lead to the lowest outcome. In a series of studies, we systematically examine how this overweighting of extreme outcomes in risky choice generalizes across two evolutionary distant species: pigeons and humans. Both species showed risky choices consistent with an overweighting of extreme outcomes when the low-value risky option could yield an outcome of zero. When all outcome values were increased such that none of the options could lead to zero, people but not pigeons still overweighted the extremes. Unlike people, pigeons no longer avoided a low-value risky option when it yielded a non-zero food outcome. These results suggest that, despite some similarities, different mechanisms underlie risky choice in pigeons and people.

4.2. **Introduction**

When confronted with uncertain rewards, humans and other animals often show similar preferences for varying levels of risk (e.g., Lagorio & Hackenberg, 2012; Ludvig, Madan, Pisklak, & Spetch, 2014; Weber, Shafir, & Blais, 2004), while other animals are often risk averse for reward amount, but risk seeking for reward delays (e.g., Bateson & Kacelnik, 1995; Mazur, 1984, 1986). These risk preferences, however, can vary with context depending on how the potential outcomes are encountered (Hertwig & Erev, 2009; Madan et al., 2019) or the range of other outcomes that are available (Ludvig, Madan, & Spetch, 2014; Stewart, 2009). For example, one recent study compared risky choice in pigeons and humans, finding that risk preferences in both species was influenced by an oversensitivity to the extreme outcomes, highest and lowest, encountered (Ludvig, Madan, Pisklak, & Spetch, 2014). In this paper, we build on that

result, by assessing risky choice with a broader range of possible outcomes, systematically assessing to what degree the risk preferences of pigeons and humans reflect common underlying psychological processes.

In non-human animals, a wide range of risk preferences for rewards have been documented: many studies have shown risk aversion as in humans (e.g., Bateson & Kacelnik, 1995), but some have found risk neutrality or even risk seeking (e.g., Barnard, Brown, Houston, & Mcnamara, 1985; Hayden, Heilbronner, Nair, & Platt, 2008). Even within a single species, such as pigeons, behavioural effects can vary across experiments (Lagorio & Hackenberg, 2012). Moreover, risk preferences may depend on factors such as energy budget (Caraco, Martindale, & Whittam, 1980; Stephens, 1981), choice framing (e.g., Constantinople, Piet, & Brody, 2018; Lakshminarayanan, Chen, & Santos, 2011; Marsh & Kacelnik, 2002), and the recent outcomes of a choice (Marshall & Kirkpatrick, 2013).

In humans, risky choice is typically assessed using scenarios in which the probabilities and outcomes are explicitly described. When people learn about the consequences of their risky choices through experience, however, different choice patterns often emerge—a discrepancy referred to as the description-experience gap (e.g., Barron & Erev, 2003; Camilleri & Newell, 2011; Hertwig & Erev, 2009; Ludvig & Spetch, 2011; Ungemach et al., 2009). For example, with described choices, people tend to be more risk seeking when the choice is between two types of losses than two types of gains (e.g., Kahneman & Tversky, 1979). In contrast, when outcomes are learned through experience, people sometimes reverse these preferences and are more risk seeking for gains than for losses (Ludvig & Spetch, 2011; Madan et al., 2017; Tsetsos, Chater, & Usher, 2012). Learning from experience is also more similar to how non-human animals make decisions (but see Constantinople et al., 2018; Heilbronner & Hayden, 2016).

One explanation for the pattern of risk preferences in experience-based choice is that the most extreme values encountered in a decision context are overweighted (e.g., Ludvig, Madan, & Spetch, 2014; Ludvig, Madan, McMillan, Xu, & Spetch, 2018). For example, consider a context where there are 4 possible options: a safe gain of +20 points, a risky 50/50 chance of +40 or 0 points, a safe loss of -20 points, and a risky 50/50 chance of -40 or 0 points. In this context, the best possible outcome is +40 and the worst possible outcome

is -40. Overweighting of these extremes would thus produce more risk seeking in choices between the gains (where the extreme is good) and risk aversion in choices between the losses (where the extreme is bad). People do indeed overweight the extremes when repeatedly choosing between options that led to safe and risky gains and losses (Ludvig & Spetch, 2011; Ludvig, Madan, & Spetch, 2014; Madan, Ludvig, & Spetch, 2014).

In a previous cross-species comparison, we found evidence suggesting that pigeons also overweight the extreme outcomes (Ludvig, Madan, Pisklak, & Spetch, 2014). Pigeons were tested in a foraging analogue of the human decision-making task in which they chose by walking behind one of two distinct panels to obtain food rewards. As in the human task above, there were four possible options: Two of the panels provided high-value outcomes: one panel led to a 50/50 chance of 2 or 4 food rewards and the other panel always led to 3 food rewards. The two other panels provided low-value outcomes: one panel led to a 50/50 chance of 0 or 2 food rewards, and the other panel always led to 1 food reward. The pigeons' choices were compared to those of humans who participated in a computer-based task with doors that similarly led to risky or safe numbers of points exchangeable for money. Both species were more risk seeking for high-value gains than for low-value gains, as though they were overweighting the extreme outcomes.

Humans consistently overweight extremes across many different sets of outcome values (see Madan et al., 2019, for a review). Pigeons, however, have only been tested with a single set of outcome values (0 to 4; Ludvig, Madan, Pisklak, & Spetch, 2014). Whether the overweighting of extremes would generalize to other values for pigeons is not known. One alternative possibility is that the seeming overweighting of extremes by pigeons was actually due in part, or even largely, to an avoidance of zero outcomes. Although the pigeons showed risk seeking for high-value choices and risk aversion for low-value choices, the deviation from indifference for the low-value choices was more pronounced (Ludvig, Madan, Pisklak, & Spetch, 2014). Moreover, in their procedure, a zero outcome represented the absence of food reward, which meant that selection of the low-value risky option was only intermittently reinforced, whereas the low-value safe option was continuously reinforced, albeit at a smaller reinforcer magnitude.

With humans, the possibility that an aversion to zero outcomes is re-

sponsible for the risk preferences was ruled out by eliminating zero outcomes and still observing the same overweighting of extremes (e.g., Ludvig, Madan, & Spetch, 2014). With pigeons, however, the contribution of zero outcomes to the extreme-outcome effect has not yet been evaluated. Indeed, animals likely process the concept of zero differently from humans (Nieder, 2016). The present experiments were thus designed to more extensively evaluate the correspondence between pigeon and human risky choice observed by Ludvig, Madan, Pisklak, and Spetch (2014) and to test whether the similar patterns of choice were produced by the same psychological mechanism. Experiment 1 was a conceptual replication of their risky-choice task using an operant task that was more similar across species. Experiment 2a tested risk preferences in the absence of zero values using an open-field task in pigeons (as in Ludvig, Madan, Pisklak, & Spetch, 2014) and with the same operant task in people. Experiment 3 directly manipulated the occurrence of zero with the operant task with pigeons. Lastly, Experiment 4 examined pigeons' risk preferences without zero outcomes after an evaluation of their capacity to discriminate between the different reward magnitudes.

4.3. Experiment 1: Zero Outcomes

The first experiment developed a new operant task for comparing risky choice across pigeons and humans (see Figure 4.1). In the task, both species selected between pairs of coloured circles that represented risky and safe outcomes of high or low value, ranging from 0 to 4 (as above; see Ludvig, Madan, Pisklak, & Spetch, 2014). Pigeons were reinforced with food and humans with points that were exchanged for money. If extremes are overweighted, then both pigeons and humans should select the risky option significantly more often when choosing between risky and safe high-value options than between risky and safe low-value options. Avoidance of zero values should also produce a difference in risk preference, driven by strong risk aversion for the low-value choices.

4.3.1. *Experiment 1a: Pigeons*

Subjects. Six adult pigeons (*Columba livia*) - four Racing Pigeons and two Silver King - selected from a University of Alberta pigeon colony were used. All pigeons had extensive learning histories, but none had served in studies of risky
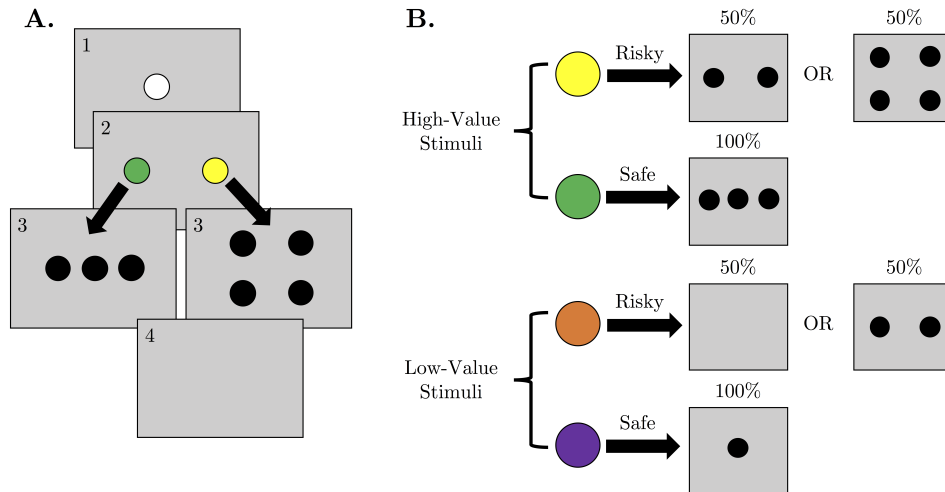
Figure 4.1: (A) A schematic of a choice trial displaying a risky high-value option and a safe high-value option, with possible outcomes in Experiment 1. (B) A schematic illustrating the various reward contingencies in effect. The colour of the choice circles relative to their outcome was counterbalanced across participants according to four different combinations of yellow, green, orange, and purple.

choice. They were individually housed in a temperature-controlled colony with a 12-hour light-dark cycle. Free access to grit and vitamin-enriched water was provided, and each pigeon was maintained at approximately 85% of its free-feeding weight by food pellets obtained during and after experimental sessions. All procedures were approved by the University of Alberta Biological Sciences Animal Care and Use Committee, following the guidelines of the Canadian Council on Animal Care. All pigeons passed the exclusion criterion and were included in the analyses (see Data Analysis section).

Apparatus. A custom-built 71.1×33.0×44.5 cm sound-attenuating operant-conditioning chamber, located in an isolated room, was used. A 17-in. View-Sonic LCD monitor was mounted centrally against the chamber's widest wall, equipped with a Carrol Touch infrared touchscreen (Elo Touch Systems Inc., Menlo Park, CA). Two feeding ports, adjacent to both sides of the monitor, provided access to food pellets via a solenoid-controlled food hopper containing Mazuri food pellets. A light within each feeding port signalled when the hopper had been raised, and an infrared beam detected entry into the port.

Each chamber was connected to a computer in an adjacent room. E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) was used to control contingencies and detect responses.

Stimuli. Experimental stimuli consisted of coloured circles (white, black, purple, green, yellow, and orange) presented against a grey background on the monitor. Each circle had a diameter of 100 pixels (approximately 2.5 cm).

Procedure. Sessions lasted for 45 min and were run six days a week at approximately the same time each day.

Training Phase 1. This phase entailed an autoshaping procedure that presented white circles either centrally on the screen or to the right or left of centre, or black circles in numerous spatial locations. Trials began with a single white or black circle (each with a 50% chance). The circle remained onscreen for 60 s or until it was pecked. It then disappeared and a randomly-selected hopper was raised allowing the pigeon 1 s of food access. Access was timed from the initial moment the pigeons head entered the feeding port. This was followed by a 20-s inter-trial interval (ITI). The spatial location of each presented circle was selected randomly without replacement from a list of locations. Autoshaping continued until the pigeons responded on more than 75% of the total stimulus presentations. Phase 1 lasted for a mean of 2.17 and range of 1-3 sessions.

Training Phase 2. This phase entailed an operant procedure in which the ITI was shortened to 2 s and a peck at the circle was required to raise a hopper. This phase continued until the pigeons completed 50 trials within a 45-min session, requiring a mean of 1.33 and range of 1-3 sessions.

Training Phase 3. In this phase, sessions consisted of only 16 trials, and the stimuli appeared in a sequence that resembled the testing procedure. First, a white circle (start stimulus) appeared centrally on the screen. A single peck at the start stimulus erased it and produced another white circle on the left or right. Pecking the new circle erased it and produced 1, 2, 3, or 4 black circles (tokens) according to the layouts depicted in Figure 4.1. A single peck at each token produced 1-s access to food. Once all tokens had been selected, a 2-s ITI ensued. Pigeons remained on this phase until their weight stabilized to approximately 85% of their free-feeding value, which required a mean of 13.5 and range of 10-15 sessions.

Training Phase 4. This phase consisted of four sessions, each containing 16 single-option trials with yellow, green, orange, and purple circles. This phase provided forced exposure to the stimuli and reward contingencies that would occur in testing, with the order and side presentation randomized across trials.

Testing. Testing consisted of 80 sessions, each lasting up to 60 minutes with 16 trials. Figure 1 shows a schematic of an example trial. Each trial began with the centrally-presented white start-stimulus. A single peck to this stimulus caused it to disappear and one or two coloured (choice) circles appeared on either side. A single peck to a choice circle erased it and produced 0, 1, 2, 3, or 4 black circles (tokens) as shown in Figure 4.1. There were four choice options. The *risky-high* option led to 2 or 4 tokens each with a 50% chance. The *risky-low* option led to 0 or 2 tokens each with a 50% chance. The *safe-high* option always led to 3 tokens, and the *safe-low* option always led to one token. A single peck to any of the tokens caused it to disappear and a randomly-selected food hopper to rise, providing 1-s of food access. The bird could then select any remaining token for another 1-s of food access until all tokens were removed. After a 2-s ITI, a new trial began.

Each session included three trial types: *Risk-preference* trials presented a choice between a risky-high and a safe-high option, or a risky-low and a safe-low option. *Catch trials* presented a choice between a high and low reward option, with all four possible choice combinations. For instance, given a choice between a risky-high and safe-low alternative (see Figure 4.1), the former provides 3 rewards *on average* and the latter only provides 1. Given that one of the catch alternatives always contains a higher average reward payout (more food reinforcement), that alternative is expected to be preferred. Because catch trials facilitate these obvious preferences, they were used as a means of assessing how well the task's contingencies are learned. *Single-option* trials presented only one option that had to be selected, ensuring that all options were occasionally chosen and all outcome contingencies were experienced. Each session included four risk-preference trials (two high and two low), eight catch trials and four single-option trials, with trial order randomized within the session and counterbalanced such that each choice stimulus appeared equally often on the left or right. All possible combinations of stimuli and reward outcomes were counterbalanced across every four sessions.

4.3.2. *Experiment 1b: Humans*

Participants. Thirty participants (aged $[M \pm SD]$ 20.5 $\pm$ 2.1 years old; 20 females) recruited from the University of Alberta undergraduate psychology participant pool participated for course credit. All participants provided informed consent, and procedures were approved by the University of Alberta research ethics board. Seven participants were excluded from the analysis due to failure to pass the catch trial criterion.

Procedure. Participants first sat as a group, and instructions were simultaneously read aloud and projected on a screen (see Appendix A). Participants then entered individual rooms to complete the computer task using a mouse to make choices. Visually, the task was identical to the pigeon testing procedure, except that a cumulative point tally was displayed at the bottom of the screen. Each click on a token raised this tally by 1 point. The experimental session was divided into 8 blocks of 64 trials. To provide a break, participants performed a maze task between blocks. The first block provided 64 single-option trials. All subsequent blocks provided 16 risk-preference trials, 32 catch trials, and 16 single-option trials. Trial types were randomly intermixed and provided every possible combination of choice stimulus, side, and reward.

After the experiment, participants' cumulative points were converted to a cash bonus of up to \$5.00. Participants were not told the conversion rate prior to the experiment. The conversion was based on a linear increase from the fewest possible points (744 points = \$0.00), through the points expected by chance (1024 points = \$2.50), to the maximum possible points (1304 points = \$5.00).

4.3.3. *Data Analysis* For consistency with prior work (Ludvig, Madan, Pisklak, & Spetch, 2014), the final third of the total collected choices were used for statistical analysis, which was conducted using R Software for statistical computing (R Core Team, 2018). As in prior studies (e.g., Ludvig & Spetch, 2011; Ludvig, Madan, Pisklak, & Spetch, 2014), pigeons or humans who showed less than 60% accuracy in choosing the high-value option across all catch-trial types were excluded from analyses. The rationale is that if subjects do not reliably choose more reward (points/money or food) over less reward, then they either did not adequately learn the reward contingencies or were not motivated by the potential rewards, making their risk preference results

difficult to interpret.

For risk-preference trials, a paired t-test was conducted on the proportion of risky option selections for the high- and low-value choice types. For catch trials, a paired t-test was conducted on the proportion of high-value option selections across trials with and without the lowest-extreme (i.e., 0) as a possible outcome. Additionally, a JZS Bayes factor (BF_{10}) with a medium prior was computed for each comparison to obtain the relative odds in favor of the alternative hypothesis over the null using the BayesFactor R package (Morey & Rouder, 2011, 2015). Effect sizes of mean differences used an unbiased estimate of Cohen's d (i.e., Hedges' g), and 95% confidence intervals were computed using the effsize package (Torchiano, 2017).

4.3.4. *Results* Figure 4.2A (pigeons) and Figure 4.2B (humans) show the mean proportion of risky option selections for high-value and low-value choices, and the paired differences between these choice types with 95% confidence intervals (Franz & Loftus, 2012). For high-value choices, risky option selection refers to choosing the stimulus leading to either 2 or 4 instead of the stimulus leading to a guaranteed 3. For low-value choices, risky option selection refers to choosing the stimulus leading to 0 or 2 instead of the stimulus leading to a guaranteed 1. On average, pigeons chose the risky option significantly more often for high-value ($M = .58$, 95% CI $[0.45, 0.72]$) than for low-value choices ($M = .10$ $[-0.02, 0.22]$), $t(5) = 11.45$, $p < .001$, $g = 4.31$ $[1.97, 6.67]$, $BF_{10} > 150$. Similarly, humans reliably selected the risky option more for high-value ($M = .57$ $[0.43, 0.71]$) than low-value choices ($M = .29$ $[0.18, 0.41]$), $t(22) = 5.09$, $p < .001$, $g = 1.04$ $[0.41, 1.68]$, $BF_{10} > 150$.
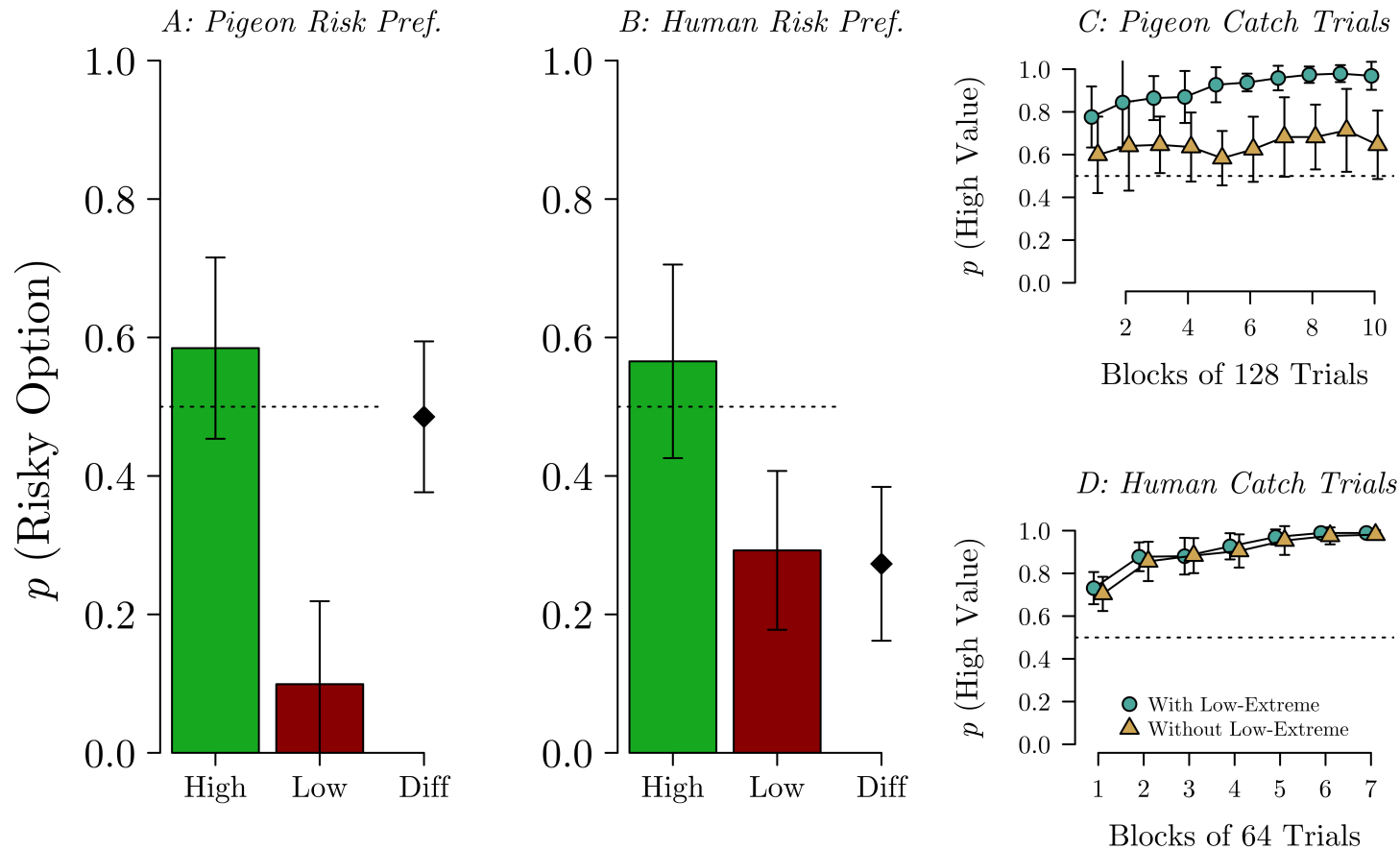
Figure 4.2: Probability of choosing the risky option for (A) pigeons and (B) humans on high- and low-value risk preference trials in Experiment 1. The mean difference is shown on the right of each plot. Catch trial learning curves for (C) pigeons and (D) humans in Experiment 1 that passed the 60% threshold. The curves depict the proportion of choices made to the high-value option for catch trials containing the low-extreme outcome (i.e., possibility of receiving zero) and catch trials without the low-extreme outcome (i.e., no possibility of receiving zero). All error bars indicate 95% confidence intervals. Dotted lines indicate chance levels.

Figure 4.2 depicts aggregated learning curves for pigeons (C) and humans (D) on catch trials. Pigeons chose the high-value option more often when the low-value option sometime led to the lowest extreme (i.e., the option leading to 0 or 2) than when that lowest extreme was not possible (i.e., the option always leading to 1). Thus, catch trials containing an option that sometimes provided no food reward were learned the most readily. For pigeons, this difference between the two types of catch trials was significant; $t(5) = 4.35$, $p = .007$, $g = 1.64$ $[0.15, 3.13]$, $BF_{10} = 8.7$. The human catch-trial data revealed no such differentiation, as near-ceiling effects were observed across all types of catch trials; $t(22) = 1.16$, $p = .257$, $g = 0.24$ $[-0.36, 0.84]$, $BF_{10} = 0.4$.

4.3.5. *Discussion* In this novel operant task, pigeons and humans exhibited similar levels and patterns of risk preference. Both species selected the risky option more often for high-value than for low-value choices, as though they were overweighting the extreme outcomes (Ludvig, Madan, & Spetch, 2014). The possibility of receiving 4 rewards seemed to pull preference towards the risky high-value option, whereas the possibility of receiving 0 rewards pushed preference away from the risky low-value option. These results replicate and extend the core findings of Ludvig, Madan, Pisklak, and Spetch (2014) to a novel experimental domain.

The catch-trial performance, however, suggests that pigeons' choices might have been driven more strongly by an avoidance of the zero reward than humans. This observation raises the possibility that the risk aversion exhibited by pigeons on low-value trials may have been driven at least in part by an explicit avoidance of a zero reward (also in Ludvig, Madan, Pisklak, & Spetch, 2014). Humans did not show the same asymmetry in their catch-trial performance, and other experiments with humans have shown that eliminating the zero, or no-reward outcomes, does not qualitatively change the results (Ludvig, Madan, & Spetch, 2014). Thus, humans may be more sensitive to any extreme outcomes, whereas pigeons may be more sensitive to zero outcomes. Experiment 2 tested this possibility.

4.4. **Experiment 2: Non-Zero Outcomes**

This experiment tested risk preferences for high- and low-value rewards, but with each value increased by one to make the lowest possible value a non-

zero outcome. The experimental design was similar to Experiment 1, but the task for pigeons differed. Because catch-trial performance of Experiment 1 and initial pilot data suggested that pigeons had difficulty learning the necessary discriminations with all non-zero values in an operant procedure, the foraging procedure described in Ludvig, Madan, Pisklak, and Spetch (2014) was used. Humans were tested with the same operant procedure as in Experiment 1.

4.4.1. *Experiment 2a: Pigeons*

Subjects. Six pigeons (*Columba livia*) (three Racers and three Silver King), all with substantial learning histories, but no experience in risky-choice experiments were housed and maintained as in Experiment 1. Each pigeon was run daily, five days a week, at the same time each day.

Apparatus. Figure 4.3 shows a photo of the experimental arena. The arena consisted of two compartments separated by a 50.8-cm central wall and enclosed by 91.4-cm front and rear walls and a single 82.5-cm side wall. The central and rear walls were built of 1.27-cm thick plywood painted white. The front and side walls were made from thin, white corrugated plastic. The arena floor was layered with aspen-chip bedding.
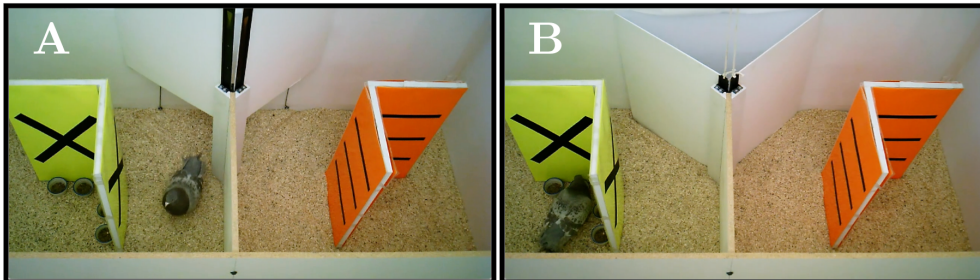
Figure 4.3: The testing arena during an example trial in Experiment 2a. (A) Pigeon entering the decision area via the open doors. (B) Pigeon eating from the food cups concealed behind the choice stimulus.

Two 44.45 × 63.50 cm white corrugated plastic entry doors, each at a 45° angle from the central wall, formed a small, triangular *decision area* at the front of the arena. Doors opened via a pulley system. A 16.51 × 20.32 cm entrance in the wall of the decision area led to a plastic 48.26 × 39.37 × 101.60 cm start box. Pigeon behaviour was monitored and recorded by a closed-circuit camera mounted centrally on the ceiling.

Stimuli. Choice panels consisted of two 30.48 × 58.42 cm planks of 2.54 cm thick plywood set at a 90° angle and covered with laminated paper of distinct colours (green, orange, purple, and yellow) and designs (a hollow black triangle, three horizontal black lines, four white squares, and a black "×"; see Figure 4.3 for an example). Two additional solid white training panels were used in Phase 2. The panels were centered inside the compartments, and food rewards were concealed behind each choice panel. Each individual food reward consisted of two Mazuri Gamebird pellets placed on top of grit inside a 6.99-cm diameter ceramic cup.

4.4.2. *Procedure.* Preliminary training occurred over several days in three phases:

Training Phase 1. Pigeons were trained via successively reinforced approximations to enter each compartment from the start-box through either entry door (randomly chosen), walk to the furthest corner to obtain three food rewards, and then return to the start box. This lasted for a mean of 4.33 and range of 2-8 sessions.

Training Phase 2. A white choice panel was placed in the centre of each compartment and pigeons were gradually shaped to walk around the panel to obtain food and return to the start-box. Phase 3 began once pigeons reliably completed 16 trials within a session, with all food rewards concealed behind the panel. This lasted for a mean of 4.83 and range of 2-8 sessions.

Training Phase 3. Pigeons completed 64 single-option trials randomly distributed over four 16-trial sessions. To start a trial, the door to one of the compartments opened, and the barrier into the decision area from the start-box was removed. Once the pigeon entered a compartment, the door was closed to prevent re-entry into the decision area until all food rewards were consumed. The door was then re-opened allowing the pigeon to return to the start-box, which now contained a food cup with 2 pellets.

The four panels indicated a risky or safe outcome of either high or low value. The safe high-value panel always indicated 4 rewards, and the risky high-value panel indicated a 50/50 chance of 3 or 5 rewards. The safe low-value panel always indicated 2 rewards, while the risky low-value panel indicated a 50/50 chance of 1 or 3 rewards. Panel presentation was counter-balanced and randomized across the 64 trials such that each panel and reward occurred

equally often on each side.

Testing. Testing lasted for 28 sessions that each contained four single-option trials, which exposed the bird once to each of the four panels, and 12 choice trials, which consisted of four *risk-preference trials* and eight *catch trials*. On risk-preference trials, pigeons chose between safe and risky options of the same expected value (e.g., risky high-value vs. safe high-value). Catch trials provide a choice between one high-value (risky or safe) and one low-value panel (risky or safe). Each session had two of each type of catch trial. Sessions were counterbalanced so that each stimulus appeared twice in both the right and left compartments on choice trials. Ordering of the trials was randomized each session.

The testing procedure was identical to Phase 3 except that both entry doors opened simultaneously on each trial. On single-option trials, one of the compartments was left empty. If the pigeon entered the empty side, both guillotine doors were left open until the pigeon entered the side with a panel.

4.4.3. *Experiment 2b: Humans*
Participants and Methods.

Group 1. Forty-four participants (aged 20.3 ± 2.3 years old; 35 females) were recruited from the same participant pool as Experiment 1b; no individuals participated in more than one experiment. In this experiment, the risky high-value option produced 3 or 5 rewards, and the risky low-value option produced either 1 or 3 rewards. The safe high-value option guaranteed 4 rewards, and the safe low-value option guaranteed 2 rewards (see Figure 4.4 for a schematic of the reward contingencies and spatial layout). Except for the outcome values, Experiment 2b employed identical methods as those used in Experiment 1b. Twenty of the forty-four participants were excluded from the analysis due to failure to obtain 60% correct on catch trials.

Group 2. A second group of 24 participants (aged 19.4 ± 1.5 years old; 13 females) was tested because so many participants in the first group failed the catch trials. These participants were given the same procedure but received more explicit instructions about the choice task (see Appendix A). All Group 2 participants passed the catch trials and were included in the analysis.

4.4.4. *Data Analysis* The data analysis was conducted as per Experiment 1.
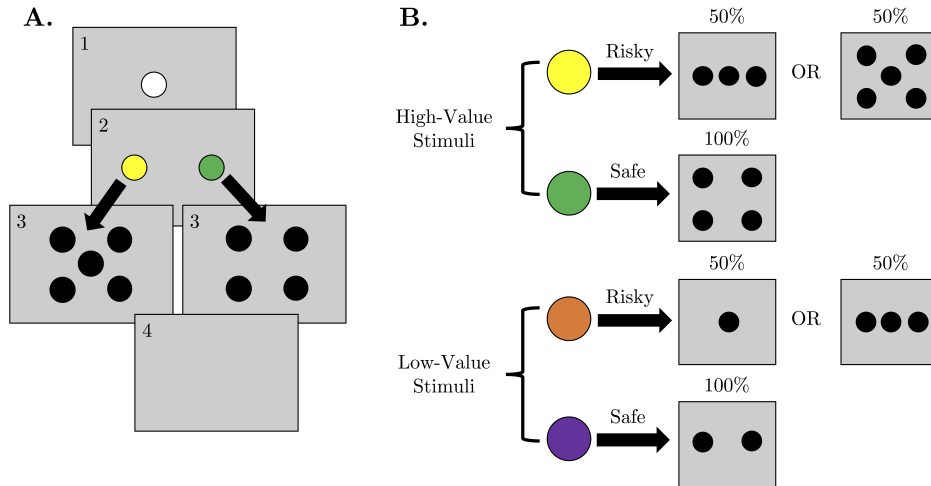
Figure 4.4: Schematics of the procedure for Experiment 2b. (A) Choice trial procedure, displaying a potential outcomes from a risky high-value choice and a safe high-value choice. (B) Illustration of the reward contingencies in effect. The colour of the choice circles relative to their outcome was counterbalanced across participants according to four different combinations of yellow, green, orange, and purple.

4.4.5. *Results*

Pigeons. Figure 4.5 shows the mean proportion of risky choices for high- and low-value rewards along with the 95% confidence interval of the paired difference. In contrast to the findings of Ludvig, Madan, Pisklak, and Spetch (2014, see Figure 4.5A) and Experiment 1, pigeons (Figure 4.5B) showed little difference in risk selection between the high-value ($M = .44$; 95% CI $[.36, .53]$) and the low-value ($M = .44$ $[.35, .53]$) choice types, $t(5) = 0.10$, $p = .928$, $g = 0.04$ $[-1.25, 1.32]$, $BF_{10} = 0.37$. Catch-trial performance (Figure 4.5D) showed no clear difference between choices with and without the low-extreme outcome, $t(5) = 1.68$, $p = .154$, $g = 0.63$ $[-0.69, 1.95]$, $BF_{10} = 0.94$.

Humans. In contrast to the pigeons, humans (see Figure 4.6) who passed the catch-trial criterion in Group 1 chose the risky option significantly more frequently for high-value choices ($M = .62$; 95% CI $[.50, .74]$) than low-value choices ($M = .39$ $[.27, .51]$), $t(23) = 2.99$, $p = .006$, $g = 0.60$ $[0.01, 1.20]$, $BF_{10} = 6.97$. Humans in Group 2 exhibited similar risk preferences, selecting the risky option more frequently for the high-value ($M = .58$ $[.42, .74]$) than the low-value choices ($M = .38$ $[.25, .50]$), $t(23) = 2.57$, $p = 0.017$, $g = 0.52$

Pigeon Risk Preference:
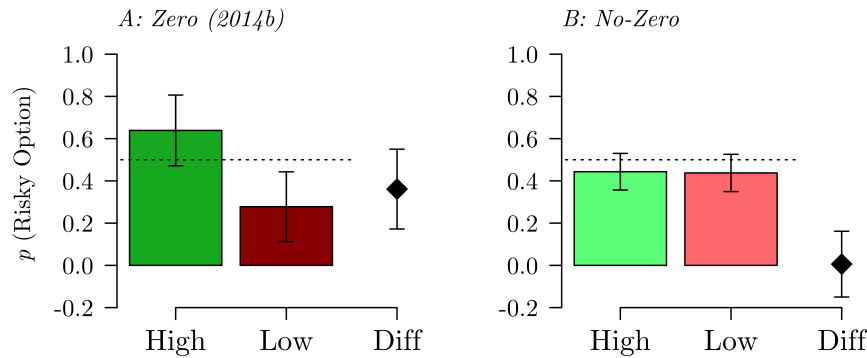
Pigeon Catch Trials:
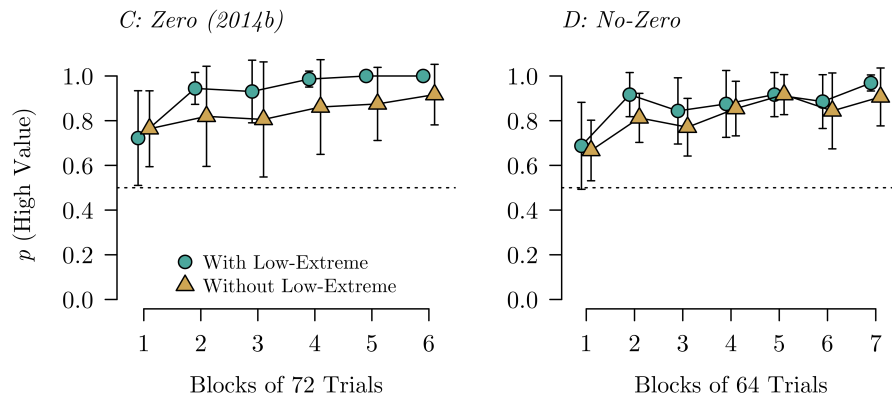
Figure 4.5: Pigeon choice data from Ludvig et al. (2014) and Experiment 2. (A-B) Proportion of risky option selections for high- and low-value choice trials. The mean difference is shown on the right of each plot. (C-D) Proportion of choices made to the high-value option for catch trials with the low-extreme and catch trials without the low-extreme. All error bars indicate 95% confidence intervals. Dotted lines indicate chance levels.

$[-0.07, 1.11]$, $BF_{10} = 3.09$. Appendix B provides further detail about catch-trial performance in Groups 1 and 2.

4.4.6. *Discussion* This experiment demonstrated a difference in how pigeons and humans make risky choices. When the outcomes did not include any zero values, humans continued to overweight the extreme outcomes (selecting

Human Risk Preference:

A: No-Zero Group 1

B: No-Zero Group 2

Human Catch Trials:

C: No-Zero Group 1

D: No-Zero Group 2
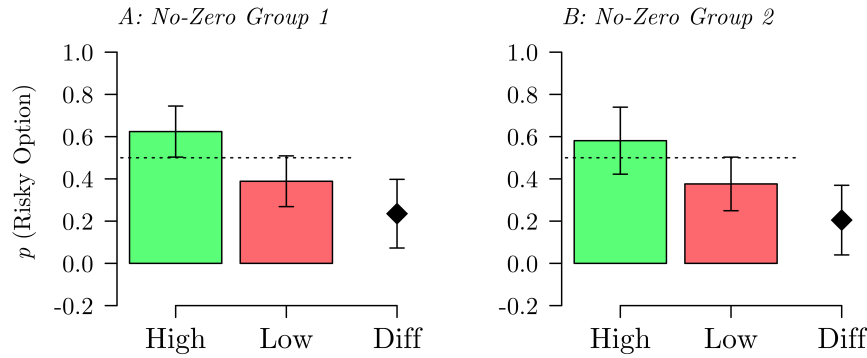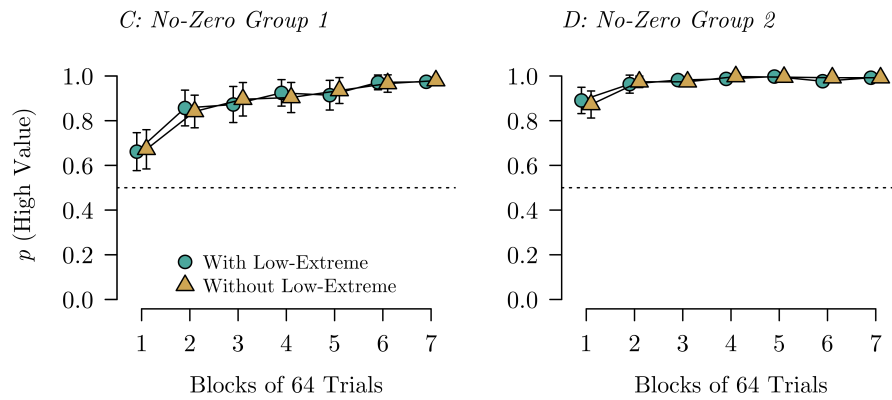
Figure 4.6: Human choice data Experiment 2. (A-B) Proportion of risky option selections for high- and low-value choice trials. The mean difference is shown on the right of each plot. (C-D) Proportion of choices made to the high-value option for catch trials with the low-extreme and catch trials without the low-extreme. All error bars indicate 95% confidence intervals. Dotted lines indicate chance levels.

the risky option more for high values), whereas the pigeons did not. Note how these pigeons were tested in the identical open-field procedure as in Ludvig, Madan, Pisklak, and Spetch (2014), who found sensitivity to extremes with zero outcomes in pigeons (see Figure 4.5A). In the current study, the pigeons (unlike humans), did not show sensitivity to these extremes when there were

no zero outcomes. This pattern of results suggests that humans may be more sensitive to extreme outcomes and pigeons may be more sensitive to zero outcomes.

4.5. **Experiment 3: Pigeons**

In view of the large methodological differences between the procedures used for pigeons in Experiment 1a (operant) and 2a (open field), Experiment 3 tested groups of pigeons with and without zero outcomes using a refinement of the operant procedure from Experiment 1a.

4.5.1. *Subjects* Subjects consisted of 16 racing pigeons (*Columba livia*). All, except one pigeon, had no prior experience in risky-choice experiments. One bird, 473, had previously been in Experiment 1. Pigeons were housed in 165 × 69 × 178 cm group cages (5-8 birds per cage) in a temperature-controlled colony room on a 12-hour light-dark cycle with free access to grit and vitamin-enriched water. Experimental sessions were run daily, six days a week, at the same time each day. Birds were maintained at approximately 80% of their free-feeding weight by post-session feeding.

4.5.2. *Apparatus* Six custom-built 29.2 × 35.6 × 38.1 cm operant-conditioning chambers, equipped with the same monitors and touch screens used in Experiment 1a, were located inside a closed room. All the chambers contained an opaque barrier that blocked the lower third of touch screen from incidental contact by the pigeon's feet and chest. White noise fed through the monitor's built-in speakers maintained the internal sound at 65db. Two feeding ports located on the side walls contained the same food hoppers described in Experiment 1a. As in Experiment 1a, each chamber was connected to a computer, located in an adjacent room, running E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA).

4.5.3. *Stimuli* Experimental stimuli were six coloured circles 110 pixels (approx. 3 cm) in diameter presented against a grey background. All circles were aligned horizontally at the same vertical location on the screen (see Figure 4.7). Circles were filled with white, black, blue with a single black vertical line, orange with a hollow black circle, green a solid black "×", or purple with three black horizontal lines.
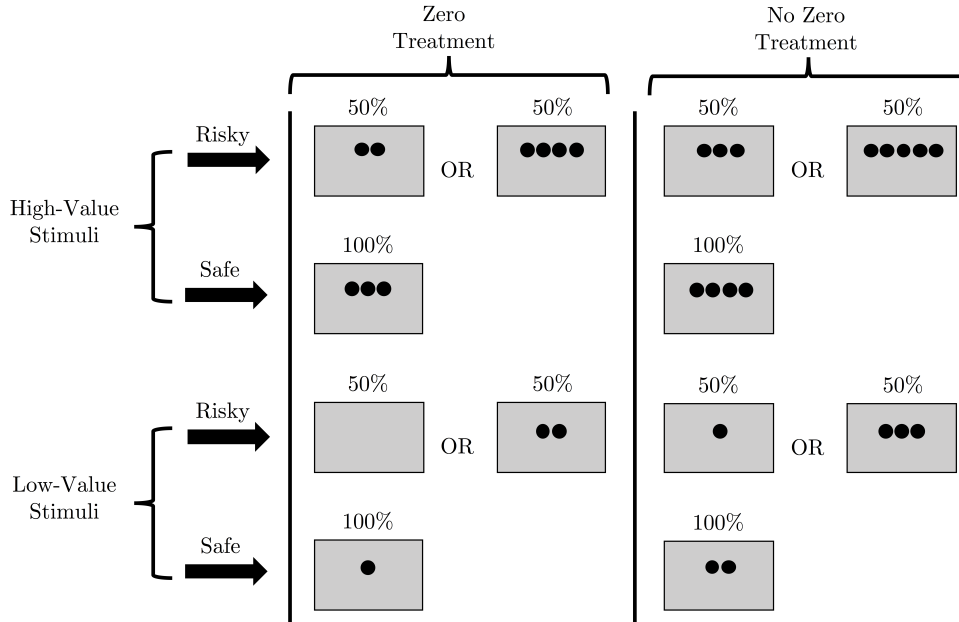
Figure 4.7: A schematic illustrating the various reward contingencies in effect for the Zero and No-Zero groups in Experiment 3.

4.5.4. *Procedure*

Training Phase 1. Pigeons were given three 90-minute sessions of an autoshaping procedure with a Fixed-Ratio (FR) 1 contingency built in. One randomly selected (without replacement) coloured circle was presented for 10 s. The circle remained on until the bird made a single peck or until 10 s elapsed at which time the circle disappeared and one randomly-selected hopper was raised, providing 1-s access to food, and then a 240-s ITI ensued.

Training Phase 2. This phase provided an FR 1 contingency with a 2-s ITI; each circle remained onscreen until it was pecked or the session timed out. Sessions contained 72 trials, and the pigeon remained in the chamber for 90 minutes. Phase 2 lasted until a pigeon completed all 72 trials, with an average response time of less than 30 s on each circle type, for a single session. Phase 2 lasted for a mean of 1.12 and range of 1-2 sessions.

Training Phase 3. Pigeons were randomly assigned to either a *Zero* group (n=8) or a *No-Zero* group (n=8). For the Zero group, the choice circles provided values equivalent to those used in Experiment 1a. For the No-Zero group,

each value was increased by 1 (see Figure 4.7) for both this phase and for the testing phase. Procedurally, this phase was the same as Training Phase 4 of Experiment 1a except that it consisted of only one session of 64 single-option trials.

Testing. Testing lasted for 30 sessions, each consisting of 16 single-option trials and 48 choice trials (16 risk-preference choices and 32 catch choices), with full counterbalancing of stimuli and values within each session presented in a randomized order. Apart from the number of trials per session, testing proceeded as in Experiment 1a.

4.5.5. *Data Analysis* Main effects and interactions between choice (high- and low-value) and Group (Zero and No-Zero) manipulations were analyzed using linear mixed-effects modelling on the mean proportion of choices made to the risky alternative during the last third of all trials. Models were fit by maximum likelihood with subjects treated as a random effect and computed with the nlme R package for R 3.5.1 statistical software (Pinheiro et al., 2018; R Core Team, 2018). An effect size in terms of r^2 is reported for each fixed-effect in the full model. Bayes Factors (BF_{10}) are also provided for each effect and tested against a null model that included only the intercept and an additive random effect of subject.

Simple-effects analysis on the difference between the high- and low-value choices for the Zero and No-Zero groups were analyzed using a paired t-test. For each simple effect, a corresponding effect-size (g) and JZS Bayes factor (BF_{10}) is provided as per Experiment 1. Similar sets of analyses are provided for catch-trial assessment.

4.5.6. *Results* Three birds in the No-Zero group failed to meet the 60% catch-trial criterion and were subsequently removed from the analysis; none of the birds in the Zero group failed to reach the criterion. Figure 4.8 shows the mean proportion of risky option selections for high- and low-value choice types along with the 95% confidence interval of the paired difference for both the Zero and No-Zero groups. Although both the Zero and No-Zero groups showed greater selection of the risky option for high-value outcomes, $M = .48$ 95% CI [.35, .61] and $M = .57$ [.35, .79], than low-value outcomes, $M = .10$ [.04, .17] and $M = .40$ [.24, .57], this difference was larger and less variable in the Zero group. Statistical analysis showed a significant main effect of both choice,

$\chi^2(1) = 15.88$, $p < .001$, $r^2 = .25$, $BF_{10} > 150$, and group, $\chi^2(1) = 9.68$, $p = .002$, $r^2 = .11$, $BF_{10} > 150$, as well as a significant interaction between these two variables, $\chi^2(1) = 4.00$, $p = .045$, $r^2 = .25$, $BF_{10} > 150$. Consistent with our previous findings, simple-effects analysis showed a significant difference between the high- and low-value outcomes in the Zero group, $t(7) = 5.33$, $p = .001$, $g = 1.78\ [0.51, 3.05]$, $BF_{10} = 38.42$, but not in the No-Zero group, $t(4) = 2.01$, $p = .115$, $g = 0.81\ [-0.71, 2.33]$, $BF_{10} = 1.24$. These results confirm that there was a larger and more reliable difference between the high- and low-value choices in the Zero group than the No-Zero group.

Figure 4.8B and Figure 4.8C show catch-trial performance for the Zero and No-Zero groups. Similar to Experiment 1a, the Zero group was more accurate for choices that contained the low-extreme value (i.e., a possibility of zero). Similar to Experiment 2a, however, this difference in catch-trial type was not observed in the No-Zero group (where zero was not a possible outcome). Statistical analysis confirmed a significant interaction between catch-trial type and group, $\chi^2(1) = 6.54$, $p = .011$, $r^2 = .40$, $BF_{10} > 10.91$. Analysis of the simple-effects revealed a large difference between choices with and without the lowest extreme for the Zero group; $t(7) = 4.08$, $p = .005$, $g = 1.13\ [-0.03, 2.28]$, $BF_{10} = 11.88$, but not for the No-Zero group; $t(4) = 0.37$, $p = .732$, $g = 0.19\ [-1.27, 1.66]$, $BF_{10} = 0.42$.

4.5.7. *Discussion* Experiment 3 manipulated the occurrence of zeros between two groups of pigeons tested with an identical operant task. Similar to Experiment 1a, the Zero group showed a pronounced difference between high- and low-value choices, whereas this difference was significantly smaller and not statistically reliable for the No-Zero group. Additionally, whereas the Zero group showed clear risk aversion for low-value choices, neither the high- nor the low-value choices in the No-Zero group differed significantly from chance levels. These observations, in conjunction with the imbalances between catch trials containing risky low-value and safe low-value outcomes, support the conclusion that pigeons are more affected by a general avoidance of zero than by an overweighting of the extremes.
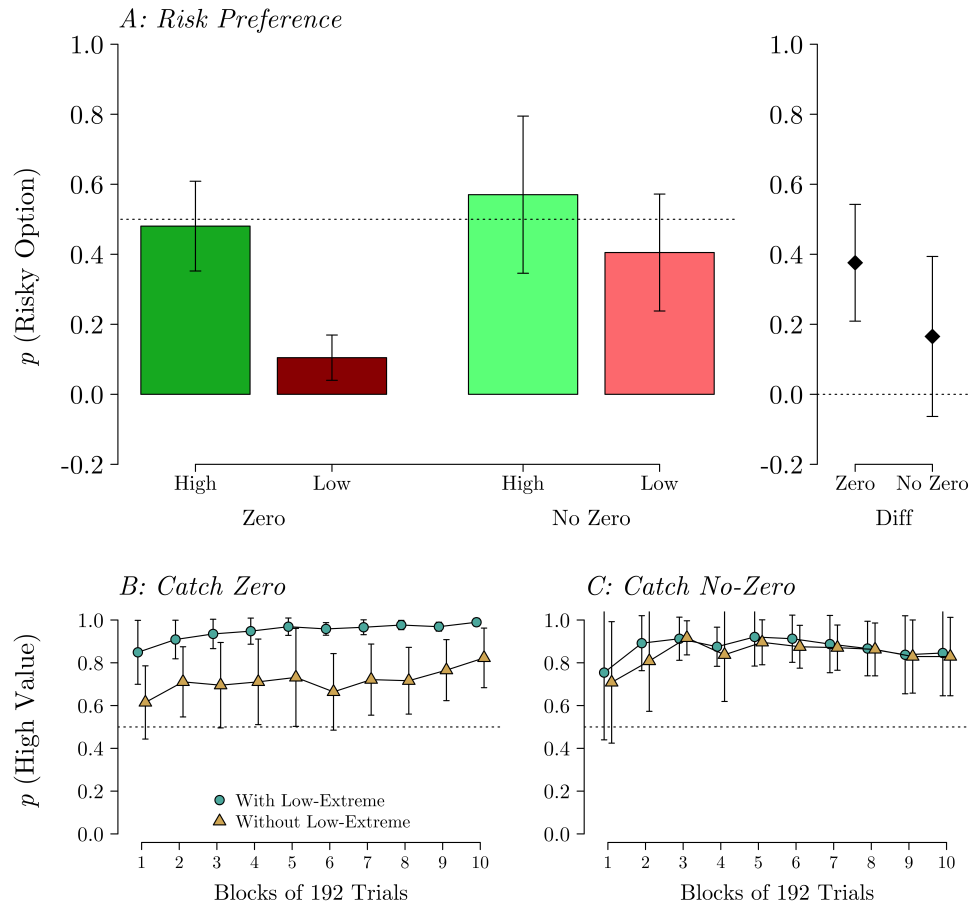
Figure 4.8: Choice data from the Zero and No-Zero groups in Experiment 3. (A) Probability of choosing the risky option for pigeons on high- and low-value risk preference trials. The right panel shows the mean difference between high- and low-value proportions for both groups. Catch trial learning curves for the (B) Zero and (C) No-Zero groups for pigeons passing the 60% threshold. The curves depict the proportion of choices to the high-value option for catch trials with and without the lowest extreme outcome. All error bars indicate 95% confidence intervals. Dotted lines indicate chance levels.

4.6. **Experiment 4: Discrimination Test**

In Experiments 2 and 3, pigeons failed to show a significant extreme-outcome effect when none of the outcomes included a zero value. This insensi-

tivity to the extremes may be due to an inability of the pigeons to discriminate between the outcomes when there are no zeros. Although the catch-trial data suggested that pigeons did not completely fail to discriminate, the catch trials did not explicitly test a pigeon's ability to discriminate between all possible outcomes (e.g., 4 versus 5). Experiment 4 sought to verify the conclusions of Experiments 2 and 3 by first pre-testing pigeons' ability to discriminate between all combinations of the No-Zero choice outcomes used in Experiment 3 (i.e., 1-5) and then testing those same pigeons using a variant of the Experiment 3 No-Zero procedure.

4.6.1. *Methods*

Subjects and Apparatus. Subjects consisted of 8 racing pigeons (*Columba livia*) with no prior experience in risky-choice experiments. Subjects were maintained as per Experiment 3 and run in the same experimental chambers described in Experiment 3.

Reinforcer Magnitude Discrimination. Similar to Experiment 3, pigeons were presented with coloured choice circles that led to a horizontal array of token circles, each of which required a single peck to obtain 1 s of access to food. The procedure was modified, however, to enhance discriminability of the alternatives. Five choice circles, 2 cm in diameter, were presented in a horizontal array just beneath the location of the token array (see Figure 4.9). Each choice circle had a fixed spatial location that was demarcated by a hollow black circle on a grey background. Each choice circle was assigned a unique colour and had a number centrally inscribed, in black Times font, that corresponded to its outcome. For instance, if a bird selected a dark green circle inscribed with a "3", that circle would disappear, and three corresponding dark green tokens would appear above it, each of which could be pecked for 1-s access to food in any order. The colour of the tokens always matched the respective choice circle that was selected, but no number appeared inside. The outcome of each choice circle was fixed to a particular reward amount (not probabilistic), and the values ranged from 1 to 5.

Two different space-colour mappings were used for the choice circles. From the leftmost circle to rightmost circle, the colours for one mapping were golden brown, light blue, dark green, dark yellow, and magenta. For the other spatial mapping, these same colours were offset by 2 positions to the right. Two
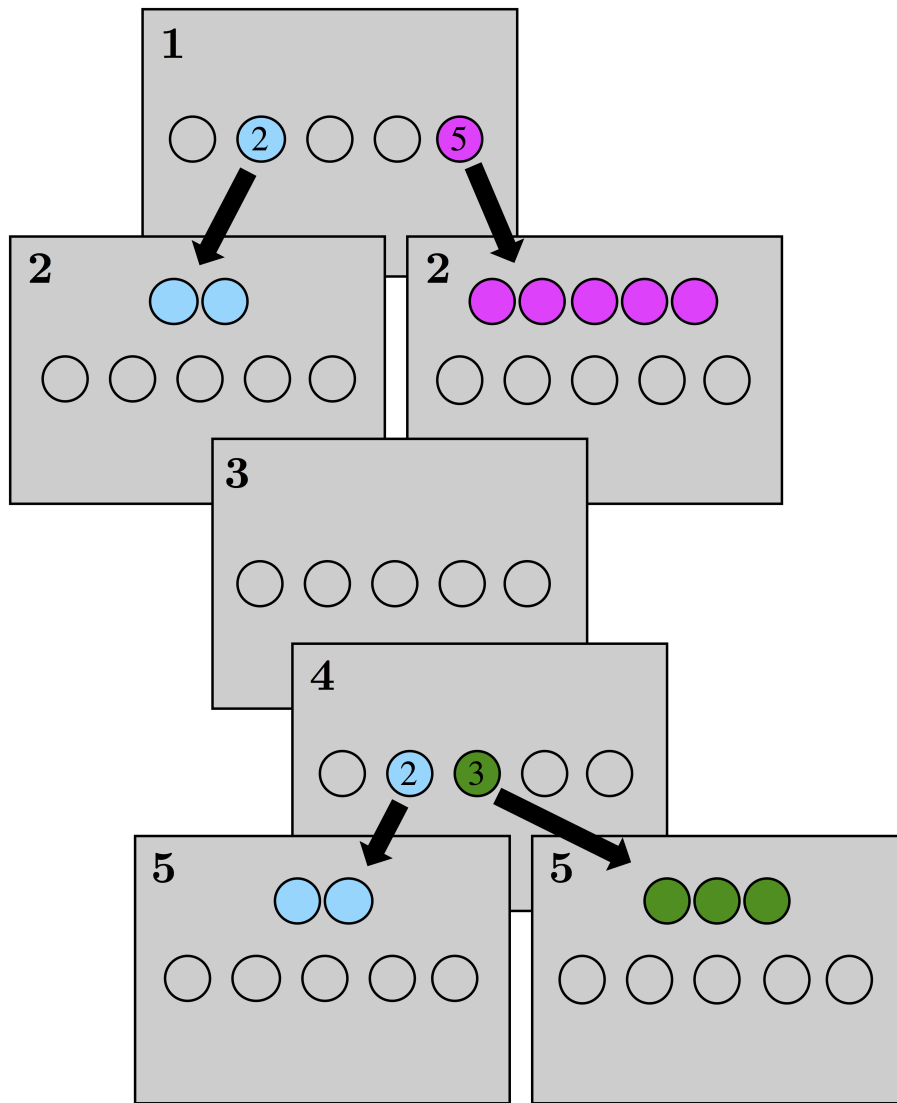
Figure 4.9: A schematic of two choice trials on Experiment 4's reinforcer amount discrimination training. Panel 1 and 4 show two of ten possible choice scenarios pigeons were given. Panel 2 and 5 show the resulting token outcomes of each choice. Panel 3 shows the inter-trial interval.

different spatial-number mappings were also used, and, from left to right, these were 1, 2, 3, 4, 5, and 4, 5, 1, 2, 3. Birds were randomly assigned to the resulting 4 colour-number combinations. Unlike the previous experiments, no

start-stimulus was used in Experiment 4.

Initially pigeons were tested on an FR1 condition that required only a single peck to select a choice stimulus. This produced weak discrimination across the full range of values (see Results section). Consequently, the birds were re-tested with the ratio requirement on the choice stimuli increased to FR40 because increased ratio requirements have been shown to improve accuracy in both simple and complex discrimination tasks (e.g., Elsmore, 1971; Wilkie & Spetch, 1978). Upon beginning the FR40 condition, all but one pigeon (bird #76, described below) were reassigned to new colour and number locations to require relearning of the contingencies. Training and testing occurred as per Experiment 3, except that Phase 2, Phase 3, and testing consisted of 65 trials per session to allow for complete within-session balancing of the stimuli. Each session consisted of 15 single-option trials and 50 dual-choice trials. All choice combinations across the outcome values were tested (e.g., 5 vs. 4, 5 vs. 3, 5 vs. 2, 5 vs. 1, 4 vs. 3, etc.) in a randomized fashion.

The FR1 condition included training Phases 1, 2, and 3; however, the FR40 condition only included training Phase 3. During Phase 3 of the FR40 condition, the ratio requirement was shaped according to a geometric progression: 1, 2, 4, 6, 9, 12, 15, 20, 25, 32, 40, with the value increasing every five trials. Pigeons remained on Phase 3 with an FR40 requirement on all trials until they completed all 65 trials within a session or three sessions had elapsed. In testing, because the high ratio requirement caused some pigeons to not complete the full 65 trials per session, each pigeon was tested until they had completed at least 1500 choice trials across the sessions. The mean number of testing sessions was 36, ranging from 30 to 55.

Risky Choice. Following the discrimination testing, pigeons were moved to the risky-choice procedure. The procedure was the same as the No-Zero group in Experiment 3, except with an FR40 response requirement. Training phases 1 and 2 were included. Four new colours were used for the choice circles and tokens: red, yellow, green, blue. Each choice circle was also inscribed with a specific letter A, B, C, or D, in black Times font. The tokens matched the colour of the selected choice stimulus and were presented above the selected circle. The choice stimuli were mapped onto a risky-high outcome, a risky-low outcome, a safe-high outcome, and a safe-low outcome. The location-colour, location-letter, and location-outcome mappings were all randomized

independently according to a Latin-square design. Each bird was randomly assigned to a specific mapping on the condition that blue and yellow circles did not correspond to the same values received during the FR40 discrimination testing. Each session consisted of a maximum of 64 trials: 16 single-option, 16 choice, and 32 catch trials in a randomized order. With the exception of one pigeon (#76), testing continued until each pigeon had completed 1440 choice trials. Bird #76, who ran for 123 sessions with a mean of 11 choice trials per session, only completed 1294 choice trials before the experiment was ended. Across all other pigeons, the mean number of sessions was 32, ranging from 30 to 40.

Data Analysis. Analyses for risky choice proceeded as in the earlier experiments. For the discrimination task, we first compared the proportion correct for the discriminations that were only a single unit apart (e.g., 3 vs. 4), which should be the hardest discriminations for pigeons. To better quantify the pigeons' discrimination across the 10 possible choice types in the FR1 and FR40 conditions, we next evaluated relative preferences in terms of the generalized matching equation (GME) as applied to reinforcer magnitude (Baum, 1974; M. Davison & Baum, 2003). Given the ubiquity of matching in choice and the straightforward application and interpretation of generalized matching relations (de Villiers, 1977), the GME offers a useful lens with which to evaluate pigeons' choices across the full range of outcomes used in Experiment 4. Equation 4.1 shows the GME applied to reinforcer magnitude:

$$\log_{10} \frac{B_1}{B_2} = a \cdot \log_{10} \frac{M_1}{M_2} + \log_{10} c \qquad (4.1)$$

where B is the number of responses emitted to a particular choice alternative, and M is the total amount of food delivered for a given choice outcome. Subscripts refer to the different options (1 or 2). The intercept parameter c, called bias, is a measure of preference for B_1, relative to B_2, independent of the reinforcer magnitudes. A positive bias indicates a general preference toward the higher reward magnitude, independent of the ratio of the items. The slope parameter a, referred to as sensitivity, indexes how the log response ratio grows with the log reinforcer magnitude. Perfect matching to the log magnitude ratios should result in bias and sensitivity values of 0 and 1 respectively.

The GME was fit to the last third of each pigeon's data for both the

FR1 and FR40 condition using least-squares linear regression with R 3.5.1 software (R Core Team, 2018). Because pigeons sometimes exclusively chose one option, a correction of 0.1 was added to the numerator and denominator of each pigeon's response ratio to prevent the occurrence of zero values. Mean bias (intercept) and sensitivity (slope) parameters are reported with corresponding 95% confidence intervals.

4.6.2. *Results* Figure 4.10A shows the probability of selecting the high-value option for the four choices that contained outcomes only 1 token apart. Across all 8 pigeons, none of the four choice types in the FR1 condition exceeded chance levels, whereas, in the FR40 condition, all but one choice type (4 vs 3) exceeded chance. Figure 10B displays the overall fit of the GME applied separately to the FR1 and FR40 conditions across all tested choices. For the FR1 condition, a very slight bias of 0.10 95% CI $[-0.40, 0.61]$ and sensitivity of 1.28 $[0.04, 2.51]$ to log reinforcer magnitude ratio was observed, neither of which exceeded what would be reasonably expected by perfect matching. By contrast, noticeable overmatching was observed in the FR40 condition, as indicated by the high sensitivity of 2.59 $[1.19, 3.99]$ to the log reinforcer magnitude ratio. Thus, as the log reinforcer magnitudes became more discrepant, the pigeons' preference became even more pronounced than would be expected had they just been matching relative responding to relative reinforcer magnitude. More critically, however, the FR40 condition also produced a heavy bias of 0.76 $[0.22, 1.30]$ towards choices containing the larger reinforcement value that was significantly larger than would be predicted by either indifference or strict matching.

During risky-choice testing, only one bird (#76) failed to reach the 60% catch-trial criterion and was excluded from the remaining analyses. Figure 10C shows the mean proportion of risky option selections for high- and low-value rewards and the 95% confidence interval of the paired difference. Similar to the No-Zero groups in Experiments 2a and 3, there was no significant difference in the proportion of risky choices between the high-value ($M = .41$; 95% CI $[0.09, 0.74]$) and low-value ($M = .34$; $[0.16, 0.52]$) choice types, $t(6) = 0.41$, $p = .697$, $BF_{10} = 0.38$, $g = 0.14$ $[-1.02, 1.31]$. Catch-trial performance (see Figure 4.10D) showed no clear difference between choices with and without the low-extreme outcome, $t(6) = 0.82$, $p = .445$, $g = 0.29$ $[-.88, 1.46]$, $BF_{10} = 0.46$.
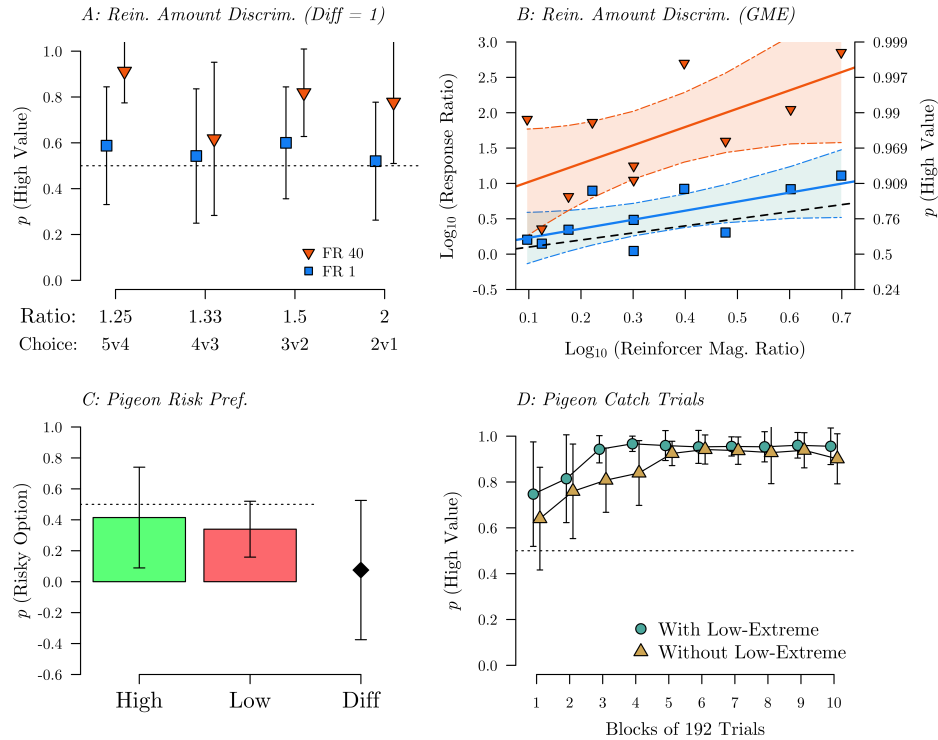
Figure 4.10: Panels A and B show the results of Experiment 4's discrimination testing in all pigeons (n=8). (A) The probability of selecting the high-value option for choices that differed by only one token. (B) The matching relation between the log response ratio (high/low) on the left y-axis and the log reinforcer magnitude ratio (high/low) on the x-axis. Right-side y-axis shows proportion of high-value responses on a log scale. Blue and orange regression lines indicate FR1 and FR40 condition respectively, dashed line indicates perfect matching. Panels C and D show the results of risky-choice testing from pigeons that passed the catch trial criterion (n= 7). Error bars and regression confidence bands are calculated at a 95% level. Dotted lines indicate chance levels.

4.6.3. *Discussion* Experiment 4 confirmed that even when pigeons can adequately discriminate between non-zero outcomes they still show little evidence for an effect of the extreme outcomes on their risky choices. These findings complement the results from Experiment 2 and 3 and suggest that pigeons are more driven by the avoidance of zero outcomes than by avoidance of a low extreme.

The ratio requirement also played a significant role in determining the degree of discrimination exhibited by the pigeons. When choice required only a single response (FR1), pigeon preferences seemed to follow basic predictions of matching, with weak discrimination at small ratios (e.g. 5:4) and progressively stronger discrimination at larger ratios (e.g. 5:1). When the response requirement was increased to 40, bias in favor of the choice alternative with the larger reward outcome became very pronounced. This increase in bias was accompanied by an additional increase in overall sensitivity to the reinforcement ratio. The enhanced discrimination of reinforcer magnitude when the response requirement was increased is consistent with previous findings that discrimination becomes more accurate as response effort increases (e.g., Elsmore, 1971; Wilkie & Spetch, 1978).

Given that pigeons demonstrated good discrimination of the reward values with the FR40 requirement, Experiment 4 evaluated the predictions of the extreme-outcome effect under this condition with no zero outcomes present. Even with a high ratio requirement and high levels of discrimination, pigeons still exhibited behaviour qualitatively similar to those seen in the No-Zero groups of Experiments 2 and 3: They had similar levels of risk preference for both high-value and low-value choices (see Figure 4.8). The clear preference for the larger reward outcome with FR40 reinforcer magnitude discriminations and the clear preference for the high-value outcome on catch trials suggest that the pattern of choice results was not due to an inability to discriminate between non-zero outcomes.

4.7. General Discussion

The present set of results demonstrate the importance of zero, or non-rewarded options, in determining risky choice in pigeons but not people. Whereas people avoided risky options that might lead to the worst outcome in a context, pigeons primarily avoided options that led to zero or no reward. When the zero value was removed as a possible outcome, humans continued to select the risky option more often for high-value options than low-value ones, consistent with previous findings (Ludvig, Madan, & Spetch, 2014). Pigeons, however, showed no consistent pattern in their risk preferences when the zero value was removed, despite accurate choice of high-value options over low-value options.

Together, these results suggest that, in contrast to our previous suggestion (Ludvig, Madan, Pisklak, & Spetch, 2014), the strikingly similar patterns of risky choice seen by pigeons and humans may be driven by different underlying mechanisms.

In humans, there is evidence that the greater risk seeking for high-value choices than for low-value choices is driven by an overweighting of extreme outcomes (Ludvig, Madan, & Spetch, 2014). For example, manipulating the set of outcomes in the decision context can shift risk preferences: for the same choice between a safe 20-point option and a risky option leading to 0 or 40 points, people's choices range from risk seeking to risk aversion depending on the other outcome values in the decision context (Ludvig, Madan, & Spetch, 2014; Madan et al., 2019). Moreover, memory tests given after the choice task reveal that people are more likely to recall the extreme outcomes first and to judge them as having occurred more frequently (Madan et al., 2014, 2017). Here, too, people were more risk seeking for high-value than low-value choices, even when no zeroes were present (Experiment 2a).

Pigeons, in contrast, were not driven by the relative extremes to the same degree. When zero (i.e., no reward) outcomes were present in Experiment 1 and 3, pigeons indeed selected the risky option more often in the high-value than low-value choices. When zero outcomes were removed in Experiments 2, 3, and 4, however, pigeons no longer showed consistently different patterns for high- and low-value choices. Thus, zero outcomes appeared to more strongly control behaviour in pigeons than people. Further research is needed to determine whether non-zero extreme outcomes play a role in pigeons' risk preferences in other situations. For example, performance could be directly compared between similar-valued choices with and without an extreme-outcome (as has been done with humans; see Madan et al., 2019). Here, in all 4 experiments, the risky options always potentially led to an extreme outcome (either high or low).

One possible reason that pigeons might be particularly driven by an avoidance of zero is suggested by the probability and delay discounting literature (e.g., Green, Myerson, & Calvert, 2010; Hayden & Platt, 2007; Mazur, 1989; Rachlin et al., 1986). When the low-value option provides either 2 or 0 food rewards, pigeons may treat this probabilistic reward as a variable delay to reward. For all other options some reward is provided on every trial, so, although

the amount of reward varies probabilistically, the occurrence of reward is never delayed. In essence, this design may resemble a self-control task whereby the pigeon chose between a smaller immediate reward, and a larger reward that sometimes occurred after a substantial delay. In these self-control tasks, pigeons typically show steep discounting functions with rewards losing most of their subjective value within a few seconds. By contrast, humans and other apes tend to exhibit much more gradual rates of discounting and are thus not subject to such a heavy early loss of value (Stevens & Stephens, 2009). This probabilistic interpretation could account for the observed differences in the pigeons' risky choice for high- and low-value options (e.g., Exp. 1a, Zero Group in Exp. 3; Ludvig, Madan, Pisklak, & Spetch, 2014). The pigeons' risk neutrality observed when all outcomes provided some reward (e.g., Exp. 2a, No-Zero Group in Exp. 3, and Exp. 4) is also consistent with this interpretation. Another possibility that merits future exploration is that zero may simply be functioning as an especially potent extreme (as opposed to delay) in certain species, such as pigeons, and certain contexts, such as all gains.

The pigeons' aversion to options that sometimes provide zero outcomes is interesting in relation to evidence from other procedures showing that pigeons are sometimes drawn to options that provide food less reliably (e.g., Belke & Spetch, 1994; Dunn & Spetch, 1990; Kendall, 1974; Pisklak et al., 2015; Stagner & Zentall, 2010). In those studies, choice was followed by a lengthy delay (10 s or greater) during which the stimuli present indicated whether the food would or would not occur. Stimuli which provided clear signals for food functioned as strong conditioned reinforcers of choice, whereas stimuli that provided ambiguous or redundant signals did not (McDevitt et al., 2016). Given this importance of the conditioned reinforcer, an interesting question is what would happen if the zero outcomes used in Experiment 1 were preceded by a unique token that provided nothing for its selection. In this case, the zero outcome would no longer be marked by the absence of any feedback cues. The strong effects of the zero outcome might become attenuated – or perhaps even reversed if the response requirement on them is greater as in studies of suboptimal choice.

One central difference between the pigeon and human experiments concerns the nature of the terminal reinforcer. For the pigeons it was food – an unconditioned reinforcer – and for the humans it was money–a highly general-

ized conditioned reinforcer. When pigeons and humans are given tokens that have to be immediately exchanged for unconditioned reinforcers, their sensitivity to risky delays is similar (Lagorio & Hackenberg, 2010). When both species are forced to accumulate tokens, thereby delaying the unconditioned reinforcer, risk sensitivity decreases. Thus, it might be predicted that forcing pigeons to accumulate tokens, as opposed to having them exchange them immediately, might bring their results into closer alignment to those observed for humans. Conversely, giving humans an immediate unconditioned reinforcer may foster preferences more like those observed for pigeons.

An interesting test of the power of zero values in driving choice would be to provide pigeons with a risky option that offers food values of 0 or 4 over a safe option that offers a value of 1. In this case, the risky option would be the better choice in terms of overall expected value, raising the question: would an avoidance of zero still dominate the pigeons' choice behaviour? In a two-alternative choice, the answer appears to be yes. For example, Menlove, Inden, and Madden (1979) gave pigeons a choice between a key that delivered a fixed 2 seconds of food and a key that gave pigeons either 8 or zero seconds of food. All pigeons preferred the fixed option, suggesting an avoidance of zero would dominate in this hypothetical experiment. What remains uncertain, however, is whether these results would generalize to a situation where the decision context contains additional higher-value choice alternatives (e.g., a choice between 0 or 4 versus 1 and a choice between 2 or 6 versus 3). Given that the current results suggest pigeons are not especially sensitive to the extreme outcomes in such contexts, avoidance of zero would likely still occur.

The present results suggest that pigeons are particularly sensitive to zero outcome values, especially in comparison to human participants, and it is interesting to speculate on why. As we have discussed, zero values could transform the task into one that taps into delay discounting, bringing pigeons' steep discounting functions into play. The use of consumable reinforcement for hungry pigeons could also make zero values particularly salient or memorable; not eating at all may be qualitatively not just quantitatively different than not eating. It is also possible that pigeons (and potentially many other species) have evolved to be highly sensitive to the absence of an important outcome, and perhaps less so to the quantity of that outcome as long as some is obtained. For example, it could be adaptive to remember choices that sometimes lead to

no food because in cases of severe food shortage, obtaining nothing could risk starvation. All of these speculations will require further experimentation with other tasks, types of reinforcement, and species.

The present research provides a striking example of how behavioural similarities across species need not be indicative of the same underlying mechanism. Although humans and pigeons chose similarly to the options presented in Experiment 1, which included a zero value (and in Ludvig, Madan, Pisklak, & Spetch, 2014), their choices were quite different when the zero outcome was removed in Experiment 2, 3, and 4. Avoiding the zero value played a very powerful role in controlling the risky choices of pigeons, but not humans. Comparative research between humans and other animal species has been remarkably successful in delineating many robust and highly important behavioural processes that generalize across numerous species - e.g., schedules of reinforcement, hyperbolic discounting, equivalence class formation, and generalized matching, to name just a few. This work, however, provides an important cautionary tale about the challenges of effective comparative research, whereby some similarities may, paradoxically, not actually be the same.

4.8. **Appendix A: Experiment 1b and 2b: Instructional Information**

The following instructions were read aloud and projected onto a lecture screen:

Instructions for Experiment 1b and Experiment 2b - Group 1

The experiment consists of a computer portion and a maze portion. For the computer portion use the mouse cursor to click on the circles that you see. Try and get as many points as possible. For the maze portion a message will appear on screen telling you when to complete a particular maze. Take a few minutes to complete the maze with a pen or pencil. All the mazes are solvable, but don't worry if you can't solve it. The experiment should take just under 2 hours to complete. You will earn 2 credits for participating, and receive a cash bonus of up to $5.00 depending upon the number of points you get in the computer portion. The more points you get the more money you will earn.

Instructions for Experiment 2b – Group 2

The experiment consists of a computer portion and a maze portion. For

the computer portion your goal is to earn as many points as you can. After
clicking on a white centre circle, you will see one or two coloured circles on
the computer screen. You choose a coloured circle by clicking on it with the
mouse. After clicking the coloured circle you will see one or more black circles,
clicking each black circle will give you one point. When there are two circles
you should choose the one you think will win you the most black circles (i.e.,
the most points). If there is only one circle on the screen, you must click on
that one circle to continue. For the maze portion a message will appear on
screen telling you when to complete a particular maze. Take a few minutes to
complete the maze with a pen or pencil. All the mazes are solvable, but don't
worry if you can't solve it. The experiment should take just under 2 hours to
complete. You will earn 2 credits for participating and receive a cash bonus
of up to $5.00 depending upon the number of points you get in the computer
portion. The more points you get the more money you will earn.

4.9. Appendix B: Experiment 2b: Group Catch Trial Results

Although risk preference was highly similar across Group 1 and Group
2, the percentage of participants who passed the catch trials was strikingly
different - only 55% of the 44 participants in Group 1 (who received minimal
instructions) compared to 100% of the 24 participants in Group 2 (who received
more detailed instructions). A Pearson's Chi-squared test showed this differ-
ence between groups to be statistically significant, $\chi^2(1) = 15.45$, $p < .001$,
$d = 1.08$ $[0.52, 1.63]$.

The subset of participants in Group 1 who passed the catch trials scored
near ceiling, while the subset of participants who failed the catch trials con-
sistently scored near chance. One possibility is that participants who failed
the catch trials had attempted to maximize rewards by responding as quickly
as possible. Indeed, those who failed the catch trials typically responded
faster ($Mdn = 488$ ms, IQR $[460, 569]$) than those who passed the catch tri-
als ($Mdn = 738$ ms $[679, 874]$). This difference was significant ($Z = 4.91$,
$p < .001$) according to a post-hoc a randomization test performing 10,000
Monte Carlo re-samplings of the median choice response times using the Coin
R package (Hothorn, Hornik, van de Wiel, & Zeileis, 2006).

Among participants who passed the catch-trial criterion, no significant

differences were observed in catch-trial performance for trials with and without the low-extreme outcome for either Group 1 ($t(23) = 1.16$, $p = .256$, $g = -0.23$ 95% CI $[-0.82, 0.35]$, $BF_{10} = 0.39$) or Group 2 ($t(5) = 1.11$, $p = .278$, $g = -0.22$ $[-0.81, 0.36]$, $BF_{10} = 0.37$).

4.10. **Author Note**

Jeffrey M. Pisklak, Department of Psychology, University of Alberta; Christopher R. Madan, School of Psychology, University of Nottingham; Elliot A. Ludvig, Department of Psychology, University of Warwick; Marcia L. Spetch, Department of Psychology, University of Alberta.

Chapter 5   General Discussion

The sections below summarize the experiments of each chapter and provide further details and considerations pertaining to the predictions of noteworthy quantitative models, some of which had not been previously discussed therein.

5.1. **Chapter 2**

5.1.1. *Summary.* Chapter 1 tested how pigeons (N = 10) allocated their responding to alternatives that varied the signalling, and thus conditionally reinforcing, function of stimuli within a concurrent-chains "suboptimal choice" task. In one condition (Sig-Both) the terminal-link stimuli for the right and left chains differentially predicted the delivery of food or extinction (blackout), with one 'suboptimal' alternative providing a lower overall probability of food reinforcement. In a control condition (Sig-Sub), the signalling function of the "optimal" alternative was removed, while holding all other parameters constant, so that its terminal links were no longer predictive of the outcome. This was done to evaluate certain key theories of suboptimal choice which suppose that a terminal-link's signal value (i.e., how strongly it is correlated with unconditional reinforcement), and not its frequency (i.e., rate), or the frequency of the terminal reinforcer, is what controls preference in suboptimal choice tasks. On the basis of these predictions, the Sig-Both condition was expected to produce indifferent responding whereas the Sig-Sub should show an increased suboptimal preference relative to that obtained in the Sig-Both condition.

The results indicated that pigeons were far from indifferent in the Sig-Both condition and in fact chose the optimal alternative at levels exceeding those predicted by chance. By contrast, the Sig-Sub condition noticeably attenuated responding to the optimal alternative. Thus, only one of the two predictions were satisfied by theories positing the exclusivity of signal value. The noticeable preference for the optimal alternative observed in the Sig-Both condition demonstrates that the frequency at which reinforcers are presented impacts preference in the suboptimal choice task. Indeed, this is a conclusion that should not be terribly surprising in light of the overwhelming evidence in favour of matching discussed in Section 1.1. Nevertheless, it is the prediction drawn from some key theories of suboptimal choice.

5.1.2. *Discussion.* While the design of Chapter 2's experiment was able to demonstrate clear control of reinforcer frequency over pigeons' choice behaviour, an open question still remains as to the nature of the frequency at play. This is because the frequency of conditional and unconditional (primary) reinforcement is necessarily confounded in the suboptimal choice task. Thus, it is impossible to say with certainty whether it was the frequency of the conditional reinforcers, the frequency of the unconditional reinforcers, or both, that drove preference towards optimality in the Sig-Both condition. The SiGN hypothesis discussed in Chapter 2, which makes successful ordinal predictions, posits a role for both types of reinforcement but has no formalized description of how their effects interact or should be weighted relative to one another. This shortcoming illustrates the necessity for a formal quantitative description of the controlling variables.

One such quantitative description is offered by the the information theoretic approach, discussed in Chapter 2. From this perspective there is an assumed competing influence of conditional reinforcement processes (i.e., temporally informative signals) and unconditional reinforcement processes (i.e., matching). This competition is determined by a complex weighting function that incorporates the ratio of the average delay to food to average delay to the temporally informative signals, alongside additional sensitivity and bias parameters (see Equation 13 of Cunningham & Shahan, 2018). On a more fundamental level, what the weighting function does is allow temporal information and matching processes to compete for control over choice responding in the suboptimal task. Based purely on the task parameters used in Chapter 2's experiment, the weighting function predicts that temporally informative signals, and not matching, should control preference in both conditions (when no sensitivity or bias adjustments are made). This leads to a prediction of indifference, similar to the signal-value and RRM hypotheses discussed in Chapter 2. However, it is perhaps not unreasonable to assume that, under conditions of equal temporal information being provided by both alternatives (as in the case of the Sig-Both condition), bias occurs in favour of unconditional reinforcement as a sort of "default" controller of choice in absence of conditional reinforcement being any use. Though, this does not appear to be an *a priori* prediction by the model's authors. In any case, the fact that the model makes no straightforward predictions in this regard seems to work against its favour.

A perspective not considered at the time of Chapter 2's publication concerns the predictions of the hyperbolic-value added (HVA) model (see Section 1.3). When value addition is considered alone, greater optimal responding is expected in the Sig-Both condition relative to the Sig-Sub, across differing sensitivities (a_t). However, the absolute difference between conditions is marginal compared to what was actually observed. As can be seen in Figure 5.1, across various levels of bias, the HVA model only predicts a difference on the order of approximately 5%. This is contrasted with the 30% difference that was actually obtained. Thus, while the ordinal direction of the HVA model's predictions were correct, it appears to provide a poor description of the results; though, a wider range of tested values would be needed for a more complete evaluation.

Another consideration concerns the predictions of delay-reduction theory (DRT; Equation 1.7). While there were some early attempts at reconciling DRT with findings of suboptimal choice (see, Dunn & Spetch, 1990; Spetch & Dunn, 1987; Spetch et al., 1990), it had been generally supposed that its strict quantitative formulation could not account for the results within the suboptimal paradigm. As originally conceived by Fantino (1969), DRT had no formal means of handling the probabilistic delivery of an unconditional reinforcer that followed a terminal link, taking as a given that unconditional reinforcers would always terminate the chain (i.e., terminal links could only ever function as a 100% signal for food). However, if one considers the *expected value* of the delays in the initial links and terminal links, while assuming that signals for extinction do not contribute to these values (a reasonable assumption given a variety of evidence, Dinsmoor, 1983; Fortes et al., 2016; Mazur, 1997; Pisklak et al., 2015; Stagner, Laude, & Zentall, 2011), then DRT seems to accommodate the suboptimal paradigm quite well, making predictions that are largely in accordance with much of its literature while still preserving the classic predictions of DRT (Fantino, 1977). Equation 5.1 shows an updated version of the Squires and Fantino (1971) DRT equation (Equation 1.7) incorporating the expected value of the delays along with additional bias (b) and sensitivity (a and k) parameters similar to those used by other contemporary models like that of the information-theoretic approach (for equations, see Cunningham &
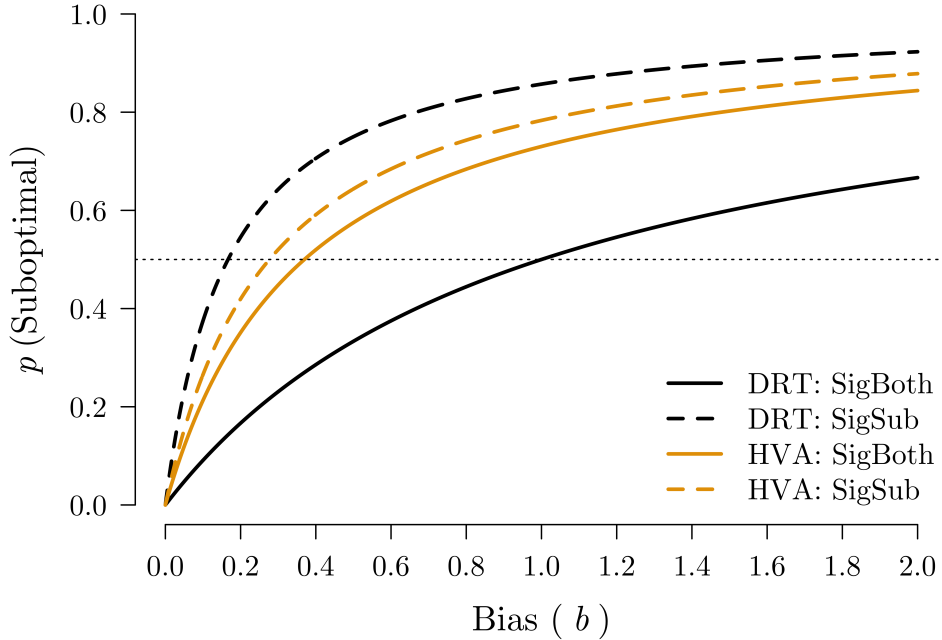
Figure 5.1: Chapter 2 predicted *a priori* suboptimal preference according to delay-reduction theory (DRT) and hyperbolic value-addition (HVA) across varying levels of bias b. When $b > 1$, preference is biased towards the suboptimal alternative. When $b < 1$, preference is biased towards the optimal alternative. DRT calculations are made independent of matching processes ($a = 0$). For HVA calculations, an undiscounted value of $A = 1$ and a discounting rate $K = 1$ are assumed. The thin dotted line indicates indifference.

Shahan, 2018) and the HVA model (Equation 1.8; see also Mazur, 2001).

$$\frac{B_1}{B_2} = b \left(\frac{R_1}{R_2} \right)^a \left(\frac{T - \mathbb{E}[t_1]}{T - \mathbb{E}[t_2]} \right)^k, \text{ when } \mathbb{E}[t_1] < T, \ \mathbb{E}[t_2] < T \qquad (5.1)$$

$\mathbb{E}$ denotes the calculation of a mathematical expectation, and reflects the basic idea that the probability of being terminally reinforced impacts how the the model treats the programmed delays of an experiment. Consider, for instance, the parameters of Chapter 2's Sig-Sub alternatives. The suboptimal alternative provided a 20% chance of food. This means that a pigeon will need to go through the initial link 5 times, on average, to obtain a single food rein-

forcement. Moreover, on 4 out of those 5 instances, the bird will encounter a terminal link that signals extinction. The remaining terminal link will be a signal that the delivery of food is guaranteed. Since the terminal-link food signal is 10 seconds long and extinction signals are not being considered in the determination of the expected delay, we can assume it will take the bird, on average, 15 seconds to be reinforced on the suboptimal alternative (assuming a FR 1 initial link lasts for 1 sec). This can be contrasted with the optimal alternative where there is an 80% chance of food delivery. In this case the bird will need to encounter the initial- and terminal-link period 1.25 times, on average, to obtain food. Since the Sig-Sub condition's terminal links do not differentially predict the presence or absence of food (i.e., are unsignalled), there are no extinction signals to disregard and thus the combined duration of each terminal link presentation needs to be considered ($10 \times 1.25 = 12.5$). Therefore, on the optimal side, the bird can reasonably be expected to have to wait, on average, 13.75 seconds to obtain a single food reinforcement. All told, when expected value is incorporated into the formal calculations for delay reduction (see Appendix A), DRT predicts approximately a 6:1 preference in favour of the suboptimal alternative for the Sig-Sub condition. By contrast, the Sig-Both condition predicts a preference of 1:1 for the suboptimal alternative (which is indifference). This means that, when expected value of the delays is considered, DRT predicts 36% difference between the Sig-Sub and Sig-Both conditions. This difference is depicted in Figure 5.1 at varying levels of bias.

Interestingly, with respect to the question pertaining to the the tradeoff between conditional and unconditional reinforcer frequency that began this discussion, Equation 5.1 makes the case that the preference observed to the optimal alternative in Chapter 2's Sig-Both condition must solely be due to the influence of unconditional reinforcement frequency. Consider that, in the Sig-Sub case, the amount of conditional reinforcement ($T - \mathbb{E}[t_x]$) is higher on the suboptimal alternative than it is on the optimal, but the optimal alternative contains a higher frequency (i.e., rate) of primary reinforcement (R_x). Thus, in the Sig-Sub case the two reinforcer types are, according to DRT, acting in opposition. By contrast, in the Sig-Both condition the amount of conditional reinforcement received on the suboptimal alternative is equal to that provided by the optimal alternative. This would foster indifference; how-

ever, the frequency of unconditional reinforcement is still higher on the opti-
mal alternative which pushes the preference towards optimality. Moreover, the
higher frequency of conditional reinforcement provided by the Sig-Both condi-
tion's optimal alternative actually causes the functional rate of unconditional
reinforcement to increase (relative to the optimal alternative on the Sig-Sub
condition). This occurs because the extinction signals are not considered to
impact the delays to food.

In summary, while the HVA model makes correct ordinal predictions, its
predicted relative difference between the conditions, as well as its predicted ab-
solute preference are far off the mark. By contrast, DRT seems to make good
ordinal predictions as well as good predictions of relative-difference between
conditions. However, its predictions in terms of absolute preference are also
lacking (even when unconditional reinforcer rates are considered), predicting
substantially greater levels of suboptimal choice than were actually observed
when no bias is assumed and no adjustments to sensitivity are made. However,
as is the case for the HVA model, a wider range of values will ultimately need
to be tested for a more complete assessment of this model's potential.

5.2. **Chapter 3**

5.2.1. *Summary.* Chapter 3 examined the role of an increased response
requirement in the choice phase of a pigeon (N = 14) suboptimal choice task.
Similar to Chapter 2, the experiment used a concurrent-chains design that, for
one "signalled" alternative, provided differential signals for food in the termi-
nal link. The other, "unsignalled", alternative provided only non-differential
food signals. The signalled (suboptimal) alternative offered a 20% chance
of food delivery, whereas the unsignalled (optimal) alternative offered a 50%
chance of food delivery. Thus, for the unsignalled alternative, the terminal
links were no more indicative of receiving food than they were of extinction.
These parameters, based on those used by Stagner and Zentall (2010), are well
known for producing high levels of suboptimal preference in pigeons and star-
lings (Vasconcelos et al., 2018; Zentall, 2016). To test the effect of increased
response requirement in the choice phase, the ratio requirement of both al-
ternative's initial links were manipulated within subjects across two blocks of
sessions. Half the pigeons received a FR 1 requirement in the initial links for

15 sessions, followed by a FR 25 requirement for a remaining 20 sessions. The other half received the same, but with the FR 25 requirement occurring first.

Past evidence had shown that increasing the length of both initial-link requirements simultaneously can have a deleterious effect on suboptimal responding (Dunn & Spetch, 1990; Kendall, 1974, 1985). However, that evidence had only been conducted in the context of VI initial-link schedules with signalled terminal links occurring on each alternative. Consequently, it was unclear how the pigeons might respond in the present case. Given that certain key theories of choice put a heavy emphasis on the impact of delay to reinforcement (Mazur, 1997; McDevitt et al., 2016), it might be surmised that increasing the ratio-requirement for the initial-links from a FR 1 to a FR 25 (which necessarily increases time spent in the initial-links) would diminish the conditionally reinforcing effect of the suboptimal alternative's terminal link, and thus suboptimal preference as a whole.

The results found that, when the initial-link requirement was small (FR 1) all pigeons exhibited a strong preference of the suboptimal alternative, irrespective of condition order. When the initial-link requirement was larger (FR 25) suboptimal preference was significantly attenuated and, for some pigeons, completely removed. In addition to the diminished suboptimal preference, the FR 25 schedule also produced a notable increase in variability that is not straightforwardly accounted for, though does mirror the high variability obtained by Zentall et al. (2017), which performed a similar manipulation with FI schedules using a slightly different procedure.

5.2.2. *Discussion.* Figure 5.2 plots the predicted preferences of DRT, HVA, and Temporal Information (vis-à-vis the information-theoretic approach), when no adjustments to bias or sensitivity are made. According to all three models, increasing initial link durations diminishes suboptimal preference; though, the degree to which this occurs varies according to model. DRT predicts the most precipitous drop, followed by the HVA, and then temporal information, which predicts almost no change across the calculated values. Of the three quantitative models, only the predictions of DRT (independent of any matching processes, $a = 0$) fall within the 95% confidence range of the obtained data for both the FR1 ($M = 0.96$, 95% CI [0.91, 1.00]) and FR25 ($M = 0.69$, 95% CI [0.51, 0.88]) conditions. However, given the inflated variability created by

the FR 25 schedule, this may not be indicative of much. Perhaps, more telling is DRT's successful prediction in the FR 1 case. That condition found strong suboptimal preference with low between-subject variability, which is a common result given the experimental parameters used (e.g., Fortes et al., 2018; Stagner & Zentall, 2010; Vasconcelos et al., 2015). Of the three quantitative models, only DRT is able to successfully predict this finding independent of free-parameter manipulation. For instance, the HVA model can be configured to make predictions equivalent to that of DRT in the FR 1 case by increasing value-addition sensitivity (a_t in Equation 1.8). Similarly, signal-value sensitivity can be adjusted when considering the predictions of temporal information under the information-theoretic approach (see Equation 10 of Cunningham & Shahan, 2018). However, apart from quantifying deviations from the predicted values of the models, there is no clear *a priori* reason why one should be inclined to do this, particularly in an instance where a sampled population is being tested (as opposed to individuals).

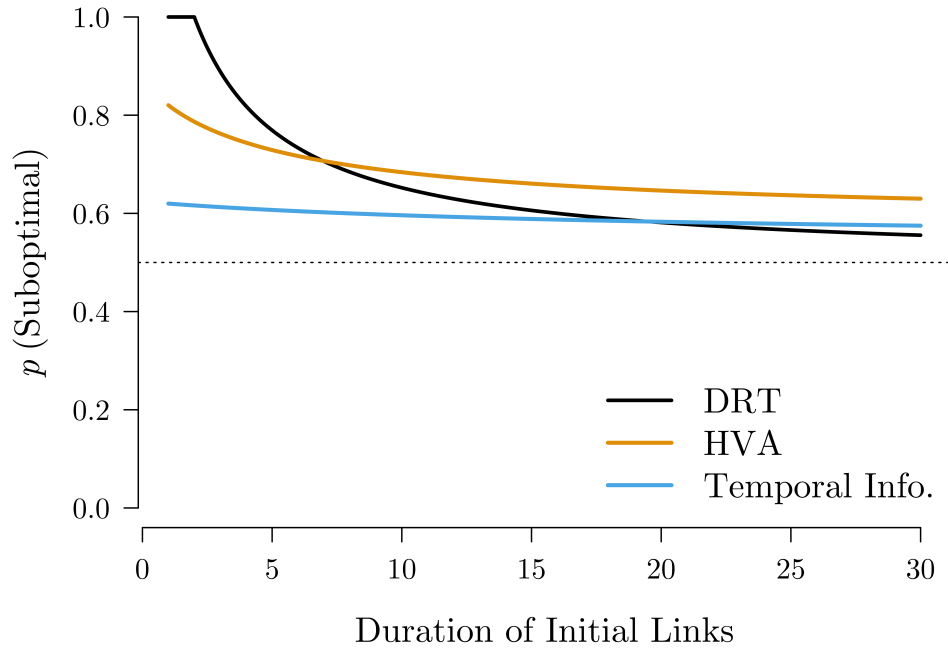Figure 5.2: Chapter 3 predicated *a priori* suboptimal preference according to theories of delay reduction (DRT), hyperbolic value addition (HVA), and temporal-information (Temporal Info.). DRT calculations are made independent of matching processes ($a = 0$). For HVA calculations, an undiscounted value of $A = 1$ and a discounting rate $K = 1$ are assumed. The thin dotted line indicates indifference.

In conclusion, it is worth stressing that evaluation of the quantitative models discussed above demands a more involved level of analysis and experimental design than has been the norm in suboptimal choice research. Typically, the focus has been placed on verbal hypotheses and unitary analyses aimed at testing the ordinal directions of those hypotheses. By contrast, the three quantitative models discussed here put forward much more detailed predictions that require a range of values to be adequately tested. This points to a notable shortcoming of the experiment described in Chapter 3 (as well as Chapter 2); specifically, only two values of the dependent variable were tested. While clear-enough effects were obtained to stress the important function of initial-link requirements, the details of how this effect varies across a series of values remains questionable (despite clear predictions being available) due to the limitations of the experimental design employed. Future research would do well to try and employ designs that can evaluate function of these controlling variables more completely than has been done here.

5.3. **Chapter 4**

5.3.1. *Summary.* Across four separate studies, Chapter 4 examined the possibility that pigeons and humans might both exhibit an *extreme-outcome effect* on their choice behaviour. The extreme-outcome effect posits that, when the probabilistic ("risky") outcomes of a choice are learned through experience (as opposed to description) of the contingencies, the most extreme outcomes in the decision context will be overweighted. This leads to the counter-intuitive prediction of an inverse reflection-effect on choice preferences: risk-seeking for gains and risk-aversion for losses (Ludvig, Madan, & Spetch, 2014). In prior work, by Ludvig, Madan, Pisklak, and Spetch (2014), it had been argued that, like humans, pigeons also exhibit this effect, and that it potentially reflects a similar mechanism of behaviour between the two disparate species. However, an alternate explanation of those findings is that the pigeons' choice behaviour is explainable in terms of the zero-value (extinction) outcome used in their design. To test this possibility, the experiments of Chapter 4 assessed both pigeons and humans on operant tasks that presented choices with differing outcomes. The outcomes varied along two dimensions: risk (*risky* and *safe*) and value (*high* and *low*). Critically, the extreme-outcome effect predicts

greater selection of a risky option, when the choice is between a risky and safe high-value outcome (of equal expected value) relative to when it is a choice between a risky and safe low-value outcome (of equal expected value).

The results obtained from human participants generally conformed to other past results examining the influence of extreme-outcomes on choice (see Madan et al., 2019, for a review). Pigeons, by contrast, only showed the effect when zero was a possible outcome for the risky low-value alternative. When all outcomes produced a reliable unconditional reinforcer (i.e., food), no extreme-outcome effects were found to occur and the pigeons simply displayed indifference between risky and safe choices of equal expected value. Further, this observed indifference did not seem to be an artefact of an inability to discriminate between the differing value's (i.e., reinforcer magnitudes) at play.

5.3.2. *Discussion.* The pigeons' contrasting behaviour when zeros are present versus absent is interesting in light of earlier discussions of choice as they pertain to reinforcer rate (see Section 1.1). In this regard, the birds' pattern of choosing may not be qualitatively different than preferring a continuous schedule of reinforcement to an intermittent schedule of reinforcement. The former provides a higher rate of reinforcement than the latter, and strong control of choice behaviour by reinforcement rate is a well known occurrence across a range of species and behaviours (de Villiers, 1977). Given that findings have established higher sensitivity towards reinforcement rate than magnitude in determining choice preference (M. Davison & Baum, 2003), the pigeons' preference of the safe low-value option, when zero is a possibility on the risky alternative, is perhaps an unsurprising result in many ways. It might also be added to this that the occasional absence of food rewards also creates a situation of intertemporal choice, whereby pigeons are choosing between rewards with different delays. In that frame, one might again predict, based on the steep hyperbolic decay rates pigeons are known to possess (Stevens & Stephens, 2009), that the smaller more certain outcome will be preferred to the larger outcome, which only occurs half as often and is therefore sometimes delayed (for example, see Mazur, 1989).

When zeros are removed from the choice context, the pigeon's overall pattern of choice falls in line with what would be predicted by mathematical expectation. Purely in terms of the values and probabilities at work, the pi-

geons behaviour appears sensible. There is no mathematical reason to prefer the risky alternative over the safe alternative across the high and low choice types since the the long run expectation is equivalent in both cases. The pigeons, on some level, seem to appreciate this; whereas the human participants, by contrast, always seem to exhibit a overweighting of the extremes for some reason. Moreover, this overweighting is manifest most strongly in terms of an aversion to risk for the low-value choices. There is no straightforward reason why this should be the case, though one proposed explanation is that humans may possess a memory-bias for extreme outcomes that causes them to be overweighted (Madan et al., 2014). This hypothesis is based upon findings that people tend to recall extreme outcomes better than non-extreme outcomes. Additionally, people will also have a tendency to overestimate the relative frequency of the extremes they have experienced. While memory, as a hypothesis, offers an appealing explanation in terms of possible intermediary (i.e., neurological) processes that are open to future corroboration and manipulation, at present it is a purported cause that is not independently verifiable from the behaviour itself; further, the functional aspects of the behaviour are still necessary to predict it. Thus, one might say that an appeal to memory still leaves something of a need for an explanation in terms of the behaviour's environmental origin/function. What history of natural selection, or what learning process, brought about this effect of overweighting the extreme outcomes (or bias in memory for extremes) and how can we control it?

Section 4.7 touched upon the possibility that there may be, across the two species, a functional difference between the outcome stimuli. The pigeons' responding was terminally reinforced with prompt differential amounts of food; whereas the humans were promptly reinforced with differential amounts of points, later exchanged for money, and then even later (one assumes) exchanged for, unconditionally reinforcing, goods and services. Thus, one might reasonably assume that the reinforcing effect of the points is some function of a pre-existing learning history with language, numbers, money, and equivalence relations. Perhaps the humans' overweighting of extreme-outcomes reflects some by-product of this very complex history. The ease with which the performance of the humans was drastically altered in Experiment 2 (see, Section 4.9) by a simple modification of the instructions seems to lend some credence to this idea. The logical consequence of this line of reasoning is that, the dif-

ferences obtained may not be indicative of learning processes distinct to each species, but rather of differences learned about the stimuli employed. In other words, the basic learning processes used by the two species could be the same, but their history of learning is very different. On this basis, one might predict that a sufficiently food-deprived human reinforced with bits of chocolate, or some other edible instead of points, would not show behaviour qualitatively different to that of the pigeons (especially if the controlling properties of the choice stimuli have been shaped prior to testing). Indeed, while studies of this kind are few and far between, research attempting to better equate animal and human task parameters has generally found good comparability (e.g., Forzano & Logue, 1994; Hayden & Platt, 2009; Lagorio & Hackenberg, 2010; Locey, Pietras, & Hackenberg, 2009; Millar & Navarick, 1984).

Overall it seems that the conclusion of Ludvig, Madan, Pisklak, and Spetch (2014) may have been premature. While it can be said that both humans and pigeons display an extreme-outcome effect when zeros are a possible extreme, when zeros are removed, only humans continue to display the effect in question. This difference between the two species could be reflective of fundamentally different processes or could be due to functional differences between the stimuli employed.

5.4. **Conclusion**

The experimental data obtained across chapters 2, 3, and 4 demonstrate that probabilistic outcomes can have wide ranging effects on choice behaviour and that those effects are a function of a wider set of variables. Further, Section 5.1.2 introduced a novel modification of DRT that seems to show some initial promise for predicting and interpreting the role of these variables in relation to one another. Future work will be need to be better tailored towards assessing this model, especially in comparison to other existing models of choice, while also trying to better bridge the gap between human and animal experimentation.

5.5. **Appendix A: Formulas for Equation 5.1**

5.5.1. *Expected Value of the Initial Link Delays.* The expected (average) time spent in a given alternative's initial link, per unconditional reinforcement,

is given by the equation:

$$\mathbb{E}[i_x] = i_x \left(\frac{1}{p_{\text{rein}_x}} \right) \tag{5.2}$$

where i_x is the programmed duration of the initial link on alternative x and p_{rein_x} is the probability of unconditional reinforcement on that same alternative.

5.5.2. *Expected Value of the Terminal Link Delays.* The expected time spent in a *signalled* terminal link (i.e., a terminal link that is is maximally correlated, $r = 1$, with the delivery of unconditional reinforcement), per unconditional reinforcement, is given by the equation:

$$\mathbb{E}[t_x] = t_x \tag{5.3}$$

The expected time spent in an *unsignalled* terminal link ($0 < r < 1$), per unconditional reinforcement, is given by the equation:

$$\mathbb{E}[t_x] = t_x \left(\frac{1}{p_{\text{rein}_x}} \right) \tag{5.4}$$

Note that Equation 5.4 is equivalent to Equation 5.3 when $p_{\text{rein}_x} = 1$.

5.5.3. *Expected Time to Unconditional Reinforcement (T).* The calculation of T is determined by summing the expected time to reach a terminal link from the onset of the initial links and the expected time to reinforcement at the onset of the terminal links. The expected time to reach a terminal link is:

$$\frac{1}{1/\mathbb{E}[i_1] + 1/\mathbb{E}[i_2]} \tag{5.5}$$

where $E[i_1]$ and $E[i_2]$ are the expected value of the initial link delays for two separate choices (as described in Section 5.5.1). The expected time to unconditional reinforcement at the onset of the terminal links is given by:

$$p_2 \cdot \mathbb{E}[t_2] + (1 - p_2) \cdot \mathbb{E}[t_1] \tag{5.6}$$

where p_2 is the probability of entering the second alternative's terminal link

and is given by:

$$p_2 = \frac{\mathbb{E}[i_1]}{\mathbb{E}[i_1] + \mathbb{E}[i_2]} \tag{5.7}$$

Therefore:

$$T = \frac{1}{1/\mathbb{E}[i_1] + 1/\mathbb{E}[i_2]} + p_2 \cdot \mathbb{E}[t_2] + (1 - p_2) \cdot \mathbb{E}[t_1] \tag{5.8}$$

5.5.4. *Rate of Unconditional Reinforcement.* The rate of unconditional reinforcement, R, for a given alternative x is:

$$R_x = \frac{n_x}{\mathbb{E}[i_x] + n_x \cdot \mathbb{E}[t_x]} \tag{5.9}$$

where n_x is the number of unconditional reinforcements obtained from one entry of alternative x's terminal link.

References

Ainslie, G. W. (1974). Impulse control in pigeons [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *21*(3), 485-489. doi: 10.1901/jeab.1974.21-485

Barnard, C. J., Brown, C. A. J., Houston, A. I., & Mcnamara, J. M. (1985). Risk-sensitive foraging in common shrews - an interruption model and the effects of mean and variance in reward rate [Journal Article]. *Behavioral Ecology and Sociobiology*, *18*(2), 139-146. doi: Doi10.1007/Bf00299042

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions [Journal Article]. *Journal of Behavioral Decision Making*, *16*(3), 215-233. doi: 10.1002/bdm.443

Bateson, M., & Kacelnik, A. (1995). Preferences for fixed and variable food sources: Variability in amount and delay [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *63*(3), 313-329. doi: 10.1901/jeab.1995.63-313

Bateson, M., & Kacelnik, A. (1996). Rate currencies and the foraging starling: The fallacy of the averages revisited [Journal Article]. *Behavioral Ecology*, *7*(3), 341-352. doi: 10.1093/beheco/7.3.341

Baum, W. M. (1974). On two types of deviation from matching law - bias and undermatching [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *22*(1), 231-242. doi: 10.1901/jeab.1974.22-231

Baum, W. M. (2010). Dynamics of choice: A tutorial [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *94*(2), 161-174. doi: 10.1901/jeab.2010.94-161

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *12*(6), 861-874. doi: 10.1901/jeab.1969.12-861

Beautrais, P. G., & Davison, M. C. (1977). Response and time allocation in concurrent 2nd-order schedules [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *27*(1), 61-69. doi: 10.1901/jeab.1977.27-61

Belke, T. W., & Spetch, M. L. (1994). Choice between reliable and unreliable reinforcement alternatives revisited: Preference for unreliable reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *62*(3), 353-66. doi: 10.1901/jeab.1994.62-353

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk [Journal Article]. *Econometrica*, *22*(1), 23-36. doi: 10.2307/1909829

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 [Journal Article]. *Neuron*, *85*(3), 602-14. doi: 10.1016/j.neuron.2014.12.050

Brownstein, A. J., & Pliskoff, S. S. (1968). Some effects of relative reinforcement rate and changeover delay in response-independent concurrent schedules of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *11*(6), 683-688. doi: 10.1901/jeab.1968.11-683

Camilleri, A. R., & Newell, B. R. (2011). Description- and experience-based choice: Does equivalent information equal equivalent choice? [Journal Article]. *Acta Psychologica*, *136*(3), 276-284. doi: 10.1016/j.actpsy.2010.11.007

Caraco, T., Martindale, S., & Whittam, T. S. (1980). An empirical demonstration of risk-sensitive foraging preferences [Journal Article]. *Animal Behaviour*, *28*(Aug), 820-830. doi: 10.1016/S0003-3472(80)80142-4

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

Catania, A. C. (1963). Concurrent performances - reinforcement interaction and response independence [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *6*(2), 253-263. doi: 10.1901/jeab.1963.6-253

Chung, S. H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *10*(1), 67-74. doi: 10.1901/jeab.1967.10-67

Conger, R., & Killeen, P. (1974). Use of concurrent operants in small group research - demonstration [Journal Article]. *Pacific Sociological Review*, *17*(4), 399-416.

Constantinople, C. M., Piet, A. T., & Brody, C. D. (2018). An analysis of decision under risk in rats [Journal Article]. *bioRxiv*.

Cunningham, P. J., & Shahan, T. A. (2018). Suboptimal choice, reward-predictive signals, and temporal information [Journal Article]. *Journal of Experimental Psychology-Animal Learning and Cognition*, *44*(1), 1-22. doi: 10.1037/xan0000160

Davison, M., & Baum, W. M. (2003). Every reinforcer counts: Reinforcer magnitude and local preference [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *80*(1), 95-129. doi: 10.1901/jeab.2003.80-95

Davison, M. C. (1972). Preference for mixed-interval versus fixed-interval schedules - number of component intervals [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *17*(2), 169-176. doi: 10.1901/jeab.1972.17-169

Davison, M. C., & McCarthy, D. (1988). *The matching law: A research review* [Book]. Hillsdale, N.J.: L. Erlbaum.

de Villiers, P. (1977). Choice in concurrent schedules and a quantitative formulation of the law of effect [Book Section]. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (p. 233-287). Englewood Cliffs, NJ: Prentice-Hall.

Dinsmoor, J. A. (1983). Observing and conditioned reinforcement [Journal Article]. *Behavioral and Brain Sciences*, *6*(4), 693-704. doi: 10.1017/S0140525x00017969

Duncan, B., & Fantino, E. (1970). Choice for periodic schedules of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *14*(1), 73-86. doi: 10.1901/jeab.1970.14-73

Dunn, R., & Spetch, M. L. (1990). Choice with uncertain outcomes - conditioned reinforcement effects [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *53*(2), 201-218. doi: 10.1901/jeab.1990.53-201

Elsmore, T. F. (1971). Effects of response effort on discrimination performance [Journal Article]. *The Psychological Record*, *21*(1), 17-24. doi: 10.1007/BF03393984

Fantino, E. (1969). Choice and rate of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *12*(5), 723-730. doi: 10.1901/jeab.1969.12-723

Fantino, E. (1977). Conditioned reinforcement: Choice and information [Book Section]. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (p. 313-339). Englewood Cliffs, NJ: Prentice-Hall.

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement* [Book]. New York: Appleton-Century-Crofts.

Fortes, I., Machado, A., & Vasconcelos, M. (2017). Do pigeons (columba livia) use information about the absence of food appropriately? a further look into suboptimal choice [Journal Article]. *Journal of Comparative Psychology*, *131*(4), 277-289. doi: 10.1037/com0000079

Fortes, I., Pinto, C., Machado, A., & Vasconcelos, M. (2018). The paradoxical effect of low reward probabilities in suboptimal choice [Journal Article]. *Journal of Experimental Psychology-Animal Learning and Cognition*, *44*(2), 180-193. doi: 10.1037/xan0000165

Fortes, I., Vasconcelos, M., & Machado, A. (2016). Testing the boundaries of "paradoxical" predictions: Pigeons do disregard bad news [Journal Article]. *Journal of Experimental Psychology-Animal Learning and Cognition*, *42*(4), 336-346. doi: 10.1037/xan0000114

Forzano, L. B., & Logue, A. W. (1994). Self-control in adult humans - comparison of qualitatively different reinforcers [Journal Article]. *Learning and Motivation*, *25*(1), 65-82. doi: 10.1006/lmot.1994.1004

Franz, V. H., & Loftus, G. R. (2012). Standard errors and confidence intervals in within-subjects designs: Generalizing loftus and masson (1994) and avoiding the biases of alternative accounts [Journal Article]. *Psychonomic Bulletin & Review*, *19*(3), 395-404. doi: 10.3758/s13423-012-0230-1

Freidin, E., Aw, J., & Kacelnik, A. (2009). Sequential and simultaneous choices: Testing the diet selection and sequential choice models [Journal Article]. *Behavioural Processes*, *80*(3), 218-223. doi: 10.1016/j.beproc.2008.12.001

Friedel, J. E., DeHart, W. B., Madden, G. J., & Odum, A. L. (2014). Impulsivity and cigarette smoking: discounting of monetary and consumable outcomes in current and non-smokers [Journal Article]. *Psychopharmacology*, *231*(23), 4517-4526. doi: 10.1007/s00213-014-3597-z

Gipson, C. D., Alessandri, J. J. D., Miller, H. C., & Zentall, T. R. (2009). Preference for 50% reinforcement over 75% reinforcement by pigeons [Journal Article]. *Learning & Behavior*, *37*(4), 289-298. doi: 10.3758/Lb.37.4.289

Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards - a life-span comparison [Journal Article]. *Psychological Science*, *5*(1), 33-36. doi: 10.1111/j.1467-9280.1994.tb00610.x

Green, L., Myerson, J., & Calvert, A. L. (2010). Pigeons discounting of probabilistic and delayed reinforcers [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *94*(2), 113-123. doi: 10.1901/jeab.2010.94-113

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities [Journal Article]. *Journal of Behavioral Decision Making*, *23*(1), 48-68. doi: 10.1002/bdm.665

Hayden, B. Y., Heilbronner, S. R., Nair, A. C., & Platt, M. L. (2008). Cognitive influences on risk-seeking by rhesus macaques [Journal Article]. *Judgment and Decision Making*, *3*(5), 389-395.

Hayden, B. Y., & Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques [Journal Article]. *Current Biology*, *17*(1), 49-53. doi: 10.1016/j.cub.2006.10.055

Hayden, B. Y., & Platt, M. L. (2009). Gambling for gatorade: risk-sensitive decision making for fluid rewards in humans [Journal Article]. *Animal Cognition*, *12*(1), 201-207. doi: 10.1007/s10071-008-0186-8

Heilbronner, S. R., & Hayden, B. Y. (2016). The description-experience gap in risky choice in nonhuman primates [Journal Article]. *Psychonomic Bulletin & Review*, *23*(2), 593-600. doi: 10.3758/s13423-015-0924-2

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *4*(3), 267-272. doi: 10.1901/jeab.1961.4-267

Herrnstein, R. J. (1964). Aperiodicity as factor in choice [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *7*(2), 179-182. doi: 10.1901/jeab.1964 .7-179

Herrnstein, R. J. (1970). On the law of effect [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *13*(2), 243-266. doi: 10.1901/jeab.1970.13 -243

Herrnstein, R. J. (1997). Melioration as behavioral dynamics [Book Section]. In H. Rachlin & D. I. Laibson (Eds.), *The matching law: Papers in psychology and economics by richard j. herrnstein* (p. 74-99). Cambridge, MA: Harvard University Press.

Herrnstein, R. J., & Loveland, D. H. (1975). Maximizing and matching on concurrent ratio schedules [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *24*(1), 107-116. doi: 10.1901/jeab.1975.24-107

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice [Journal Article]. *Psychological Science*, *15*(8), 534-539. doi: 10.1111/j.0956-7976.2004.00715.x

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice [Journal Article]. *Trends in Cognitive Sciences*, *13*(12), 517-523. doi: 10.1016/j.tics.2009 .09.004

Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2006). A lego system for conditional inference. *The American Statistician*, *60*(3), 257–263.

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *77*(2), 129-146. doi: 10.1901/jeab.2002 .77-129

Kacelnik, A., Vasconcelos, M., Monteiro, T., & Aw, J. (2011). Darwin's "tug-of-war" vs. starlings' "horse-racing": how adaptations for sequential encounters drive simultaneous choice [Journal Article]. *Behavioral Ecology and Sociobiology*, *65*(3), 547-558. doi: 10.1007/s00265-010-1101-2

Kahneman, D., & Tversky, A. (1979). Prospect theory - analysis of decision under risk [Journal Article]. *Econometrica*, *47*(2), 263-291. doi: 10.2307/1914185

Kelleher, R. T., & Gollub, L. R. (1962). A review of positive conditioned reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *5*, 543-97. doi: 10.1901/jeab.1962.5-s543

Kendall, S. B. (1974). Preference for intermittent reinforcement [Journal Article]. *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 [Journal Article]. *Behavioural Processes*, *10*(4), 399-413. doi: 10.1016/0376-6357(85)

90040-3

Killeen, P. (1972). The matching law [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *17*(3), 489-495. doi: 10.1901/jeab.1972.17-489

Lagorio, C. H., & Hackenberg, T. D. (2010). Risky choice in pigeons and humans: A cross-species comparison [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *93*(1), 27-44. doi: 10.1901/jeab.2010.93-27

Lagorio, C. H., & Hackenberg, T. D. (2012). Risky choice in pigeons: preference for amount variability using a token-reinforcement system [Journal Article]. *J Exp Anal Behav*, *98*(2), 139-154. doi: 10.1901/jeab.2012.98-139

Lagorio, C. H., & Madden, G. J. (2005). Delay discounting of real and hypothetical rewards iii: Steady-state assessments, forced-choice trials, and all real rewards [Journal Article]. *Behavioural Processes*, *69*(2), 173-187. doi: 10.1016/j.beproc .2005.02.003

Lakshminarayanan, V. R., Chen, M. K., & Santos, L. R. (2011). The evolution of decision-making under risk: Framing effects in monkey risk preferences [Journal Article]. *Journal of Experimental Social Psychology*, *47*(3), 689-693. doi: 10 .1016/j.jesp.2010.12.011

Locey, M. L., Pietras, C. J., & Hackenberg, T. D. (2009). Human risky choice: Delay sensitivity depends on reinforcer type [Journal Article]. *Journal of Experimental Psychology - Animal Behavior Processes*, *35*(1), 15-22. doi: 10.1037/a0012378

Ludvig, E. A., Madan, C. R., McMillan, N., Xu, Y. Q., & Spetch, M. L. (2018). Living near the edge: How extreme outcomes and their neighbors drive risky choice [Journal Article]. *Journal of Experimental Psychology-General*, *147*(12), 1905-1918. doi: 10.1037/xge0000414

Ludvig, E. A., Madan, C. R., Pisklak, J. M., & Spetch, M. L. (2014). Reward context determines risky choice in pigeons and humans [Journal Article]. *Biology Letters*, *10*(8). doi: 10.1098/rsbl.2014.0451

Ludvig, E. A., Madan, C. R., & Spetch, M. L. (2014). Extreme outcomes sway risky decisions from experience [Journal Article]. *Journal of Behavioral Decision Making*, *27*(2), 146-156. doi: 10.1002/bdm.1792

Ludvig, E. A., & Spetch, M. L. (2011). Of black swans and tossed coins: Is the description-experience gap in risky choice limited to rare events? [Journal Article]. *Plos One*, *6*(6). doi: ARTNe2026210.1371/journal.pone.0020262

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2014). Remembering the best and worst of times: Memories for extreme outcomes bias risky decisions [Journal Article]. *Psychonomic Bulletin & Review*, *21*(3), 629-636. doi: 10.3758/s13423 -013-0542-9

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2017). The role of memory in distinguishing risky decisions from experience and description [Journal Article]. *Quarterly Journal of Experimental Psychology*, *70*(10), 2048-2059. doi: 10.1080/17470218.2016.1220608

Madan, C. R., Ludvig, E. A., & Spetch, M. L. (2019). Comparative inspiration: From puzzles with pigeons to novel discoveries with humans in risky choice [Journal Article]. *Behavioural Processes*, *160*, 10-19. doi: 10.1016/j.beproc.2018.12.009

Madden, G. J., Begotka, A. M., Raiff, B. R., & Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards [Journal Article]. *Experimental and Clinical Psychopharmacology*, *11*(2), 139-145. doi: 10.1037/1064-1297.11.2.139

Madden, G. J., Raiff, B. R., Lagorio, C. H., Begotka, A. M., Mueller, A. M., Hehli, D. J., & Wegener, A. A. (2004). Delay discounting of potentially real and hypothetical rewards: Ii. between- and within-subject comparisons [Journal Article]. *Experimental and Clinical Psychopharmacology*, *12*(4), 251-261. doi: 10.1037/1064-1297.12.4.251

Marsh, B., & Kacelnik, A. (2002). Framing effects and risky decisions in starlings [Journal Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(5), 3352-3355. doi: 10.1073/pnas.042491999

Marshall, A. T., & Kirkpatrick, K. (2013). The effects of the previous outcome on probabilistic choice in rats [Journal Article]. *Journal of Experimental Psychology-Animal Behavioral Processes*, *39*(1), 24-38. doi: 10.1037/a0030765

Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays [Journal Article]. *Journal of Experimental Psychology-Animal Behavior Processes*, *10*(4), 426-436. doi: 10.1037//0097-7403.10.4.426

Mazur, J. E. (1986). Fixed and variable ratios and delays - further tests of an equivalence rule [Journal Article]. *Journal of Experimental Psychology-Animal Behavior Processes*, *12*(2), 116-124. doi: 10.1037/0097-7403.12.2.116

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

Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement [Journal Article]. *Animal Learning & Behavior*, *25*(2), 131-147. doi: 10.3758/BF03199051

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

Mazur, J. E., & Romano, A. (1992). Choice with delayed and probabilistic reinforcers - effects of variability, time between trials, and conditioned reinforcers [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *58*(3), 513-525. doi: 10.1901/jeab.1992.58-513

McDevitt, M. A., Dunn, R. M., Spetch, M. L., & Ludvig, E. A. (2016). When good news leads to bad choices [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *105*(1), 23-40. doi: 10.1002/jeab.192

McDevitt, M. A., Pisklak, J. M., Spetch, M., & Dunn, R. (2018). The influence of outcome delay on suboptimal choice [Journal Article]. *Behavioural Processes*, *157*, 279-285. doi: 10.1016/j.beproc.2018.10.008

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

Mcdiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference [Journal Article]. *Psychonomic Science*, *2*(7), 195-196. doi: 10.3758/Bf03343402

Mcsweeney, F. K. (1975). Matching and contrast on several concurrent treadle-press schedules [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *23*(2), 193-198. doi: 10.1901/jeab.1975.23-193

Menlove, R. L., Inden, H. M., & Madden, E. G. (1979). Preference for fixed over variable access to food [Journal Article]. *Animal Learning & Behavior*, *7*(4), 499-503. doi: 10.3758/Bf03209709

Millar, A., & Navarick, D. J. (1984). Self-control and choice in humans - effects of video game playing as a positive reinforcer [Journal Article]. *Learning and Motivation*, *15*(2), 203-218. doi: 10.1016/0023-9690(84)90030-4

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? [Journal Article]. *Learning & Behavior*, *40*(4), 439-447. doi: 10.3758/s13420-012-0065-7

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses [Journal Article]. *Psychological Methods*, *16*(4), 406-419. doi: 10.1037/a0024377

Morey, R. D., & Rouder, J. N. (2015). Bayesfactor: Computation of bayes factors for common designs [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=BayesFactor` (R package version 0.9.12-2)

Neuringer, A. J. (1969). Delayed reinforcement versus reinforcement after a fixed interval [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *12*(3), 375-383. doi: 10.1901/jeab.1969.12-375

Nieder, A. (2016). Representing something out of nothing: The dawning of zero [Journal Article]. *Trends in Cognitive Sciences*, *20*(11), 830-842. doi: 10.1016/j.tics.2016.08.008

Odum, A. L., Baumann, A. A. L., & Rimington, D. D. (2006). Discounting of delayed hypothetical money and food: Effects of amount [Journal Article]. *Behavioural Processes*, *73*(3), 278-284. doi: 10.1016/j.beproc.2006.06.008

Ojeda, A., Murphy, R. A., & Kacelnik, A. (2018). Paradoxical choice in rats: Subjective valuation and mechanism of choice [Journal Article]. *Behavioural Processes*, *152*, 73-80. doi: 10.1016/j.beproc.2018.03.024

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2018). nlme: Linear and nonlinear mixed effects models [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=nlme` (R package version 3.1-137)

Pisklak, J. M., McDevitt, M. A., Dunn, R. M., & Spetch, M. L. (2015). When good pigeons make bad decisions: Choice with probabilistic delays and outcomes [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *104*(3), 241-251. doi: 10.1002/jeab.177

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rachlin, H. (1971). On the tautology of matching law [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *15*(2), 249-251. doi: 10.1901/jeab.1971.15-249

Rachlin, H. (2000). *The science of self-control* [Book]. Cambridge, Mass.: Harvard University Press.

Rachlin, H., Battalio, R., Kagel, J., & Green, L. (1981). Maximization theory in behavioral psychology [Journal Article]. *Behavioral and Brain Sciences*, *4*(3), 371-388. doi: 10.1017/S0140525x00009407

Rachlin, H., Logue, A. W., Gibbon, J., & Frankel, M. (1986). Cognition and behavior in studies of choice [Journal Article]. *Psychological Review*, *93*(1), 33-45. doi: 10.1037/0033-295x.93.1.33

Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective-probability and delay [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *55*(2), 233-244. doi:

10.1901/jeab.1991.55-233

Roper, K. L., & Zentall, T. R. (1999). Observing behavior in pigeons: The effect of reinforcement probability and response cost using a symmetrical choice procedure [Journal Article]. *Learning and Motivation*, *30*(3), 201-220. doi: 10.1006/lmot.1999.1030

Rosati, A. G., Stevens, J. R., Hare, B., & Hauser, M. D. (2007). The evolutionary origins of humanpatience: Temporal preferences in chimpanzees, bonobos, and human adults [Journal Article]. *Current Biology*, *17*, 1663–1668. doi: 10.1016/j.cub.2007.08.033

Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. (1950). A preliminary study of training conditions necessary for secondary reinforcement [Journal Article]. *Journal of Experimental Psychology*, *40*(1), 40-45. doi: 10.1037/h0062153

Schroeder, S. R., & Holland, J. G. (1969). Reinforcement of eye movement with concurrent schedules [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *12*(6), 897-903. doi: 10.1901/jeab.1969.12-897

Shahan, T. A., & Cunningham, P. (2015). Conditioned reinforcement and information theory reconsidered [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *103*(2), 405-418. doi: 10.1002/jeab.142

Shapiro, M. S., Siller, S., & Kacelnik, A. (2008). Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the european starling (sturnus vulgaris) [Journal Article]. *Journal of Experimental Psychology-Animal Behavior Processes*, *34*(1), 75-93. doi: 10.1037/0097-7403.34.1.75

Sherman, J. A., & Thomas, J. R. (1968). Some factors controlling preference between fixed-ratio and variable-ratio schedules of reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *11*(6), 689-702. doi: 10.1901/jeab.1968.11-689

Shimp, C. P. (1969a). Concurrent reinforcement of two interresponse times - relative frequency of an interresponse time equals its relative harmonic length [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *12*(3), 403-411. doi: 10.1901/jeab.1969.12-403

Shimp, C. P. (1969b). Optimal behavior on free-operant experiments [Journal Article]. *Psychological Review*, *76*(2), 97-112. doi: 10.1037/h0027311

Silberberg, A., & Fantino, E. (1970). Choice, rate of reinforcement, and changeover delay [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *13*(2), 187-197. doi: 10.1901/jeab.1970.13-187

Skinner, B. F. (1981). Selection by consequences [Journal Article]. *Science*, *213*(4507), 501-504. doi: 10.1126/science.7244649

Smith, A. P., Bailey, A. R., Chow, J. J., Beckmann, J. S., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Stimulus value predicts choice over frequencies [Journal Article]. *Plos One*, *11*(7), e0159336. doi: 10.1371/journal.pone.0159336

Smith, A. P., Beckmann, J. S., & Zentall, T. R. (2017). Gambling-like behavior in pigeons: 'jackpot' signals promote maladaptive risky choice [Journal Article]. *Scientific Reports*, *7*, 6625. doi: 10.1038/s41598-017-06641-x

Smith, A. P., & Zentall, T. R. (2016). Suboptimal choice in pigeons: Choice is primarily based on the value of the conditioned reinforcer rather than overall reinforcement rate [Journal Article]. *Journal of Experimental Psychology-Animal Learning and Cognition*, *42*(2), 212-220. doi: 10.1037/xan0000092

Spetch, M. L., Belke, T. W., Barnet, R. C., Dunn, R., & Pierce, W. D. (1990). Suboptimal choice in a percentage-reinforcement procedure - effects of signal condition and terminal-link length [Journal Article]. *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 Article]. *Journal of the Experimental Analysis of Behavior*, *47*(1), 57-72. doi: 10.1901/jeab.1987 .47-57

Spetch, M. L., Mondloch, M. V., Belke, T. W., & Dunn, R. (1994). Determinants of pigeons choice between certain and probabilistic outcomes [Journal Article]. *Animal Learning & Behavior*, *22*(3), 239-251. doi: 10.3758/Bf03209832

Squires, N., & Fantino, E. (1971). Model for choice in simple concurrent and concurrent-chains schedules [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *15*(1), 27-38. doi: 10.1901/jeab.1971.15-27

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 [Journal Article]. *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 Article]. *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 [Journal Article]. *Psychonomic Bulletin & Review*, *17*(3), 412-416. doi: 10.3758/Pbr.17 .3.412

Stephens, D. W. (1981). The logic of risk-sensitive foraging preferences [Journal Article]. *Animal Behaviour*, *29*(May), 628-629. doi: 10.1016/S0003-3472(81) 80128-5

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory* [Book]. Princeton, N.J.: Princeton University Press.

Stevens, J. R., & Stephens, D. W. (2009). *The adaptive nature of impulsivity* (G. J. Madden & W. K. Bickel, Eds.) [Book Section]. Washington, DC: American Psychological Association.

Stewart, N. (2009). Decision by sampling: The role of the decision environment in risky choice [Journal Article]. *Quarterly Journal of Experimental Psychology*, *62*(6), 1041-1062. doi: 10.1080/17470210902747112

Stubbs, D. A. (1971). Second-order schedules and problem of conditioned reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *16*(3), 289-313. doi: 10.1901/jeab.1971.16-289

Torchiano, M. (2017). effsize: Efficient effect size computation [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=effsize` (R package version 0.7.1)

Tsetsos, K., Chater, N., & Usher, M. (2012). Salience driven value integration explains decision biases and preference reversal [Journal Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(24), 9659-9664. doi: 10.1073/pnas.1119569109

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted when rare outcomes are experienced (rarely)? [Journal Article]. *Psychological Science*, *20*(4), 473-479. doi: 10.1111/j.1467-9280.2009.02319.x

Vasconcelos, M., Machado, A., & Pandeirada, J. N. S. (2018). Ultimate explanations and suboptimal choice [Journal Article]. *Behavioural Processes*, *152*, 63-72. doi: 10.1016/j.beproc.2018.03.023

Vasconcelos, M., Monteiro, T., Aw, J., & Kacelnik, A. (2010). Choice in multi-alternative environments: A trial-by-trial implementation of the sequential choice model [Journal Article]. *Behavioural Processes*, *84*(1), 435-439. doi: 10.1016/j.beproc.2009.11.010

Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information [Journal Article]. *Scientific Reports*, *5*. doi: ARTN1387410.1038/srep13874

Weber, E. U., Shafir, S., & Blais, A. R. (2004). Predicting risk sensitivity in humans and lower animals: Risk as variance or coefficient of variation [Journal Article]. *Psychological Review*, *111*(2), 430-445. doi: 10.1037/0033-295x.111.2.430

Wilkie, D. M., & Spetch, M. L. (1978). Effect of sample and comparison ratio schedules on delayed matching to sample in pigeon [Journal Article]. *Animal Learning & Behavior*, *6*(3), 273-278. doi: 10.3758/Bf03209613

Williams, B. A. (1994). Conditioned reinforcement - experimental and theoretical issues [Journal Article]. *Behavior Analyst*, *17*(2), 261-285. doi: 10.1007/Bf03392675

Williams, B. A., & Dunn, R. (1991). Preference for conditioned reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *55*(1), 37-46. doi: 10.1901/jeab.1991.55-37

Williams, B. A., & Fantino, E. (1978). Effects on choice of reinforcement delay and conditioned reinforcement [Journal Article]. *Journal of the Experimental Analysis of Behavior*, *29*(1), 77-86. doi: 10.1901/jeab.1978.29-77

Zentall, T. R. (2016). Resolving the paradox of suboptimal choice [Journal Article]. *Journal of Experimental Psychology-Animal Learning and Cognition*, *42*(1), 1-14. doi: 10.1037/xan0000085

Zentall, T. R. (2017). An animal model of human gambling based on pigeon suboptimal choice [Journal Article]. *Research & Reviews: Neuroscience*, *1*(2), 27-37.

Zentall, T. R., Andrews, D. M., & Case, J. P. (2017). Prior commitment: Its effect on suboptimal choice in a gambling-like task [Journal Article]. *Behavioural Processes*, *145*, 1-9. doi: 10.1016/j.beproc.2017.09.008