

CROSS-SPECIES ANALYSIS OF CHOICES BETWEEN REINFORCER SEQUENCES

By

LEONARDO ANDRADE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT  
OF THE REQUIREMENTS FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

© 2010 Leonardo Andrade

To Patricia

## ACKNOWLEDGMENTS

I would like to thank my advisor, mentor, and friend, Timothy Hackenberg, for providing me the opportunity to study in the United States, and for his valuable guidance. I would like to thank my committee members, Marc Branch, Timothy Vollmer, David Smith, and Drake Morgan for their recommendations and insights. I would like to thank my friends Jennifer Rusak, Carla Lagorio, and Rachelle Yankelevitz for conducting many of the experimental sessions and for their input on this study. I would like to thank my friends Anne Macaskill and Katie Salsgiver for helping me with computer programming and data analysis essential in completing the experiments described here. I would like to thank Anthony DeFulio and Jorge Reyes, two wonderful people who have always been there for me. Their friendship was essential during these five years. I would like to particularly thank my father and my mother for their unconditional support, encouragement, and love. Without them, none of my accomplishments would have been possible. I would like to thank my brother Rodrigo, and my sisters Mariana and Joana for being a constant source of reinforcement in my life. Finally, I would like to deeply thank my best friend and loving wife Patricia. I would never have gone this far if she was not by my side.

## TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	7
LIST OF FIGURES.....	8
ABSTRACT.....	9
CHAPTER	
1 GENERAL INTRODUCTION.....	11
Introduction.....	11
Sequences of Reinforcers.....	15
Sequences of Reinforcers and the Hyperbolic-Decay Model.....	19
Cross-Species Comparison.....	22
Goals of the Present Study.....	24
2 EXPERIMENT 1.....	29
Introduction.....	29
Methods.....	30
Subjects.....	30
Apparatus.....	30
Training.....	31
Token-Exchange Training.....	31
Token-Production Training.....	31
Experimental Procedure.....	32
Results and Discussion.....	34
3 EXPERIMENT 2.....	44
Introduction.....	44
Methods.....	45
Subjects.....	45
Material, Location, and Equipment.....	45
Procedure.....	46
Sequences of Reinforcement.....	49
Delay Sensitivity Test.....	49
Questionnaire.....	50
Results and Discussion.....	50
Choice Patterns.....	51
Questionnaire Results.....	54

4	EXPERIMENT 3 .....	64
	Introduction .....	64
	Methods .....	65
	Subjects.....	65
	Apparatus .....	65
	Procedure.....	65
	Results and Discussion.....	67
5	GENERAL DISCUSSION .....	78
	Cross-species Analysis of Choice.....	79
	Limitations and Future Directions .....	82
	APPENDIX: QUESTIONNAIRE .....	84
	LIST OF REFERENCES .....	86
	BIOGRAPHICAL SKETCH.....	92

## LIST OF TABLES

<u>Table</u>		<u>page</u>
2-1	Mean proportion of choices (initial-link) and mean number of responses on the terminal-link for each alternative in Experiment 1. ....	43
3-1	Sequence of conditions for each subject in Experiment 2. ....	61
4-1	Experimental conditions implemented in Experiment 3. ....	74

## LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1	Hyperbolic-decay value of WOR, IMP, and STD sequences of reinforcement ... 28
2-1	Diagram of the terminal-links implemented for each sequence in Experiment 1. .... 39
2-2	Schematic of the Immediate Consumption condition ..... 40
2-3	Schematic of the Delayed Consumption condition. .... 41
2-4	Mean proportion of choices for each alternative in Experiment 1 ..... 42
3-1	Picture of the screen with all the visual components used in the concurrent chain schedule..... 58
3-2	Screen shot of the choice phase (initial-link).. ..... 58
3-3	Screen shot of the token exchange phase. .... 59
3-4	Diagram of the terminal-links implemented for each sequence in Experiment 2. .... 59
3-5	Diagram of the terminal-links implemented during delay sensitivity test..... 60
3-6	Mean proportion of choices for each alternative in Experiment 2.. ..... 62
3-7	Percentage of answers A and B given to the questionnaire questions. .... 63
4-1	Mean proportion of choices for each alternative in Experiment 3 for P883 and P942.. ..... 75
4-2	Mean proportion of choices for each alternative in Experiment 3 for P702 and P894..... 76
4-3	Mean proportion of choices on the richer alternative in function of its relative value..... 77

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

## CROSS-SPECIES ANALYSIS OF CHOICES BETWEEN REINFORCER SEQUENCES

By

Leonardo Andrade

May 2010

Chair: Timothy D. Hackenberg  
Major: Psychology

When given choices between sequences of reinforcers distributed over time, humans and non-humans show different patterns of preference: humans tend to prefer sequences that improve over time whereas non-humans tend to prefer sequences that worsen over time (i.e., sequences with more-highly valued reinforcers available early). The differences in performance may be due to fundamental differences in the way reinforcer sequences contribute to reinforce value. Alternatively, the differences may be attributed to methodological differences, specifically in the nature of the reinforcers used with humans (hypothetical or conditioned reinforcers) and non-humans (consumable reinforcers). A total of three experiments were conducted. Experiments 1 and 2 were aimed at comparing choice patterns across pigeons and humans between sequences of token and consumable reinforcers that provided the same overall rate, delivered at different delays. The results obtained in Experiment 1 and 2 showed generally similar choice patterns across species, with both preferring sequences with the shortest delay to the initial reinforcer in the series. Experiment 3 was an extension of Experiment 1 and was aimed at further assessing pigeon's sensitivity to selectively delayed reinforcers in

a sequence. Overall, results obtained in this study are broadly consistent with models of temporal discounting expanded to include the impact of sequences of delayed reinforcers acting in parallel from the time of the choice.

## CHAPTER 1 GENERAL INTRODUCTION

### **Introduction**

How organisms allocate behavior among different sources of reinforcement has been the subject matter of a wide range of scientific endeavors. From ethologists studying animals in naturalistic settings deciding among alternative food patches to economists analyzing consumers preferences, the quest for unraveling the effects of relevant variables upon an organism's choices has been an important issue for behavioral, biological, and social sciences in general.

The allocation of behavior among concurrent sources of reinforcement can be conceptualized as choice (Herrnstein, 1970; Pierce & Cheney, 2008; Skinner, 1950). From this point of view, choice is not mental mechanism; instead, choice is simply a description of how behavior is allocated among concurrent sources of reinforcement (Herrnstein, 1970; Pierce & Cheney, 2008; Skinner 1950). Herrnstein (1970), for example, argues that even in relatively simple environments, such as a Skinner box with a single key, the organism is choosing between the reinforcers produced by pecking the key and the other reinforcers derived from all other possible actions. In his words:

Even in a simple environment like a single-response operant-conditioning chamber, the occurrence of the response is interwoven with other, albeit unknown, responses, the relative frequencies of which must conform to the same general laws that are at work whenever there are multiple alternatives...No matter how impoverished the environment, the subject will always have distractions available, other things to engage its activity and attention, even if these are no more than its own body, with its itches, irritations, and other calls for service. (Herrnstein, 1970, p. 254-55)

Choice is an important topic and has received attention from many different scientific disciplines because choices (or decisions) typically produce important and often conflicting delayed consequences – not only for the individual but for society as

well. Studying, saving money for retirement, recycling, exercising, abstaining from drugs; all of these are examples of choices that involve important delayed consequences that are often in conflict with more immediate reinforcer alternatives. The term intertemporal choice is frequently used to refer to these types of choices which involve some kind of trade-off between costs and benefits occurring at different points in time (Frederick, Loewenstein, & O'Donoghue, 2002). A critical component in this type of choice is delay discounting, the process typically invoked to explain why organisms seem to disproportionately overvalue reinforcers arranged closer in time (Berns, Laibson, & Loewenstein, 2007; for a review refer to Frederick, et al., 2002).

The interval between a response and the contingent reinforcer has been acknowledged as fundamental in learning, discrimination, and choice since the inception of Psychology as a scientific field. The importance of reinforcer delay probably dates since Thorndike (1913), as can be noted from the following quote:

As a corollary to the law of effect we have the fact that the strengthening effect of satisfyingness varies with its intimacy with the bond in question as well as with the degree of satisfyingness. Such intimacy, or closeness of connection between the satisfying state of affairs and the bond it affects, may be due to close temporal sequence or attentiveness to the situation and response. Other things being equal, the same degree of satisfyingness will act more strongly on a bond made two seconds previously than on one made two minutes previously, – more strongly on a bond between a situation and a response attended to closely than on a bond equally remote in time in an unnoticed series. (Thorndike, 1913, p. 172-173)

The relevance of reinforcer delay was elaborated by Hull (1943), assuming a central role in his theory of reinforcement generalization and goal gradients. Although Hull did not use the term delay discounting, he used empirical data from experiments involving manipulation of the delay to reinforcement on choice (e.g., Anderson, 1932;

Perin, 1943; Wolfe, 1934; in Hull 1943) in support of what was essentially an exponential discounting function.

The term delay discounting is generally used today to refer to the decrease in subjective value of a reinforcer (or commodity in Economics) as a function of time. As the delay to contact the reinforcer increases, the present value of a given reinforcer decreases (i.e., is discounted). Economists refer to this functional relation as 'positive time discounting', one part of the Discounted Utility (DU) model (Samuelson, 1937) used to make predictions of intertemporal choices in relation to economic assets. It is beyond the purposes of this paper to provide a full description of the DU model; suffice to say that it is an exponential model that incorporates a free parameter representing a fixed discount rate. As a result, it predicts that the subjective value of a reinforcer (or commodity) always decreases in the same proportion as its delay. Among the assumptions of the DU model, it is important to highlight for the purpose of the present paper, the following two: (1) Preferences do not change over time; they are irreversible. That is, if an organism shows preference for outcome *a* at time *x*, over outcome *b* at *y*, it is assumed that preference for *a* over *b* will remain constant regardless of any increment or decrement in time, as long as this change in time is common to both outcomes. (termed 'time consistency' by economists); (2) The overall value of a sequence of outcomes is equal to the sum of the present discounted value of each member of the sequence; that is, each outcome in the series is viewed as *independent* and *additive* (for a review, see Frederick, et al., 2002),

Although the DU model remains the standard model in economics, empirical findings have shown violations of some of the key predictions of the model. Economists

call these violations 'anomalies'. Probably the most robust such anomaly comes from the empirical findings related to assumption (1) above, which show that preference is not static or time consistent, and that it does reverse when a common delay is added or subtracted from the choice alternatives. Interestingly, these findings were first obtained not by economists but by experimental psychologists using non-human subjects (Ainslie, 1974, 1975; Rachlin & Green, 1972). Within the Rachlin and Green study, pigeons given a choice between a smaller sooner reinforcer (SSR) and a larger later reinforcer (LLR) predominantly chose the SSR alternative. However, when an additional constant delay was added to both alternatives, preference shifted toward the LLR alternative. This finding has been replicated with pigeons (Ainslie & Herrnstein, 1981; Green, Fisher, Perlow, & Sherman, 1981) and humans (Green, Fristoe, & Myerson, 1994; Kirby & Herrnstein, 1995; Millar & Navarick, 1984; Solnick, Kannenberg, Eckerman, & Waller, 1980).

As a result, additional models of delay discounting have been proposed, the most popular of which are ones based on a hyperboloid function. Such functions can accommodate preference reversals, and have been shown to provide a good fit to the data (Green & Myerson, 2004; Kirby, 1997; Mazur, 2001). The main feature of the hyperbolic models is that the discounting rate of the function decreases more sharply in the short run compared to the long run. In other words, the curve follows a non-linear-decreasing function that gets steeper with the proximity of the reinforcer. A hyperbolic model that has been widely used is the following model proposed by Mazur (1984):

$$V = \frac{A}{1 + KD} \quad (1-1)$$

where  $V$  is the present value of a reinforcer delivered at delay  $D$ ,  $A$  represents the amount or the undiscounted value of a reinforcer, and  $K$  is a free parameter, a constant that determines how sharp the value of a reinforcer decreases over time. According to Mazur (2001), the term value refers to the strength of a reinforcer, and the model predicts that when faced with a choice between reinforcer alternatives, organisms will choose the alternative that has the highest value at the instant of choice.

### **Sequences of Reinforcers**

The great majority of studies on choice have been focused on single outcomes. However, as noted by Kirby (2006), “rather than occurring all at once at a single moment in time, the consequences of our choices are more often scattered over future time periods” (p.273). Interest in the systematic investigation of the effects of multiple outcomes on choice is fairly recent. With a few exceptions, it was not until the late 1980’s and early 1990’s that publication of research involving sequences of outcomes began to appear in the literature. The general lack of interest in reinforcer sequences is likely due to the assumption that the principles related to individual outcomes were extendable to series of outcomes as well (Ariely & Loewenstein, 2000). That is, the generality or external validity was assumed rather than empirically tested. As a result, not much is known about sequences of reinforcers and the possible interactive effects that the reinforcers that are part of a sequence may have with each other (Kirby, 2006).

A temporal sequence can be defined as “a series of outcomes spaced over time” (Loewenstein & Prelec, 1993). The manner in which events/outcomes are conceptualized and separated in time varies across studies. It may therefore be more appropriate to view changes of events in time as lying on a continuum. For instance, a sequence can be conceptualized as a series of discrete events (or segments) clearly

separated by time (e.g., specific amounts of annual income to be received in the following 5 years), or it can be viewed as a more cohesive or unitary segment on which the patterns of a given stimulus dimension change continuously over time (e.g., patterns or streams of varying sound intensities, or continuous change in water temperature).

A significant study involving choices between sequences of outcomes was carried out by Loewenstein and Prelec (1993), in which human subjects answered a questionnaire containing sequences of qualitatively different hypothetical reinforcers, arranged in different orders. The authors manipulated the number of outcomes on each sequence (from 2 up to 5 outcomes) as well as the inter-event duration. In one set of questions, for example, subjects were first asked whether they preferred to have dinner at a fancy French restaurant or at a local Greek restaurant. Subsequently, the participants who preferred the French restaurant were asked the following two questions: (1) whether they would prefer to have dinner at the French restaurant on Friday in 1 month or on Friday in 2 months; and (2) whether they preferred to have dinner at the French restaurant on Friday in 1 month and dinner at the Greek restaurant on Friday in 2 months; or dinner at the Greek restaurant on Friday in 1 month and dinner at the French restaurant on Friday in 2 months. (Those who initially preferred the Greek restaurant received the converse choices.)

The results of this study showed an interesting and seemingly dichotomous pattern of responses. When the question involved a single outcome, the majority of the subjects preferred to have the French dinner sooner rather later (question numbered (1) above). However, when the question involved two outcomes, the Greek and the French restaurants (question numbered (2) above), the same participants chose to defer their

favorite restaurant (French) and have the Greek dinner sooner. This pattern of choices favoring sequences that improve in time was replicated in the same study with questions involving weekends with friends versus weekends with an abrasive aunt, and other questions containing up to 5 outcomes with varying degrees of hypothetical pleasure.

There have been other studies with humans which provide further support for preference for sequences with improving trends. The great majority of studies in this line of research involve some kind of hypothetical monetary outcomes, such as hypothetical payment (Loewenstein & Sicherman, 1991; Hsee, Abelson, Salovey, 1991; Matsumoto, Peecher, & Reech, 2000; Guyse, Keller, & Eppel, 2002; Read & Powell, 2002), gambling (Ross & Simonson, 1991; Read & Powell, 2002), and stock markets (Ariely & Zauberman, 2000; Matsumoto et al., 2000). In addition to money, research on sequences of outcomes have also assessed the effects of other types of hypothetical scenarios, such as hypothetical grades (Hsee et al., 1991); vacations and meals in restaurants (Loewenstein & Prelec, 1993; Matsumoto et al., 2000; Montgomery & Unnava, 2009); trends in the quality of the environment (Guyse et al., 2002); health (Chapman, 2000; Guyse et al., 2002); and subjective experiences of discomfort (Varey & Kahneman, 1992). The few studies using real outcomes involved earning points exchangeable for money (Schmitt & Kemper, 1996), playing video games (Ross & Simonson, 1991); listening to music (Montgomery & Unnava, 2009); aversive noise (Ariely & Zauberman, 2000; Schreiber & Kahneman, 2000; Ariely & Loewenstein, 2000); aversive cold temperatures (Kahneman, Fredrickson, Schreiber, & Redelmeier, 1993); and aversive heat and mechanical pressure (Ariely, 1998).

In general, the results of the studies cited above indicate that human subjects prefer sequences comprised of outcomes containing an improving/increasing trend over sequences of outcomes containing worsening/declining trends or sequences in which the components remain constant through time. It is worth emphasizing this result because it indicates that humans discount sequences of outcomes in a different fashion than they do single outcomes. Contrary to single-outcome contingencies in which individuals typically show positive time discounting, preference for improving sequences may actually show no discounting, or even negative discounting (Chapman, 2000; Loewenstein & Prelec, 1993). Loewenstein and Prelec (1993) argue that preference for improvement is an overdetermined phenomenon in humans. They claim that when the context of the sequences is highlighted prior to the choice, and the individual's attention is drawn to the sequential nature of events, an effect of the interaction between the components – the gestalt properties – emerge. Thus, their results and interpretation challenge separable formulations such as the weighted utility model common in economics (Samuelson, 1937) because they indicate that the overall value of a sequence is not equal to the sum of individual components.

The results of the studies involving intertemporal sequences described thus far in the present study should be interpreted with caution due to the experimental protocol adopted by them. The stimulus presented to the participants (i.e., the independent variable) is typically hypothetical situations described in sentences or depicted via some kind of graphic representation or diagram, in which the size of the geometric figures represent the intensity of some property/dimension of stimulus events. The dependent measure is comprised of choices, ranking, or rating of these stimuli. When real stimuli

are used, subjects are typically exposed to the stimulus first and then are required to rank or rate it on some arbitrary scale after the exposure. In some studies, the participant is also required to provide ratings on-line (continuously) in addition to the posterior rating. This is done so that the experimenter can compare the participant's "global" and "local" evaluations of the events. When choice is used as a dependent variable with real events, choice is accessed via deception. That is, after being exposed to some intertemporal sequences, participants are asked to choose a sequence to be re-exposed at the end of the experiment, but in fact they are never exposed to the choice they commit (e.g., Ariely & Loewenstein, 2000; Schreiber & Kahneman, 2000; Kahneman et al., 1993). Therefore, most of the time, what some authors call choice or preference should be seen more as kind of "pseudo-choice" or "pseudo-preference".

### **Sequences of Reinforcers and the Hyperbolic-Decay Model**

Preference for an improving over a worsening sequence thus runs counter to predictions made by most psychological and economic theories based on delay discounting (Brunner, 1999). These theories predict that individuals should always choose the alternative that produces the best outcome first, followed by reinforcers with decreasing values (i.e., worsening/declining sequence), rather than the reverse.

The following hyperbolic equation, proposed by Mazur (1986) has been used in choice studies with multiple reinforcers with humans and non-humans (Brunner & Gibbon, 1995; Brunner, 1999; Kirby, 2006; Mazur, 1986, 2007; Shull, Mellon, & Sharp, 1990). This model is an extrapolation of the Equation 1-1, and is represented algebraically as

$$V = \sum_{i=1}^n \frac{A}{1 + KD_i} \quad (1-2)$$

where  $V$  represents the value of the sequence of reinforcers delivered after some delay  $D$ ;  $n$  represents the number of reinforcers in the sequence,  $A$  is the undiscounted value of a reinforcer (the value of the reinforcer if it was delivered immediately), and  $K$  is a free parameter that determines how sharply  $V$  decreases in direct function of  $D$ .

In short, according to this hyperbolic model, the discounted value of a sequence of reinforcers is equal to the sum of the present discounted values of each individual reinforcer in the sequence (this form of calculating the hyperbolic value of sequence has been called parallel discounting by Brunner & Gibbon, 1995). In other words, the value of a sequence is simply the sum of the value of the individual members of the sequence, as if they were separable and independent (see Brunner & Gibbon, 1995, for alternative hyperbolic delay discounting models within sequences). In a sequence comprised of four reinforcers, for instance, the value of the sequence would be calculated as follows:

$$V = \frac{A_1}{1 + KD_1} + \frac{A_2}{1 + KD_2} + \frac{A_3}{1 + KD_3} + \frac{A_4}{1 + KD_4} \quad (1-3)$$

This model is very similar to the one proposed earlier by McDiarmid and Rilling (1965), in which they calculated the value of a sequence as the sum of the immediacies of each reinforcer in the series (i.e., the reciprocal of the delays). It is important to note that both McDiarmid and Rilling (1965) and the DU model in economics makes the same assumption of the hyperbolic-decay model in relation to the independent contribution of each member of the sequence.

This hyperbolic-decay equation, has been tested with multiple reinforcers in rats (Brunner, 1999; Brunner & Gibbon, 1995; Mazur, 2007), pigeons (Mazur, 1986, Shull et al., 1990), and humans (Kirby, 2006), and the results have shown that it provides a very good fit to the data (for contrasting results see Moore, 1979, 1982). The studies conducted by Mazur (1986, 2007), Brunner and Gibbon (1995), and to some extent Kirby (2006) are important to emphasize because they tested the predictions of the parallel model using a titration procedure, a technique very suitable for testing the quantitative properties and predictions of mathematical models. It is also important to note, however, that with the exception of Brunner and Gibbon (1995) and Brunner (1999, Exp.1), all the other studies which evaluated the predictions of the parallel model cited above, analyzed preference among sequences that contained unequal number of reinforcers. This is important to note because although these studies have shown evidence that the value of subsequent elements in the sequence is added (and independent), the number of reinforcers or rate of reinforcers delivered was not equal between the alternative sequences and these variables could have had some confounding effect on choice that was not accounted for by this model.

Research with non-human animals with multiple and equal numbers of reinforcers is relatively scarce. In one of the few studies along these lines, Brunner (1999, Exp. 1) investigated rats' choices between sequences that improved or worsened *in time*. More specifically, rats were given repeated choices between a sequence in which the inter-pellet delay increased over time (worsening sequence) and an alternative sequence in which the inter-pellet delay decreased over time (improving sequence). Contrary to the results found with humans and qualitatively different

hypothetical reinforcers (e.g., Loewenstein & Sicherman, 1991; Loewenstein & Prelec, 1993), but consistent with delay discounting models such as Equations 1-2 and 1-3, rats strongly preferred the worsening sequence.

### **Cross-Species Comparison**

Although an increasing number of studies have shown that humans, too, discount delayed events (e.g., Green, Fry, & Myerson, 1994; Solnick, et. al, 1980; Estle, Green, Myerson, & Holt, 2007; Logue, Peña-Correal, Rodriguez, & Kabela, 1986; Kirby, 2006), some data indicate that human choice may be less delay sensitive than choice in other species. Other results suggest that the degree to which discounting occurs varies depending on the kind of reinforcer used (i.e., consumable reinforcer, points exchanged for money, hypothetical reinforcers; see Navarick, 2004, for a discussion). In general, an estimated discount rate (i.e., the  $K$  value) of approximately 0.014 is found with humans (Rachlin, Raineri, & Cross, 1991), whereas a value of approximately 1 is reported with pigeons (Mazur, 1984, 1987).

In studies of “self-control” type choices (i.e., studies involving choice between a smaller sooner reinforcer and a larger later reinforcer), non-humans (typically birds and rats) have been shown to be highly impulsive (delay sensitive), predominantly choosing smaller-sooner reinforcers (Ainslie, 1974; Logue & Peña-Correal, 1984; Mazur & Logue, 1978). Conversely, humans tend to choose the larger-later reinforcer (Logue et al., 1986; Logue, King, Chavarro, & Volpe, 1990; Flora & Pavlik, 1992; Hyten, Maden, & Field, 1994; see review by Logue, 1988). In general, human performance is well accounted by a rate maximizing rule of the following form:

$$V = \frac{A}{T} \quad (1-4)$$

Where  $V$  represents the value of a given course of action,  $A$  is the size of a reinforcer, and  $T$  is the time between reinforcers.

The main difference between the hyperbolic discount and the global maximization model is in the way reinforcer variables are averaged. For the hyperbolic discount model, value is a decreasing nonlinear function of reinforcer delay, while that for global maximization model, value is a nondiscounted arithmetic average of aggregate reinforcer per unit time.

The underlying cause for the distinct results obtained across humans and non-humans on choice studies is an important question that is still unanswered. The following reasons may be raised to account for that: species qualitative differences, such as human higher cognitive abilities or verbal behavior that makes them capable of making more rational decisions (Horne & Lowe, 1993); or methodological differences, more precisely, lack of methodological equivalence in studies across species. One procedural difference is the type of reinforcer used. In studies with humans, hypothetical reinforcers are typically used, while with non-humans, actual reinforcers (typically, food) are used. Another procedural difference is the format for presenting choices. With humans, hypothetical choices are presented once and the subject is not exposed to real outcomes, while non-humans, on the hand, are repeatedly exposed to choices that produce real outcomes. This repeated exposure, then, allows for choices to be affected by experience with the task.

Instead of hypothetical reinforcers, some experiments with humans have used points exchangeable for money at the end of the session (e.g., Flora & Pavlik, 1992; Hyten, et. al, 1994; Logue et. al, 1986). In this type of procedure, some of the

methodological differences pointed out above, such as the consumable nature of the reinforcers and the lack of opportunity to be affected by the experience with the task, may not apply. However, one significant difference remains: the delay to the consumption of the reinforcer. In non-humans experiments, reinforcer access typically occurs at the end of each choice trial, while in human experiments that incorporate the point-money system, the access to the reinforcer usually occurs at the end of the session or at the end of the experiment. In other words, the moment on which the opportunity to consume the reinforcer is made available differs in studies conducted in humans and non-humans. With humans, the opportunity to consume the reinforcer is usually more delayed in the session compared with non-human experiments. Empirical evidence of the strong control over choice responses exerted by the time frame over which reinforcers are accessed has been amply demonstrated by Hyten et al. (1994) with humans and by Jackson and Hackenberg (1996) with pigeons. Jackson and Hackenberg (1996), for instance, have shown that pigeon's choice was largely controlled by the moment on which conditioned reinforcers (tokens) could be exchanged within the session. When tokens could be exchanged immediately after they were obtained, pigeons showed the pattern of choices typically found in non-humans, that is, they exhibited strong preference for the smaller-sooner reinforcer. However, when the opportunity to exchange tokens was made more delayed in the session (condition that served as analogue to human experiments) pigeons showed preference for the larger-later reinforcer, a pattern consistent with the typical performance found in humans.

### **Goals of the Present Study**

In order to make valid animal-human extensions or comparisons, it is essential that the experimental contingencies are methodologically equivalent. This

methodological standardization is one of the broad aims of the present study. Here, we attempt to bring the choice procedures used with humans and pigeons into better alignment. To do so, we made the typical experimental procedure used with pigeons more “human-like” by introducing conditioned reinforcers (tokens) that were exchangeable for consumable reinforcers (food) at different times of the session. To facilitate inter-species comparisons, we also implemented tokens as conditioned reinforcers in the human experiment that were also exchangeable at different times of the session for consumable reinforcers. But rather than food, humans exchanged their tokens for video clips from popular television shows, a reinforcer that has proven quite effective in laboratory research with humans (Hackenberg & Pietras, 2000; Lagorio & Hackenberg, 2010; Locey, Pietras, & Hackenberg, 2009; Navarick, 1996, 1998).

Three experiments were conducted to assess the effects of intertemporal delays of a sequence of multiple reinforcers on choice. In the first two experiments, we investigated the pattern of choice across species (pigeons and humans) between sequences of token and consumable reinforcers that provided the same overall rate, delivered at different temporal patterning. A sequence with increasing inter-reinforcement delays, Worsening (WOR), a sequence with decreasing inter-reinforcement delays, Improving (IMP), and a sequence comprised of fixed inter-reinforcement delays, Standard (STD), were implemented in Experiment 1 (pigeons) and 2 (humans). In addition, two economic contexts under which the tokens earned could be exchanged for consumable reinforcers were also incorporated in the study. In one economic context, Delayed Consumption (DC), token exchange opportunities were made available after the delivery of the last reinforcer in the sequence. In the other

context, Immediate Consumption (IC), token exchange opportunities were made available immediately after each token delivery. Thus, two independent variables were manipulated across conditions in Experiment 1 and 2: the intertemporal delay within the sequence, and the delay to the exchange period. In addition, this study was also aimed to evaluate the ordinal predictions of the hyperbolic decay model and the global maximization.

To better illustrate the predictions of the hyperbolic model, the value of individual reinforcers in each sequence, as well as their total sum in the IC condition is shown in Figure 1-1. These values were plotted assuming the free parameter  $K$  value being equal to 1, a value typically reported with pigeons (Mazur, 1984, 1987). For simplicity, the parameter  $A$  was also arbitrarily set equal to 1. (This was done because reinforcer amount was not manipulated and remained constant throughout this study.) As depicted in the bottom right graph, the hyperbolic delay discounting predicts differential preference in the IC condition. According to the predictions of Equation 1-2, the WOR sequence should be preferred over both of the other sequences, and the STD sequence should be preferred over the IMP sequence. In the DC conditions, on the other hand, this model predicts indifference because regardless of the specific sequence, food is made available at the same time (at the end of the terminal link). Global maximization model (Equation 1-4) predicts indifference across all conditions because reinforcer variables are averaged arithmetically.

As can be seen in Figure 1-1, the hyperbolic-decay model predicts that the first reinforcer in the series exerts a strong control over preference. The extent to which the delays to subsequent reinforcers in a sequence control choice is still largely unknown

and this was the purpose of Experiment 3. Using a similar procedure of Experiment 1, Experiment 3 was aimed at assessing pigeon's sensitivity to selectively delayed reinforcers in a sequence.

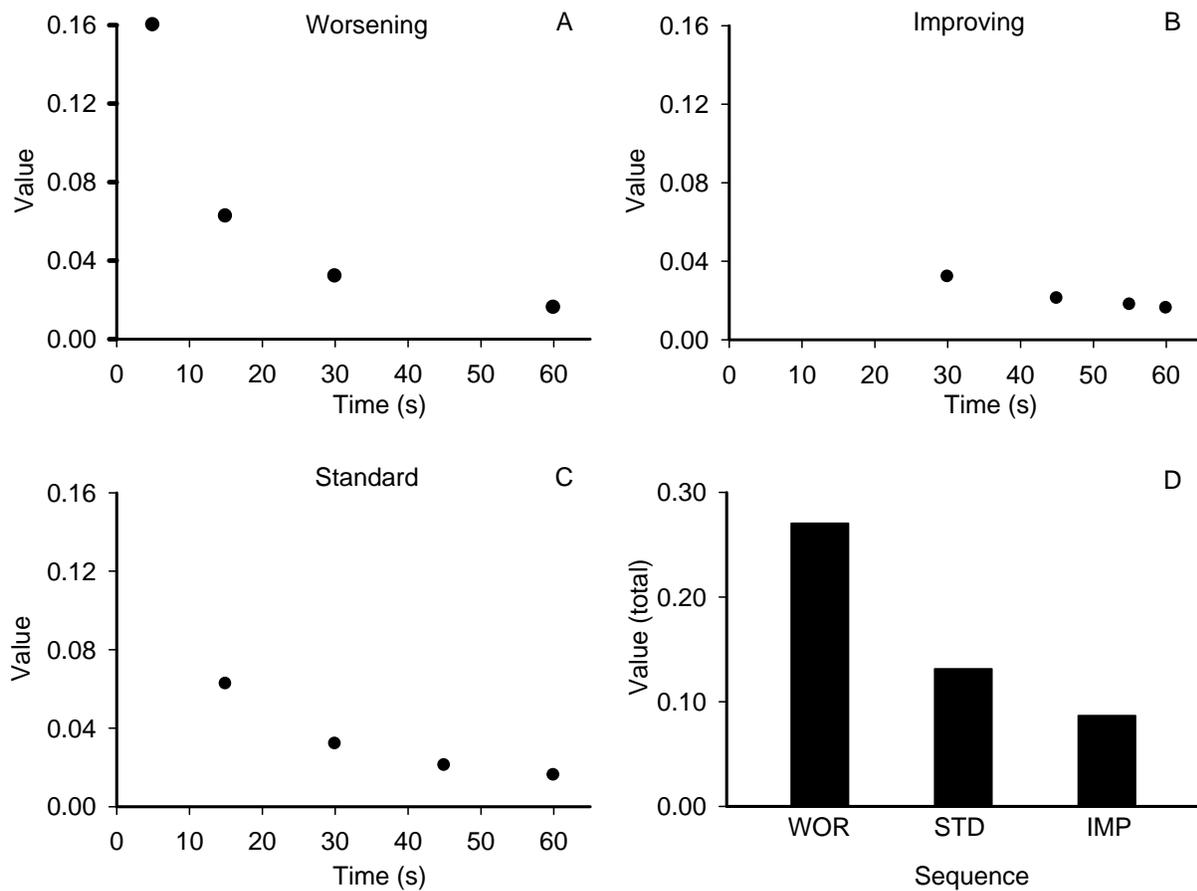


Figure 1-1. Hyperbolic-decay value of WOR, IMP, and STD sequences of reinforcement. A) Hyperbolic value of each reinforcer in the Worsening sequence; B) Hyperbolic value of each reinforcer in the Improving sequence; C) Hyperbolic value of each reinforcer in the Standard sequence; D) Total value of each sequence (in the IC condition). In A, B, and C, the x-axis refers to the delay to each reinforcer in the series.

## CHAPTER 2 EXPERIMENT 1

### **Introduction**

Brunner (1999, Exp. 1) found that when rats are given repeated choices between a sequence in which the inter-pellet delay increased over time and an alternative sequence in which the inter-pellet delay decreased over time, rats preferred the former. Using pigeons as subjects, the present experiment also investigated choices between sequences of equal number of reinforcers in which the inter-reinforcement delay increased (worsening sequence) or decreased (improving sequence). In addition, a third sequence comprised of reinforcers delivered at fixed delays (standard sequence) was implemented. The choices occurred in the context of a discrete-trial procedure. Choices in the initial link produced sequences comprised of token and consumable reinforcers delivered at different temporal patterning but the same overall rate in the terminal-link.

Another variable of interest was the economic context, which determined when the tokens could be exchanged for consumable reinforcers. In the Immediate-Consumption conditions, token-exchange opportunities were given immediately after the delivery of each token, whereas in the Delayed-Consumption conditions, token-exchange opportunities were given after the delivery of the last token in the sequence.

In addition to the manipulation of the inter-reinforcement delay and exchange opportunity, the broader goal of this study was to compare choices between reinforcer sequences across species. In order to make valid animal-human extensions or comparisons, it is essential that the experimental contingencies are methodologically as equivalent as possible. The experimental procedures adopted in this experiment with

pigeons and the following one with humans was an attempt to bring the methods into greater alignment.

## **Methods**

### **Subjects**

Four naïve male White Carneau pigeons (*Columba livia*) served as subjects. The birds were housed individually in a humidity and temperature-controlled colony room where they had continuous access to water and grit. The lights in the colony room were on from 7:00 am to 11:00 pm. The pigeons were maintained at approximately 83% of their free-feeding weights via additional post-session feeding when necessary.

### **Apparatus**

A standard operant conditioning chamber with a modified stimulus panel served as experimental location. The working space measured 35 cm high by 35 cm wide by 30.5 cm long. The stimulus panel contained three horizontally-aligned plastic keys, located 8.7 cm from the top. Each key had a circumference of 7.85 cm and were located 5.7 cm apart from each other. The center key was trans-illuminated red or white, while the two side keys were trans-illuminated white, green, or yellow. The minimum force required to operate each key was approximately 0.3 N. A row of 12 horizontally-aligned and evenly spaced red stimulus lights was inserted in the panel at 4.5 cm below the ceiling and served as tokens. The circumference of each token was approximately 4.71 cm. Only the four centermost token lights were used and were lit sequentially from left to right or from right to left contingent on response on the left or right key, respectively. A white houselight centrally located above the row of tokens remained on throughout the session. A food hopper delivered the primary reinforcer, mixed grain, accessed through an aperture measuring 5.8 cm by 5 cm, and centrally

located 10.5 cm above the floor. A white light illuminated the hopper while it was suspended and food was being presented. The experiment was controlled through a microcomputer and MED-PC interface located in an adjacent room.

## **Training**

All pigeons were first exposed to 1 or 2 days of adaptation to the chamber, followed by sessions of magazine training, key-peck shaping, token-exchange training, and token-production training. After adaptation, pigeons were magazine trained, in which the food hopper was raised at irregular intervals, until they reliably approached and ate from the hopper. At this point, the center key was illuminated red and pecking was reinforced with food via the method of successive approximations.

## **Token-Exchange Training**

When key-pecking had been established, the pigeons were exposed to sessions in which the tokens were paired with food. This was accomplished by illuminating the center key and the four centermost tokens. Each peck on the center red (exchange) key produced 3-s access to food and a short (.04 s) beep, while turning off the center key and a single token. Immediately after the food hopper was lowered, the red center key was re-illuminated and a new cycle began. Each subsequent exchange response produced the same events, until all 4 tokens had been exchanged for food. The tokens were always exchanged in sequence (either from left to right, or vice versa), with the starting position determined randomly each 4-token cycle. Sessions lasted for 12 trials (48 reinforcers).

## **Token-Production Training**

Following token-exchange training, the pigeons were trained to produce tokens by pecking the side (white) keys. On each cycle, the left or right key was illuminated

white, and a single peck produced 4 tokens simultaneously, a short beep, and the center (exchange) key. Each peck on the exchange key produced food, as described above, until all 4 tokens had been exchanged, after which a new cycle began with one of the side keys lit white. The position of the active key (left or right) was determined randomly in each cycle, with the restriction that each occurred 6 times in each session. Sessions lasted for 12 cycles.

### **Experimental Procedure**

Following training, the pigeons were given repeated choices between sequences of tokens and food using a concurrent-chains schedule with two links. The initial link consisted of a concurrent fixed-ratio 1 fixed-ratio 1 (Conc FR1 FR1) in the presence of white keys. Thus, a single peck on either white side key produced one of two terminal-link stimuli—a green key or a yellow key. The terminal link was comprised of a sequence of delays to each of 4 tokens and an exchange schedule. Each token presentation was accompanied by a brief beep. The tokens were presented from left to right (if the initial-link response had been on the left key) or from right to left (if the initial-link response had been on the right key). Each terminal link was followed by a 5-s inter-trial interval (ITI) during which only the houselight remained on. At the beginning of each choice cycle, the center key was lit white, a single peck on which produced the initial-link stimuli. This initiation response was implemented to reduce the likelihood of position biases by ensuring that the initial links began with a response that was equidistant to the side keys.

Figure 2-1 shows a schematic of the terminal-link events. Three different sequences providing the same overall rate of reinforcement but different temporal patterning were used. All sequences included four tokens in the terminal link spread

over the same overall time-span, timed from terminal-link onset. More specifically, the tokens were presented response-independently in the terminal-link and the total time on which all four were presented was held constant at 60 s. The *Standard* (STD) sequence contained four tokens presented at equal inter-temporal delays of 15 s. In the *Worsening* (WOR) sequence the delays between successive tokens increased (5 s, 10 s, 15 s, and 30 s), whereas in the *Improving* (IMP) sequence the delays between successive tokens decreased (30 s, 15 s, 10 s, and 5 s). It is important to note that the inter-temporal delay between token presentation, as well as overall duration of the terminal link depicted in Figure 2-1 below, does not include the token exchange and food consumption periods.

Besides manipulation of the inter-temporal delay of token presentation, the other major independent variable was the scheduling of the token-exchange period. Tokens were either exchanged for food after the delivery of the fourth token in the sequence – *Delayed Consumption* (DC) – or immediately after the delivery of each individual token within the sequence – *Immediate Consumption* (IC). The token-exchange schedule was signaled by the darkening of the side keys and the illumination of the center red key. During the IC conditions, when the token and exchange key was presented, the delay timer for the next token delivery in the sequence was stopped until the token earned was exchanged. After the completion of the token exchange (precisely after the food hopper was lowered), the timer that controlled the delay to the next token was reset and the key pecked in the initial-link was re-illuminated until the presentation of the next token. The exchange schedule was identical to the exchange schedule used during training, except that each token was exchanged for 2.5 s rather than 3 s access

to food to maintain stable running weights. Refer to Figures 2-2 and 2-3 below for a schematic of the IC and DC conditions, respectively.

Sessions were scheduled once per day, seven days per week. A session was composed of 12 cycles, 2 forced-choice and 10 free-choice cycles. On forced-choice cycles, only one of the two initial-link keys was lit. Such cycles were implemented to ensure adequate exposure to both alternatives of the concurrent pair. The order of presentation of each forced-choice alternative was randomly determined but both alternatives were always presented once during the initial two cycles each session. The final 10 cycles each session were free-choice cycles, in which both alternatives were available.

Table 2-1 shows the order of conditions and the number of sessions conducted at each. Some conditions were replicated to assess reliability of preferences. There were also frequent reversals of the contingencies to assess position and color biases. In addition, the position of each alternative sequence was counter-balanced across subjects. Because Pigeon P894 developed a side bias during the experiment, it was exposed to two additional conditions with a higher ratio of forced-choice: free-choice trials aimed to rectify the bias.

Conditions remained in effect for a minimum of 12 sessions and until the proportion of initial-link responses was deemed stable according to the following criteria: (a) absence of increasing or decreasing trend across 5 consecutive sessions; and (b) absence of the highest or lowest point in the condition.

## **Results and Discussion**

Figure 2-4 shows the mean proportion of choices for the stable (last 5) sessions of each experimental condition. Each graph represents the performance of a single

pigeon, whose number is located at the top-left corner. Bars on the left and right side of each graph represent the choices allocated to the left and to the right side alternatives, respectively. The filled bars depict the proportion of responses during the IC conditions and the unfilled bars the proportion of responses during the DC condition. The labels on the left correspond to the actual sequence allocated on the left choice alternative, and the labels on the right indicate the sequence allocated to the right alternative. During the experiment each sequence was paired with the other two sequences and many reversals and replications were conducted. To facilitate the visual analysis, each pairwise comparison was grouped and separated by the dotted lines. This is important to note because the order of conditions displayed in Figure 2-4 do not reflect the exact order on which the pigeons were exposed to some of the experimental conditions. (The order of conditions is shown in Table 2-1.) Error bars indicate standard deviations from the means of the last 5 sessions of each experimental condition. The labels in bold show which sequence of the pair contained the highest hyperbolic value.

In the first pairwise comparison – STD versus IMP condition – Pigeons P883 and P942 strongly preferred (>.90 choice proportions) the STD sequence. Pigeon P702 showed preference for the IMP condition during the first exposure, but preferred the STD condition in the subsequent two side reversal conditions. Subject P894 showed a strong bias toward the right key in the first two experimental conditions, so it was exposed to two additional conditions with a higher ratio of forced-choice: free-choice trials to rectify the bias. After exposure to these training conditions, Pigeon P894 preferred the STD sequence when it was allocated on the left and on the right side (third and fourth conditions). It is important to note that this subject was exposed to fewer

numbers of sessions during the second and third condition—5 and 6 sessions, respectively.

When the WOR was pitted against the STD sequence, all pigeons exhibited strong preferences for the WOR sequence. For Pigeons P702, P894, and P942 such preferences were seen in every condition and replication; for Pigeon 883, preference for WOR was seen in three of four conditions. Specifically, the occasion on which P883 preferred STD over WOR was in the condition immediately following the DC condition.

In the third pairwise comparison – WOR versus IMP – all subjects showed a strong preference for the WOR sequence on both occasions in which they were exposed to these conditions. This was the critical comparison, as far as reconciling prior results is concerned. The results are extremely clear and consistent with prior results conducted with nonhuman subjects: Strong preference for the sequence with the shorter initial delay to reinforcement.

In general, during the DC conditions pigeons distributed choices more equally across alternative sequences. There was also more variability within the last five sessions, as seen by the larger error bars when compared to all the other experimental conditions. In the fourth pairwise comparison – WOR versus STD (DC condition), Pigeons P883, P702, and P894 emitted a larger proportion of responses toward the WOR sequence, whereas P942 allocated a larger proportion of responses toward the STD sequence. When IMP was pitted against the WOR sequence, Pigeon P883 showed preference for WOR whereas the other birds distributed their responses more equally between both alternatives (i.e., proportions were approximately .5).

Preference for the WOR sequence during DC conditions might be explained based on the conditioned reinforcer functions of the tokens. However, it is important to note that no replications or reversals were conducted during the DC conditions. Since there was no replication, it becomes important to analyze performance during the DC conditions in the transition from the immediately preceding condition (Table 2-1). The only difference from the preceding condition was the moment on which the exchange period was made available, even the sides on which each sequence was located was identical. It was decided not to change the location of the alternatives during the transition from the IC to DC conditions because we wanted to introduce a single additional variable (the DC condition), instead of two additional variables (the DC condition and the side reversal), when moving to a new condition. The results show that with the exception of Pigeon P942, there were clear carryover effects—a tendency to allocate more responses on the side associated with the preferred alternative from the previous condition. Daily analysis of the performance during this condition indicated that the contingency was exerting progressively less control over the behavior. It was decided not to replicate the DC condition with the sides reversed because of the risk of losing experimental control in subsequent conditions. Therefore, it is possible that the preference for the WOR sequence shown by some subjects during the DC conditions was due to the conditioned reinforcer effects of the tokens. However, the results are confounded with carryover effects from previous conditions. Future studies might try to disentangle the effects produced from the preceding conditions in order to assess the degree to which responses during the DC conditions was indeed being controlled by the conditioned reinforcers.

An analysis of the responses on the key that remained illuminated in the terminal-link was conducted and the results are depicted in Table 2-1. Table 2-1 shows the mean proportion of choices (initial-link) and the mean number of responses on the illuminated key during the terminal-link. The means are from last 5 sessions of each experimental condition. All subjects responded in the illuminated key during the terminal-link, but the number of responses was largely variable across conditions and across subjects. P942 was the subject who most frequently responded during the terminal-link, emitting up to 693 responses in preferred alternative.

An analysis of choice latencies in relation to preference and the sequence relative value was also conducted for each of the subjects. No orderly relation was found, so the data are not shown.

According to the predictions of the hyperbolic model, pigeons should prefer the WOR sequence over the other two sequences, and should prefer the STD sequence over the IMP sequence. Excluding the DC conditions, pigeons showed preference for the WOR sequence in 20 out of 21 occasions and showed preference for the sequence with the higher hyperbolic value in 29 out of 33 occasions. Therefore, the results obtained in Experiment 1 provide strong support for the ordinal predictions of the hyperbolic decay model over the predictions of the global maximization model, which predicts indifference across all experimental conditions.

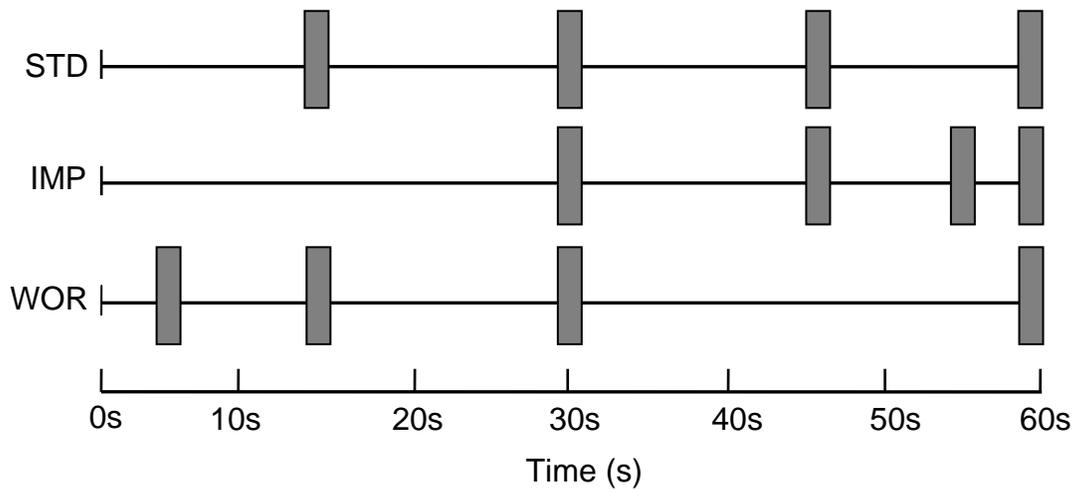


Figure 2-1. Diagram of the terminal-links implemented for each sequence in Experiment 1. The horizontal lines show the terminal link with time going from left to right. Each vertical bar represents the temporal placement of tokens timed from terminal-link onset: 15 s, 30 s, 45 s and 60 s in the Standard sequence, 5 s, 15 s, 30 s, and 60 s in the Worsening sequence, and 30 s, 45 s, 55 s and 60 s in the Improving sequence.

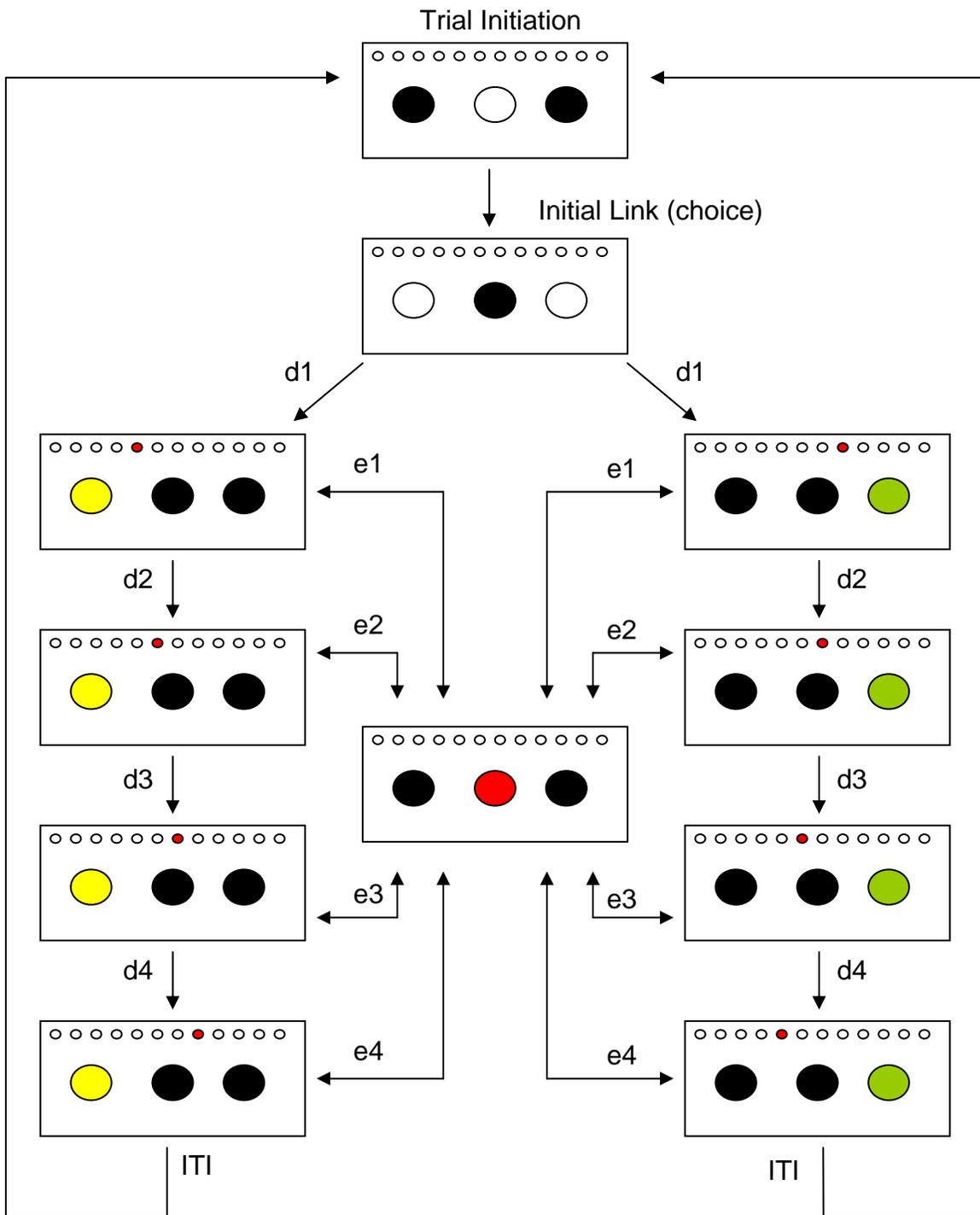


Figure 2-2. Schematic of the Immediate Consumption condition. The letters d1, d2, d3, d4 refer to the delays to the presentation of the first, second, third, and fourth token; and e1, e2, e3, e4, refer to the exchange of the first, second, third, and fourth token in the sequence.

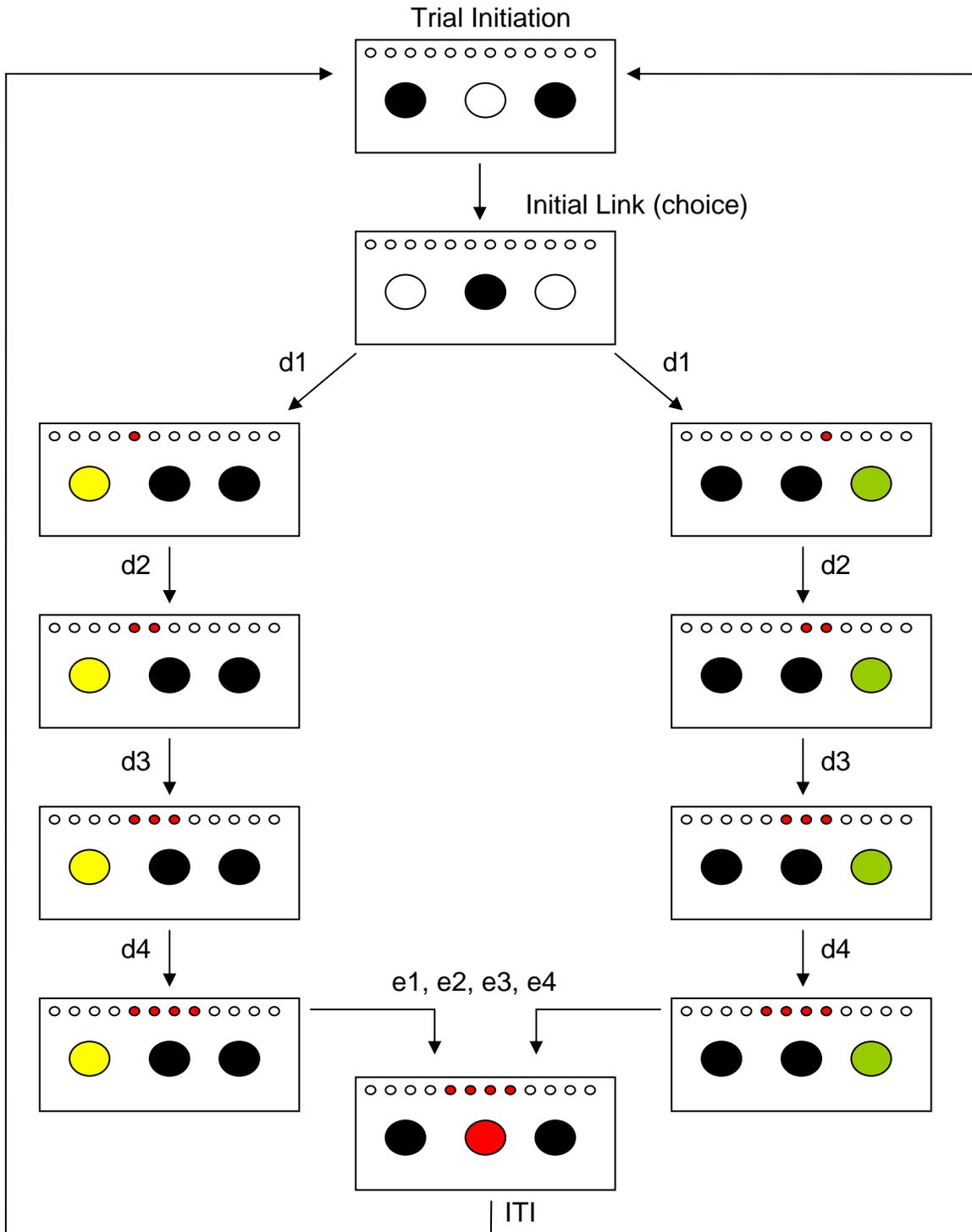


Figure 2-3. Schematic of the Delayed Consumption condition. The letters d1, d2, d3, d4 refer to the delays to the presentation of the first, second, third, and fourth token; and e1, e2, e3, e4, refer to the exchange of the first, second, third, and fourth token in the sequence.

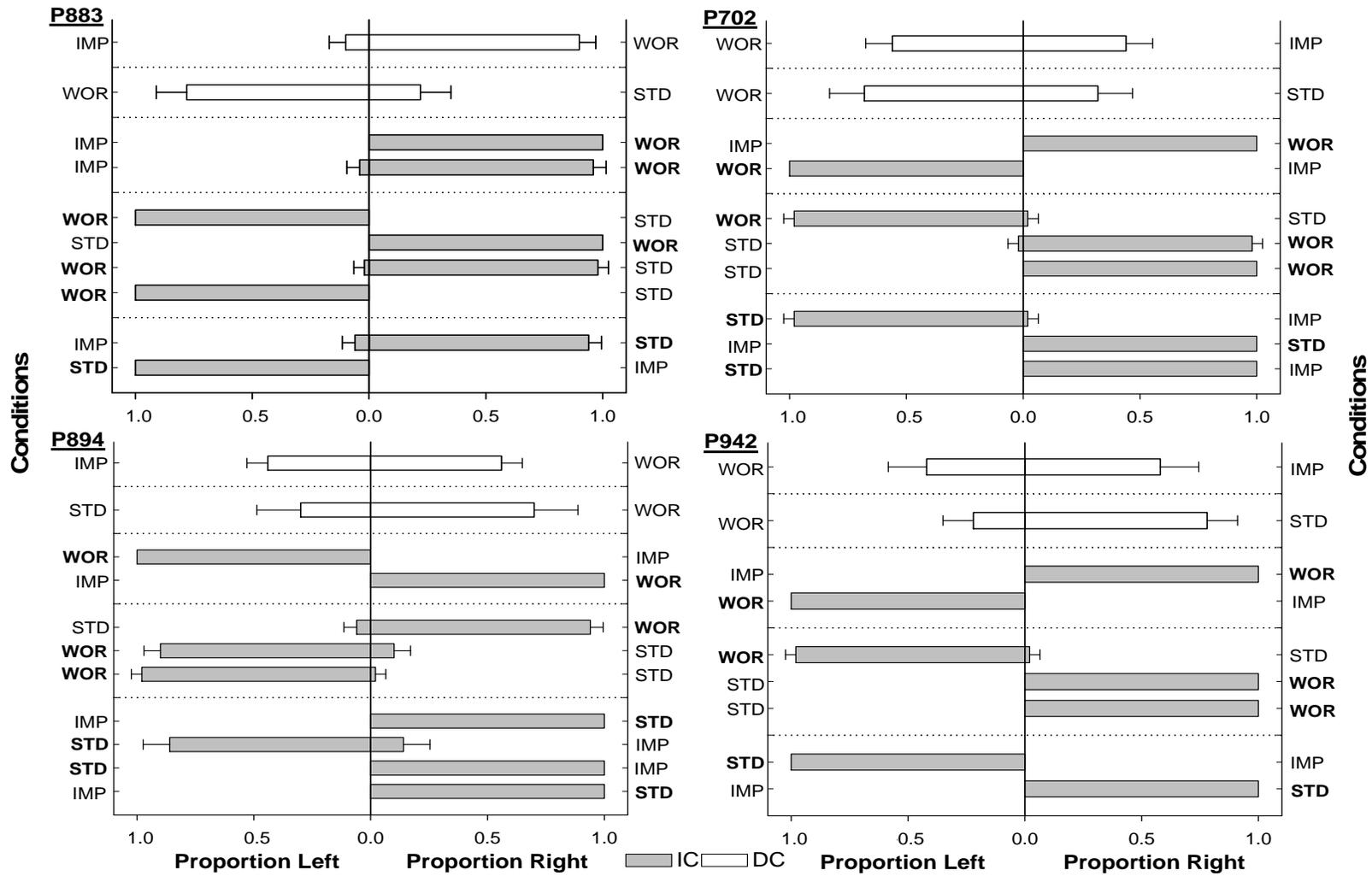


Figure 2-4. Mean proportion of choices for each alternative in Experiment 1. Bars on the left and on the right side depict the proportions from the last 5 sessions on the left and right side, respectively. Filled bars depict IC conditions and unfilled bars depict DC conditions. See text for further details.

Table 2-1. Mean proportion of choices (initial-link) and mean number of responses on the terminal-link for each alternative in Experiment 1. Data obtained from the last 5 sessions of each condition.

Sub. Cond.	Sess.	Left Alternative			Right Alternative			
		Sequence	Prop. Choices	Resp.T-L	Sequence	Prop. Choices	Resp T-L	
883	1	22	STD	1	5.4	IMP	0	0
	2	19	IMP	0.06	0.4	STD	0.94	2.2
	3	12	WOR	1	7.8	STD	0	0
	4	16	IMP	0.04	0	WOR	0.96	0.6
	5	31	IMP(DC)	0.1	8	WOR(DC)	0.9	4.2
	6	12	WOR	0.02	0.2	STD	0.98	0
	7	13	STD	0	0	WOR	1	2.2
	8	19	WOR	1	0.6	STD	0	0
	9	19	WOR(DC)	0.88	13.6	STD(DC)	0.22	1.2
	10	18	IMP	0	0	WOR	1	0.2
942	1	18	IMP	0	0	STD	1	694
	2	26	STD	1	686	IMP	0	0
	3	13	STD	0	0	WOR	1	306
	4	12	WOR	1	471	IMP	0	0
	5	25	WOR(DC)	0.42	21.6	IMP(DC)	0.58	518
	6	18	STD	0	0	WOR	1	656
	7	15	WOR	0.98	176	STD	0.02	10.6
	8	18	WOR(DC)	0.22	2	STD(DC)	0.78	129
	9	13	IMP	0	0	WOR	1	422
702	1	12	STD	0	0	IMP	1	4.2
	2	13	IMP	0	0	STD	1	2.2
	3	17	STD	0.98	3	IMP	0.02	0
	4	12	STD	0	0	WOR	1	0.2
	5	16	WOR	1	5.8	IMP	0	0
	6	41	WOR(DC)	0.56	3.4	IMP(DC)	0.44	3.6
	7	23	STD	0.02	0	WOR	0.98	0
	8	21	WOR	0.98	0	STD	0.02	0
	9	36	WOR(DC)	0.68	0.8	STD(DC)	0.32	0.6
	10	16	IMP	0	0	WOR	1	0.2
894	1	12	IMP	0	0	STD	1	9.6
	2	5	STD	0	0	IMP	1	15.8
	3	6	STD	0.86	30.4	IMP	0.14	1.4
	4	18	IMP	0	0	STD	1	12
	5	20	WOR	0.98	0.8	STD	0.02	0
	6	18	IMP	0	0	WOR	1	2.2
	7	16	IMP(DC)	0.44	49.2	WOR(DC)	0.56	5
	8	14	WOR	0.9	26.4	STD	0.1	0.8
	9	17	STD	0.06	0	WOR	0.94	13.8
	10	20	STD(DC)	0.3	2.4	WOR(DC)	0.7	11.4
	11	16	WOR	1	0.2	IMP	0	0

## CHAPTER 3 EXPERIMENT 2

### Introduction

Studies have shown that humans prefer improving rather than worsening sequences. This finding presents formidable theoretical and empirical challenges. Theoretically, the findings run counter to most psychological and economic models based on delay discounting; empirically, they are inconsistent with the results found with non-human subjects (Brunner, 1999; Exp. 1). Although preference for improving sequences has been consistently reported in the literature using humans as subjects, the vast majority of these studies used hypothetical outcomes (e.g., Chapman, 2000; Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991). Within the few studies that involved real events (e.g., Ariely & Loewenstein, 2000; Schreiber & Kahneman, 2000; Kahneman et al., 1993), choice was assessed via deception. That is, participants were led to believe that they were going to be exposed to sequences of events at the end of experiment, but they were never exposed to the outcomes of the choices they made. One of the main goals of the experiment presented in this section was to analyze human preference using a procedure that allowed subjects to *repeatedly* choose between sequences of outcomes and be repeatedly exposed to the contingent outcomes of their responses. The procedure implemented in this experiment was analogous to Experiment 1 so that performance across species could be compared. The subjects chose between sequences in which the inter-reinforcement delay increased (worsening), decreased (improving), or remained fixed (standard) in the terminal link. As in Experiment 1, choices produced tokens and consumable reinforcers in the terminal-link, but instead of food, popular TV shows were used as reinforcers. In

addition, the economic context—defined in term of token-exchange opportunities—was also manipulated as they were in Experiment 1.

## **Methods**

### **Subjects**

Four adult humans (two male and two female) were hired to serve as participants after signing an informed consent. All were recruited via local newspaper advertisement or flyers spread out in the University of Florida campus, and had no prior experience with similar experiments. The total number of sessions subjects were exposed to until the completion of the study ranged from 24 to 54, and they earned between \$5 to 6 dollars per/hour.

### **Material, Location, and Equipment**

Two small rooms, each containing a chair, a desk, a computer, a pair of speakers, a keyboard, and a mouse, served as experimental location. Both rooms were used during the experiment, but a given subject was always studied in the same room. During sessions, subjects remained seated in front of the computer monitor and responded to the visual stimuli presented on the screen by “clicking” with the computer mouse. The computers were IBM-compatible, and the visual interface displayed on the screen, as well as data collection were controlled via Visual Basic 6.0 software program. The monitor screen measured approximately 36.5 cm wide by 27.5 cm high, and was placed on the desk at approximate eye-level height of the subject when seated.

A picture of the visual interface displayed on the screen of the computer is shown in Figure 3-1. For the purpose of clarification, all the stimulus components used in the concurrent-chain schedule are presented together. The visual interface was comprised of four aligned red circles that served as tokens, and three aligned colored rectangles.

The circles had a circumference of 7.85 cm and were aligned at approximately 3.5 cm from the top of the screen. They were equally distanced from each other (2 cm) and centered on the screen. The left-most and right-most tokens were located at approximately 9.6 cm from the left and right side, respectively. The three aligned rectangles representing the choice alternatives and token exchange response were centered on the screen, and located approximately 4.5 cm below the tokens. The rectangles measured 7.7 cm wide by 7 cm high and were equally distant at approximately 3.2 cm. The center rectangle was colored red, white, or gray; and the side rectangles were colored blue, green, yellow, orange, or gray, depending on the experimental condition or the specific link within the chained schedule. The background screen color was gray. When inactive, the tokens and the rectangles were also colored gray (the exact same color of the screen background) but were outlined in a manner that they remained slightly visible on the screen.

### **Procedure**

Two sessions lasting approximately 50 min each, were scheduled per day, five days a week. Sessions occurred successively and were intercalated by a five-min break. In each session, participants were instructed to leave their personal belongings in a safe place, and were escorted by the research assistant to the testing room. No timing device of any sort was allowed to be brought into the experimental room.

A variety of popular TV shows were recorded and used as backup reinforcers throughout the experiment. The videos were converted to AVI format and stored at the hard disk of both computers. Each episode was divided into segments of approximately 30 s and was played each time the subject exchanged a token. A total of 48 video segments, which corresponded to a full show episode, were played each session. The

particular episode played during a session was selected by the subject prior to the beginning of that session from among ten available options: (1) Friends (season 6); (2) Friends (season 7); (3) Family Guy (season 1); (4) Looney Tunes; (5) Seinfeld (season 4); (6) Simpsons (season 2); (7) Simpsons (season 3); (8) Sports Bloopers; (9) Will and Grace (season 1); and (10) Wallace and Gromit.

To avoid repeated episodes, the program would automatically play the next available episode in the sequence. In case the subject watched all episodes available of a given show, the program would prompt the subject to choose a different show prior to the initiation of the session. After choosing a given show, and clicking on the “continue” button, two additional messages were displayed: “You will need to use only the mouse for this part of the experiment” and “when you are ready to begin, click the “begin” button below”. Immediately after subjects clicked on the buttons following the prompts, the experimental session started and subjects were exposed to the experimental choice contingency.

Subjects did not receive any instruction about the experimental contingencies during the experiment. The only instructions were to follow the prompts displayed on the screen, use the computer “mouse” during the session, and use the keyboard only at the end of the session (when prompted to rate the videos). During each session, the subject remained alone in the experimental room while the research assistant stayed in an adjacent room until the completion of the sessions.

The experimental choice contingency implemented here is analogous to the one used in Experiment 1. Subjects were given repeated choices between sequences of token and video clips using a concurrent-chains schedule with two links. The initial link

consisted of a Conc FR1 FR1 and the terminal link was comprised of a sequence of delays to each of 4 tokens and an exchange schedule. In the initial link, the token and center rectangle were inactive, while the two side rectangles were colored yellow or green and remained flashing on the screen (Figure 3-2). A single response on either rectangle produced the terminal-link stimuli, which was signaled by the following events: (1) The clicked alternative stopped flashing and became inactive; (2) the other alternative also became inactive but colored with the gray background color; and (3) initiation of the timer that controlled the delivery of the tokens. The terminal link was comprised of a sequence of delays to each of 4 tokens and exchange schedule associated with the alternative clicked. When the token was presented a brief beep was emitted. Similar to Experiment 1, the tokens were also presented from left to right or from right to left depending on whether the choice occurred on the left or right alternative, respectively. No ITI was implemented in this experiment. At the beginning of each cycle, the visual display showed a flashing white centered rectangle, click upon which produced the initial-link stimuli (i.e., trial-initiation response).

Similarly to Experiment 1, tokens were exchanged either after the delivery of the fourth token in the sequence (DC Condition) or immediately after the delivery of each individual token in the sequence (IC Condition). The timer that controlled the delivery of the tokens during DC and IC conditions worked in a similar manner as described in Experiment 1. Each token was exchanged for approximately 30-s of video segment, and the exchange schedule was signaled by flashing the red center rectangle (exchange rectangle) and the deactivation and darkening of the choice-rectangle alternatives. A single click on the exchange rectangle produced a brief beep and the

token exchange. Figure 3-3 is a picture of the computer screen during the exchange schedule.

### **Sequences of Reinforcement**

Figure 3-4 depicts the sequences of terminal-link events implemented in this experiment. The sequences followed the same rationale of the sequences used in Experiment 1, but the overall time-span of the delivery of all tokens (i.e., the terminal-link duration) was 2 min rather than 1 min. The duration of the terminal-link was extended in relation to Experiment 1 because previous unpublished work in our laboratory have shown that average delays of approximately 30 seconds were well suited in exerting control over the behavior of human subjects. More specifically, the *Standard* (STD) sequence contained four tokens presented at equal inter-temporal delays of 30 s, the *Worsening* (WOR) sequence the delays between successive tokens increased (10 s, 20 s, 30 s, and 60 s), and the *Improving* (IMP) sequence the delays between successive tokens decreased (60 s, 30 s, 20 s, and 10 s).

Experimental conditions were in effect for a minimum of four sessions and until choice proportions were deemed stable via visual inspection.

### **Delay Sensitivity Test**

Prior to exposure to the main experimental conditions described above, all subjects were exposed to a contingency aimed to assess delay sensitivity to the video reinforcer. During this pre-experimental phase, the contingency involved choices between alternatives that produced a single reinforcer delivered after different delays. One alternative produced a video clip segment after 5 s, while the other alternative produced the same outcome after a delay of 30 s. The duration of the video clip segment was the same, only the delay varied between the alternatives. To maintain a

constant reinforcement rate between both alternatives, a 25-s post-reinforcer delay followed the 5-s reinforcer delay option. Figure 3-5 shows a schematic of the contingency implemented in this phase.

The experimental procedure adopted in this phase was very similar to the main choice procedure already described above, except for the following differences: (1) The color of the choice rectangles were orange and blue, instead of green and yellow; (2) there was no token delivered in the terminal link; (3) once the programmed delay had expired, the centered red rectangle was presented, and a single click would produce the video; (4) the overall duration of the terminal link was 30 s instead of 120 s; (5) each session was comprised of the total of 8 forced and 40 free-choice trials.

Subjects received a minimum of 10 sessions. Only those who showed strong and unambiguous preference for the shorter delay were invited to continue in the experiment.

### **Questionnaire**

Subjects were given a questionnaire at the end of the experiment that included questions involving hypothetical sequences of two outcomes in which the delay to the outcomes were manipulated across questions. These questions were taken from the article published by Loewenstein and Prelec (1993) and are shown in Appendix A.

### **Results and Discussion**

Out of the 11 subjects who were exposed to the 10 minimum sessions, five did not show preference for the shorter delay alternative during the delay sensitivity test and were not invited to continue in the experiment. Six subjects demonstrated sensitivity to the shorter delay alternative during the delay sensitivity phase, but two decided to drop

the experiment shortly after this phase was completed. The results of the present experiment are based upon the performance of the four remaining subjects.

### **Choice Patterns**

Figure 3-6 is similar to Figure 2-4, and shows the mean proportion of choices over the last 3 sessions on the left and right side alternatives across experimental conditions. One subject, H148, was exposed to all experimental conditions, including the reversals and replications. Subject H146 was exposed to all experimental conditions but was not exposed to all reversals and replications, and Subjects H154 and H161 finished only 2 and 3 conditions, respectively.

When the STD was pitted against IMP, subjects H161 and H148 showed preference for IMP, whereas H146 and H154 showed preference for STD predominantly. In the WOR versus STD condition, all subjects showed strong preference for the WOR sequence. In the third pairwise comparison – WOR versus IMP, the three subjects who were exposed to this condition also showed preference for the WOR sequence. Note that H148, showed preference for IMP in one out of three occasions in which he was exposed to this pair of sequences.

Only subjects H146 and H148 were exposed to the DC conditions, and both subjects showed larger proportion of responses toward the WOR sequence over the STD and IMP sequences. Because there was not a replication of the DC conditions, it is important to analyze participant's performance at the preceding condition. (Note that the data shown in Figure 3-6 was grouped and thus do not reflect the exact order on which subjects were exposed to the experimental conditions.) The exact order of exposure to the conditions is shown in Table 3-1. As in Experiment 1, the conditions that preceded the DC conditions differed only in that they were IC conditions. In other

words, the comparisons involved the same pair of sequences, located at the same side; the only difference was that the token-exchange was made available after the delivery of each token. In 3 out of 4 exposures to the delayed conditions, H146 and H148 preferred the alternative on the side consistent with the preferences in the immediately prior condition. Therefore, results during DC conditions suggest carryover effects.

Considered as a whole, the human participants showed preference for the WOR sequence in 13 out of 14 opportunities under IC conditions, and showed preference for the sequence with higher hyperbolic value in 16 out of 20 occasions. These results support the ordinal predictions of the hyperbolic decay model, and are consistent with the results obtained with rats (Brunner, 1999, Exp 1) and with pigeons (Experiment 1 of the present study). They are inconsistent, however, with results obtained in prior research with humans, which report preference for improving sequences (Ariely, 1998; Ariely & Loewenstein, 2000; Chapman, 2000; Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991; Schreber & Kahneman, 2000).

A possible explanation for the contrasting results may be attributed to the specific nature of the reinforcer used or the specific reinforcer parameter being manipulated across experiments. In the present experiment, the value of a sequence of reinforcers was measured as a function of the delay to each member in the series, whereas other studies typically include magnitude manipulations, as well: either in relation to some qualitative property of a hypothetical stimulus (Chapman, 2000; Loewenstein & Sicherman, 1991; Loewenstein & Prelec, 1993); or in relation to some actual sensorial experience, such as aversive noise or temperatures (Ariely, 1998; Ariely & Loewenstein, 2000; Schreber & Kahneman, 2000). Therefore, it is possible that the manner in which

organisms discount delayed events alone differ in some important ways to the manner in which organisms discount delayed events that also vary in intensity/magnitude. In addition, it is also plausible, that the manner in which organisms discount different outcomes varies in accord to the specific nature of the reinforcer. Here, the reinforcer implemented was segments of popular TV shows, whereas previous studies used either qualitative different hypothetical events or aversive stimulation. The assumption that a single fixed discount parameter holds across different outcomes needs to be empirically tested.

Another significant procedural difference was the manner in which choices were presented and the dependent variable measured across studies. In previous experiments, participants made a single choice, and the dependent measure was a verbal response – rating or marking the preferred sequence. In the present experiment, participants were exposed to the choice contingency multiple times, which presumably allowed them to learn from direct experience with the task and its contingent consequences; and performance was measured based on the relative response allocation when it was deemed stable. Thus, while the present procedures differed in some significant ways from previous methods used with humans, they are in greater alignment with methods typically used with nonhumans. This makes them more suitable in cross-species comparisons, a major aim of the present investigation. Future research might profitably be directed toward approximating the more-typical human methods to determine more precisely the conditions under which the results begin to depart from the temporal discounting reported here.

## Questionnaire Results

Figure 3-7 shows the results obtained from the questionnaire given at the end of the experiment (Appendix A). The x-axis refers to the question number, and the y-axis refers to the percentage of subjects who selected answer A (black bar) or B (gray bar). To facilitate comparison with the previous literature, the results reported by Loewenstein and Prelec (1993) with the same questions were plotted and are shown in the top graph. The middle graph shows the results for the four participants who completed the present experiment.

Question 2 involved a single outcome delivered sooner (answer A) or later (answer B); and questions 3-6 involved sequences comprised of 2 outcomes with decreasing/worsening trend (answer A) and increasing/improving trend (answer B). It is important to note that the questions 3-6 differed from each other in relation to the time-frame events occurred; specially the time interval between the 2 outcomes. The time separation between the 2 outcomes were longer in questions 3 and 5 compared to questions 4 and 6. The results of question 1 are irrelevant and are not shown in the graphs.

All participants from this experiment (middle graph) preferred to have dinner at their preferred restaurant sooner, rather than later (Question 2). Results obtained from question 3-6 showed that the patterns of choices between improving and worsening trends were not constant among all hypothetical scenarios. The majority of subjects preferred the worsening sequence in the scenarios described in questions 3 and 5, but demonstrated preference for improving sequences in the scenarios described in question 4 and 6.

The results reported by Loewenstein and Prelec (1993) showed a large percentage of subjects preferring the improving alternative in questions 4 and 6, and slight preference for this alternative (close to indifference) in questions 3 and 5. Therefore, compared to the results obtained via questionnaire in this study, results reported by Loewenstein and Prelec (1993) were similar in respect to question 2, 4, and 6, but were different with respect to question 3 and 5. This different pattern of preference across different questions suggests that if trend has an effect on choice, this effect seems to be partly dependent upon the proximity of the outcomes embedded in the series. This confirms Loewenstein and Prelec's claim that the spread of the outcomes is an important variable determining preference for improving sequences. To reiterate, the four subjects of the present experiment showed a strong preference for worsening over improving sequences when making choices with actual outcomes spread over 2 minutes in the terminal-link. If the temporal proximity of the events plays a role, as speculated above based on the hypothetical scenarios, one would expect different response patterns if the intervals between the actual outcomes were changed.

The relation between choices involving one and two outcomes is captured in the analysis of questions 2 and 3. Although results were slightly different across this and the earlier experiment, an interesting similar pattern can be observed: some subjects who chose to have the preferred outcome sooner rather than later when faced with choices with single outcome; choose to postpone the preferred outcome when a less preferred outcome was embedded in the series.

It is important to note that there were some methodological differences that may account for the different results obtained in the present study and the one published by

Loewenstein and Prelec (1993). The first difference was the sample of subjects who were recruited to answer the questions. In the present experiment, each subject (University of Florida students) answered all the questions; whereas in Loewenstein and Prelec's study, one group of subjects comprised of Harvard students (N=82) answered questions 1-3, and another group comprised of visitors of the Museum of Science and Industry in Chicago (N=48) answered questions 4-6. Another difference may be attributed to fatigue. Here, subjects answered the questionnaire after being exposed to 2 experimental sessions lasting approximately 2 hours, whereas subjects of the earlier study were recruited to answer the questionnaire alone. Thus, it is possible that subjects from the present study were tired and did not pay enough attention to the questions. In addition, it is also possible that prior exposure to the experimental contingency may have had some effect on the subjects' evaluation of the scenarios laid before them via the questionnaire.

The results obtained via questionnaire here are supported by the results published by Matsumoto et al. (2000, Exp.2). In that study, business students were also given a choice between restaurant sequences (identical to question 3) and a larger percentage of subjects selected the worsening sequence over the improving.

The validity of the results obtained via the questionnaire in this experiment is limited due to small number of subjects. To increase the sample size, another figure was created adding the answers obtained from the subjects who dropped the present experiment, as well as the results obtained from a pilot study whose subjects answered the exact same questionnaire. The results are shown in the bottom graph of the figure (N=58). The results displayed in this graph show a general similar pattern. The main

difference is that the percentage of subjects who preferred worsening over improving in question 3 and 5 became much less disparate when compared to the four subjects from the present experiment alone.

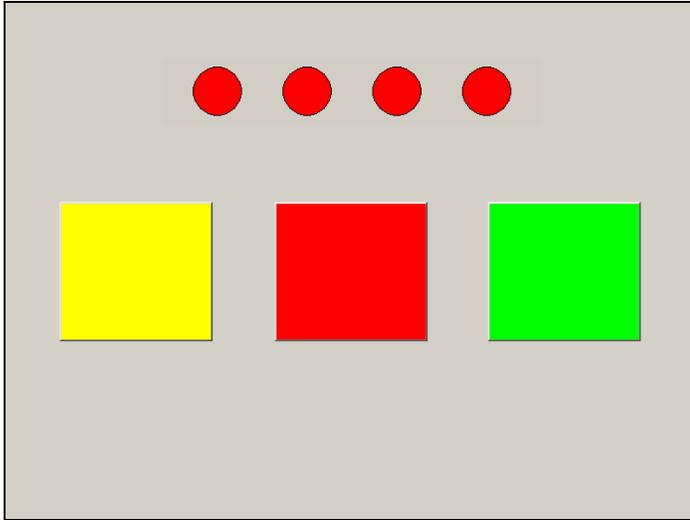


Figure 3-1. Picture of the screen with all the visual components used in the concurrent chain schedule. The red circles represent tokens and the side rectangles represent the choice alternatives. The centered rectangle represents the token exchange or trial initiation response when colored red or white, respectively.

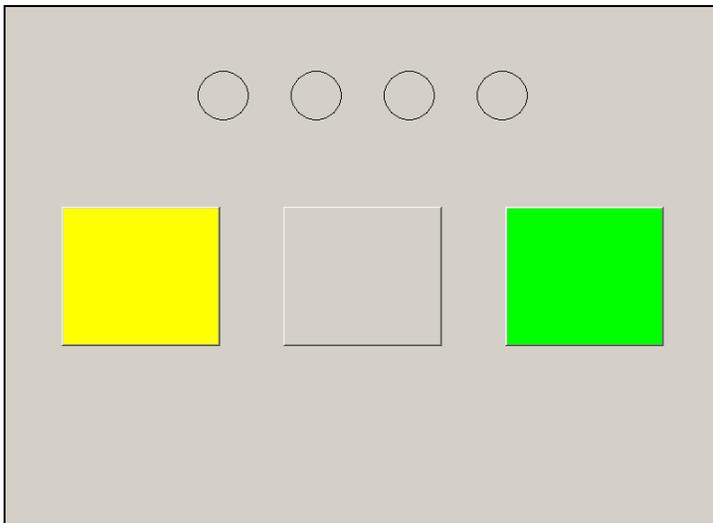


Figure 3-2. Screen shot of the choice phase (initial-link). The left and right rectangles represent the choice alternatives and remained flashing on the screen until a response was made. The center rectangle was visible, but inactive.

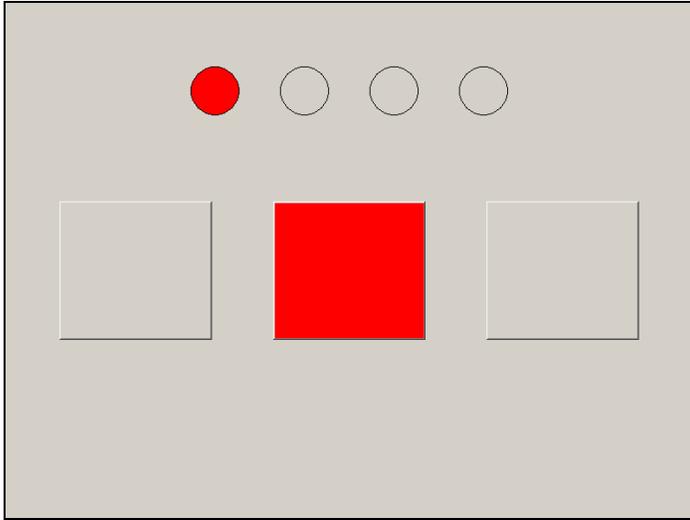


Figure 3-3. Screen shot of the token exchange phase.

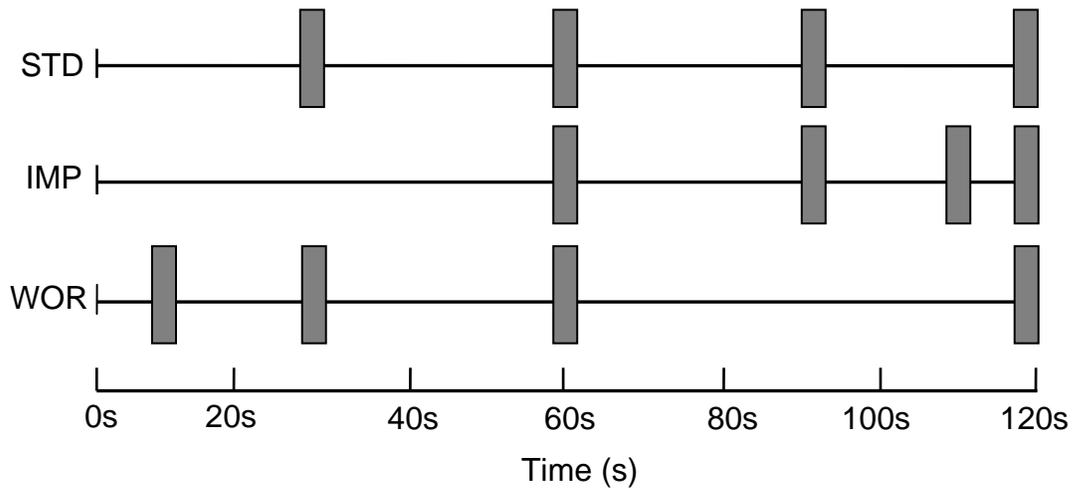


Figure 3-4. Diagram of the terminal-links implemented for each sequence in Experiment 2. The horizontal lines show the terminal link with time going from left to right. Each vertical bar represents the temporal placement of tokens timed from terminal-link onset: 30 s, 60 s, 90 s and 120 s in the Standard sequence; 10 s, 30 s, 60 s, and 120 s in the Worsening sequence, and 60 s, 90 s, 110 s and 120 s in the Improving sequence.

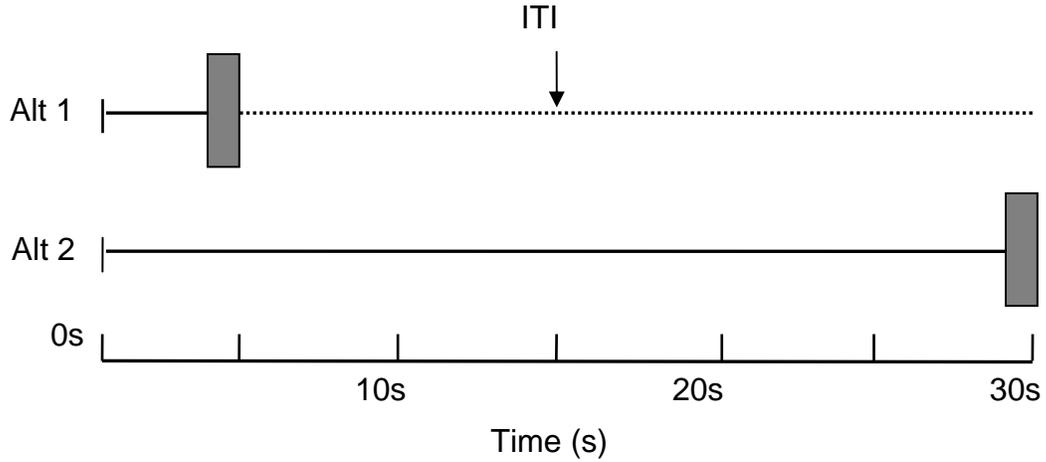


Figure 3-5. Diagram of the terminal-links implemented during delay sensitivity test. Responses on alternative 1 produced a video segment after a delay of 5 sec, whereas responses on alternative 2 produced video after a delay of 30 seconds. In addition to the delay to video, Alternative 1 terminal link also contained an ITI of 25 seconds.

Table 3-1. Sequence of conditions for each subject in Experiment 2. DST refers to the delay sensitivity test phase. Long and short refer to the alternatives that produced the reinforcer after 30 and 5 sec, respectively. The letters L and R in parenthesis indicate the location (left or right) of each sequence. IC refers to Immediate Consumption, whereas DC refers to Delayed Consumption condition.

Subject	Condition	Sessions
H146	DST - Long(R) X Short(L)	4
	STD(R) X IMP(L) - IC	12
	STD(R) X WOR(L) - IC	4
	WOR(R) X IMP(L) - IC	6
	WOR(R) X IMP(L) - DC	4
	STD(R) X WOR(L) - IC	4
	STD(R) X WOR(L) - DC	6
H148	DST - Long(R) X Short(L)	2
	STD(R) X IMP(L) - IC	8
	STD(R) X WOR(L) - IC	4
	STD(L) X WOR(R) - IC	4
	WOR(R) X IMP(L) - IC	10
	WOR(R) X IMP(L) - DC	4
	STD(R) X WOR(L) - IC	4
	STD(R) X WOR(L) - DC	4
	STD(L) X WOR(R) - IC	6
	WOR(L) X IMP(R) - IC	4
	WOR(R) X IMP(L) - IC	4
H154	DST - Long(R) X Short(L)	4
	STD(R) X IMP(L) - IC	6
	STD(L) X IMP(R) - IC	4
	STD(R) X IMP(L) - IC	10
	STD(R) X WOR(L) - IC	4
H161	DST - Long(R) X Short(L)	6
	WOR(R) X IMP(L) - IC	4
	STD(R) X WOR(L) - IC	4
	STD(R) X IMP(L) - IC	6
	WOR(R) X IMP(L) - IC	4

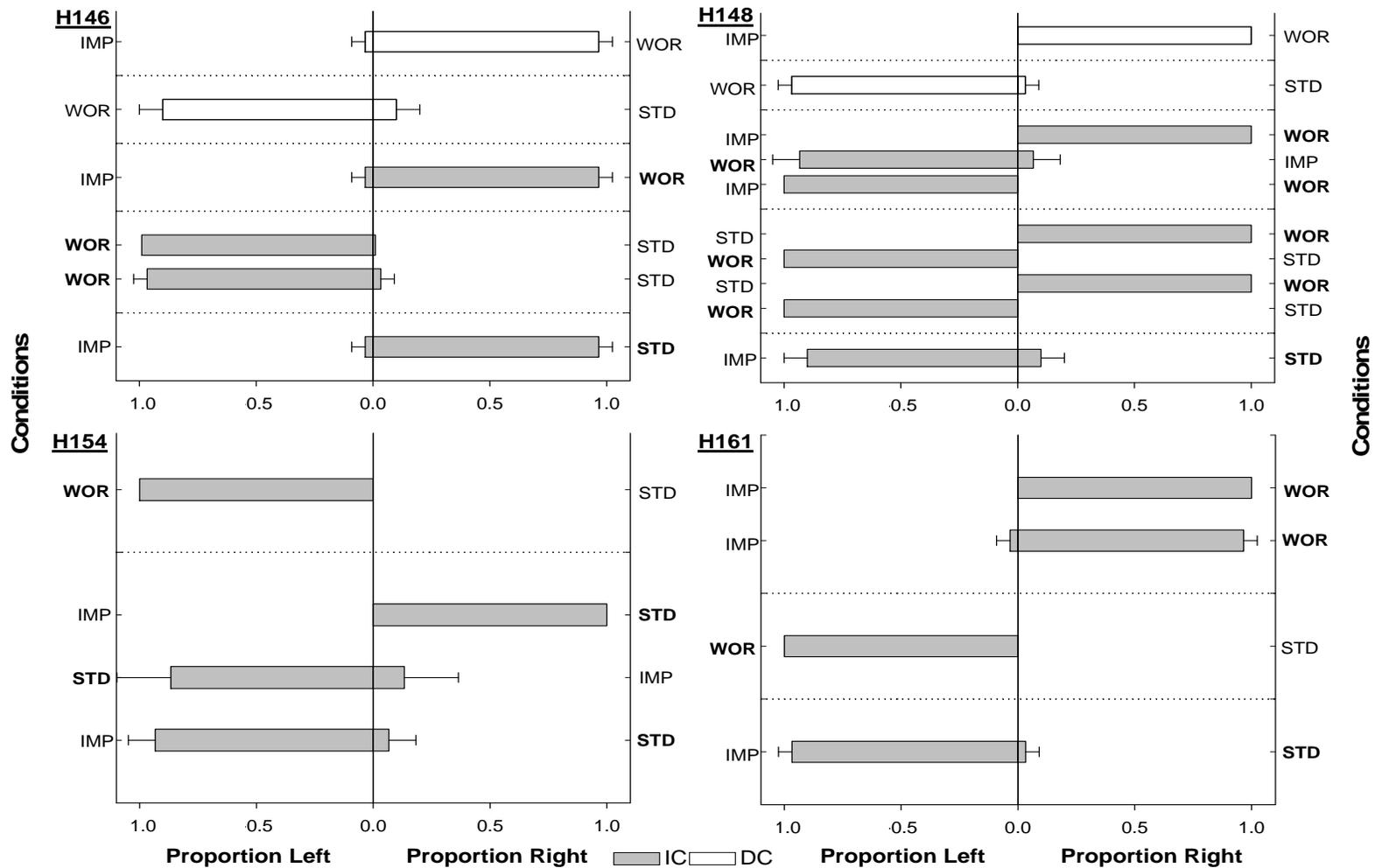


Figure 3-6. Mean proportion of choices for each alternative in Experiment 2. Bars on the left and on the right side depict the proportions from the last 3 sessions on the left and right side, respectively. Filled bars depict IC conditions and unfilled bars depict DC conditions. See text for further details.

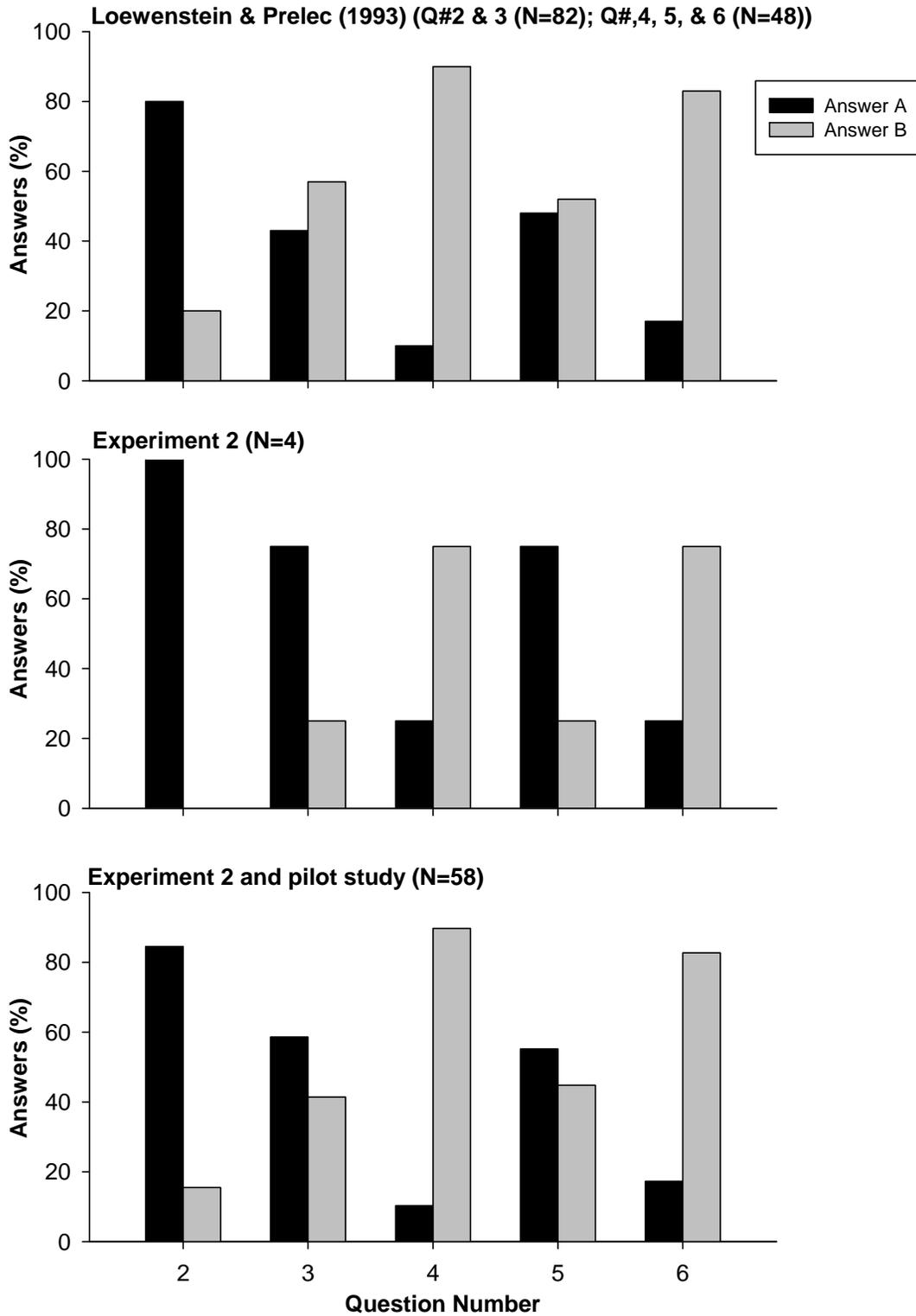


Figure 3-7. Percentage of answers A and B given to the questionnaire questions. Black bars refer to answer A and gray bars refer to answer B

## CHAPTER 4 EXPERIMENT 3

### **Introduction**

The results obtained in Experiment 1 showed that pigeons' choices were largely controlled by the delay to the first reinforcer in the series. The extent to which the delay to subsequent reinforcers exerts control over choice is still largely undetermined. Prior research on the topic of sensitivity to multiple reinforcers in a sequence is somewhat mixed. Moore (1979, 1982), for instance, found evidence that pigeons' choices between sequences of reinforcers are governed primarily by the delay to the first reinforcer; subsequent reinforcers in the series have little or no effect on choice. On the other hand, Mazur (1986, 2007), Brunner (1999), Brunner and Gibbon (1995), and Shull et al. (1990), found evidence that animal's choices (pigeons and rats) are sensitive to the delay to subsequent reinforcers, and that the data are well accounted by the predictions of the hyperbolic-decay model iterated across multiple reinforcers.

Although prior results suggest that choices are affected by temporally remote reinforcers in a sequence, little is currently known about the number and range of reinforcers to which choices are sensitive. Unlike prior studies in which animals chose between sequences differing in both the number of reinforcers and the delays to individual reinforcers in a sequence, the present study arranged choices between sequences of reinforcers delivered at the same overall rate but at different temporal patterning. The number of reinforcers in the sequences ranged from 2 to 4, but always included the same number and overall rate within a block of sessions. Delay sensitivity was assessed by selectively manipulating the delay to a single reinforcer while holding constant the delay to the other reinforcer(s) in the sequences.

## Methods

### Subjects

Four pigeons from Experiment 1 served as subjects.

### Apparatus

The same apparatus of Experiment 1 was used.

### Procedure

The procedure implemented in Experiment 3 was similar to the experimental procedure used in Experiment 1 under IC conditions, except for the following: (1) The number of forced-choice trials was increased from 2 to 4; (2) Access to food was decreased from 2.5 to 2.25 seconds per token exchange; (3) The inter-trial interval (ITI) following the last reinforcer of a sequence was varied across conditions to maintain constant the reinforcement rate between both alternatives. In Experiment 1, the last reinforcer was always delivered at the same moment in the terminal-link (i.e., 60 sec), and each terminal link was always followed by a constant 5-s ITI. Thus, the total duration of the terminal-link plus ITI was held constant at 65 seconds. In the present experiment, the total duration of the terminal-link plus ITI was also held constant at 65 seconds, but due to unequal moment on which the last reinforcer within a pair of sequences was delivered in some manipulations, different ITI were implemented following the delivery of the last reinforcer.

Across phases of the experiment, the sequences consisted of either 2, 3, or 4 food reinforcers, but always included the same number and overall rate of food reinforcers across blocks of sessions (conditions) within a phase. Timed from the terminal-link onset, the delay to an individual reinforcer in a sequence was selectively manipulated while maintaining constant the delay(s) to the other reinforcers in the

sequence. In this way, sensitivity to a particular delay within a reinforcer sequence was assessed across phases.

Table 4-1 depicts the conditions implemented in this experiment. It shows the delays to individual reinforcers on each sequence, the overall hyperbolic value of each sequence, the difference in value between the two sequences, and the relative value of one sequence over the other. To ease the visual inspection, the higher value sequence, the one containing the shorter delay to a specific reinforcer in the series, was always placed in column S (where S refers to short); whereas the sequence containing the longer delay to a specific reinforcer was always placed in column L. The delays to each reinforcer in columns L and S are timed in seconds from the moment of choice, and the underlined values in these columns specify the individual reinforcer manipulated within a condition. The hyperbolic value of each sequence was measured assuming the parameters  $K$  and  $A$  equal to 1. The difference in value among the two sequences was calculated by subtracting the sequence with lower hyperbolic value from the higher-value sequence ( $\text{Value L} - \text{Value S}$ ), and the relative value was calculated by dividing the value of sequence with higher value by the sum of the values of both sequences ( $\text{Value S} / (\text{Value S} + \text{Value L})$ ). The relative value index has a value of .5 when the hyperbolic values of both alternatives are equal. When this index has a value above .5 it indicates that the hyperbolic value of the sequence located in the S column had a higher value; when the relative value is below .5 it indicates that the sequence placed in the L column had a higher value. To clarify the metric value ( $V$ ) as used here with this particular delay discounting model (i.e., hyperbolic-decay model), note that with the parameter values  $K$  and  $A$  set equal to 1, the maximum  $V$  that a *single* reinforcer can

have is 1; which would essentially be equal to the undiscounted value of the reinforcer (i.e., it means that there would be no discounting at all—it reflects the value if the reinforcer was delivered at a delay equal to zero—which is technically impossible). The closer  $V$  is to zero, the more complete discounting has occurred.

The last condition of Experiment 3 (Condition 15 in Table 4-1) differed from all other conditions of the experiment. Contrary to the previous conditions, in which the delay of a single reinforcer was selectively manipulated, this condition was comprised of a comparison between two sequences of reinforcers in which the delay to all reinforcers differed. This condition was implemented to assess whether preference for the sequence with higher hyperbolic value would hold even when that sequence contained a longer delay to the first reinforcer in the sequence.

### **Results and Discussion**

Figures 4-1 and 4-2 provide a detailed characterization of the preference profiles for each subject. These figures show the mean proportion of choices on each alternative across all experimental conditions for Pigeons P883 and P942 (Figure 4-1) and P702 and P894 (Figure 4-2). The experimental conditions are displayed in the order that they occurred in the experiment (starting from the bottom). Each pairwise comparison, including the replications and reversals, if any, are grouped and separated by the dotted lines. To facilitate the description of the results, each comparison (i.e., each condition or set of conditions on which the delay to a specific reinforcer was manipulated) is numbered, the location of which is inside each graph. The transition from one condition to the next always involved arranging the richer sequence on the side opposite the preferred alternative from the previous condition.

Note that the order of conditions, as well as the specific delay values implemented, was not identical for each pigeon. To guide the reader, the analysis of the results starts with the individual description of the performance of P883. In the first set of experimental conditions – comparison number 1 (C1) in the graph (first and second condition starting from the bottom), the delay to the 2<sup>nd</sup> reinforcer in a 4-reinforcer sequence was manipulated, and subject P883 did not show systematic preference for the shorter alternative (S ALT henceforth). More specifically, when the sequence with the shorter 2<sup>nd</sup> delay was located on the left side, P883 preferred the alternative on the left side, a pattern that remained nearly identical in the following condition. In a subsequent set of conditions, comparison 6 (C6), sensitivity to the 2<sup>nd</sup> reinforcer in a 4-reinforcer sequence was re-assessed using more discrepant parameter values. More specifically, the delay to the 2<sup>nd</sup> reinforcer of one sequence was decreased from 15 to 10 s whereas the delay to the 2<sup>nd</sup> reinforcer of the other sequence was increased from 35 to 40 s. When these delay values were implemented, P883 showed differential preference for the S ALT. In C2, the delay to the 2<sup>nd</sup> reinforcer in a 2-reinforcer sequence was manipulated, and P883 showed preference for the S ALT irrespective of position. In C3 and C4, when sensitivity to the 3<sup>rd</sup> reinforcer in a 3-reinforcer sequence was assessed, P883 did not show sensitivity for the sequence with shorter 3<sup>rd</sup> reinforcer in C3 but did in C4 (when the delay parameter of the 3<sup>rd</sup> reinforcer was decreased from 25 s to 20 s). P883 did not show sensitivity to the 4<sup>th</sup> and 3<sup>rd</sup> reinforcer in a sequence comprised of 4 reinforcers (C5 and C7, respectively). In the last set of conditions (C8), the one in which the sequence with highest hyperbolic value was the one with longer delay to the first reinforcer, P883 showed preference for the highest value sequence.

Following is an analysis of the performance of all subjects together. In relation to sequences comprised of 4 reinforcers, all four subjects were exposed to at least one condition in which the 2<sup>nd</sup> reinforcer was manipulated, but only P942 (C1) and P883 (C7) showed preference for the S ALT. Out of the subjects exposed to the manipulation involving the 3<sup>rd</sup> and 4<sup>th</sup> reinforcer in a 4-reinforcer sequence, none of the subjects showed strong preference for the S ALT. Nevertheless, note that performance of P942 when the 4<sup>th</sup> reinforcer in the series was manipulated (C6) indicates at least some delay sensitivity, as shown by the exclusive preference for the S ALT when this sequence was located on the right side and the near indifference when this sequence was located on the left. When the last reinforcer in a 2-reinforcer sequence was manipulated, all subjects exposed to this manipulation (P883 (C2), P702 (C3), and P894 (C3)) showed preference for the S ALT. In regards to the last reinforcer of a 3-reinforcer sequence, only two subjects showed preference for the S ALT when the delay values were made more extreme: P883 (in C4), and P702 (in C5). When 2<sup>nd</sup> reinforcer in a 3-reinforcer sequence was manipulated, the two subjects exposed to this condition, P702 and P894, also showed preference for the S ALT when the values were made more extreme (C9 for P702, and C7 for P894). In the last set of conditions, the one in which the sequence with highest hyperbolic value was the one with longer delay to the first reinforcer, 3 out of the 4 subjects showed preference for the highest value sequence.

An analysis of the performance of the subjects across experimental conditions shows that subjects often responded more exclusively on one side alternative irrespective of the location of S ALT. This similar pattern of responses toward one specific side alternative across conditions, in particular during the side reversal

conditions, indicates lack of sensitivity to the delay manipulated as well as side bias. In general, performance of P883 and P894 indicates a left side bias, whereas performance of P942 and P702 indicates a right side bias.

Figure 4-3 shows for each pigeon the mean proportion of choices on the higher-value sequence (y-axis) as a function of its relative value (x-axis), computed as in Table 4-1. Data are from the last five sessions in each condition, with different symbols representing different reinforcer manipulations. The first number (from the left) written inside the legend indicates which reinforcer was manipulated in the series, and the second number indicates how many reinforcers the sequence was comprised of in each comparison. For instance, the label 2<sup>nd</sup>- 3R indicates that the manipulation involved the second reinforcer in a sequence comprised of 3 reinforcers; and the label 3<sup>rd</sup>-4R indicates that the third reinforcer in a 4-reinforcer sequence was manipulated.

To enhance visual inspection, only one data point is shown per condition for each subject. In cases in which the subject was exposed to the same condition more than once (i.e., when there were replications and reversal conditions), the data were taken from the condition in which preference toward the higher-value sequence was lowest. In other words, in these cases the data were taken from the condition in which less sensitivity to the manipulation was seen. This method was the most conservative measure and seemed most appropriate in light of the persistent side bias shown by all subjects (Figures 4-1 and 4-2). The lines connecting some of the plots serve the function of highlighting the conditions that involved the same manipulation but differed in terms of the relative hyperbolic value (i.e., those comparisons in which the delay to a single reinforcer was made more extreme/discrepant within the pair of sequences). The

horizontal dotted line along each graph indicates proportion of responses equal to .5 (indifference). Proportion of responses above this line reflects preference for the higher-value sequence, whereas below the line indicates preference for the lower-value alternative.

In general, data suggest that preference is affected by the relative value of a sequence. In addition, preference also seems to be differentially affected by the number of reinforcers that comprise the sequence, as well as the specific location of the reinforcer being manipulated. The relationship between preference and relative value is more clearly observed by analyzing performance under the conditions involving sequences with the same number of elements and same reinforcer manipulation, but different relative values. This relationship is shown in the plots connected with a line. The connected points show that when the relative value of a sequence is further increased by making a specific reinforcer delay more extreme, preference toward that sequence typically increases. With two exceptions (connected closed circles from P702 and connected opened circles from P894), preference for the higher-value sequence increased substantially in direct relation to its relative value. In nearly every case, the change in relative value within the same manipulation produced a reversal of preference from the lower-value alternative to the higher-value one.

It is important to note that across all the conditions of this experiment, the difference in overall value between the pair of sequences being compared was very small. This small difference is reflected in the low relative value shown in the x-axis across the comparisons conducted in this experiment. The relative values ranged from .51 to .57 (assuming  $K=1$ ), which shows that when the delay to the first reinforcer is held

constant and the terminal link length is fixed at 1 minute, the manipulation of the subsequent reinforcers does not have a large impact on the overall value of the sequence. This is especially the case for the reinforcers delivered later in series, as the 3<sup>rd</sup> or 4<sup>th</sup> reinforcer. Nevertheless, the differential choice patterns exhibited when a given delay parameter value was made more discrepant indicates that choices were indeed sensitive to the changes, even in cases when preference for the richer sequence was not obtained.

In summary, the results of Experiment 3 show that choices were sensitive to the delay to the 2<sup>nd</sup> reinforcer in a sequence of 3, and sometimes to the 2<sup>nd</sup> reinforcer in a sequence of 4. For only one pigeon was there evidence of sensitivity to the final reinforcer in a sequence of 4. For no subjects was there sensitivity to the 3<sup>rd</sup> reinforcer in a 4-reinforcer sequence. In short, choices were sensitive to the delay to individual reinforcers in sequences comprised of 2 and sometimes 3 outcomes, but not in sequences of 4 outcomes. In the final condition, in which the first reinforcer delay was pitted against the value of the 4-reinforcer sequence in its entirety, three of four pigeons preferred the sequence with higher value despite a longer delay to the first reinforcer of that sequence (final comparison).

Overall, the results are broadly consistent with models of temporal discounting expanded to include the impact of sequences of delayed reinforcers. The differential pattern of responses in function of the delay manipulations implemented in this study, suggests that the delay to each individual reinforcer in the sequence has an effect on the overall value of the sequence as a whole. Therefore, the results obtained in the present experiment are inconsistent with the results reported by Moore (1979, 1982)

who found that choices is only affected by the delay to the first reinforcer in the series, but are consistent with the previous studies showing that the delay to subsequent reinforcers is an important variable determining choices between sequences of reinforcers (Brunner, 1999; Brunner & Gibbon, 1995; Mazur, 1986; 2007; Shull et al., 1990).

Table 4-1. Experimental conditions implemented in Experiment 3. Column L refers to the sequences that contained the individual longer reinforcer delay, and column S refers to the sequences that contained the individual shorter delay. The delays to each reinforcer are timed in seconds from the terminal-link onset. The underlined values specify the individual reinforcer that is being manipulated within a condition. The value difference was measured by subtracting Value S from Value L, and the relative value was calculated by using the following formula:  $V_s/(V_s+V_l)$ , where V refers to overall hyperbolic value, and the subscripts identify the short and long columns.

Cond	Delays (sec)		ITI		Hyperbolic Value (V)				Subject
	L	S	L	S	L	S	L-S	Relative	
1	5, <u>35</u>	5, <u>15</u>	30	50	0.194	0.229	0.035	0.541	883, 702, 894
2	5, <u>35</u>	5, <u>10</u>	30	55	0.194	0.258	0.063	0.57	702, 894
3	5, 15, <u>45</u>	5, 15, <u>20</u>	20	45	0.251	0.277	0.026	0.525	883, 942
4	5, 15, <u>45</u>	5, 15, <u>25</u>	20	40	0.251	0.268	0.017	0.516	883
5	5, 15, <u>45</u>	5, 15, <u>30</u>	20	30	0.251	0.261	0.011	0.51	942
6	5, 10, <u>45</u>	5, 10, <u>15</u>	20	50	0.279	0.32	0.041	0.534	702, 894
7	5, 10, <u>60</u>	5, 10, <u>15</u>	5	50	0.274	0.32	0.046	0.539	942, 702, 894
8	5, <u>40</u> , 60	5, <u>10</u> , 60	5	5	0.207	0.274	0.067	0.569	702, 894
9	5, <u>55</u> , 60	5, <u>10</u> , 60	5	5	0.201	0.274	0.073	0.577	702, 894
10	5, 15, 20, <u>60</u>	5, 15, 20, <u>25</u>	5	40	0.293	0.315	0.022	0.518	883, 942, 702
11	5, 15, <u>45</u> , 60	5, 15, <u>30</u> , 60	5	5	0.267	0.278	0.011	0.51	942
12	5, 10, <u>55</u> , 60	5, 10, <u>15</u> , 60	5	5	0.292	0.336	0.045	0.536	883, 942
13	5, <u>35</u> , 45, 60	5, <u>15</u> , 45, 60	5	5	0.233	0.267	0.035	0.535	883, 942, 702, 894
14	5, <u>40</u> , 45, 60	5, <u>10</u> , 45, 60	5	5	0.229	0.296	0.067	0.563	883, 702
15	10, 40, 50, 60	15, 20, 25, 30	5	35	0.151	0.181	0.03	0.544	883, 942, 702, 894

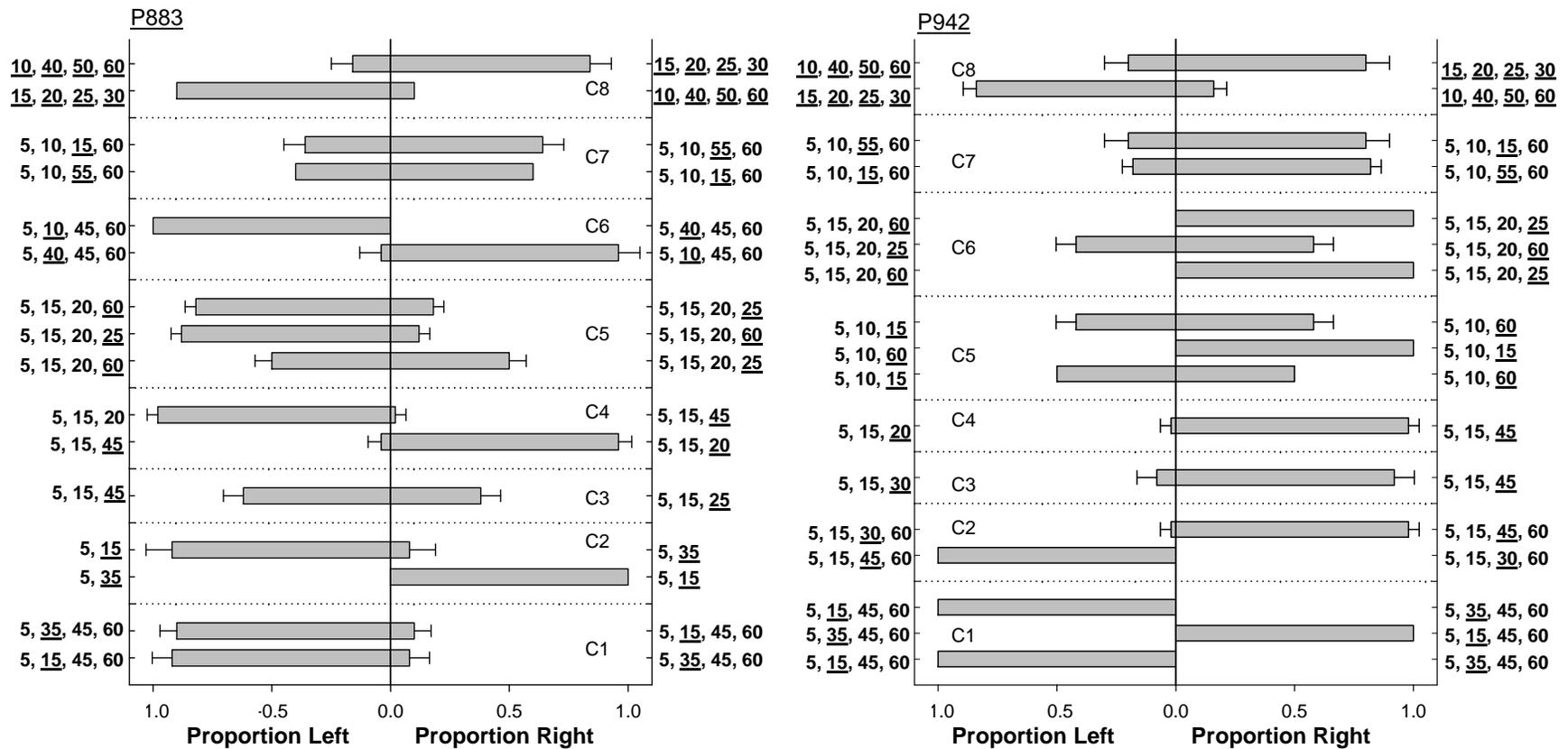


Figure 4-1. Mean proportion of choices for each alternative in Experiment 3 for P883 and P942. Bars on the left and on the right side depict the proportions from the last 5 sessions on the left and right side, respectively. The underlined number in each label shows the specific reinforcer manipulated in each comparison. Error bars indicate standard deviation from the means of the last 5 sessions of each condition.

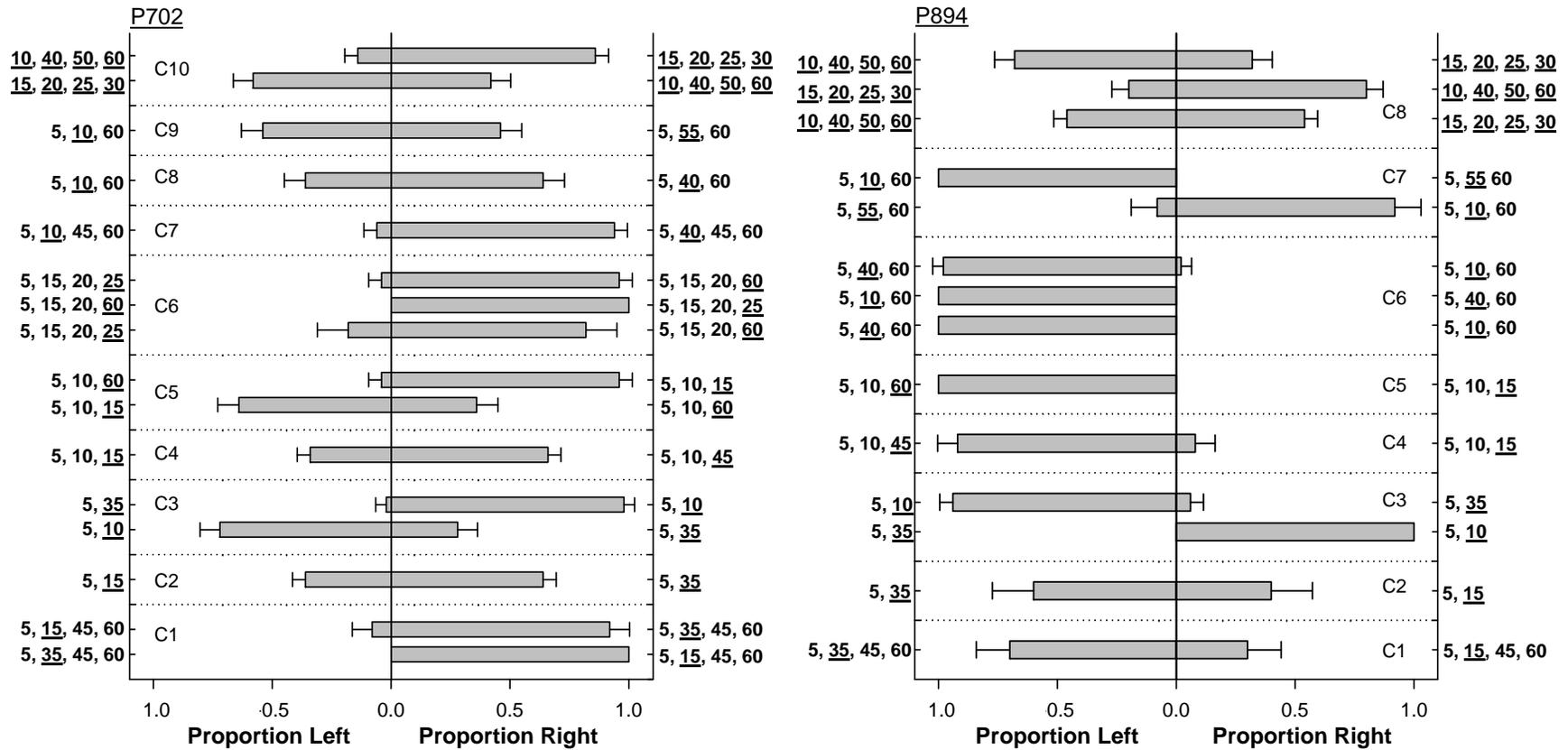


Figure 4-2. Mean proportion of choices for each alternative in Experiment 3 for P702 and P894. Bars on the left and on the right side depict the proportions from the last 5 sessions on the left and right side, respectively. The underlined number in each label shows the specific reinforcer manipulated in each comparison. Error bars indicate standard deviation from the means of the last 5 sessions of each condition.

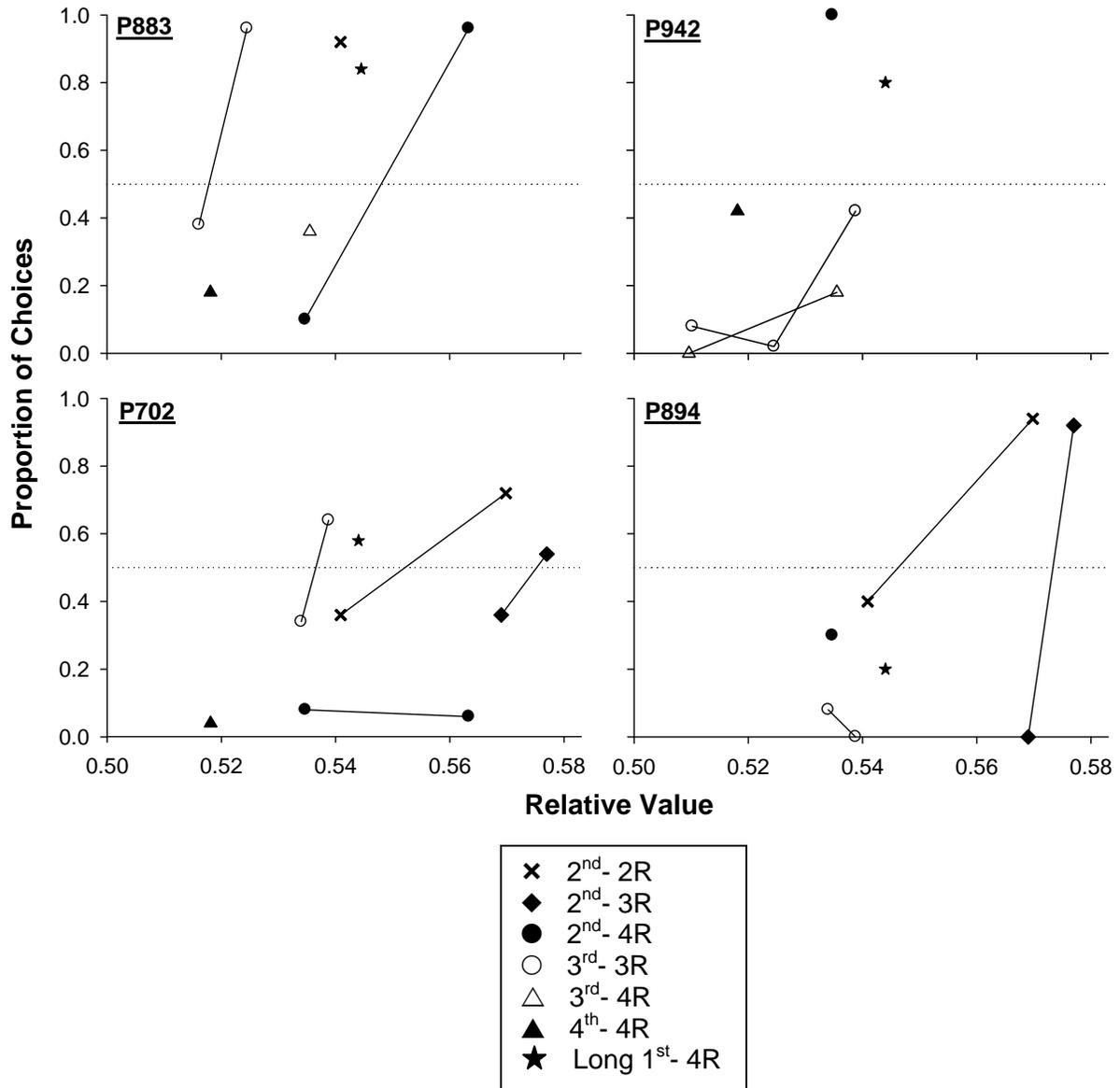


Figure 4-3. Mean proportion of choices on the richer alternative in function of its relative value. The first number (from the left) written inside the legend indicates which reinforcer was manipulated in the series, and the second number indicates how many reinforcers the sequence was comprised of in each comparison. See text for further details.

## CHAPTER 5 GENERAL DISCUSSION

One of the main goals of this study was a comparative analysis of choices between reinforcer sequences in humans and pigeons. Experiments 1 and 2 were aimed to bring procedures used with pigeons and humans into better alignment through the token and consumable-type reinforcers and the manipulation of the economic context. In general, a similar pattern of choices was found across species. Data obtained during IC conditions, in which tokens were immediately exchangeable for food, showed that both pigeons and humans tended to prefer sequences with the shortest delay to the initial reinforcer. In DC conditions, in which tokens could not be exchanged until the end of the trial, preferences still generally favored the sequences with the shorter initial reinforcer delay, though the more variable and more susceptible to carryover effects from the immediately prior condition. Both the differences across the IC and DC conditions and the general similarities across species are consistent with the general thrust of the present research and the broader program of which it is a part (Hackenberg, 2005, 2009), in showing that species differences often reflect differences in methods used to study different species.

Experiment 3 was an extension of Experiment 1, aimed at further assessing sensitivity to a particular delay within a reinforcer sequence. Pigeons' choices showed sensitivity to the delay to each reinforcer in a sequence comprised of 2 and 3 outcomes, but not with a sequence comprised of 4 outcomes. In relation to 4 reinforcer sequences, none of the pigeons showed sensitivity to the 3<sup>rd</sup> reinforcer, and only 2 showed sensitivity to the 2<sup>nd</sup> reinforcer in the series. These results indicate that the control of the behavior diminishes when sequences are comprised of more than 3

reinforcers. Further evidence of sensitivity to multiple reinforcers arrayed over time was obtained in the last condition, in which a majority of subjects preferred a sequence containing a longer delay to the first, but a shorter delay to 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> reinforcer in the series. Together with the differential preference toward the sequence with the shorter delay after the values were made more extreme, these results indicate that the value of each reinforcer in the series is independent and additive, with value computed according to the hyperbolic discounting equation. And along with the results obtained in the IC conditions of Experiment 1 and 2, these data provide further empirical support for the predictions of Equation 1-1.

It is important to mention that the DU model used in Economics also incorporates the assumption that the value of a sequence of reinforcers is equal to the present discounted value of each member of the sequence added together (that is, each reinforcer is viewed as independent and additive). Therefore, the results presented here are also broadly consistent with the predictions of the DU model. This study was not designed to differentiate between the hyperbolic-decay model and the DU model; to do so it would be necessary to implement a procedure that allowed a more precise quantitative analysis that would make it possible to distinguish the hyperboloid function of the former from the exponential function of the latter.

### **Cross-species Analysis of Choice**

The results obtained in Experiment 1 and 2 are consistent with prior data on temporal discounting with non-humans (Brunner, 1999, Exp.1) but differ from prior data with humans (e.g., Ariely, 1998; Chapman, 2000; Loewenstein & Prelec, 1993; Loewenstein & Sicherman, 1991; Schreiber & Kahneman, 2000). This suggests that

previously reported differences between species may be due in part to procedural discrepancies rather than to more fundamental differences in behavioral process.

Some important procedural differences include: (1) the nature of the reinforcer—actual outcomes rather than hypothetical ones; (2) repeated exposure to the choice contingency and its consequences under steady-state conditions; (3) parametric manipulation of reinforcer variables on a within-subject basis; (4) measurement of preference via actual choices rather than verbal ratings.

The last point raised above (4) is a critical one, raising fundamental questions of response definition. The literature in choice between sequences of reinforcement has traditionally been dominated by studies in which subjects report what they think they would do it in a given hypothetical scenario. But an essential question remains unanswered: Would they actually do as they say when exposed to actual events in life? Verbal statements about imagined behavior and actual response allocation among alternative sources of reinforcement (i.e., choices) are essentially two different classes of operant behavior potentially controlled by different classes of stimulus events.

One reason hypothetical scenarios presented via questionnaire has been the method of choice in Psychology and Economics is practicality. It is far easier to provide subjects with a brief questionnaire than to study their choices repeatedly over time and in relation to a parametric range of reinforcement variables. This is not to diminish the importance of research that uses hypothetical choice as dependent variables. This type of research has led to interesting findings and has shed light in many interesting issues over the years. The point being raised here is that as any other research, reliability or

external validity of the findings obtained has to be tested, rather than assumed; and this is especially the case with research that uses imagined scenarios as research tools.

In the present human experiment, subjects *behaviorally* demonstrated preference for the worsening over improving sequence, but when given the questionnaire preferred the improving sequence in 2 of the 4 scenarios described. Although the results obtained here were a bit different from Loewenstein and Prelec (1993) in that the latter reported a higher percentage of subjects preferring the IMP sequence across all four scenarios, they confirmed the authors claim that the “spread” (to use their term) of the outcomes within the sequence is a variable that affects preference for improving.

Preference for IMP presents formidable theoretical and empirical challenges because it runs counter to most psychological, economic, and evolutionary models based on delay discounting. From an evolutionary standpoint, it is possible that the strong sensitivity to reinforcer immediacy has been selected over time due to its importance to survival (Moore, 1988). Discounting of future events makes adaptive sense when environmental factors such as uncertainty or competition for resources are taken into consideration (Kagel, Green, & Caraco, 1986). From this evolutionary perspective, then, preference for improving sequences is challenging because “it carries with it the risk of losing the best part of the reward sequences” (Brunner, 1999, p. 96).

Although the world we (humans) inhabit today is much more stable than in previous epochs, it is highly unlikely that organisms with different tendencies to respond to future events have been selected through evolution. There has been insufficient time for such process to occur. Therefore, in accounting for human preference for improving over worsening sequence, one would have to look into the ontogenetic history of

individuals in order to find factors that potentially can counter-balance this more fundamental (phylogenetic) tendency to choose sequences containing the best outcome first (i.e., worsening). One such factor (or mechanism) that seems theoretically plausible is that choosing improving sequences is rule-governed behavior. Rules are defined as verbal antecedents or contingency-specifying verbal stimulus, and rule-governed behavior is behavior under the control of this verbal antecedent stimulus (Skinner, 1966). Thus, it is possible that the verbal community teaches its members to choose sequences that improve rather than sequences that worsen in much the same way they seem to teach its members to choose larger later reinforcers (self-control) over smaller immediate ones (impulsivity).

### **Limitations and Future Directions**

Perhaps the most significant limitation across all three experiments was that the procedure implemented here often produced extreme preference for a single alternative. This preference exclusivity was likely due to the concurrent FR1FR1 schedule arranged in the choice phase (i.e., discrete-trial in the initial-link). This feature of the procedure makes it difficult to capture different degrees of preference. It would be informative to see by how much subjects prefer one alternative over the other, instead of only which one they prefer. In addition, this procedure makes it challenging to dissociate response bias from indifference. To demonstrate genuine preference, frequent side reversals were necessary. Future research might employ longer initial-link choice periods, which may produce more graded preferences between the reinforcer alternatives.

Future studies involving choice with actual reinforcers might manipulate variables other than delay, such as quality, magnitude, or probability of reinforcement. And in

addition to positive reinforcers, it would also be of great interest to examine preference for sequences of negative reinforcers. Note that there have been studies published using aversive events, but the dependent measure used was mainly rating (e.g., Ariely, 1998; Ariely & Zauberman, 2000; Schreibber & Kahneman, 2000); and when subjects were required to chose a sequence, they were not exposed to the sequence they selected.

To illustrate how such research might proceed with actual choices between actual negative reinforcer sequences, consider a recent pilot study from our lab. With video clip reinforcers, we manipulated the quality of the videos by inserting frequent interruptions while the video was being watched. The worsening alternative led to a video segment with increasing number of interruptions, whereas the improving alternative led to a segment with decreasing number of interruptions. Unfortunately, due to time constraints, the experiment was not completed. Nevertheless, the study suggests a useful method for approaching the important issue of choice between streams of aversive events.

APPENDIX  
QUESTIONNAIRE

Question 1: Which would you prefer if both were free?

- a. Dinner at a fancy French restaurant
- b. Dinner at a local Greek restaurant.

<b><i>If you prefer French</i></b>	<b><i>If you prefer Greek</i></b>
<p><u>Question 2:</u> Which would you prefer?</p> <ul style="list-style-type: none"> <li>a. Dinner at the French restaurant on Friday in 1 month</li> <li>b. Dinner at the French restaurant on Friday in 2 months.</li> </ul>	<p><u>Question 2:</u> Which would you prefer?</p> <ul style="list-style-type: none"> <li>a. Dinner at the Greek restaurant on Friday in 1 month</li> <li>b. Dinner at the Greek restaurant on Friday in 2 months.</li> </ul>
<p><u>Question 3:</u> Which would you prefer?</p> <ul style="list-style-type: none"> <li>a. Dinner at the French restaurant on Friday in 1 month and dinner at the Greek restaurant on Friday in 2 months</li> <li>b. Dinner at the Greek restaurant on Friday in 1 month and dinner at the French restaurant on Friday in 2 months</li> </ul>	<p><u>Question 3:</u> Which would you prefer?</p> <ul style="list-style-type: none"> <li>a. Dinner at the Greek restaurant on Friday in 1 month and dinner at the French restaurant on Friday in 2 months</li> <li>b. Dinner at the French restaurant on Friday in 1 month and dinner at the Greek restaurant on Friday in 2 months</li> </ul>

Question 4:

Imagine you must schedule two weekend outings to a city where you once lived. Suppose one outing will take place this coming weekend, the other the weekend after.

- a. This weekend: friends and  
Next weekend: unpleasant Aunt
- b. This weekend: unpleasant Aunt  
Next weekend: friends

Question 5:

Suppose one outing will take place this coming weekend, the other in 6 months (26 weeks)

- a. This weekend: friends  
26 weeks from now: unpleasant Aunt
- b. This weekend: unpleasant Aunt  
26 weeks from now: friends

Question 6:

Suppose one outing will take place in 6 months (26 weeks from now), the other the weekend after (27 weeks from now)

- a. 26 weeks from now: friends  
27 weeks from now: unpleasant Aunt
  
- b. 26 weeks from now: Unpleasant Aunt  
27 weeks from now: friends

## LIST OF REFERENCES

- Anderson, A. C. (1932). Time discrimination in the white rat. *Journal of Comparative Psychology*, 13, 27-55.
- Ainslie, G. (1974). Impulse control in pigeons. *Journal of the Experimental Analysis of Behavior*, 21, 485-489.
- Ainslie, G. (1975). Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 82, 463-496.
- Ainslie, G., & Herrnstein, R. (1981). Preference reversal and delayed reinforcement. *Animal Learning & Behavior*, 9, 476-482.
- Ariely, D. (1998). Combining experiences over time: The effects of duration, intensity changes and on-line measurements on retrospective pain evaluations. *Journal of Behavioral Decision Making*, 11, 19-45.
- Ariely, D., & Loewenstein, G. (2000). When does duration matter in judgment and decision making? *Journal of Experimental Psychology: General*, 129, 508-523.
- Ariely, D., & Zauberman, G. (2000). On the making of an experience: The effects of breaking and combining experiences on their overall evaluation. *Journal of Behavioral Decision Making*, 13, 219-232.
- Berns, G., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice – toward an integrative framework. *Trends in Cognitive Sciences*, 11, 482-488.
- Brunner, D. (1999). Preference for sequences of rewards: further tests of a parallel discounting model. *Behavioral Process*, 45, 87–99.
- Brunner, D., & Gibbon, J. (1995). Value of food aggregates: parallel versus serial discounting. *Animal Behavior*, 50, 1627–1634.
- Chapman, G. B. (2000). Preferences for improving and declining sequences of health outcomes. *Journal of Behavioral Decision Making*, 13, 203-18.
- Estle, J. S, Green, L., Myerson, J., & Holt, D. (2007). Discounting of monetary and directly consumable rewards. *Psychological Science*, 18, 58-63.
- Flora, S. R., & Pavlik, W. B. (1992). Human self-control and the density of reinforcement. *Journal of the Experimental Analysis of Behavior*, 57, 201-208.
- Frederick, S. Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature* XL (June): 351-401.

- Green, L., Fisher, E., Perlow, S., & Sherman, L. (1981). Preference reversal and self control: Choice as a function of reward amount and delay. *Behaviour Analysis Letters*, 1, 43-51
- Green, L., Fristoe, N., & Myerson, J. (1994). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychonomic Bulletin & Review*, 1, 383-389.
- Green, L., Fry, A. F., & Myerson, J. (1994). Discounting of delayed rewards: A life-span comparison. *Psychological Science*, 5, 33-36.
- Green, L., Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychological Bulletin*, 130, 769-792.
- Guyse, J., Keller, L., & Eppel, T. (2002). Valuing environmental outcomes: Preferences for constant or improving sequences. *Organizational Behavior and Human Decision Processes*, 87, 253-277.
- Hackenberg, T. D. (2005). Of pigeons and people: Some observations on species differences in choice and selfcontrol. *Brazilian Journal of Behavior Analysis*, 1, 135-147.
- Hackenberg, T. D. (2009). Token reinforcement: A review and analysis. *Journal of the Experimental Analysis of Behavior*, 91, 257-286.
- Hackenberg, T. D., & Pietras, C. J. (2000). Video access as a reinforcer in a self-control paradigm: A method and some data. *Experimental Analysis of Human Behavior Bulletin*, 18, 1-5.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Horne, P. J., & Lowe, C. F. (1993). Determinants of human performance on concurrent schedules. *Journal of the Experimental Analysis of Behavior*, 59, 29-60.
- Hsee, C. K., Abelson, R. P., & Salovey, P. (1991). The relative weighting of position and velocity in satisfaction. *Psychological Science*, 2, 263-266.
- Hull, C. (1943). *Principles of Behavior: An Introduction to Behavior Theory*. Oxford England: Appleton-Century.
- Hyten, C., Madden, G. J., & Field, D. P. (1994). Exchange delays and impulsive choice in adult humans. *Journal of the Experimental Analysis of Behavior*, 62, 225-233.
- Jackson, K., & Hackenberg, T. D (1996). Token reinforcement, choice, and self-control in pigeons. *Journal of the Experimental Analysis of Behavior*, 66, 29-49.

- Kagel, J. H., Green, L., & Caraco, T. (1986). When foragers discount the future: Constraint or adaptation? *Animal Behaviour*, 271-283.
- Kahneman, D., Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: Adding a better end. *Psychological Science*, 4, 401-405.
- Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, 126, 54-70.
- Kirby, K. N. (2006). The present values of delayed rewards are approximately additive. *Behavioural Processes*, 72, 273-282.
- Kirby, K. N., & Herrnstein, R. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychological Science*, 6, 83-89.
- Lagorio, C. H., & Hackenberg, T. D. (2010). Risky choice in pigeons and humans: A cross-species comparison. *Journal of the Experimental Analysis of Behavior*, 93, 27-44.
- Locey, M. L., Pietras, C. J., & Hackenberg, T. D. (2009). Human risky choice: Delay sensitivity depends on reinforcer type. *Journal of the Experimental Analysis of Behavior*, 35, 15-22.
- Loewenstein, G., & Prelec, D. (1993). Preferences for sequences of outcomes. *Psychological Review*, 100, 91-108.
- Loewenstein, G., & Sicherman, N. (1991). Do workers prefer increasing wage profiles? *Journal of Labor Economics*, 9, 67-84.
- Logue, A. W. (1988). Research on self-control: An integrating framework. *Behavioral and Brain Sciences*, 11, 665-679.
- Logue, A. W., King, G. R., Chavarro, A., & Volpe, J. S. (1990). Matching and maximizing in a self-control paradigm using human subjects. *Learning and Motivation*, 21, 340-368.
- Logue, A. W., & Peña-Correal, T. E. (1984). Responding during reinforcement delay in a self-control paradigm. *Journal of the Experimental Analysis of Behavior*, 41, 267-277.
- Logue, A. W., Peña-Correal, T. E., Rodriguez, M. L., & Kabela, E. (1986). Self-control in adult humans: Variation in positive reinforcer amount and delay. *Journal of the Experimental Analysis of Behavior*, 46, 159-173.

- Matsumoto, D., Peecher, M., & Rich, J. (2000). Evaluations of outcome sequences. *Organizational Behavior and Human Decision Processes*, 84, 331-352.
- Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 426-436.
- Mazur, J.E. (1986). Choice between single and multiple delayed reinforcers. *Journal of Experimental Analysis of Behavior*, 46, 67–77.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value* (Vol.5, pp. 55-73). Hillsdale, NJ: Erlbaum.
- Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, 108, 96-112.
- Mazur, J. E. (2007). Rats' choices between one and two delayed reinforcers. *Learning & Behavior*, 35, 169-176.
- Mazur, J. E., & Logue, A. W. (1978). Choice in a “self-control” paradigm: Effects of a fading procedure. *Journal of the Experimental Analysis of Behavior*, 30, 11-17.
- McDiarmid, C. G., & Rilling, M. E. (1965). Reinforcement delay and reinforcement rate as determinants of schedule preference. *Psychonomic Science*, 2, 195-196.
- Millar, A., & Navarick, D. (1984). Self-control and choice in humans: Effects of video game playing as a positive reinforcer. *Learning and Motivation*, 15, 203-218.
- Montgomery, N., & Unnava, H. (2009). Temporal sequence effects: A memory framework. *Journal of Consumer Research*, 36, 83-92.
- Moore, J. (1979). Choice and number of reinforcers. *Journal of the Experimental Analysis of Behavior*, 32, 51-63.
- Moore, J. (1982). Choice and multiple reinforcers. *Journal of the Experimental Analysis of Behavior*, 37, 115-122.
- Moore, J. (1988). Evolution and impulsiveness. *Behavioral and Brain Sciences*, 11, 691.
- Navarick, D. J. (1996). Choice in humans: Techniques for enhancing sensitivity to reinforcement immediacy. *The Psychological Record*, 46, 539–554.
- Navarick, D. J. (1998). Impulsive choice in adults: How consistent are individual differences? *The Psychological Record*, 48, 665–674.

- Navarick, D. J. (2004). Discounting of delayed reinforcers: Measurement by questionnaires versus operant choice procedures. *The Psychological Record*, 54, 85-94.
- Perin, C. T. (1943). The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *Journal of Experimental Psychology*, 32, 95-109.
- Pierce, W. D., & Cheney, C.D. (2008). *Behavior Analysis and Learning*. New York, NY: Psychology Press.
- Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, 17, 15-22.
- Rachlin, H., Raineri, A., Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233-44.
- Read, D., & Powell, M. (2002). Reasons for sequence preferences. *Journal of Behavioral Decision Making*, 15(5), 433-460.
- Ross, W. T., & Simonson, I. (1991). Evaluations of pairs of experiences: A preference for happy endings. *Journal of Behavioral Decision Making*, 4, 273-282.
- Samuelson, P. (1937). A note on measurement of utility. *Review of Economic Studies*, 4, 155-161.
- Schmitt, D., & Kemper, T. (1996). Preference for different sequences of increasing or decreasing rewards. *Organizational Behavior and Human Decision Processes*, 66, 89-101.
- Schreiber, C. A., & Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General*, 129, 27-42.
- Shull, R.L., Mellon, R., Sharp, J.A (1990). Delay and number of food reinforcers: Effects on choice and latencies. *Journal of Experimental Analysis of Behavior*, 53, 235–246.
- Skinner, B. F. (1950). Are theories of learning necessities? *Psychological Review*, 57, 193-216.
- Skinner, B. F. (1966). An operant analysis of problem solving. In B. Kleinmuntz (Ed.), *Problem solving: Research, method and theory* (pp. 225-257). New York: Wiley.
- Solnick, J. V., Kannenberg, C. H., Eckerman, D. A., & Waller, M. B. (1980). An experimental analysis of impulsivity and impulse control in humans. *Learning and Motivation*, 11, 61-77.

Thorndike, E. (1913). *Educational Psychology. Vol. I: The Original Nature of Man.* Columbia Univ.: New York.

Varey, C., & Kahneman, D. (1992). Experiences extended across time: Evaluation of moments and episodes. *Journal of Behavioral Decision Making*, 5, 169-185.

Wolfe, J. B. (1934). The effect of delayed reward upon learning in the white rat. *Journal of Comparative Psychology*, 17, 1-21.

## BIOGRAPHICAL SKETCH

Leonardo was born in Boston (MA) in 1974, when his father was doing his PhD in Economics. He left New England when he was three years of age, and lived in Brazil until coming to Florida in 2005. He did his undergraduate studies at the University Center of Brasilia (UniCeub), a private university more oriented toward service-delivery (application) than toward research. While studying there, he read Skinner's *Science and Human Behavior*, a book that had a profound impact in his views about Psychology. After reading Skinner, he decided that he needed to learn more about Behavior Analysis and obtain more experience with research, an area that was clearly lacking in his education. So, after Leonardo graduated in Psychology in 2002, he decided to apply for a place in the master's degree program at the University of Brasilia; a top, public, research-oriented university in Brazil. He was accepted in 2003 and finished his master's degree in 2005.

The two years Leonardo spent in the University of Brasilia were very fruitful for his academic purposes. While studying there he had the privilege of working with Elenice Hanna. He soon started some experiments involving self-control in pigeons and children, one of which would become his master's thesis. More specifically, his master's thesis described a research project aimed at assessing the effects of activities made available during the delay to reinforcement on self-control choices in children.

Shortly after Leonardo obtained his master's degree, he decided he wanted to do his PhD in the United States. He contacted Timothy Hackenberg, a professor he had been very fortunate to have met in Brazil, and who had indicated that he shared some of the same research interests Leonardo did, such as token reinforcement and choice.

Leonardo was accepted at University of Florida and started his doctoral studies in the fall of 2005.

While studying at University of Florida, Leonardo has conducted research mainly examining choice and conditioned reinforcers. He has been involved in three long research projects. One project involved the assessment of token value. In this study he investigated how the amount of food for which a token was exchangeable affected subjects' demand functions, breakpoints, and preferences. The second project involved the assessment of the reinforcer functions of generalized and conditioned reinforcers. The third project, the one he has chosen as his dissertation, involved a cross-species comparison of choices between sequences of reinforcers.

Leonardo passed his qualification exam in the fall of 2009, and graduated from University of Florida in the spring of 2010.