

THE EFFECTS OF CONCURRENT REINFORCEMENT OF
VARIABILITY AND DIFFERENTIAL OUTCOMES
ON PIGEONS' ACQUISITION RATE
LEARNING A DIFFICULT
TARGET SEQUENCE

A Thesis Presented to the Faculty
of
California State University, Stanislaus

In Partial Fulfillment
of the Requirements for the Degree
of Master of Science in Psychology

By
Meloney A. West
November 2018

CERTIFICATION OF APPROVAL

THE EFFECTS OF CONCURRENT REINFORCEMENT OF VARIABILITY
AND DIFFERENTIAL OUTCOMES ON PIGEONS' ACQUISITION RATE
LEARNING A DIFFICULT TARGET SEQUENCE

By
Meloney A. West

Signed Certification of Approval page is on file with the University Library.

Dr. Bruce Hesse
Professor of Psychology

Date

Dr. William Potter
Professor of Psychology

Date

Dr. Victor Luevano
Professor of Psychology

Date

© 2018

Meloney A. West
ALL RIGHTS RESERVED

DEDICATION

This work is dedicated to my husband, my mom, and my children for their support during the long process of completing this thesis. I'd also like to give special thanks to my wonderful husband, Russ West, for always believing in me even when I didn't believe in myself!

ACKNOWLEDGEMENTS

I'd like to acknowledge the help of my thesis committee, Dr. William Potter, Dr. Victor Luevano, and most of all Dr. Bruce Hesse., for their help in the process of completing this thesis. Their time and thought into helping me edit my draft writings and suggestions and questions are greatly appreciated. I'd also like to give special thanks to Dr. Bruce Hesse for teaching me during my first behavior analysis class and helping me to fall in love with the experimental analysis of behavior.

TABLE OF CONTENTS

	PAGE
Dedication.....	iv
Acknowledgements.....	v
List of Figures.....	vii
Abstract.....	viii
CHAPTER	
I. Introduction.....	1
Operant Conditioning and Shaping New Responses	2
Methods to Increase Variability.....	3
The Differential Outcomes Effect.....	7
The Current Study.....	8
II. Methodology.....	10
Subjects.....	10
Apparatus.....	10
Design.....	11
Procedure.....	12
III. Results.....	16
Bird 1.....	16
Bird 2.....	17
Bird 3.....	19
Bird 4.....	20
IV. Discussion.....	22
Limitations and Possible Confounds	25
References.....	28

LIST OF FIGURES

FIGURE	PAGE
1. Figure 1	17
2. Figure 2	18
3. Figure 3	20
4. Figure 4	21

ABSTRACT

This study investigated if the addition of a differential outcomes arrangement to direct reinforcement of variability would further facilitate learning a difficult target sequence. Four White Carneau pigeons (*Columba livia*) of undetermined sex using an alternating treatment design were used in this experiment. Differential outcomes included differences in food hopper light color and amount of time to food reinforcement between two concurrently available reinforcement schedules. Three of four birds learned the target sequence sooner and more consistently when differential outcomes were implemented compared to when they were not (i.e. equal outcomes). Findings add to the literature by showing that differential outcomes can be used in addition to concurrent reinforcement of variability to further expedite pigeons learning difficult sequences. The importance of both variability and repetition (stereotypy) are discussed in real-world situations and in their role while learning difficult target sequences.

CHAPTER I

INTRODUCTION

Have you ever wondered how a person learned to solve a really difficult problem? Understanding how we learn new things is important not only for our own well-being and survival, but also for teaching others. According to B. F. Skinner's selectionist theory of operant conditioning (1938, 1953), the consequences of behavior are a crucial part of understanding learning. Responses are selected by their consequences and in order for new operant behaviors to be shaped, variations in responses are necessary. The shaping of new responses requires both a repertoire with variable behaviors to be selected from and reinforcement to maintain responding until the desired target behavior is established (Skinner, 1981).

Although there is inherent variability in all behavior, sometimes more is needed to successfully shape a new response (Grunow & Neuringer, 2002). Withholding reinforcement and reducing the frequency of reinforcement have been shown to increase variability, but can cause problems when being used to train a new response (Antonitis, 1951; Neuringer, Kornell, & Olufs, 2001). Directly reinforcing variability has been shown not only to effectively increase variable responding while avoiding aforementioned problems, but also to facilitate learning new responses, especially those responses that are particularly difficult to learn.

Neuringer, Deiss, and Olson (2000) compared three groups of rats learning difficult target sequences of lever presses. Results showed that only rats concurrently reinforced for responding variably learned difficult target sequences. Rats that either

did not receive concurrent reinforcement at all or did receive concurrent reinforcement that was not contingent on responding variably did not learn the difficult target sequences. Although concurrent reinforcement of variability facilitated learning difficult target sequences in the Neuringer et al (2000) study, learning seems like it could be further expedited if the animals were given more information as to which schedule of reinforcement they were responding on. Studies using differential outcomes procedures, in which a correct response to each of multiple available contingencies results in its own unique reinforcement outcome, have resulted in faster learning during discrimination tasks (see Goeters, Blakely, & Poling, 1992, for a review). Unique response-outcome associations are thought to be learned that provide information that help guide operant behavior during such tasks (Urcuioli, 2005). The present study was designed to find out if, in addition to concurrent reinforcement of variable responding, a differential outcomes procedure would further expedite learning target sequences.

Operant Conditioning and Shaping New Responses

One way to consider learning is from the perspective of B. F. Skinner's selectionist theory of operant conditioning (1938, 1953). Operant conditioning is the process by which humans and animals learn from the consequences of their behavior. Operant behavior is explained by considering its functional relationship to the consequences produced by it. For example, if you want to know why a child throws tantrums, it's important to analyze the function of that tantrum throwing behavior by considering what effect or consequence that behavior has produced in the past. If the

child's tantrum behavior has resulted in something rewarding to the child, like attention, that behavior is strengthened and will occur more often. In this way operant responses are selected by their history of reinforcing consequences.

Shaping is a method used to teach a new response that an animal or person is not currently able to emit (Pierce & Cheney, 2008). To establish a new response, closer and closer approximations to the target or terminal behavior are reinforced. For example, shaping a rat to press a lever with its paw would require reinforcing instances of behavior that come closer and closer to the final performance of lever pressing. This may begin with reinforcing the rat when it comes closer and closer in proximity to the actual lever, then reinforcing when the rat's paw comes closer to the lever. This reinforcement of closer and closer approximations would continue in order to develop the new response of lever pressing. The shaping of new responses is only possible because animals and people respond variably. In other words, no two emitted responses are exactly alike and this behavioral variation is what makes it possible for new responses to be shaped using reinforcement.

Methods to Increase Variability

Although there is some level of variation in all operant behavior, sometimes it is necessary to increase baseline levels of variability in order to teach new behaviors (Grunow & Neuringer, 2002).

Withholding reinforcement has been shown to increase variability (Antonitis, 1951; Neuringer et al., 2001). However, there are problems with using extinction procedures to increase variability to shape new responses. Rates of responding

decrease consistently during periods of extinction, sometimes even to zero levels (Grunow & Neuringer, 2002). During shaping, responding must be provided at least occasionally. Without enough responses to select from, shaping becomes difficult, if not impossible to accomplish. It has also been shown that previously reinforced responses persist when reinforcement is withheld (Shwartz, 1981; Myerson and Hale, 1988). This also makes it difficult to select new responses during shaping because variation is needed to choose responses that come closer to the desired terminal behavior.

Decreasing reinforcement frequency without completely removing it has been suggested as an alternative to extinction procedures to increase variability (Grunow & Neuringer, 2002). Some studies have shown that by lowering frequencies of reinforcement as is done using intermittent schedules (Boren, Moerschbaecher, & Whyte, 1978; Lachter & Corey, 1982), and by decreasing the amount of reinforcement (Tatham, Wanchisen, & Hine, 1993), variability increases. However, other studies report little or no effects when reinforcement is decreased (Eckerman & Lanson, 1969; Machado, 1989).

Direct Reinforcement of Variability

Variability produced from direct reinforcement has been demonstrated in many studies (see Neuringer, 2009 for a review). Early studies revealed that high levels of variability in porpoises' behavior (Pryor, Haag, & O'Reilly, 1969) and pigeons' pecking behavior (Blough, 1966) could be produced by directly reinforcing variable behaviors. Later studies showed that high levels of sequence variability could

be produced when relatively infrequent sequences of lever presses in rats or key presses in pigeons were reinforced (Machado, 1989; Neuringer, 1991). Precise control of various levels of variability (i.e. low, medium, high) by directly reinforcing each level has also been accomplished (Machado, 1994; Neuringer, 1992; Page & Neuringer, 1985).

Neuringer et al. (2000) examined the role of direct reinforcement of variability in operant learning by comparing three different procedures used to train response sequences with rats. The rats were required to emit left (L) and right (R) lever presses in a specified sequence to receive food pellet reinforcement. Five experimental phases were examined in which predetermined target sequences ranged from easy (i.e. LR, RLL) to difficult (i.e. RRLR, LLRRL). Rats were divided into three groups that were all reinforced on a continuous schedule of reinforcement (CRF) for emitting the predetermined target sequence. For one group, the variability-contingent reinforcement (VAR) group, concurrent reinforcement for responding variably was provided on a VI60s schedule. A second group, the reinforcement of any sequence (ANY) group, also received concurrent reinforcement on a VI60s schedule, but for any sequence of responses other than the target sequence, regardless of whether or not the emitted sequence was considered variable. For the third control (CON) group, concurrent reinforcement was not available; this group was only reinforced if the target sequence was emitted. Results indicated that response rates of VAR and ANY groups did not differ from that of the CON group for easier 2- and 3-response sequence targets, but were significantly higher for more difficult 4- and 5-response

sequence targets. This suggests that available concurrent reinforcement on the VI60s schedule helped to maintain responding while learning difficult target sequences.

More important to the current study, only the VAR group, that received reinforcement for variability, showed evidence of learning the more difficult four- and five-response sequences as indicated by both a higher percentage of targets emitted in sessions and higher absolute rate of target emissions. These results led Neuringer et al. (2000) to conclude that concurrent reinforcement of variability facilitated learning difficult target sequences.

In the Neuringer et al. (2000) experiment two schedules of reinforcement were simultaneously available for the animals in the VAR group. Correctly responding either variably or by emitting the difficult target sequence was reinforced in exactly the same way with one food pellet. The fact that no schedule-specific discriminative stimuli were available and that correct responses to either available schedule resulted in the exactly same reinforcement leads one to wonder how an animal was to differentiate which schedule it was responding to. With this setup, an animal would have to learn over time that the target response results in immediate reinforcement every time it is emitted. This arrangement seems to provide few cues to the animal, thus leading to slower learning of the difficult sequence. If cues were available, indicating which of the two schedules the animal was responding to, it seems like learning would be expedited.

The Differential Outcomes Effect

The differential outcomes effect (DOE) is demonstrated when an increased rate and accuracy of learning are observed when each of two or more discriminative stimuli is correlated with its own distinct outcome (e.g., type of reinforcer) during discrimination tasks. This differential outcome (DO) arrangement is unlike standard operant procedures in which any of multiple correct responses results in the same outcome or equal outcomes (EO). Trapold (1970) first investigated this phenomenon in an attempt to find out if different reinforcers establish distinctively different expectations which can function as additional cues for each correct response. He subjected experimentally naïve rats to a 2-choice conditional discrimination task in which a response to the right lever (R1) in the presence of one stimulus (tone) and a response to the left lever (R2) in the presence of a different stimulus (click) was required. When he reinforced the correct responses to each R1 and R2 with different types of reinforcement (i.e. R1: food; R2: sucrose), he found that rats learned the correct responses significantly faster than if the type of reinforcement were the same (i.e. R1: food; R2: food).

Since Trapold's (1970) original investigation, the DOE has been consistently replicated in many experiments in different ways. The DOE has been demonstrated using rats (Fedorchak & Bolles, 1986; Papini & Silingardi, 1989), dogs (Overmier, Bull, & Trapold, 1971), pigeons (Urcuioli & Zentall, 1990; Williams, Butler, & Overmier, 1990), and humans (Miller, Waugh, & Chambers, 2002; Mok & Overmier, 2007). Both within-subject (Honig, Matheson, & Dodd, 1984; Kruse & Overmier,

1982) and between-subject (DeLong & Wasserman, 1981; Friedman & Carlson, 1973) experimental designs have been used to reproduce the DOE. Outcomes have differed by type of reinforcement (e.g. food versus water), when reinforcement has been provided (e.g. immediately versus delayed), and by amount of reinforcement given (e.g. 1 versus 5 pellets of food).

Enhancement of learning by differential outcomes has been shown to be one of the most consistent and powerful effects on the learning and retention of conditional discriminations (Urcuioli, 1990a). However, since Traphold's (1970) first demonstration of DOE using a two-choice successive discrimination procedure, differential outcomes experiments have been limited to the use of variations of two general procedures: the matching-to-sample (MTS) task and the two-choice successive discrimination task (Goeters et al., 1992).

The Current Study

Concurrent reinforcement of variable responding has been shown to facilitate learning difficult target sequences with rats (Neuringer et al., 2000). The purpose of the current research was to find out if, in addition to concurrent reinforcement of variability, differential outcomes would further facilitate learning a difficult target sequence with pigeons. It was hypothesized when pigeons experienced differential outcomes between target response sequences and variable response sequences, the target response sequence would be learned sooner (i.e. earlier session as evidenced by graphed results) than when they experienced the same outcomes for both the target and variable sequences. It was also hypothesized that when pigeons experienced

differential outcomes of reinforcement, the target sequence would be learned more consistently as evidenced by visual analysis of trend of learning curve.

CHAPTER II

METHODOLOGY

Subjects

Four White Carneau pigeons (*Columba livia*) of undetermined sex were used in this study. All pigeons had previous experience learning pecking sequences in operant chambers. Pigeons were maintained at 85%-95% or more of their free-feeding weight with free access to water in their home-cages. Food was obtained during experimental sessions and additional food was given post-session if needed to maintain the bird's weight criterion or to compensate for a bird losing weight between sessions. Each pigeon was housed individually in a home-cage located in a colony room with other individually-housed pigeons on a 16:8 hr light-dark cycle.

Apparatus

Four BRS/LVE operant chambers were used in this experiment. Within each operant chamber, a work panel on one wall consisted of three response keys and an automated food hopper. Response keys were circular, horizontally arranged, and located 25.4 cm above the floor, and 6 cm apart. The food hopper was centered and positioned 12.7 cm below response keys. A feeder light illuminated food within the food hopper when it became available. A white house light illuminated the inside of each chamber during experimental sessions. A ventilation fan attached to each chamber ran continuously during experimental sessions, also served to dampen external sounds.

Computer Programming and Data Recording

An IBM compatible computer used MED-PC interface and software to arrange the production of stimuli and record responses to keys in the chambers. The left and right response keys were illuminated with colored light from small, 12 light projectors using colored photographic filters. Pecks to these keys closed a micro-switch and were recorded as responses by the computer. The center key was dark and not used in the current study.

Design

An alternating treatments design was used in this study over the course of 40 sessions (i.e. 20 sessions for each condition). Each of four birds was exposed to both the equal outcomes (EO) and differential outcomes (DO) arrangement, which was alternated by session. Sessions occurred daily for most of the experiment, and then occurred every-other day for the last 1/3 of experiment to adjust for bird satiation from food reinforcement.

In order to control for order effects, counterbalancing was achieved by randomly assigning birds to begin the experiment (i.e. first session) by experiencing either the EO or DO condition. This resulted in two birds experiencing the EO condition during the first session and two birds experiencing the DO condition during the first session. Chosen target sequences for each condition were also counterbalanced according to whether the required initial key peck was to the left or right key. This led to two birds experiencing RRLR as the EO target and LLRL as the

DO target and the other two birds experiencing LLRL as the EO target and RRLR as the DO target.

Target Sequences

Birds responded by pecking the right and left response keys. The order of pecking responses resulted in a sequence. During baseline, target sequences were determined to be difficult when they reached four responses in length (i.e. RLRR). There were 16 possible four-response sequences that could be emitted during any given trial, and of these, 10 were initially excluded as possible choices for targets based on not meeting one or more of the following three requirements: exactly two switches between keys, both keys being used, and no more than two repeats of either key. Of the six remaining sequences, two (RRLR and LLRL) were chosen for the EO and DO conditions of the experiment because they were inverted for counterbalancing purposes and similar in that they both had a 2-key repetition at the beginning of the sequence.

Procedure

Preliminary Training

All birds were first hand-shaped to eat food pellets from the food hopper. Birds were then shaped by the method of successive approximation to peck each of the two response keys that would be used in the experiment when illuminated. After birds successfully learned to peck each response key, they were reinforced with food in the food hopper for pecking any two-key sequence. After which, a switch between keys was required to gain reinforcement if the bird had not already been reinforced

for doing so. After birds learned to switch between two keys, three-response, followed by four-response sequences were required. If a bird did not meet the sequence requirement (2-, 3-, or 4-response sequence), a 4s time-out began during which the first 3s the house light and key lights were dark, followed by the house light coming back on for the remaining 1s. If a response was made during the last second of the time-out, the 1s interval was reset.

Experimental Phase

The experimental procedure involved two conditions (EO and DO) and began after preliminary training. Daily experimental sessions included a series of trials. Each trial began with both response keys illuminated in the same color. When a key was pecked, a short-0.2s darkening of it occurred, after which it relit. This was to confirm the peck was detected by the computer. Four key-peck responses made up a sequence and were included in each trial. Trials were separated by 3s inter-trial intervals (ITI) in which response keys were dark, but the house light remained on. Responses during this period reset the interval and were not included in data analyses. Sessions were conducted daily, seven days a week at approximately the same time each day. Each session ended either when a bird obtained 50 reinforcers or after 50 min had elapsed.

Intervention Phase. On alternating treatment days (i.e. sessions), each bird experienced either the EO or DO condition. Birds were randomly assigned as to which condition would be experienced on the first session. Two birds (birds 1 and 3) began the experiment by experiencing the EO condition during the first session, and

then experienced the DO condition during the following session (i.e. alternating day). The other two (birds 2 and 4) experienced these two conditions in reverse order (i.e. DO then EO). For two birds (birds 1 and 3), the target sequence was RRLR for the EO condition, and LLRL for the DO condition. For the other two birds (birds 2 and 4), the target sequences were reversed (i.e. LLRL for EO and RRLR for DO).

During each experimental session, two schedules of reinforcement were concurrently available. In addition to reinforcement being given for emitting a difficult target sequence (Target) on a CRF schedule as in the earlier training phase, reinforcement was also concurrently available for responding variably (VAR) on a VI60s schedule. Only responses that were infrequently emitted met the VAR contingency. A bird's current response sequence was compared to all its past response sequences across all previous sessions to determine if that response was infrequently emitted. The relative frequency of the response needed to be less than a threshold value of .067 as in the Neuringer et al. (2000) experiment to be considered infrequently emitted. If the target contingency was met, the VI60s schedule for the VAR reinforcement was reset. The target sequence was considered to be mastered when a bird obtained 80% or more of the 50 available reinforcers for the target sequence in a session for three consecutive sessions.

Equal outcomes. If either the CRF target contingency or the VI60s variability contingency was met, reinforcement was immediately available for 4.5s in the food hopper illuminated by a white feeder light (for two birds and a blue light for the other two). Reinforcement, the length of time food was available and the presence of the

white feeder light (or blue for two birds), stayed the same regardless of which schedule was met. In other words, a correct response to either of the two available schedules was reinforced in the same way. The EO phase of this experiment comprised the traditional, equivalent way that reinforcement has been provided in past experiments that concurrently reinforced variable responding when learning difficult target sequences (Maes & van der Goot, 2006; Neuringer et al., 2000).

Differential outcomes. If the VI60s variability contingency was met, reinforcement was immediately available for 3s in the food hopper illuminated by a white feeder light (for two birds and a blue light for the other two). However, if the CRF target contingency was met, reinforcement was immediately available for 6s in the food hopper illuminated by a different color light than that used for the other contingency (e.g. blue feeder light if white for other). Reinforcement for correct responses to the VI60s variability schedule was provided in the same, common way as both schedules in the EO phase, but reinforcement for meeting the CRF target contingency was distinct and provided longer access to food (6s) and different colored food hopper illumination. During the DO phase of the experiment, reinforcement was different depending on which of the two schedules were met.

CHAPTER III

RESULTS

Three of four birds met criterion for learning the target sequence during the 20 sessions of the experiment. Of the three birds that learned the target, all three met criterion sooner during the DO condition in comparison to the EO condition. Visual analysis indicated a greater learning curve during the DO condition when compared to the EO condition for three of the four birds. The fourth bird's data showed a similar positive trend in learning during both the EO and DO conditions.

Bird 1

Figure 1 illustrates the percentage of reinforcement Bird 1 acquired from the target sequence during EO condition when RRLR was the target sequence and DO condition when LLRL was the target sequence. Bird 1 met criterion for learning the target sequence during both EO and DO conditions within the 20 sessions of the experiment. During both conditions, Bird 1 began the experiment receiving reinforcement for emitting the target sequence at a low percentage. During the first eight sessions, Bird 1 showed an increasing trend for emitting the target sequence for both conditions with a greater increase in percentage during the DO condition when compared to the EO condition, with the exception of one outlier (i.e. session 6). Over the course of the sessions 9-20, Bird one showed a steadily increasing trend of emittance of the target sequence during the DO condition, but showed a decreasing trend in learning over the course of several sessions (i.e. sessions 9-14). Bird one met

criterion for learning the target sequence four sessions sooner during the DO condition (i.e. by session 14) than the EO condition (i.e. by session 18).

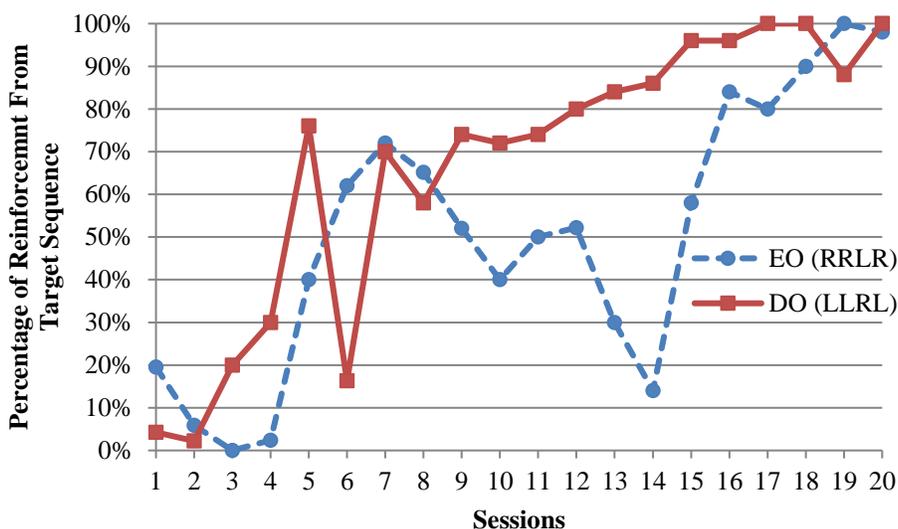


Figure 1. The percentage of reinforcement Bird 1 acquired from the target sequence during EO condition when RRLR was the target sequence and DO condition when LLRL was the target sequence. Percentage was calculated by dividing the total reinforcement gained by emitting the target sequence divided by the total amount of reinforcement gained during a session.

Bird 2

Figure 2 illustrates the percentage of reinforcement Bird 2 acquired from the target sequence during EO condition when LLRL was the target sequence and DO condition when RRLR was the target sequence. Bird 2 did not meet criterion for learning the target sequence during either the EO or DO conditions during the 20 sessions of the experiment. During the EO condition, Bird 2 showed a higher

percentage of target response during the initial session (i.e. 60%) compared to the DO condition (i.e.13%). During the EO condition, the level of responding showed moderate variability in pattern of learning during both the EO and DO conditions, with an overall higher level of responding to the target during the EO condition in comparison to the DO condition. Bird 2 showed a steady upward trend in learning during the DO condition compared to the EO condition in which the trend was much flatter.

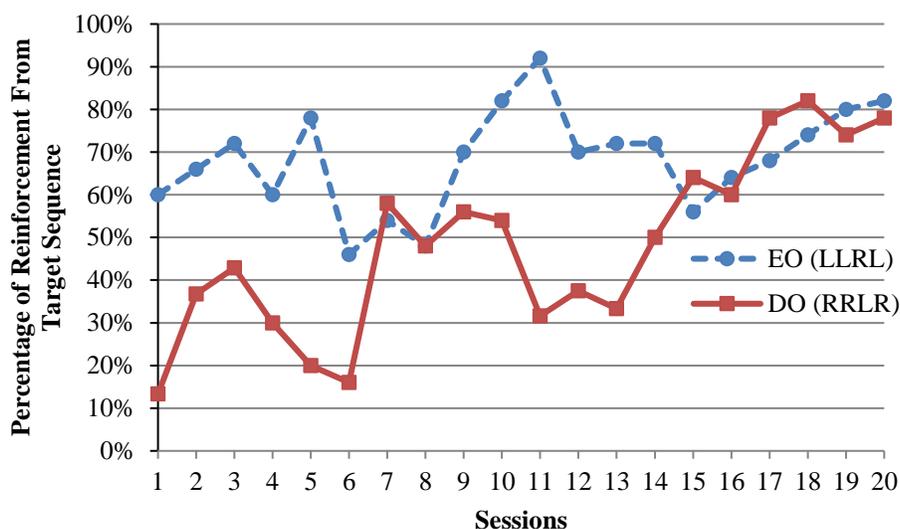


Figure 2. The percentage of reinforcement Bird 2 acquired from the target sequence during EO condition when LLRL was the target sequence and DO condition when RRLR was the target sequence. Percentage was calculated by dividing the total reinforcement gained by emitting the target sequence divided by the total amount of reinforcement gained during a session.

Bird 3

Figure 3 illustrates the percentage of reinforcement Bird 3 acquired from the target sequence during EO condition when RRLR was the target sequence and DO condition when LLRL was the target sequence. Bird 3 met criterion for learning the target sequence during both EO and DO conditions within 20 sessions. Bird 3 began the experiment by emitting the target sequence at about the same percentage (i.e. EO at 40% and DO at 34%). After this, Bird 3 showed an increasing trend in learning during the DO condition for the remainder of the sessions, with a sharp incline in positive trend during sessions 1-6. During the EO condition, Bird 3 also showed a positive trend in learning to a lesser degree with some decline in learning during sessions 4-7 emitting 46%, 78%, 56%, and 44% of reinforcement for emitting the target sequence. Bird 3 showed little to no variability during the DO condition and greater variability from trend during the EO condition. Bird 3 met criterion for learning the target sooner during the DO condition six sessions sooner than during the EO condition.

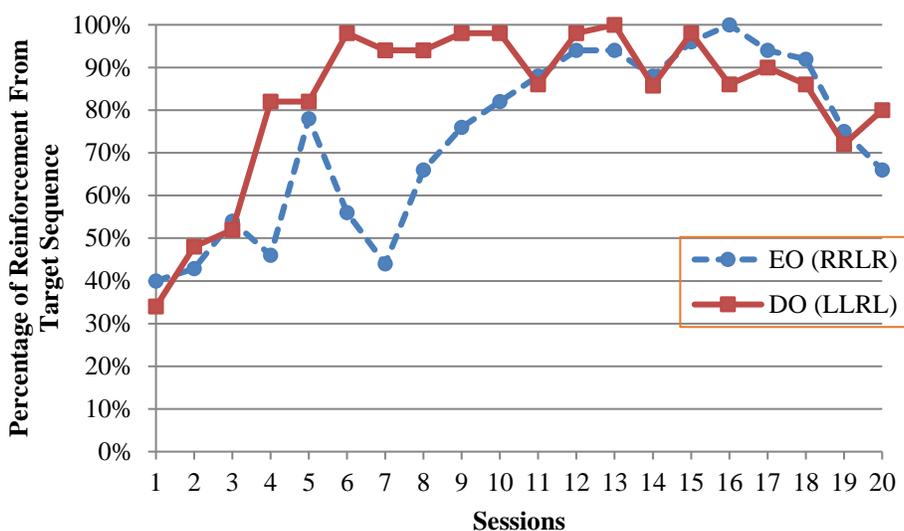


Figure 3. The percentage of reinforcement Bird 3 acquired from the target sequence during EO condition when RRLR was the target sequence and DO condition when LLRL was the target sequence. Percentage was calculated by dividing the total reinforcement gained by emitting the target sequence divided by the total amount of reinforcement gained during a session.

Bird 4

Figure 4 illustrates the percentage of reinforcement Bird 4 acquired from the target sequence during EO condition when LLRL was the target sequence and DO condition when RRLR was the target sequence. Bird 4 met criterion for learning the target sequence during both EO and DO conditions within 20 sessions. Bird 4 began the experiment by emitting the target sequence at the same percentage during both conditions (i.e.46%). Bird 4 showed a similar positive trend in learning during both the EO and DO conditions, but more variability and a decline in percentage of target emitted during sessions 12-14 of the DO condition. Bird 4 met criterion for learning

the target sequence during the DO condition one session sooner than during the EO condition.

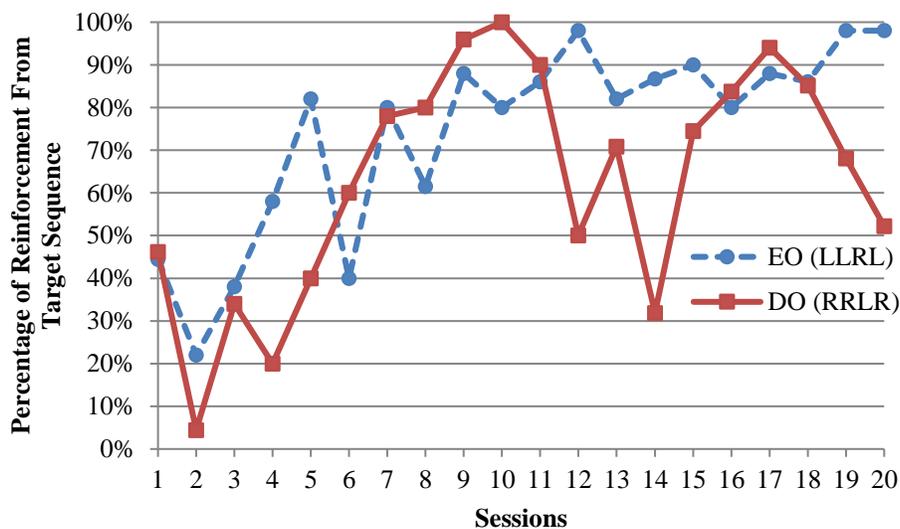


Figure 4. The percentage of reinforcement Bird 4 acquired from the target sequence during EO condition when LLRL was the target sequence and DO condition when RRLR was the target sequence. Percentage was calculated by dividing the total reinforcement gained by emitting the target sequence divided by the total amount of reinforcement gained during a session.

CHAPTER IV

DISCUSSION

It was hypothesized that pigeons would learn a target sequence sooner when reinforcement was provided using differential outcomes versus equal outcomes between target sequences and variable sequences. Three of four birds in this experiment provided support for this hypothesis. It was also hypothesized that pigeons would learn a target sequence more consistently, as evidenced by visual analysis of data, when target and variable responses were reinforced using differential outcomes compared to when they were reinforced using equal outcomes. Three of four birds provided support for this hypothesis.

Bird 1 learned the target sequence more quickly and steadily when differential outcomes of reinforcement were provided in comparison to when they were not. Bird 1 met criterion for learning the target sequence during the DO condition four sessions sooner than during the EO condition (i.e. by session 14 vs. 18). During sessions 13-15 of the EO condition, Bird 1 appeared to also show a decline during learning. Bird 1 showed a greater positive trend in learning during the DO condition in comparison to the EO condition providing evidence that the addition of differential outcomes helped facilitate learning the target sequence.

Bird 2 did not meet criterion for learning the target sequence in either the EO or the DO condition, but did show a greater trend in learning during the DO condition in comparison to EO condition. Although Bird 2 showed a greater overall level of

learning during the EO condition, this was likely due to Bird 2 having a previous learning history in which reinforced sequences began with pecking the left key leading to Bird 2 preferring to begin each sequence with a left key peck. Bird 2 showed a positive upward trend in learning during the DO condition and an almost flat trend during the EO condition. This provides evidence that Bird 2 was aided in learning the target sequence due to the addition of differential outcomes of reinforcement.

Bird 3 learned the target sequence more quickly and steadily during the DO condition in comparison to the EO condition. Bird 3 met criterion for learning the target sequence during the DO condition six sessions sooner than during the EO condition (i.e. by session 6 vs. 12). During the first few sessions of both the EO and DO conditions, bird 3 appeared to be learning the target sequence at similar rates, but showed a decline in learning during sessions 6 and 7 of the EO condition, and a steep positive trend in learning during the first six sessions of the DO condition. Bird 3 also showed less variability in learning during the DO condition indicating learning was smoother and more consistent with the addition of differential outcomes.

Bird 4 learned the target sequence during both the EO and DO conditions. Bird 4 met criterion for learning the target sequence during the DO condition one session sooner than during the EO condition (i.e. by session 10 vs. 11). A decline in target percentage during sessions 11-14 and sessions 18-20 of the DO condition was evident that directly correlates with Bird 4 showing a decline in responding during those same sessions. Bird 4 was probably satiated as evidenced by an increase in

documented weight during those sessions and the documented notation of Bird 4 stopping pecking early during those sessions. Bird 4 only responded throughout the entire session during the EO condition every other day. After sessions were changed from daily to every other day, Bird 4 began responding again during both conditions. Bird 4 showed similar trends in learning both the EO and DO conditions, but if the data during sessions when Bird 4 stopped responding during the DO condition are excluded, Bird 4 data showed a greater trend in learning during the DO condition than the EO condition.

Results showing the addition of a differential outcomes arrangement enhances learning are consistent with previous research using MTS and two-choice successive discrimination tasks (Goeters et al., 1992), but also show the ability of differential outcomes to be used in different learning scenarios. This research shows that differential outcomes can also be used between sessions rather than only within sessions as previous research has done. Previous differential outcomes research is limited to use of differential outcomes within sessions as is the traditional arrangement for MTS and two-choice successive discrimination tasks.

This experiment demonstrates the importance of both responding in variable or creative ways and in repetitive or precise ways. Variation is necessary for new behaviors to occur as in shaping a new behavior. In this experiment, the concurrent reinforcement of variable responding (i.e. VAR contingency) aided birds in learning a target sequence by preventing extinction and providing the necessary variation of key pecks to figure out or learn the target sequence. Stereotypy or responding precisely is

also desirable in various situations. A basketball player benefits by responding in a precise way when that behavior results in succeeding in making a free-throw shot. In this experiment, reinforcement of the target sequence was also important and using differential outcomes facilitated learning the target both sooner and more consistently.

In addition to reinforcement of variable responding, differential outcomes aided pigeons in learning target sequences. This makes sense because in order to figure out a difficult sequence, different or variable responses are required. Once the target sequence is emitted, differential reinforcement likely aids by providing additional information as to which schedule is being reinforced.

Limitations and Possible Confounds

There were some limitations and confounds in this study that need to be mentioned. Birds used in this experiment had previous learning histories in which they learned similar-type sequences. These learning histories may have led some birds to prefer beginning to peck either the left or right key or prefer a particular sequence of key pecks. Future research may benefit from using naïve birds as experimental subjects.

The design of the differential outcomes arrangement could have been a confounding variable in this experiment. One part of the differential outcomes arrangement included a difference in amount of food (i.e. VAR contingency met = 3s of access to food and CRF contingency met = 6s of access to food). The amount of food, rather than differential outcomes, could have been the factor influential in faster

and more consistent learning. In order to ensure that the differential outcome is the factor that established experimental control, future research should consider using counterbalancing. One suggestion would be to counterbalance birds to experience either less or more amounts of food during the differential outcomes condition. For example, two birds would receive 3s access to food for meeting VAR contingency and 6s access to food for meeting the CRF contingency, and the other two birds would experience the opposite (i.e. VAR = 6s access to food and CRF = 3s access to food). Another alternative could be that each bird experiences three conditions (i.e. one EO condition and two DO conditions) with two DO arrangements counterbalanced in the same way as mentioned in the first suggestion.

The alternating treatment design in which a bird alternated days learning different target sequences may have made it more difficult for some birds to learn each day's target sequence. Some birds showed greater variability in responding which might be the result of the alternating method in which conditions were only identified by the key colors. The alternating day's reinforced target may have influenced the birds learning for the current day. Future research testing differential outcomes may want to consider using a different treatment design where each target sequence is trained to criterion before moving on to the next target sequence. Another alternative may be to use a larger subject pool using group design in which each group of birds only learns one target.

Another possible issue was regarding each bird's difference in amount of food required to become satiated. Some birds simply became satiated and stopped

responding during the experiment when they began to get more food reinforcement after learning the target sequence better. Due to birds becoming satiated and stopping any responding, sessions were run every other day rather than daily halfway through this experiment. Future experiments may want to consider beginning the experiment by running sessions every other day rather than daily to prevent bird satiation that leads to no responses during sessions. It also may be helpful to collect preliminary data regarding each bird's difference in appetite or rate of food satiation prior to planning the session timeframes for the experiment. When using birds that become satiated sooner, amount of time to food reinforcement or maximum number of available reinforcement during each session may need to be lowered.

The present study extended previous research by applying differential outcomes to the task of learning a target sequence. To this researcher's knowledge, the DOE has not yet been examined in published literature using this type of learning procedure. Overall, the results from 3 of 4 birds indicated using differential outcomes does facilitate learning as previous differential outcomes research has previously shown. This research adds to the literature by demonstrating that differential effects can be used to facilitate learning outside of MTS and two-choice successive discrimination tasks.

REFERENCES

REFERENCES

- Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology*, *42*, 273-281. doi:10.1037/h0060407
- Bennett, A. T., & Cuthill, I. C. (1994). Ultraviolet vision in birds: What is its function? *Vision Research*, *34*, 1471-1478. doi:10.1016/0042-6989(94)90149-X
- Blough, D. S. (1966). The reinforcement of least frequent interresponse times. *Journal of the Experimental Analysis of Behavior*, *9*, 581-591. doi:10.1901/jeab.1966.9-581
- Boren, J. J., Moerschbaecher, J. M., & Whyte, A. A. (1978). Variability of response location on fixed-ratio and fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *30*, 63-67. doi:10.1901/jeab.1978.30-63
- Carlson, J. G., & Wielkiewicz, R. M. (1972). Delay of reinforcement in instrumental discrimination learning of rats. *Journal of Comparative and Physiological Psychology*, *81*, 365-370. doi:10.1037/h0033531
- Delius, J. D. (1968). Color preference shift in hungry and thirsty pigeons. *Psychonomic Science*, *13*, 273-274. doi:10.3758/BF03342520

- Delong, R. E., & Wasserman, E. A. (1981). Effects of differential reinforcement expectancies on successive matching-to-sample performances in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 394-412.
- Eckerman, D. A., & Larson, R. N. (1969). Variability of response location for pigeons responding under continuous reinforcement, intermittent reinforcement and extinction. *Journal of Experimental Analysis of Behavior*, 12, 73-80. doi:10.1901%2Fjeab.1969.12-73
- Fedorchak, P. M. & Bolles, R. C. (1986). Differential outcome effect using a biologically neutral outcome difference. *Journal of Experimental Psychology: Animal Behavior Processes*, 12, 125-130. doi:10.1037/0097-7403.12.2.125
- Friedman, G. J., & Carlson, J. G. (1973). Effects of a stimulus correlated with positive reinforcement upon discrimination learning. *Journal of Experimental Psychology*, 97, 281-286. doi:10.1037/h0034123
- Goeters, S., Blakely, E., & Poling, A. (1992). The differential outcomes effect. *Psychological Record*, 4, 389-424.
- Grunow, E. M., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review*, 9, 250-258. doi:10.3758/BF03196279
- Honig, W. K., Matheson, W. R., & Dodd, P. W. D. (1984). Outcome expectancies as mediators for discriminative responding. *Canadian Journal of Psychology*, 38, 196-217. doi:10.1037/h0080827

- Kruse, J. M., & Overmier, J. B. (1982). Anticipation of reward omission as a cue for choice behavior. *Learning and Motivation, 13*, 505-525. doi:10.1016/0023-9690(82)90007-8
- Lachter, G. D., & Corey, J. R. (1982). Variability of the duration of an operant. *Behavioral Analysis Letters, 2*, 97-102.
- Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 52*, 155-166. doi:10.1901/jeab.1989.52-155
- Machado, A. (1994). Polymorphic response patterns under frequency-dependent selection. *Animal Learning & Behavior, 22*, 53-71. doi:10.3758/BF03199956
- Maes, J. H. R., & van der Goot, M. (2006). Human operant learning under concurrent reinforcement of response variability. *Learning and Motivation, 37*, 79-92. doi:10.1016/j.lmot.2005.03.003
- Miller, O. T., Waugh, K. M., & Chambers, K. (2002). Differential outcomes effect: Increased accuracy in adults learning kanji with stimulus specific rewards. *The Psychological Record, 52*, 315-324.
- Mok, L. W., & Overmier, J. B. (2007). The differential outcomes effect in normal human adults using a concurrent-task within-subjects design and sensory outcomes. *The Psychological Record, 57*, 187-200.
- Myerson, J., & Hale, S. (1988). Choice in transition: A comparison of melioration and the kinetic model. *Journal of the Experimental Analysis of Behavior, 49*, 291-302. doi:10.1901/jeab.1988.49-291

- Neuringer, A. (1991). Operant variability and repetition as functions of interresponse time. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 3-12. doi:10.1037/0097-7403.17.1.3
- Neuringer, A. (1992). Choosing to vary and repeat. *Psychological Science*, *3*, 246-250. doi:10.1111/j.1467-9280.1992.tb00037.x
- Neuringer, A. (2009). Operant variability and the power of reinforcement. *The Behavior Analyst Today*, *10*, 319-343. doi:10.1037/h0100673
- Neuringer, A., Deiss, C., & Olsen, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 98-111. doi:10.1037/0097-7403.26.1.98
- Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes*, *27*, 79-94. doi:10.1037//0097-403.27.79
- Overmier, J. B., Bull, J. A., & Trapold, M. A. (1971). Discriminative cue properties of different fears and their role in response selection in dogs. *Journal of Comparative and Physiological Psychology*, *76*, 478-482. doi:10.1037/h0031403
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 429-452. doi:10.1037/0097-7403.11.3.429

- Papini, M. R. & Silingardi, C. A. (1989). Mechanisms of acquisition of behavioral sequences in rats. *Learning and Motivation*, 20, 73-86. doi:10.1016/0023-9690(89)90033-7
- Pierce, W. D., & Cheney, C. D. (2008). Shaping: The method of successive approximation. (4th Ed.), *Behavior Analysis and Learning* (pp. 77-78). New York, NY: Psychology Press.
- Pryor, K. W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, 12, 653-661. doi:10.1901/jeab.1969.12-653
- Sahgal, A., & Iversen, S. D. (1975). Colour preferences in the pigeon: A behavioural and psychopharmacological study. *Psychopharmacologia (Berl.)*, 43, 175-179. doi:10.1007/BF00421021
- Schwartz, B. (1981). Reinforcement creates behavioral units. *Behaviour Analysis Letters*, 1, 33- 41.
- Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Macmillan.
- Skinner, B. F. (1981). Selection by consequences. *Science*, 213, 501-504.
- Straka, J. A. (1966). Color preference in the pigeon, utilizing drinking as the response measure. *The Psychological Record*, 16, 203-205.

- Tatham, T., Wanchisen, B., & Hinline, P. (1993). Effects of fixed and variable ratios on human behavioral variability. *Journal of the Experimental Analysis of Behavior, 59*, 349-359. doi:10.1901/jeab.1993.59-349
- Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation, 1*, 129-140. doi:10.1016/0023-9690(70)90079-2
- Urcuioli, P. J. (1990a). Differential outcomes and many-to-one matching: Effects of correlation with correct choice. *Animal Learning & Behavior, 18*, 410-422. doi:10.3758/BF03205323
- Urcuioli, P. J. (1990b). Some relationships between outcome expectancies and sample stimuli in pigeons delayed matching. *Animal Learning & Behavior, 18*, 302-314. doi:10.3758/BF03205290
- Urcuioli, P. J. (1991). Retardation and facilitation of matching acquisition by differential outcomes. *Animal Learning & Behavior, 19*, 29-36. doi:10.3758/BF03197857
- Urcuioli, P. J. (2005). Behavioral and associative effects of differential outcomes in discrimination learning. *Learning & Behavior, 33*, 1-21. doi:10.3758/BF03196047
- Urcuioli, P. J. & Zentall, T. R. (1990). On the role of trail outcomes in delayed discriminations. *Animal Learning & Behavior, 18*, 141-150.
- Williams, D. A., Butler, M. M., & Overmier, J. B. (1990). Expectancies of reinforce location and quality as cues for a conditional discrimination in pigeons.

Journal of Experimental Psychology: Animal Behavior Processes, 16, 3-13.

doi:10.1037/0097-7403.16.1.3