THE UNIVERSITY OF MANITOBA

BEHAVIORAL REQUIREMENT AS A DETERMINANT OF PREFERENCE FOR DELAYED VERSUS IMMEDIATE REINFORCEMENT

BY

RON COLES

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF PSYCHOLOGY

OCTOBER 1976

WINNIPEG, MANITOBA

"BEHAVIORAL REQUIREMENT AS A DETERMINANT OF PREFERENCE FOR DELAYED VERSUS IMMEDIATE REINFORCEMENT"

by

RON COLES

A dissertation submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF ARTS

ABSTRACT

Pigeons' preference for small immediate reinforcement versus larger but more delayed reinforcement was tested under two main conditions. In the first assessment, Rachlin and Green's (1972) model of self-control was used. This model incorporates the phenomenon of commitment which is defined as the relative number of responses emitted to a key that provides access only to large delayed reinforcement. Commitment is said to be controlled by temporal variables. Specifically, the probability of commitment should increase as the time between the response and the presentation of reinforcement increases.

The second setting in which preference was tested involved a modified version of Rachlin and Green's (1972) procedure. A schedule requirement was inserted into the procedure subsequent to the emission of the response which defined commitment or noncommitment. (Noncommitment was defined as response which occurred to a key leading to a choice between small immediate and large delayed reinforcement.)

Several problems arose during the first assessment of preference. These involved an inability to maintain the behavior of some of the birds at the larger temporal values of the independent variable, and an inability to shift preference away from small immediate reinforcement for some subjects at the original values of the long delay. Manipulating

the value of these variables resulted in a resolution of both problems. The results of this portion of the study indicate that preference for large delayed reinforcement increased in a manner consistent with Rachlin and Green's (1972) model, however the effect was more evident in some subjects than it was in others.

The schedule requirement was introduced following completion of the first assessment. Except for the addition of a fixed interval (FI) schedule, all other aspects of the procedure remained constant between the last phase of the first test situation and the first phase of the second testing situation. This allowed direct comparison of relative commitment across testing procedures. The results show that addition of the schedule requirement changed preference away from large delayed reinforcement. This effect was strong in two subjects and minimally present in a third subject. A floor effect prevented assessment of this variable in the fourth subject.

The results of the study are discussed in terms of two possible interpretations. The first interpretation involves the notion that the frequency of choosing small immediate reinforcement is increased, relative to choosing large delayed reinforcement, because the immediacy of the small immediate reinforcer provides a more powerful reinforcing effect than is possible under the delayed contingency. This position, however, is not consistent with previous work

on FI schedules. A second interpretation of the data is that they reflect a tendency of pigeons to prefer the shorter of two FI schedules. Although no rationale can be given to explain this preference, the phenomenon is consistent with previous literature.

It is concluded that in the absence of a specific behavioral requirement leading to reinforcement, Rachlin and Green's (1972) model of self-control is valid. However it is also concluded that the addition of a specific behavioral requirement is sufficient to change preference levels, and therefore any complete account of the process of self-control must incorporate the effect due to behavioral requirement as well as the effect due to temporal variables.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my committee members, Dr. J. J. Pear, Dr. G. R. Norton, and Dr. F. A. Starke; to my advisor, Dr. R. Steven Heaps; and to Barb and Robbie. Each has provided a unique and valuable contribution to this thesis.

I would also like to thank the Department of Psychology, University of Winnipeg, for the use of their facilities.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Over twenty years ago Skinner (1953) provided one of the first, widely publicized, behaviorally oriented examinations of the concept of self-control. He began with the premise that man controls his own behavior in the same manner he controls the behavior of others - through the manipulation of variables which affect that behavior.

Consistent with this interpretation, Skinner (1953) maintained that occasions for the necessity of self-control arise only when behavior results in conflicting contingencies. The term conflicting contingencies refers to the fact that the behavior has both positive and negative consequences. These consequences, in turn, generate what Skinner termed the controlling response and the controlled response. The relationship between the two responses is described as follows: "one response, the controlling response, affects variables in such a way as to change the probability of the other, the controlled response" (Skinner, 1953, p.231). It is quite obvious that the relative probabilities of either response is a fundamental issue. Therefore to provide an effective technology of self-control it is necessary to identify the variables that affect the responses in question.

Speaking somewhat non-technically, many people have

pointed out that the relative probability of occurrence for either the controlling response or the controlled response is often determined by preference. Although the term preference seems to have been reified, close analysis has shown it to be a function of other controlling variables. Aside from any subjective influence that may be a result of some peculiar reinforcement history, preference among various positive consequences of behavior can be controlled by manipulation of amount, delay and rate of reinforcement (Chung & Herrnstein, 1967; Herrnstein, 1964a; Logan, 1965).

In terms of empirical support, the present literature is fairly consistent with respect to the three variables just mentioned. Choice or preference and rate of reinforcement has been extensively investigated. Some early work in this area was carried out by Autor (1969) and Herrnstein (1964a). Both conducted investigations into the strength of conditioned reinforcement using an experimental procedure known as a concurrent chains technique. In this procedure the S (in both cases pigeons were used) responds to one of two simultaneously available keys. Pecking is maintained by intermittent presentation of a second key colour which has been paired with a certain schedule of primary reinforcement. When different colours were presented on alternate keys in the terminal link, with each colour having been paired with a different schedule of reinforcement, both researchers found that the relative rate of responding in the initial

link (a common index of preference) was related to the frequency with which primary reinforcement occurred in the terminal link. In other words, relative rate of responding in the initial link matched the relative rate of reinforcement in the terminal link. Here the term matched refers to a proportional relationship between a measure of the independent variable and the dependent variable. Several mathematical formulas have been generated to describe this function (Baum & Rachlin, 1969; Herrnstein, 1970; Rachlin, 1971).

It should be noted that both Autor (1969) and Herrnstein (1964a) also measured strength of conditioned reinforcement in terms of probability of reinforcement (i.e., the number of responses divided by the number of reinforcements). Using a variable ratio (VR) schedule in the terminal link, Autor (1969) showed that relative rate of responding in the initial link matched relative probability of reinforcement in the terminal link. Although matching in terms of probability of reinforcement was obtained in both Autor's (1969) work and in Herrnstein's (1964a) work, Herrnstein concluded that frequency of reinforcement was a more appropriate measure because its regression line gave a closer fit to the line depicting matching than did probability of reinforcement.

Logan (1965) used a discrete trials procedure with a T maze to study the effects of amount and delay of reinforcement upon choice behavior in rats. He conducted six massed

trials per day with trials one to four being forced choice, trial five being free choice, and trial six being forced in the opposite direction of trial five. The relative frequency of going left or right on trial five was the measure of choice. Logan showed that both amount and delay exerted an effect upon choice, with relative values of some combinations of the independent variables yielding a preference for one alternative, other relative values resulting in a choice of the other alternative, and thirdly, some relative values yielding indifference or no preference.

Rachlin and Baum (1969) manipulated amount of reinforcement (seconds access to grain hopper) on one key of a two key concurrent situation. Variable-interval (VI) schedules were programmed independently, with reinforcement on one key being signalled by the presentation of a red light. At all other times the key was dark and inoperative. When the number of seconds access to the hopper was allowed to vary on the signalled key, while remaining constant on the standard key, the results showed that relative rate of responding on the unsignalled key varied inversely with the duration of reinforcement on the signalled key. Rate of responding on the signalled key remained constant throughout the study and in general was approximately equal to the number of reinforcers dispensed in a session (20 responses per hour).

There are several studies which have dealt solely

with delay of reinforcement. Chung (1965a) and Chung and Herrnstein (1967) have shown that in a two key concurrent paradigm, relative rate of responding bore an inverse relationship to relative delay of reinforcement. This may be restated as follows: relative rate of responding matched relative immediacy of reinforcement, where immediacy is defined as the reciprocal of delay.

Although not denying the importance of delay of reinforcement as a variable which influences behavior, Neuringer and Schneider (1968) and Neuringer (1969b) questioned the relative value of certain factors that could be involved in that type of delay.

Neuringer and Schneider (1968) maintained that behavior may be influenced by the time between reinforcements, the number of inter-reinforcement responses, or a combination of both. The procedure to test which of the three possibilities was the important feature involved placing half their Ss on a fixed-interval (FI) schedule and half on a fixed-ratio (FR) schedule. By programming a brief black-out period after each response, the authors were able to vary the inter-reinforcement interval and the number of responses independently. Due to the fact that post-black-out and post-reinforcement pausing increased as black-out duration increased on the FR schedule but not on the FI schedule, the authors concluded that "the time between reinforcement controls responding independently of the number of responses

emitted during that time" (Neuringer & Schneider, 1968, p. 666).

Neuringer (1969b) stated that past research raised the possibility that indifference may result from procedures in which delay of reinforcement would be temporally equated with an FI response requirement. The experiment (Neuringer, 1969b) involved a concurrent chains technique in which the terminal link was either an FI schedule or a black-out condition of the same time length. With terminal links equated temporally, Neuringer's data show only a slight preference for the FI schedule (less than 5%). This preference was later shown to be due to the black-out. Taking the bias due to black-out into account, it was concluded indifference resulted.

In a second portion of the study Neuringer (1969b) systematically varied the lengths of the terminal link components. His findings were that preference for a "standard" key increased as the length of the component of a "comparison" key went from shorter to longer, regardless of whether the standard was the FI component or the delay component. This conclusion was similar to Neuringer and Schneider (1968) in that it was felt that the time (delay) between a response and reinforcement for that response controls the probability of that response regardless of whether other responses intervene.

These findings, especially that of Neuringer (1969b),

have implications for preference procedures, for if the response to which Neuringer refers is a choice response, then a logical deduction is that choice is independent of subsequent behavior. Some evidence in support of this position is found in the writings of Rachlin and Green (1972) and Ainslie (1974). Both studies are direct tests of self-control paradigms, and as such are highly relevant to the current issue.

Rachlin and Green (1972, p. 16) take as a basic premise that preferences are ". . . governed by the reward actually obtained and are independent of behavior per se between the choice and the reward". Here the term reward refers to a value that is the product of a given delay and a given magnitude of reinforcement.

Their experimental procedure was designed to analyze commitment as a form of self-control. In relation to self-control, commitment refers to the emission of a response which forces the organism to bypass one reward in favor of obtaining another. For example, placing one's alarm clock on the other side of the room is a commitment response which forces the individual to get up. In the morning the possibility of choosing has been forfeited by the prior commitment.

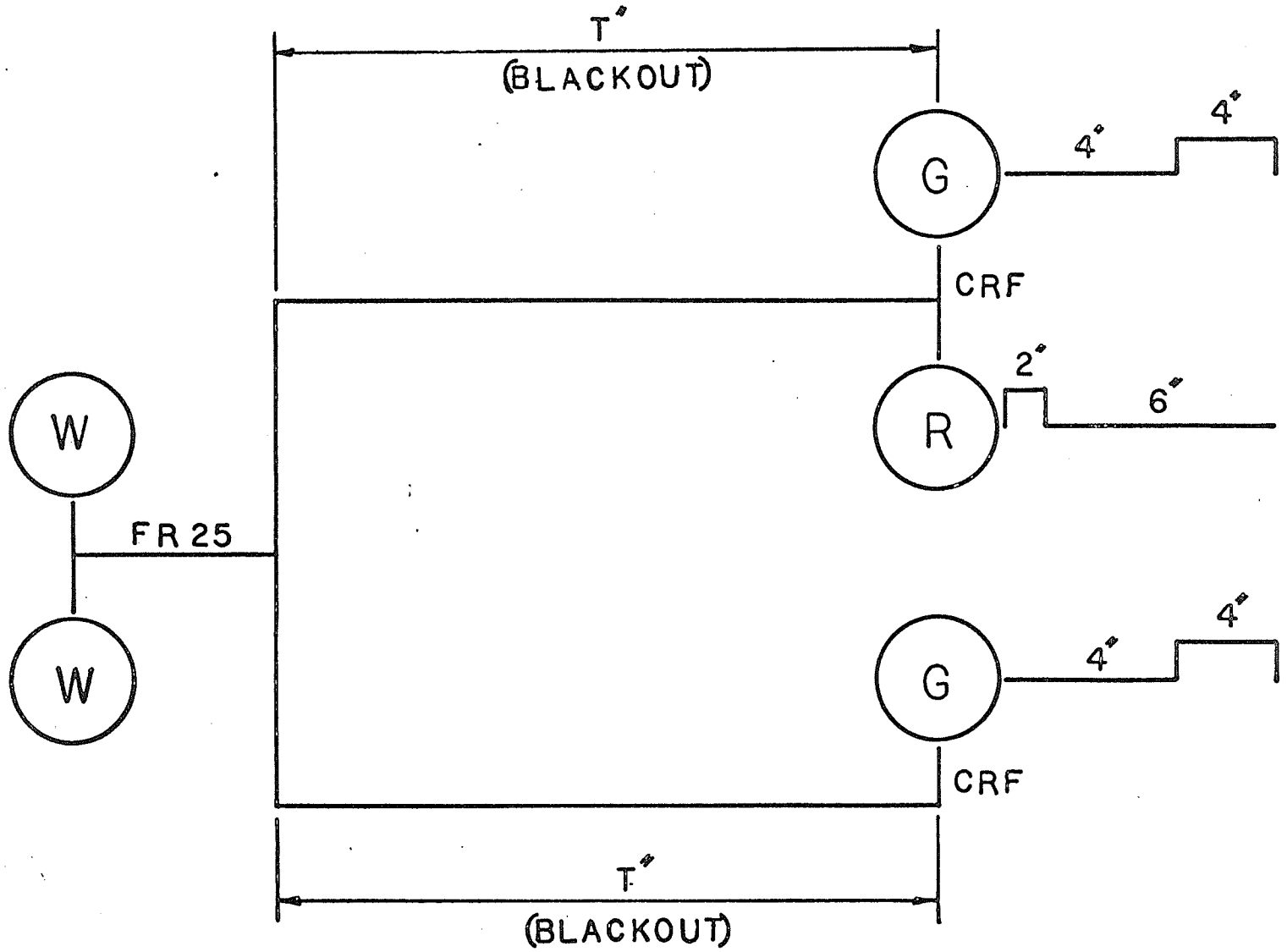
The experimental setting involved a modified concurrent chains procedure. The initial link involved two keys illuminated with a white light. The initial link could be completed by an FR 25, distributed in any manner across

the two keys. The location of the 25th peck determined the nature of the second link. A peck to the right hand key resulted in a black-out period of t seconds, which terminated with the appearance of a red and a green coloured key. One key peck to the red key resulted in immediate access to two seconds of reinforcement followed by six seconds of black-out. One peck to the green key resulted in an additional four second black-out and then four seconds of reinforcement. Termination of the initial link FR 25 on the left hand key also resulted in a t second black-out. However, here the period defined by t was followed by the appearance of the green key alone which had the same schedule consequence as previously described. The overall procedure is depicted in Figure 1.

Insert Figure 1 about here

Rachlin and Green (1972) hypothesized and subsequently confirmed that at low values of t there would be a preference for immediate reward; however as t increased in value, preference would shift to the larger but more delayed reward. Thus in the initial link, the left hand key would be preferred since it provided the only viable means of obtaining this reward. (The left key is designated as the only means of gaining access to the larger reward because of the fact that given a choice between immediate reinforcement and delayed reinforcement, the Ss virtually always chose immediate reinforce-

Figure 1. Schematic presentation of the procedure used by Rachlin and Green (1972). In the initial link, both keys are white. A black-out period of t seconds separates the initial link from the terminal link. In the terminal link, the red key is paired with the small immediate reinforcer, while the green key is paired with large delayed reinforcement.


T"
(BLACKOUT)
G
4"
4"
CRF
2"
R
6"
W
FR 25
W
G
4"
4"
CRF
T"
(BLACKOUT)

ment).

Ainslie's (1974) study of self-control parallels Rachlin and Green's (1972) work in that both models indicate that preference is a function of delay of reinforcement. Indeed, Rachlin (1970) ties the experiments together when he discusses Ainslies' original work in terms of "gradients of delay of reinforcement".

Ainslie's (1974) procedure involved the presentation of two stimulus colours and a delay period in which the key was dark. On each trial a green light initially illuminated a single key for 7.5 seconds. One key peck during this period resulted in a delay period which lasted until the trial terminated with 4 seconds access to grain. If no peck occurred during green, the key went dark for 4.5 seconds (the key was also inoperative). This was followed by a 3 second presentation of a red key light. A key peck during red resulted in 2 seconds immediate access to reinforcement. If no pecks occurred to red, it was followed with 4 seconds reinforcement.

Ainslie's data show that some of his birds learned to peck the green key and thus received the larger amount of reinforcement. In a subsequent control procedure, Ainslie extinguished responding to the green stimulus by presenting red regardless of what the organism did. He also extinguished responding to red by making a 2 second blank in illumination the only consequence for responding.

Several other studies have addressed themselves to the issue of self-control in non-human subjects (Fantino, 1966; Neuringer, 1969a; Mahoney & Bandura, 1972; Lambert, Bersh, Hineline & Smith, 1973). The Fantino (1966) study and the Lambert, Bersh, Hineline and Smith (1973) study resemble Rachlin and Green's (1972) and Ainslie's (1974) work. The common feature of all four studies is the attempt to train a response that will provide maximum benefit in situations involving more than one source of reinforcement. The other studies (Neuringer, 1969a; Mahoney & Bandura, 1972) relate more to the notion of self-reinforcement.

Fantino (1966) was able to show that pigeons could withhold responding for brief periods in order to obtain greater reinforcement. In Fantino's study, responding in the presence of a red key light resulted in immediate reinforcement followed by a period of extinction. If the organism did not respond to the red key, it eventually turned green. The consequence of responding in green was the presentation of primary reinforcers in the absence of any extinction period. Following a training period of six months, some of the pigeons in this study were able to withhold responding for a period of three seconds in order to obtain the greater amount of reinforcement.

The Lambert et al. (1973) study is the negative version of Rachlin's (1974) self-control paradigm. In this study, rats had to emit a response, which was immediately

followed by shock, in order to prevent the delayed onset of five identical shocks spaced one second apart. A lever press and a shuttle response were tested in this study. The authors also tested subjects when escape from the delayed shocks was prevented and when it was possible. The data revealed that the rats were able to learn the punished response which provided a means of avoiding larger delayed aversive stimulation.

Neuringer (1969a) and Mahoney and Bandura (1973) were able to train pigeons to peck a key before allowing themselves to eat from a hopper that was continuously available during any given trial. Mahoney and Bandura (1973) defined the behavior as self-reinforcement because it met the criteria of involving the free availability of the reinforcer and completion of a "required" response prior to delivery of the reinforcer. Catania (1975), however, has pointed out that a behavioral sequence involving two responses, the second of which involves the delivery of the reinforcer, does not define self-reinforcement per se. He concludes that even though Mahoney and Bandura (1973) have demonstrated the operation of self-reinforcement, they failed to demonstrate the process of self-reinforcement. In other words, they did not demonstrate that the delivery of the reinforcer increased the probability of the response.

Burns and Powers (1975) are apparently the first researchers to have published a systematic replication (across

species) of Rachlin and Green's (1972) work. The authors used two male humans age 9 and 10 years as subjects. Tokens which were exchangeable for money were used as reinforcers. The procedure closely paralleled Rachlin and Green (1972) with some minor modifications. The children were run at t values of 0, 4, 8, 12, 16, 32 and 64 seconds. The data obtained in this experiment did show that preference reversed as t was manipulated, however the direction of the relationship was the opposite of Rachlin and Green's (1972) findings. For the children in the study, increases in t resulted in increasing preference for immediate reinforcement.

In many of the studies just cited (i.e., Ainslie, 1974; Burns & Powers, 1975; Fantino, 1966; Lambert et al., 1973; Rachlin & Green, 1972), it was assumed that time was the sole controlling variable regarding preference for delayed versus immediate reinforcement. This implies that any behavior involved will have a sum zero effect -- that is any behavior involved in the situation, neither adds nor detracts from the reinforcement value of the situation. While this position is consistent with Neuringer (1969b) and would seem to be supported, at least indirectly by Rachlin and Green (1972) and Ainslie (1974), the procedural shortcomings of these experiments make their position hard to defend. None of these experiments (Ainslie, 1974; Neuringer, 1969b; Rachlin & Green, 1972) utilized response rate or response requirement as an independent or dependent variable and