AN EXPERIMENTAL STUDY OF THE EFFECT OF POSITION REVERSAL AFTER ONE OR TWO REINFORCEMENTS ON SIMPLE T-MAZE LEARNING IN THE RAT

Thesis for the Degree of M. A.

MICHIGAN STATE COLLEGE

Dorothy Nell Moore

1949

This is to certify that the

thesis entitled

An Experimental Study of the Effect
of Position Reversal After One or
Two Reinforcements on Simple T-Maze
Learning in the Rat

presented by

Dorothy Nell Moore

has been accepted towards fulfillment
of the requirements for

M.A. degree in Psychology

M. Ray Denny
Major professor

Date August 25, 1949

AN EXPERIMENTAL STUDY OF THE EFFECT OF POSITION REVERSAL AFTER ONE OR TWO REINFORCEMENTS ON SIMPLE T-MAZE LEARNING IN THE RAT

by

Dorothy Nell Moore

A THESIS
Submitted to the Graduate School of Michigan
State College of Agriculture and Applied
Science in partial fulfilment of the
requirements for the degree of

MASTER OF ARTS

Department of Psychology

1949

THESIS

ACKNOWLEDGMENT

Grateful acknowledgement is made to Dr. M. Ray Denny for his willing guidance and encouragement, and for his ever-present assistance in the study reported here.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

The complexity of adaptive phenomena has long offered a challenge to the theoretical psychologist in the field of learning. Although the concept of reinforcement[1] occupies a central position in learning theory, psychologists are in considerable disagreement as to the nature of reinforcement. The attempt to give a single description of the reinforcement process - one which will embrace the many diverse learning situations - has initiated much experimental research and much theorizing.

From these attempts it appears that the crucial question is "What is the critical factor for the strengthening of an adaptive response?", or, simply, "What is the essential condition for learning?"

Of the many theories which have been designed to answer this question, Thorndike's (13) 'law of effect' and Hull's (7) principle of reinforcement have probably wielded greatest influence. For convenience, these two theories may be roughly classed together as 'drive satisfaction' theories, as they both designate reward as the essential condition for learning.

In opposition to the drive satisfaction or drive reduction theories, is Tolman's (14) principle of expectancy. In his system, reward or drive satisfaction does not strengthen a response

---

1 'Reinforcement', as here used, is defined as a state of affairs which incrementally strengthens a response.

tendency, but serves to keep the organism motivated or in a state of expectancy for a particular goal object or behavior consequence. Tolman's theory and others which generally adhere to this principle of anticipation or expectancy may be classed as 'expectancy theories'.

Some effort has been made to integrate these two viewpoints into a single theory. These will be examined in the following section.

II. THEORETICAL BACKGROUND

Of the current theories concerning the nature of reinforcement, those of Thorndike, Hull, Tolman, Mowrer, White and Denny are of special concern in the present study.[1]

A. Edward L. Thorndike - Law of Effect

One of the most influential principles describing the nature of reinforcement has been the 'law of effect' proposed by Thorndike (13). Essentially Thorndike holds that the satisfying outcome of a response tends automatically to strengthen the association between the stimulus and the response. Thorndike (13) presents his law as follows:

> When a modifiable connection between a situation and a response is made and is accompanied or followed by a satisfying state of affairs, that connection's strength is increased.....
>
> By a satisfying state of affairs is meant roughly one which the animal does nothing to avoid, often doing such things as attain and preserve it. (Thorndike, 13, p. 176)

Thus for Thorndike reward becomes the essential condition for learning.

Hilgard and Marquis (4) point out that the law of effect is somewhat of a misnomer in that it does not require that the behavior sequences strengthened by reward should necessarily be instrumental in securing the reward. The effective factor in determining the

---

1 The attempt is made to give an outline of only the concepts and principles contained in these theories which are relevant to this study; the pertinent experimental data is too extensive for reporting here.

selection of the correct response is its proximity in time to the reinforcement, i.e., the last response which occurs prior to the reward is the one most strongly reinforced.

There is no implication of purposive behavior or of insight contained in Thorndike's formulations. Reward or goal satisfaction acts directly on neighboring connections to strengthen them, without mediation by ideas or consciousness on the part of the organism.

Originally Thorndike held that if the stimulus response connections were followed by an annoying state of affairs or punishment, the connection would be weakened. However, because of experimental evidence which contradicts this phase of the theory, his most recent formulation of the law of effect omits the consequences of punishment.

B. Clark L. Hull - Principle of Reinforcement

Hull's (5, 6, 7) theoretical interpretation of learning is the most systematic of the current theories. This view of primary reinforcement, although more quantitative and particularized, is basically the same as Thorndike's law of effect. In both the reward acts directly and mechanically on cue-response connections. Thorndike (12) defines reward or reinforcement in terms of the satisfying consequences of a response, while Hull (7) thinks of reinforcement as drive-reduction or the decrement in a physiological need.

In his book, <u>Principles of Behavior</u>, Hull (7) states that when a condition of need exists, random and variable behavior is evoked, and the following chain of events could result:

> In case one of these random responses, or a sequence of them, results in the reduction of a need dominant at the time, there follows as an indirect effect what is known as reinforcement (G). This consists in (1) a strengthening of the particular receptor-effector connections which originally mediated the reaction and (2) a tendency for all receptor discharges ($\dot{s}$) occurring at about the same time to acquire new connections with the effectors mediating the response in learning.......... As a result, when the same need again arises in this or in a similar situation, the stimuli will activate the same effectors more certainly, more promptly and more vigorously than on the first occasion. (Hull, 7, p. 386)

Hull also stresses the point that this increment in habit strength occurs only when the receptor and effector activities are in close temporal contiguity and are closely followed in time by a reinforcing state of affairs (drive-reduction). He states:

> Whenever a reaction (R) takes place in temporal contiguity with an afferent receptor impulse ($\dot{s}$) resulting from the impact upon the receptor of a stimulus energy (S), and this conjunction is followed closely by a diminution in the drive, D, and in the drive receptor discharge (s_D), there will result an increment, Δ ($\dot{s} \dashrightarrow R$), in the tendency for that stimulus on subsequent occasions to evoke that reaction. (Hull, 7, p. 71)

It is postulated that habit strength grows simply as an increasing exponential function of the number of reinforcements.

Hull (5, 6) also sets forth a concept involving fractional anticipatory responses (r_G), which, he maintains, become conditioned stimuli to adaptive behavior. These fractional anticipatory responses are small parts or fractions of the more complex goal response. For example, salivating, chewing, and swallowing, which are fractionations of the complete eating process and do not interfere with most overt motor responses, constitute what Hull means by fractional anticipatory response. Such responses with their accompanying stimulus components function as behavior-directing stimuli or provide the physical basis for purposive behavior.[1]

The role of fractional responses is elaborated by Hull (5) in the following way:

> The drive stimulus accounts very well for the random seeking reactions of a hungry organism, but alone it is not sufficient to produce the integration of complex behavior sequences such as is involved in maze learning. There must always be a reward of some kind. Once the reward has been given, however, the behavior undergoes a marked change most definitely characterized by evidences of actions anticipatory of the goal, which actions tend to appear as accompaniments to the sequence ordinarily leading to the full overt goal reaction.
>
> It is shown how these fractional anticipatory reactions could be drawn to the beginning of the behavior sequence and maintained throughout by the action of the drive stimulus (S_D). The kinaesthetic stimulus resulting from this persistent anticipatory action should furnish a second stimulus (s_G) which would persist very much like S_D. These two persisting stimuli alike should have the capacity of forming multiple excitatory tendencies to the evocation of every reaction within the sequence. (Hull, 5, p. 504)

---

1 It is important to note, however, that Hull does not posit any causal relationship between fractional anticipatory reactions or their completion (for instance, in the complete goal response of eating) and reinforcement, as such. Within the Hullian framework need-reduction remains as the only and essential principle of primary reinforcement.

C. Edward C. Tolman - Principle of Expectancy

Tolman's (14) sign-gestalt or expectancy theory offers an alternative to rewarded response learning. It postulates that the organism follows 'signs' which mark the 'behavior route' leading to the 'significate' or goal; a behavior route instead of a movement pattern is learned.

The interpretation emphasizes the perceptual or cognitive capacities of the animal. Rewarding states of affairs operate mainly to specify and maintain expectancies. Expectancies or cognitive maps, rather than specific motor responses, are learned primarily through simple contiguity.

Although Tolman does not provide a structural basis for expectations, he does not deny that there is one, and attempt has been made by some to tie the principle of expectancy to an objective basis (e.g., White (15), with his fractional anticipatory response analysis).

D. O. H. Mowrer - Two-Factor Theory[1]

The two-factor theory proposed by Mowrer (9) was primarily developed in order to account for the phenomenon of avoidance learning which, according to Mowrer, is not adequately explained by Hull's reinforcement theory. This interpretation divides learning into two types, that for (1) skeletal muscle responses and

---

1 A discussion of the implications which Mowrer's two-factor theory holds for psychotherapy is presented by Shoben (10), pp. 135-144.

(2) smooth muscle or autonomous responses. The former is accounted for by the principle of drive decrement, while vicero-motor activity is said to be learned through simple association or contiguity.

For example, anxiety or fear, which is the visceral aspect of pain, is conditioned to the cues associated with the onset of the drive stimulus (pain). The anxiety so established acts as a secondary drive which then allows the avoidant skeletal responses to be strengthened according to the principle of reinforcement as formulated by Hull (7).

Mowrer, in other words, has ignored parsimonious considerations and postulated two classes of learning which must be accounted for by two different principles.

E. R. K. White - Completion Hypothesis

An attempt has been made by White (15) to integrate Hull's theoretical constructs with those of Tolman and Thorndike by utilizing the concept of fractional anticipatory reactions. In the completion hypothesis, it is proposed that the completion of fractional anticipatory responses constitutes a reinforcement act.

White states:

> The fractional anticipatory reaction is an incipient response or an 'activity in progress' in the literal sense that it is a specific physical act which has been started and not completed. The goal situation makes possible the rounding-out of a coordinated activity-pattern, or the finishing of a complex act in the same manner in which it has been finished on previous occasions. If, then, completion means the transition from an incipient reaction to the complete

> reaction of which it was previously a part (in Hull's symbolism, the transition from r_G to R_G), our translation of the satisfaction hypothesis can be expressed as follows: The completion of a fractional anticipatory reaction tends to reinforce recent and concomitant S-R connections. (White, 15, p. 399)

In this manner, White, by proposing Hull's (5, 6) concept of fractional anticipatory response as the unifying principle, has encompassed Thorndike's (13) law of effect by interpreting the completion of fractional anticipatory responses as satisfying states of affairs, and incorporated the directing influence present in Tolman's (14) expectancy principle.

It is unfortunate, however, that White has not advanced his hypothesis beyond the most tentative stage. It would seem that with further development such a hypothesis might offer a promising approach to understanding the nature of reinforcement.

F. M. Ray Denny - Pertinent Response Hypothesis

An interpretation of reinforcement which is somewhat similar to White's (15), although independently developed from his, has been proposed by Denny (3) in a series of unpublished lectures.[1]

According to Denny's theoretical analysis, the so-called skeletal and autonomic types of learning can be subsumed under one principle of reinforcement. In the case of the classical or respondent conditioning of eye blink, pupilary reflex, knee jerk, leg withdrawal, galvanic skin response, etc. (not strictly in the autonomic

---

1 The writer is indebted to Dr. Denny for the use of his unpublished lecture material, from which this outline of the pertinent response hypothesis is directly derived, and for his cooperation in personally clarifying the principles which he proposes.

category, it should be added), the principle of contiguity seems to account satisfactorily for the establishment of the conditioned response. Also in the case of anxiety or fear it is presumably the presentation of shock or noxious stimulation, not its cessation, that sets up the secondary drive of fear. Yet not any pairing of stimuli or of response and stimuli will bring about learning.

Denny states:

> In instrumental learning or operant conditioning it is well known that drive satisfaction or reward must also be present. But what is this so called drive reduction? Is it actually different from jerking one's knee when the appropriate stimulus is given to the appropriate structure? Is there <u>actually</u> drive reduction when a rat gets a tiny pellet of food in a maze or a Skinner box situation? The organism is so structured originally to respond in a fairly consistent and specific way, say to a blow on the patellar tendon, and to respond grossly or emotionally to a noxious stimulus. When it eats a piece of food it also makes certain original responses such as chewing, salivating, and swallowing. When an organism gives an innate or reflexive response in a neutral stimulus situation, this particular stimulus situation acquires the property to evoke this response. That is, when an organism gives the responses it it supposed to make, or is so structured to make, to a prepotent stimulus the remaining stimuli in the context acquire the property to evoke this response. (Denny, 3)

It is postulated that the animal can make responses that it is supposed to make under two main sets of conditions, which may by no means be mutually exclusive: (1) <u>Permanent</u> - when its innate structure so dictates, (2) <u>Temporary</u> - when the momentary state of the organism, primarily its current response organization or set, presupposes the organism toward one type of response rather than another. In other words when an organism is set to eat food, (has fractional anticipatory responses in terms of making incipient and implicit eating responses) the appropriate or pertinent act for the animal is to eat the food in the goal box and it learns to do that faster with succeeding trials.

It also learns to make the more successful instrumental responses leading up to the eating of the food. It is proposed (1) that all responses occurring in immediate temporal contiguity with the consummatory or pertinent response are also being established and strengthened and (2) that any stimulus which tends to increase or confirm the anticipatory response acts to strengthen any concurrent response.[1]

Denny's reinforcement hypothesis is then as follows: If the organism makes the response it is permanently or temporarily structured to make, then that response and others very close to it in time become hooked up and fixated to the present stimuli. In instrumental learning this amounts to saying that the instrumental response in order to be learned, must occur concurrently with the fractional anticipatory responses.

Presumably the fractional anticipatory responses which constitute the organism's set acquire some habit strength to the

---

1 Unlike Tolman's (14) analysis, Denny's position accounts for the learning of appropriate action as well as the expectancy.

maze situation as soon as the first goal response is made; on each subsequent trial an increase in the anticipatory reaction occurs with the making of responses leading up to the goal response, and, in turn, responses instrumental in supplying to the goal object are strengthened. After the first trial subsequent responses are learned in essentially the same way as Hull (7) proposes that responses which lead to drive reduction are learned.[1]

---

1 In Hullian terms, all learning, according to Denny, takes place by means of secondary reinforcement.

III. STATEMENT OF THE PROBLEM

Explicit in the theory of reinforcement forwarded by Denny is the principle that the organism must be 'set' to make the responses which lead to reward or consummation before learning can take place. Thus in a completely new situation, responses removed to any degree in time from the goal response cannot receive an increment in habit strength until after the anticipatory set has been established.

According to this interpretation, it could be assumed that all that is learned on the first and perhaps second trial by the hungry rat which finds food in a new maze is an anticipatory response. Instrumental responses are not strengthened on these trials because there is as yet no anticipatory set to be consummated.

The hypothesis to be tested to support this theory is as follows: If, in a simple T-maze learning situation, the goal boxes are reversed in position from left to right, and vice versa after one and possibly two rewarded trials, we should expect no difference in the learning of these animals and animals trained consistently to one side.[1]

---

1 This position reversal technique is similar to that used by Spence (12) in a discrimination learning test of the continuity and non-continuity theories.

IV. EXPERIMENTAL PROCEDURE

A. Apparatus

The apparatus consisted of a simple T-maze, made up of a starting box, a combined stem and choice point section, a pair of arms, and two goal boxes. The ground plan of the maze is presented in Figure 1 (p. 15).

With the exception of the goal boxes the inside alley width was 5"; the height of all parts of the maze measured 11-1/2". The sides and floor of the maze proper were made of 3/4" plywood, and main sections of the maze were moveable. The interior of the maze was painted a uniform gray throughout.

The roof of the starting box was wood, the stem was covered with a fine screening which was difficult to see through. The choice point section, the two arms and the two goal boxes had a hardware cloth roof of 1/2" mesh.

Vertical sliding doors were placed at the entrance to each goal box, in the choice point section, and at the exit of the starting box. They were painted the same gray as the maze interior and were made of 1/4" plywood. An inverted T-shaped door was placed in the choice point section as shown in the ground plan (Figure 1). This door was so constructed to prevent the animal from retracing its steps once it had made a choice.

With the exception of the starting box door which was opened from two to five seconds after the animal was placed in the box, all doors in the maze were open at the beginning of each trial. The doors at

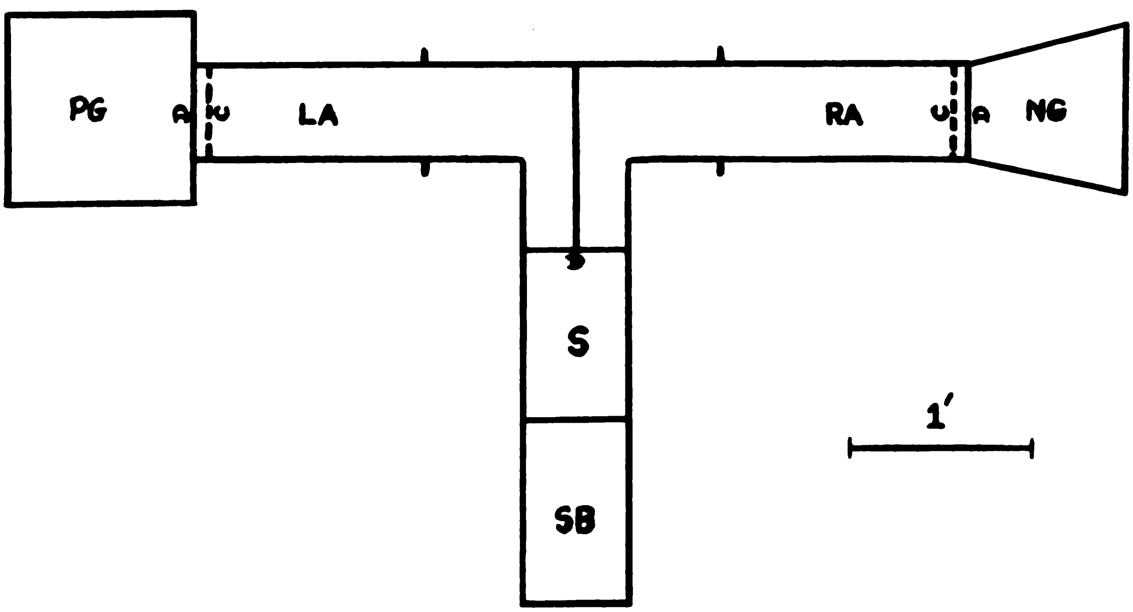

Figure 1. General plan of apparatus. SB--starting box; S--stem; LA--left arm; RA--right arm; PG--positive goal box; NG--negative goal box; D--doors; C--curtains

the entrance of the goal boxes were closed immediately after the animal entered. All moveable doors worked by a system of strings and weights. A block of wood of the same size and color as the other doors was inserted flush with the stem on forced trials.

One inch from the entrance to each goal box, curtains of black material were suspended from a cross bar so as to obscure the goal portions of the maze lying beyond.

The negative goal box was trapezoidal in shape, was painted white, and had a smooth sheet metal floor. The inner dimensions of this box were 5-1/4" at the entrance, 9-1/4" at the extreme end, and 9-3/4" long.

The positive goal box was approximately square in shape, having inside dimensions of 11-1/4" x 11-3/4". It was painted black throughout, and had a 1/4" plywood floor which was covered with hardware cloth. A round coaster-like glass food dish, approximately 2" in diameter, was placed on the floor opposite the entrance to the box.

B. Subjects

The subjects were albino rats from the animal colony of the psychology department at Michigan State College. A total of 68 animals which had no prior experimental experience were used. The ages ranged from 108 to 190 days. A total of 33 female rats and 35 male rats were used. They were placed in the groups so as to approximately equalize the number of males and females in each sub-group.

C. Preliminary Training

Seven days of preliminary training was given to all animals. For the first four days, this consisted of handling and petting by E.

On the fifth, sixth, and seventh days of the preliminary period, the animals were placed in a straight alley, which consisted of the starting box and one of the arms from the maze already described. Each S received four trials per day, making a total of 12 preliminary trials in the straight alley. The S's received no food reward on these trials. They were retained in the second (or arm) section for fifteen seconds.

Beginning on the fifth day of the preliminary training period all of the animals were placed on a food regimen of nine gms. per day. They received this at regular feeding time for the remainder of the preliminary period.

At no time during these seven days did E feed the S's, and all preliminary handling, and training was carried on at least three to four hours from the time of feeding.

D. Method

Upon the completion of preliminary training, S's were placed at random in one of three groups. There were 24 animals in Group X-2, 20 animals in Group X-1, and 24 animals in Group C. Each main group was subdivided into two equal sub-groups.

Group X-2 animals were goal-reversed after two rewarded trials. For half the animals the positive goal was on the right for the pre-reversal trials, and on the left following position reversal. The other half was trained left for the first two trials and to the right for the remaining trials.

Group X-1 had one rewarded trial before goal reversal was imposed. Except for this condition, the sub-groups were set up in the same manner as in Group X-2.

Group C served as a control group and received no position reversal. One sub-group was rewarded for running right, while the second sub-group was trained to the left alley throughout the entire experimental period of five days.

This procedure with sub-groups was followed in order to randomize position preference in the animals.

Food reward, of one large pellet of dog food (approximately .40 gms.), was given in the black goal box at the end of the correct alley. Animals remained in the positive goal box until the food was eaten. Each S was kept twenty seconds in the negative white goal box which was always at the end of the incorrect alley. The end boxes were changed in position in order to conform to the experimental design for each sub-group. In the event an animal refused to enter the goal box within forty-five seconds, he was removed from the maze. A record was kept of all refusals to enter.

Each S had four trials per day for a period of five days. With the exception of the trials of the first day for Group X-1, the first two trials of each day allowed the animal free choice, while the remaining two trials were forced in such a manner as to equalize the number of correct and incorrect trials each day. The last two trials were forced in a way to make possible only four combinations of responses on the four trials. These were RRLL, LLRR, RLLR, and LRRL. This pattern was adhered to in order to discourage alternation of response.

Since position reversal was carried out in Group X-1 following one rewarded trial, the first day trials for this group were alternately free and forced, in order that one correct and one incorrect trial could precede reversal. A period of 45 minutes elapsed before the second set of free and forced trials was run. On all succeeding days the pattern of running was the same for Group X-1 as for Group C and X-2, i.e., two free trials and two forced trials per day in that order.

Animals were run in blocks of from six to ten in number. The order of running changed from day to day, but remained constant for successive trials on the same day.

At the end of each days run, the S's were fed nine gms. of Purina dog chow checkers in individual feeding cages. Thus a food deprivation of from 22 to 23 hours preceded each day's trials.

V. RESULTS

The results in terms of the per cent of correct responses for Groups X-1, X-2, and C are graphically presented in Figures 2 (p. 21) and 3 (p. 23).[1] The curves in Figure 2 are based upon the per cent of correct responses for the initial trials of each day, while the learning curves of Figure 3 are plotted in terms of the per cent of correct responses for the first two free trials of each day.[2] The initial trial data is considered superior in this study because (1) it is not affected by the tendency toward alternation which persisted despite the precautions taken to prevent it, and (2) it provides an equilization of the number of previous food reinforcements.

The results shown in Figure 2 indicate that the non-reversal group (C) and the group on which goal reversal was imposed after one rewarded trial (X-1) are well matched in performance. There is practically no difference anywhere along the learning curves. However, the X-2 group, when compared with the X-1 and C groups shows retardation on all days. The greatest retardation is shown on the initial trials of the second and third days. There is over an 80 per cent drop in the per cent of correct responses on the second day. Group C and Group X-1 both attained a performance level of 75 per cent on the second day whereas Group X-2 had only 20.8 per cent correct responses.

---

1 The data from which these curves are plotted is presented in Tables A and B of the Appendix.

2 The per cent of correct responses for the first day is calculated in terms of the correct position for trials following reversal in order to show the relationship between initial position preference and post-reversal results.

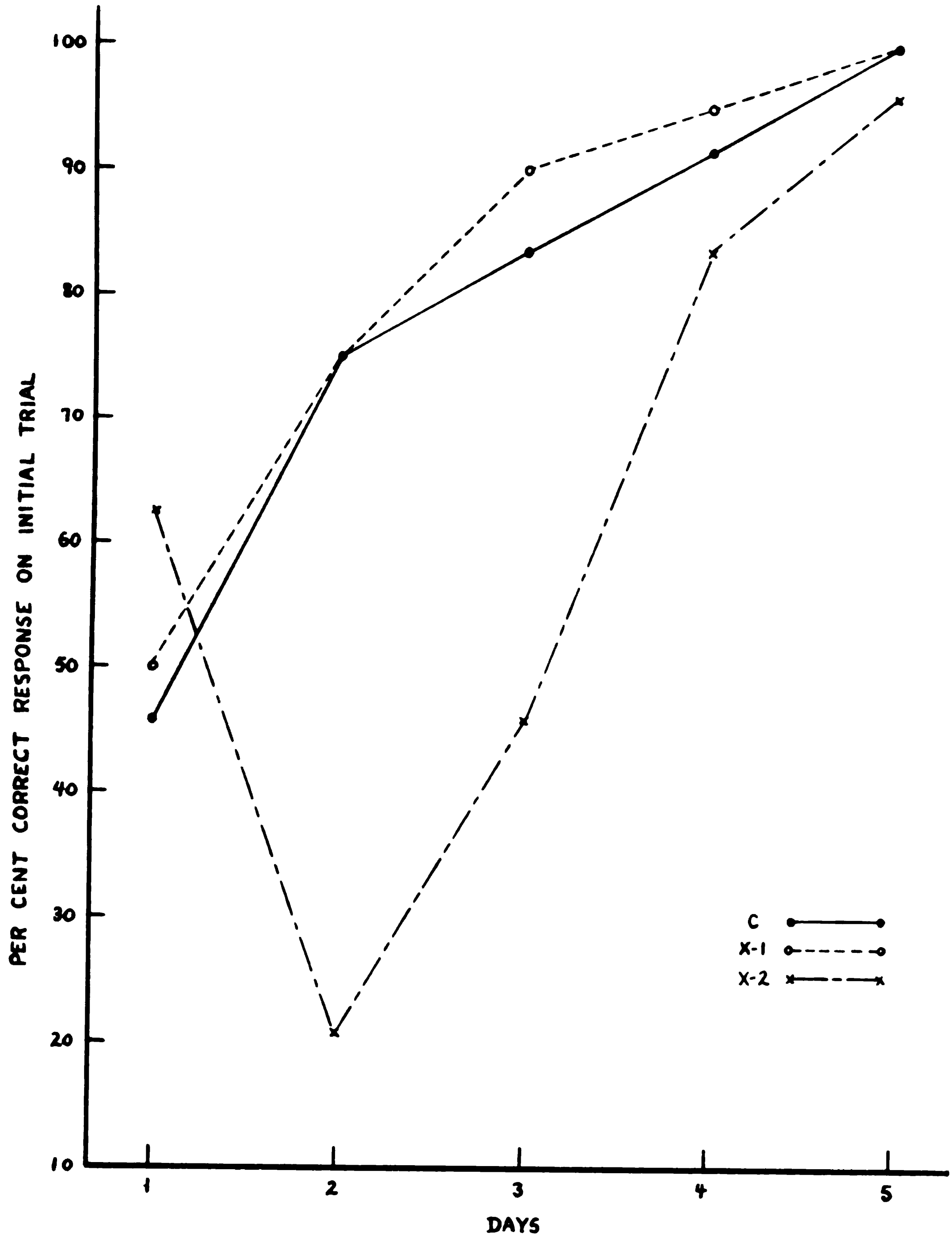

Figure 2. Learning Curves of Groups C, X-1, and X-2, as Based upon the per Cent of Correct Responses on the Initial Trial of the Day

Essentially the same trends are shown by the learning curves based on two trial data in Figure 3 (p. 23). The somewhat smaller difference shown between the X-2 group and the other groups in the two trial data is expected, at least in part, for the reasons previously mentioned. The performance of the X-1 and C groups is superior to the X-2 group on the second and third days and the drop in the per cent of correct responses for the X-2 group on the second day is again in evidence, with no corresponding decline in the other two groups. However, no consistent difference between X-2 and the other two groups exists on the fourth and fifth days with the two-trial data, indicating probably the slight effect of the variable being manipulated in a long run learning analysis.

According to Hull's (7) description of the growth of habit strength in which the first increments to the habit are the largest we would expect that any difference which would appear between the 'reversal' and the control groups would be greatest on the days immediately following goal reversal, rather than on later trials or in overall performance. This is evident in the learning curves of Figures 2 and 3.

Table 1 presents an overall comparison of Groups C and X-2 and Groups X-1 and X-2 for the first two free trials of days 2-5.[1]

---

1 The data of the first day is excluded because the variable of position reversal did not operate for all groups until the second day.

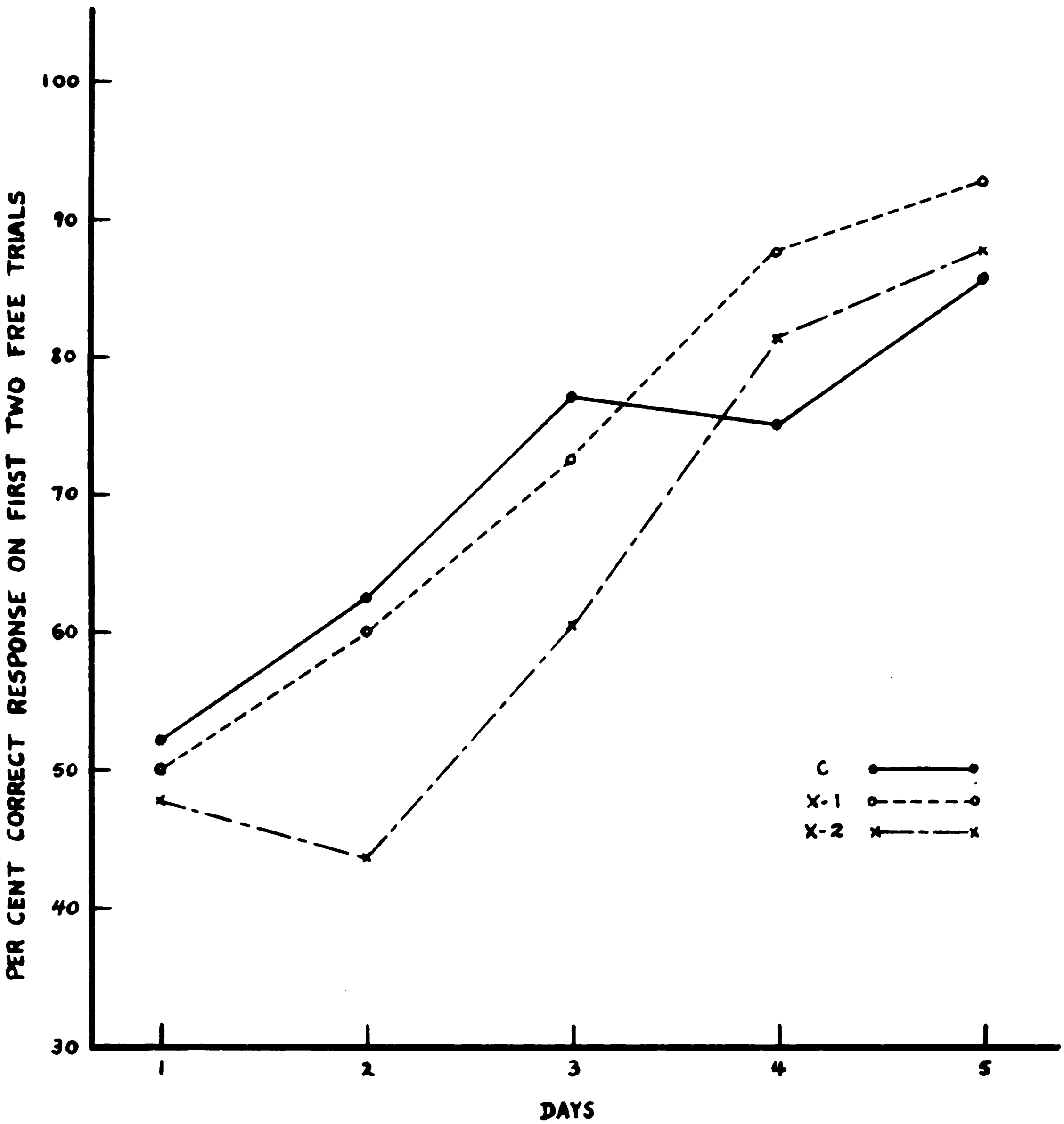

Figure 3. Learning Curves of Groups C, X-1, and X-2, as Based upon the Per Cent of Correct Responses on the Two Free Trials of the Day

TABLE I

Comparison of Groups C and X-2 and of Groups X-1 and X-2 in Terms of the Mean Number of Correct Responses on Both Free Trials of Days 2-5

Group	N	No. Correct Responses Mean	S. D.	Diff.	t	P
C	24	6.00	1.19	.540	1.61	.20
X-2	24	5.46	1.08			
X-1	20	6.25	.99	.790	2.47	<.02
X-2	24	5.46	1.08			

Here we see that the difference between the means of Group C and Group X-2 yields a t of only 1.61 which is significant only at about the 20 per cent level of confidence. The difference between X-1 and X-2 groups, however, yields a fairly large t of 2.47 which is significant beyond the two per cent level of confidence.

TABLE II

A Comparison of Groups C and X-2 and of Groups X-1 and X-2 on the Second and Third Days of Learning in Terms of the Number of Correct Responses on the Initial Trial*

		Day 2			Day 3		
Group	N	No. Correct Responses	Chi Square	P	No. Correct Responses	Chi Square	P
C	24	18	12.02	<.001	20	5.82	<.02
X-2	24	5			11		
X-1	20	15	10.78	<.002	18	7.54	<.01
X-2	24	5			11		

* The Yates correction for continuity, i.e., a deduction of .5 from each of the discrepancy values, has been made in the calculation of chi square to allow for the small frequencies.

When we turn to Table II for a comparison in terms of the number of correct responses on the initial trials of the second and third days of learning we see much more significant differences as obtained by the chi square test.

The X-2 group gives significantly fewer responses than either Group C or Group X-1 on both the second and third days, i.e., on the days when a predicted difference should show up. The null hypothesis of no difference between the X-2 group and the other two groups on the initial trial of the second and third days can therefore be rejected.[1]

1 The chi square test cannot be legitimately employed with the two-trial data.

VI. DISCUSSION OF RESULTS

The results of the present study show clearly that the X-1 group which was goal-reversed after one rewarded trial shows no retardation on subsequent trials when compared with the control group which received no position reversal, and, presumably, there is no difference between these two groups. However, learning by the X-2 group, which had position reversal carried out following two rewarded trials, was significantly slowed up following reversal, or, in other words, had learned to some extent to turn in the direction of the first two reinforcements.

This seems definitely to suggest that under the conditions prevailing in this study, no learning of the instrumental response took place on the first rewarded trial, while on the second rewarded trial some increment to the habit was effected.

In the light of Denny's (3) hypothesis this may be interpreted as meaning (1) until a fractional anticipatory response is set up, no instrumental learning is possible, and (2) that an anticipatory set starts to build up in one trial. A confirmation of this partial set on the second trial is effective in bringing about some increment to the instrumental response, causing a decrement in performance when the goals are reversed after two reinforcements.

However an, as yet, undiscussed aspect of the results prevents such an interpretation from being conclusive enough to reject the drive-satisfaction theories of Hull (7) and Thorndike (13). Since the animal had no expectation of finding food in the maze on the first trial, random and exploratory behavior was elicited, and a lapse of approximately 30 to 120 seconds occurred between the response of turning right or left and finding and eating the food. According to Hull's (7) principle of the reinforcement gradient, such a long delay between response and reward would probably not allow for an increment to the habit on this first trial; therefore, a decrement in performance following goal reversal after one rewarded trial would not be expected.

Whichever of these interpretations is accepted, some kind of expectancy set or anticipation would seem to be essential for instrumental response learning. Unless there is a set which acts in a behavior-directing capacity immediate reward in the maze situation is impossible.

The question then becomes, "What is the role of the expectancy set?" Does it function to bring about (1) more immediate drive reduction, or (2) a more immediate increase in the making of fractional (implicit) or complete (overt) consummating responses? The evidence for the concept of secondary reinforcement (1, 2, 8, 11,) which does not involve drive reduction, and the improbability of one small pellet of food reducing the hunger drive militates against the first alternative. Nevertheless, further research must be carried out to determine which of these interpretations, if any, is correct.

An experiment similar to this study is suggested in which the preliminary training includes having the reversal-after-one-reinforcement-group eat out of the glass dish that is used in the positive goal box. This might serve to reduce the lapse in time between the instrumental response and the goal response on the initial experimental trial, making it similar to the time taken by the control group on the second rewarded trial. Precautions would still have to be taken to insure the absence of an expectancy set on the first trial, although under these conditions there might be elicited anticipatory reactions upon the perception of the glass dish in the goal box.

Results from additional studies such as the one suggested above will serve to make the exploratory findings of this study more complete and make possible a better analysis of the nature of reinforcement.

VI. SUMMARY AND CONCLUSIONS

This study was designed to test the hypothesis that when expectancy of reward is absent, learning, under the condition of position reversal following one and possibly two rewarded trials, will proceed in much the same manner as under conditions of non-reversal of goals. There were two experimental groups: one, of 20 animals, which was goal-reversed following one rewarded trial, and another, of 24 animals which was reversed following two rewarded trials. A control group (N=24) received no reversal.

The apparatus consisted of a simple T-maze. All animals received seven days of handling prior to five days of learning trials, on which days they received two food reinforcements per day and an equal number of trials to each side.

The results revealed no differences between the non-reversed group and the group which was reversed following one rewarded trial. A significant difference was obtained on the second and third days between the group which was goal-reversed following two trials and the other two groups.

On the basis of results found in this study the following conclusions may be drawn:

1. In the absence of expectancy on the first rewarded trial, no learning of the instrumental response takes place; therefore an expectancy or anticipatory set seems essential for instrumental response learning.

2. Some expectancy seems to be built up in one trial.

3. On the basis of this study alone the specific function of the expectancy set cannot be determined.

BIBLIOGRAPHY

LITERATURE CITED

1. Bugelski, R. Extinction with and without sub-goal reinforcement. *J. Comp. Psychol.*, 1938, 26, 121-133.

2. Denny, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. *J. Exp. Psychol.*, 1946, 36, 373-389.

3. Denny, M. R. Unpublished lecture series (untitled). Michigan State College, East Lansing, 1948.

4. Hilgard, E. R. and D. G. Marquis *Conditioning and learning.* New York: D. Appleton-Century Co., 1940, 429 pp.

5. Hull, C. L. Goal attraction and directing ideas conceived as habit phenomena. *Psychol. Rev.*, 1931, 38, 487-506.

6. Hull, C. L. Mind, mechanism and adaptive behavior. *Psychol. Rev.*, 1937, 44, 1-32.

7. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century Co., 1943, 422 pp.

8. Mowrer, O. H. A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, 46, 553-565.

9. Mowrer, O. H. On the dual nature of learning - a reinterpretation of "conditioning" and "problem-solving". *Harv. Educ. Rev.*, 1947, 17, 102-148.

10. Shoben, E. J. A learning-theory interpretation of psychotherapy. *Harv. Educ. Rev.*, 1948, 18, 129-145.

11. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century Co., 1938, 457 pp.

12. Spence, K. W. An experimental test of the continuity and non-continuity theories of discrimination learning. *J. Exp. Psychol.*, 1945, 35, 253-266.

13. Thorndike, E. L. *The fundamentals of learning.* New York: Bureau of Publications, Teachers College, Columbia University, 1932, 638 pp.

14. Tolman, E. C. *Purposive behavior in animals and men.* New York: D. Appleton-Century Co., 1932, 463 pp.

15. White, R. K. The completion hypothesis and reinforcement. *Psychol. Rev.*, 1936, 43, 396-404.

ADDITIONAL REFERENCES

1. Denny, M. R. The effect of using differential end boxes in a simple T-maze learning situation. J. Exp. Psychol., 1946, 36, 245-249.

2. Hilgard, E. R. Theories of learning. New York: D. Appleton-Century Co., 1948, 409 pp.

3. Hull, C. L. Thorndike's fundamentals of learning. Psychol. Bull., 32, 807-823.

4. Hull, C. L. The goal-gradient hypothesis and maze learning, Psychol. Rev., 1932, 39, 25-43.

5. Krechevsky, I. 'Hypotheses' in rats. Psychol. Rev., 1932, 39, 516-532.

6. Mowrer, O. H. Preparatory set (expectancy) - a determinant in motivation and learning. Psychol. Rev., 1938, 45, 62-91.

7. Spence, K. W. Theoretical interpretations of learning. In F. A. Moss, edit., Comparative psychology (revised edition), New York: Prentice-Hall, Inc., 1946, 280-329.

8. Tolman, E. C. Law of effect. Psych. Rev., 1938, 44, 200-203.

APPENDIX

TABLE A

Comparison of Groups X-1, X-2, and C in Terms of the Per Cent of Correct Responses on the Initial Trial of Each Day

Day	Group C		Group X-1		Group X-2	
	No. Correct	Per Cent	No. Correct	Per Cent	No. Correct	Per Cent
1	11	45.9	10	50.0	15	62.5*
2	18	75.0	15	75.0	5	20.8
3	20	83.3	18	90.0	11	45.9
4	22	91.7	19	95.0	20	83.3
5	24	100.0	20	100.0	23	95.8
Mean Per Cent	95	79.2	82	82.0	68	56.6

* This per cent is calculated in terms of the correct position for trials <u>following</u> reversal.

TABLE B

Comparison of Groups C, X-1, and X-2 in Terms of the Per Cent of Correct Responses on the Two Free Trials Each Day

	Group C		Group X-1		Group X-2	
Day	No. Correct	Per Cent	No. Correct	Per Cent	No. Correct	Per Cent
1	25	52.1	20	50.0	23	47.9*
2	30	62.5	24	60.0	21	43.7
3	37	77.1	29	72.5	29	60.4
4	36	75.0	35	87.5	39	81.2
5	41	85.4	37	92.5	42	87.5
Mean Per Cent	169	70.4	145	72.5	156	65.0

* This per cent is calculated in terms of the correct position for trials following reversal.

Behavior Data of Animals in Group C on Ten Free Trials
Given at the Rate of Two Trials Per Day.
"X" Represents a Correct Response;
"O" Represents an Incorrect Response

Trials

	1	2	3	4	5	6	7	8	9	10	Total Correct
Animal #1	O	X	X	O	O	X	X	X	X	X	7
Animal #2	O	X	O	X	X	O	X	O*	X	X	6
Animal #3	O	X	X	O	X	X	X	X	X	X	8
Animal #4	X	O	X	X	X	X	X	O*	X	O*	7
Animal #5	O	O	O	O*	O	X	X	O	X	O*	3
Animal #6	O	X	X	O	X	X	X	X	X	X	8
Animal #7	X	O	X	O	O*	X	X	O*	X	O*	5
Animal #8	O	X	X	O	X	X	O	X	X	X	7
Animal #9	X	X	X	O	X	X	X	X	X	X	9
Animal #10	O	X	O	X	X	X	X	X	X	X	8
Animal #11	X	X	X	O	X	X	X	O	X	O	7
Animal #12	X	O	X	X	X	O	X	O	X	X	7
Animal #13	O	X	O	X	X	O	X	X	X	X	7
Animal #14	O	X	X	X	X	X	X	X	X	X	9
Animal #15	O	X	X	X	X	O	X	O*	X	X	7
Animal #16	X	X	X	X	X	X	X	O*	X	O*	8
Animal #17	X	O	X	X	X	X	X	X	X	X	9
Animal #18	X	X	X	X	X	O	X	X	X	X	9
Animal #19	X	O	X	O	X	O	X	X	X	X	7
Animal #20	X	O	O	X	X	O	O*	X	X	O	5
Animal #21	O	X	X	O	X	X	X	O	X	X	7
Animal #22	O	O	O	X	X	X	X	O	X	X	6
Animal #23	O	O	X	O	O*	X	X	X	X	O*	5
Animal #24	X	O	X	O	X	X	X	X	X	X	8
Total Correct Ea. Trial	11	14	18	12	20	17	22	14	24	17	169

* Denotes refusal to enter goal box within 45 seconds.

Behavior Data of Animals in Group X-1 on Ten Free Trials
Given at the Rate of Two Trials Per Day.
"X" Represents a Correct Response;
"O" Represents an Incorrect Response

Trials

	1	2	3	4	5	6	7	8	9	10	Total Correct
Animal #1	X	X	X	X	X	O	X	O	X	O	7
Animal #2	O	X	X	O*	X	X	X	X	X	X	8
Animal #3	O	X	X	X	X	X	X	X	X	X	9
Animal #4	O	X	O	X	X	O	X	O*	X	X	6
Animal #5	X	O	X	O	X	X	X	O	X	X	7
Animal #6	O	O	O	O	X	X	X	X	X	X	6
Animal #7	O	O	X	O	X	O	X	X	X	O	5
Animal #8	O	X	X	X	X	X	X	X	X	O	8
Animal #9	X	X	X	O	X	O	O*	X	X	X	7
Animal #10	X	X	X	O	X	O	X	X	X	X	8
Animal #11	O	X	X	O	X	X	X	X	X	X	8
Animal #12	O	O	O	O	X	O	X	X	X	X	5
Animal #13	X	O	X	X	X	O	X	X	X	X	8
Animal #14	X	X	X	X	X	O*	X	X	X	X	9
Animal #15	X	O	O	O	X	X	X	O	X	X	6
Animal #16	X	O	X	O	O	X	X	X	X	X	7
Animal #17	X	O	O	X	O	X	X	X	X	X	7
Animal #18	O	O	X	O	X	X	X	X	X	X	7
Animal #19	X	O	X	X	X	O	X	X	X	X	8
Animal #20	O	X	X	X	X	X	X	X	X	X	9
Total Correct Ea. Trial	10	10	15	9	18	11	19	16	20	17	145

* Denotes refusal to enter goal box within 45 seconds.

Behavior Data of Animals in Group X-2 on Ten Free Trials
Given at the Rate of Two Trials Per Day.
"X" Represents a Correct Response;
"O" Represents an Incorrect Response

Trials

	1	2	3	4	5	6	7	8	9	10	Total Correct
Animal #1	X	X	O	X	O	X	X	X	X	O*	7
Animal #2	O	X	O	O	X	O	X	X	X	X	6
Animal #3	O	X	O	X	O	X	X	X	X	X	7
Animal #4	O	X	X	X	X	X	X	O	X	O	7
Animal #5	O	X	O	X	X	O	X	X	X	X	7
Animal #6	O	X	O	X	X	X	X	X	X	X	8
Animal #7	O	X	O	X	O*	X	X	X	X	X	7
Animal #8	O	X	O	X	O	X	O*	X	X	O*	5
Animal #9	O	X	X	O	O	X	O	X	X	X	6
Animal #10	O	O	X	O	X	O	X	X	X	X	6
Animal #11	X	O	O	O*	O*	X	X	X	X	X	6
Animal #12	O	X	O	X	X	X	X	O	X	O	6
Animal #13	X	O	O	O*	O	O*	X	X	X	X	5
Animal #14	X	O	O	X	O*	X	X	O	X	X	6
Animal #15	O	X	O	X	X	X	X	X	X	X	8
Animal #16	O	X	O	X	O*	X	O*	X	X	X	6
Animal #17	X	O	O	X	O*	X	X	X	X	X	7
Animal #18	X	O	O	X	O	X	O*	X	X	X	6
Animal #19	O	X	X	O	O	X	X	O	X	O	5
Animal #20	O	X	O	X	X	X	X	X	X	X	8
Animal #21	O	X	O	X	X	X	X	X	X	X	8
Animal #22	X	O	X	O*	X	X	X	X	X	X	8
Animal #23	X	X	O	O	X	O	X	O	O*	X	5
Animal #24	X	O	O	X	O*	O	X	X	X	X	6
Total Correct Ea. Trial	9	16	5	16	11	18	20	19	23	19	156

* Denotes refusal to enter goal box within 45 seconds.