

AN EXPERIMENTAL STUDY OF THE EFFECT OF THE INITIAL REINFORCEMENT ON SUBSEQUENT RESPONSE TENDENCY UNDER DIFFERENTIAL CONDITIONS OF FOOD EXPECTANCY

> Thesis for the Degree of M. A. MICHIGAN STATE COLLEGE Robert Lloyd Martindale 1950

This is to certify that the

thesis entitled

"An experimental study of the effect of the initial reinforcement on the subsequent tendency under differential conditions of food expectancy."

presented by

Mr. Robert Martindale

has been accepted towards fulfillment of the requirements for

M.A. degree in Psychology

M. Ray Denny Major professor

Date May 25, 1950

O-169



AN EXPERIMENTAL STUDY OF THE EFFECT OF THE INITIAL REINFORCEMENT ON SUBSEQUENT RESPONSE TENDENCY UNDER DIFFERENTIAL CONDITIONS OF FOOD EXPECTANCY

by

Robert Lloyd Martindale

A THESIS

Submitted to the School of Graduate Studies of Michigan State College of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

MASTER OF ARTS

Department of Psychology

. K. .

· ·

. .

ACKNOWLEDGEMENT

Grateful acknowledgement is made to Dr. M. Ray Denny, under whose supervision this study was carried out, for his guidance and participation in the work herein reported. The author also extends his gratitude to Dr. Leo Katz for outlining the complex statistical procedures and to Richard A. Behan for his indispensable aid in carrying out the statistical analyses.

TABLE OF CONTENTS

		Page
LIST OF T	ABLES	iii
LIST OF F	IGURES	iii
I.	INTRODUCTION	1
II.	THEORETICAL ORIENTATION	3
III.	EXPERIMENTAL HYPOTHESES	14
IV.	EXPERIMENTAL PROCEDURE	15
	Experimental Apparatus Subjects Habituation Period Preliminary Training Training Extinction	
۷.	RESULTS AND DISCUSSION	26
VI.	SUB-EXPERIMENT	39
	Procedure Results and Discussion	
VII.	SUMMARY AND CONCLUSIONS	44
BIBLIOGRA	РНҮ	46
APPENDIX		47

LIST OF TABLES

Table		Page
I.	Analysis of Variance of Training Data for Preliminary Training Groups and Training Groups in Terms of the Number of Correct Responses per Animal	27
II.	Group Comparisons in Terms of the Mean Number of Correct Responses for the Total Training Period	28
III.	Comparison of Preliminary Training Groups and Training Groups in Terms of the Number of Correct Responses on Pertinent Training Trials	29
IV.	Analysis of Variance of Extinction Data for Preliminary Training Groups in Terms of the Number of Correct Responses per Animal	33
۷.	Analysis of Variance of Eating Latency Data for Preliminary Training Conditions and Training Conditions in Terms of the Mean Latency per Day per Animal	34

LIST OF FIGURES

Figur	e	Page
I.	Experimental Apparatus Floor Plan	16
II.	Learning Curves of Composite Groups I, II, and III in Terms of Per Cent of Ultimate Correct Responses on Each Trial	31
III.	Latency of Eating Response Curves for Composite Groups \blacktriangle and C in Terms of the Mean Time per Trial per Day	35

зъ.

•••••

•••••

•

• • • • • • • • • • •

I INTRODUCTION

The theoretical background of the present study is oriented around basic philosophic concepts most generally referred to as scientific mechanism or determinism. It allows no consideration of mystical or metaphysical concepts or of any viewpoint that postulates anything that is not ultimately observable by the scientific community. All constructs (mechanisms of observable relationships) must be defined in terms of the observable either directly or indirectly. A similar viewpoint is expressed by Hull as follows:

> whereas argument reaches belief in its theorems because of antecedent belief in its postulates, scientific theory reaches belief in its postulates to a considerable extent through direct or observational evidence of the soundness of its theorems. (1)

(1) Hull, Clark L., <u>Principles</u> of <u>Behavior</u>, p. 9, 1943. This work will hereafter be referred to as P. B.

A living organism is defined as any relatively unitary collection of matter that is capable of differential adaption to an external stimulus situation (physical energy external to the organism) or internal stimulus situation (physical energy

(1)

components of a need state necessary to the maintenance of the organism in its essential unitary state). In order to meet this requirement the living organism is essentially a communications or feedback system in the process of a single continuous operation that transmits physical stimulus energies. The nervous system with its receptor beginning and its effector termination is in essence just this sort of communications mechanism. This viewpoint of the living organism is derived almost totally from the implications of the field of "cybernetics".(2) The organism can neither add nor

(2) Wiener, Norbert, Cybernetics, 1948.

subtract from these stimulus energies but only receives, reintegrates, and distributes them in a manner most expedient to its maintenance in so far as these energies are within the reception and dissemination range of the organism. This point of view bears some similarity to Hull's conception of organisms as "self-maintaining mechanisms" (3) but our viewpoint is elaborated on a

(3) P. B., p.384.

more molecular level.

From this it is assumed that the organism is a stimulus-response mechanism. On a molar level such a

mechanism consists of: (1) a complex physical structure for communications, and (2) a totality of all stimulus energies that have impinged on the receptors and have not been dissipated by the effectors. Since the subject matter we are dealing with is the molar, observable components of the organism's adjustment to its environment, we can postulate theoretical constructs that allow us to predict lawful relationships between afferent and efferent aspects of the maintenance of the organism.

Watson conceived of an identity between the stimulus and the response such that either could be predicted from the other.(4) This identity hypothesis

(4) Watson, John B., The Ways of Behaviorism, p. 2, 1928.

is essentially congruent with the presented point of view but Watson minimized the obvious complexities of the relationship. However, Hull, although he implicitly accepts this identity hypothesis, further elaborates the necessity for the postulation of the intervening constructs logically deduced from the observable, molar antecedents and consequents.(5)

(5) Hull, Clark L., "Mind, Mechanism, and Adaptive Behavior", <u>Psychological Review</u>, vol. 44 (1937), p. 1-32.

II THEORETICAL ORIENTATION

Learning of an instrumental act is dependent upon the establishment of bonds or relations between an instrumental act and a stimulus complex in a rewarding situation. A rewarding situation consists of: (1) reduction of a physiological need that is based on the innate structure of the organism, or (2) a stimulus situation that has been in temporal contiguity with physiological need reduction and therefore represents actual need reduction. This viewpoint of reinforcement is almost identical with Hull's latest revision of his postulate and corollary concerned with primary and secondary reinforcement.(6)

(6) Hull, Clark L., <u>Behavior Postulates</u> and <u>Corollaries</u> - <u>November</u>, <u>1949</u>, p. 2.

Denny suggests that all reinforcement of an instrumental act is, in Hullian terms, secondary reinforcement. Because actual need reduction does not take place in sufficiently close temporal contiguity with an instrumental act to establish learning bonds, secondary reinforcement is what is actually present. This would be particularly true at the beginning of training. Denny also postulates a view of reward similar to the above. He classifies reward in terms of the

(4)

demands of the organism's <u>permanent</u> or innate structure and <u>temporary</u> structure (momentary state of the organism). These states predispose the organism toward a particular response, and when such a response occurs a reinforcing state of affairs exists. Fractional-anticipatoryresponses are included in the latter category.(7)

(7) Denny, M. Ray, Unpublished lecture series (untitled), 1950.

It is implicit in the preceeding postulations that all rewarding situations must occur within the temporal duration of a <u>superthreshold</u> neural stimulus trace in order for that stimulus to get hooked up to the rewarding situation. The shorter the time between the onset of a neutral afferent neural impulse and the onset of an afferent neural impulse of a rewarding stimulus situation, the stronger will be the learning bond thus established. <u>Therefore, if a neutral</u> <u>instrumental act occurs in sufficiently close temporal</u> <u>contiguity with a rewarding stimulus situation, it can be</u> <u>postulated that learning of the instrumental act will</u> <u>occur with the first presentation of the rewarding</u>

Hull presents a similar viewpoint to this in his "law of primary reinforcement" without the specific reference to the effect of the first reinforcement as follows:

Whenever an effector activity occurs in temporal contiguity with the afferent impulse, or the perseverative trace of such an impulse, resulting from the impact of a stimulus energy upon a receptor, and this conjunction is closely associated in time with the diminuation in the receptor discharge characteristic of a need, there will result an increment to the tendency for that stimulus on subsequent occasions to evoke that reaction. (8)

(8) P. B., p. 80.

A similar interpretation is also inherent in Hull's postulation of the "gradient of reinforcement" as different from the "goal gradient". The "gradient of reinforcement" represents an interval of time after an instrumental response in which secondary reinforcement is not necessary for an increment in that response tendency. After this primary temporal duration secondary reinforcement of some type is necessary to produce an increment.(9) This postulation seems to imply that what

(9) P. B., p. 158.

we are dealing with in the case of the "gradient of reinforcement" is a perseverative stimulus trace.

The experimental evidence from which Hull measures the "gradient of reinforcement" has been attacked by Spence and Grice on the grounds that all secondary reinforcement had not been eliminated during the time delay and for that matter could not be eliminated because of the secondary reward value of the instrumental act itself. Spence states his position as follows:

> The interpretation that learning under conditions of delay of primary reward involves a backward action of the goal object on the preceeding stimulus-response event is rejected. The hypothesis suggested as an alternative to this conception is that all such learning occurs as a result of the development of secondary reinforcement, the action of which is conceived to take place immediately upon occurance of the response.(10)

(10) Spence, K. W., "The Role of Secondary Reinforcement in Delayed Reward Learning", <u>Psychological Review</u>, vol. 54 (1947), p. 1-8.

Grice concludes from experimental results that when all secondary reward is eliminated subsequent to the instrumental act, learning may still occur within a substantial time interval. He states further that "as long as the trace of the correct response is within the range of the generalization gradient of the proprioceptive pattern stimulating the organism at the time of the response in the (correct) alley, the proprioceptive pattern will acquire secondary reinforcing properties".(11)

(11) Grice, G. Robert, "The Relation of Secondary Reinforcement to Delayed Reward in Visual Discrimination Learning", Journal of Experimental Psychology, vol. 38 (1948), p. 15.

The author and Denny have both interpreted these formulations as meaning that secondary reinforcement value of the instrumental response is a necessary condition to learning when all other forms of secondary reward have been eliminated during the delay period. However, before the first reinforcement the instrumental act has no reward value. Therefore, if there is an increment to response tendency immediately after the first reinforcement, it must be due to a primary "gradient of reinforcement". The goal gradient cannot be operative at this point because the instrumental response could never have been experienced in absolute contiguity with the secondary reward of the goal gradient until after the second reinforcement.

In a more recent revision of his postulates Hull explicitly postulates the nature of the stimulus trace as a function of intensity and time. The action of the stimulus in the afferent nervous system after its physical termination can still serve as a learning stimulus when brought into contiguity with primary or secondary need reduction. Although, again, there is no specific reference to the first reinforcement, it would seem inconsistent, if the stimulus trace of the instrumental response contiguous with the first reinforcement of that response did not result in an increment to the tendency for that response to be evoked upon a future presentation of those stimuli.(12)

(12) Hull, Clark L., <u>Behavior Postulates and Corollaries</u> - <u>November</u>, <u>1949</u>, p. 2.

The author regards behavioral oscillation as the behavioral component of a physiological need state and as such anticipatory response is not believed to mediate learning on a molar level. In as much as the physiological need state of the organism may vary from moment to moment it follows, if oscillatory behavior is correlated with the need state, that this behavior will vary from moment to moment. In other words continuous change in the structure of the organism is presumably the basis of behavioral variation or oscillation. Thus we have a need state basis for oscillation rather than attributing oscillation to a "little understood physiological process" as Hull does.(13)

(13) P. B., p. 393.

Exploratory responses as behavioral manifestations of oscillation would necessarily have both external and internal stimulus correlates. Therefore, it would seem on the basis of former postulations that exploratory stimulus components because of their close temporal contiguity with the rewarding situation would gain secondary reward value. However, if we examine this a

little closer we see that such is not the case. It is true that the close temporal contiguity of these stimuli to the rewarding situation would provide optimum conditions for learning, but these stimuli and their responses are present at other times when they are <u>not</u> rewarded and, therefore, extinguish. In fact, it is obvious that these responses are unrewarded in most instances of their occurrence, as for example, with a hungry animal in an empty home cage. In view of this interpretation it appears highly doubtful that there is any learning of an exploratory response on a molar level.

Actually Hull's interpretation does not differ from this to any great extent. The primary need state initiates and maintains anticipatory responses. Although he states that the resulting stimuli would establish learning bonds to rewarding situations, he does not regard them as an important factor in the learning of a response sequence.(14) However, it should

(14) Hull, Clark L., "Goal Attraction and Directing Ideas as Habit Phenomenon", <u>Psychological Review</u>, vol. 38 (1931), p. 504.

be further noted that Hull regards the "fractional component of the goal response" as something quite different from the anticipatory response. This "fractional component of the need reduction process" is regarded only as a possible mechanism of reinforcement

generalizing back along the goal gradient. (15)

(15) P. B., p. 100.

It is postulated that behavior is specified in terms of the need state by the need state itself. It specifies behavior to the extent that non-rewarded exploratory responses are more quickly inhibited and rewarded responses are more quickly facilitated. The greater the deprivation of an organism in terms of the particular physiological need, the less will be the variability of behavior in a rewarding stimulus situation. Therefore, response latency is postulated as a negatively correlated function of the degree of deprivation, and is not a direct measure of habit strength. The degree of deprivation does not facilitate learning directly but only provides more optimum conditions for learning by decreasing the time interval between response and reward. With the degree of deprivation varied and the time interval between an instrumental act and reward held constant, there should be no difference in the probability of occurrence of the instrumental response.

From the writer's point of view behavior is specified in terms of the particular <u>learning situation</u> by a successive spatial movement of the rewarding stimulus situations back to previously neutral stimuli

until the instrumental act is reached. It should be noted that this is a molar interpretation and that all stimuli in the experimental situation would establish learning bonds to the rewarding situation at least on a subliminal level in terms of the perseverative stimulus trace. This process continues infinitely back from the initial rewarding situation so that behavior becomes specified in terms of more remote and more general stimuli on a more and more molar level.

Therefore, if an animal is rewarded in a stimulus context spatially antecedent to an instrumental act, then the animal's future learning of the instrumental act should not be facilitated. An opposing viewpoint would be that if an animal is rewarded previous to the instrumental act, then fractional-anticipatoryresponses or expectancies would be set up to the stimulus pattern. These expectancies would specify behavior in terms of the relative need state and this specification of behavior would facilitate learning of a subsequent instrumental response. For example Dorothy N. Moore in a Master's thesis concludes on the basis of her experimental results that "in the absence of expectancy on the first rewarded trial, no learning of the instrumental act takes place; therefore an expectancy or anticipatory set seems essential for instrumental response learning".

This conclusion appears to be in antithesis to the present theoretical framework.(16) However, an

(16) Moore, Dorothy Nell, <u>An Experimental Study of the</u> <u>Effect of Position Reversal after One or Two Reinforcements</u> <u>on Simple T-Maze Learning in the Rat, Thesis for the</u> <u>Degree of M. A.</u> (Michigan State College, 1949), p. 29.

examination of her experimental apparatus reveals that the instrumental act was temporally so distant from the rewarding stimulus situation that the stimulus trace of the instrumental act was probably subliminal at the time of reward.(17) In this case, therefore, no observable

(17) <u>Ibid.</u>, p. 15.

evidence of learning would have been detected after the first reinforcement.

III EXPERIMENTAL HYPOTHESES

The previous postulations assume that expectancies do not mediate learning. Therefore, the following hypotheses are evident in the experimental design:

A. If, in a T-maze learning situation designed so as to reduce the time interval between the instrumental act and the reinforcement to minimum, the end boxes associated with reward and non-reward are reversed after one reinforcement, the animals' learning of the instrumental act should be retarded as compared with the animals' having no reversal.

B. If, in the same T-maze situation, the end boxes are reversed after the second reinforcement, the animals' learning of the instrumental act should be retarded as compared with animals having no reversal and this retardation should be greater than in hypothesis A.
C. If animals are trained so as to set up food expectancies to the starting box and straight alley, their subsequent learning of the instrumental act should not be in any way facilitated as compared with naive animals as long as both have approximately immediate secondary reinforcement in the subsequent learning situation.

(14)

IV EXPERIMENTAL PROCEDURE

A. Experimental Apparatus

The apparatus was a modified, enclosed T maze (see floor plan in Fig. 1). It was specifically designed to make the delay of reward of the instrumental act as short as possible. The response from the choice point was a continuously similar turning response either to the right or the left. The entire apparatus was covered with a $\frac{1}{4}$ in. hardware cloth except the starting box which had an opaque wood cover. The floors were 3/4in., the sides $\frac{1}{2}$ in., and the doors $\frac{1}{4}$ in. plywood. The curved surfaces of the maze were constructed of heavy gauge galvanized metal. The inside depth of all parts of the maze was 8 in. All interior surfaces of the maze except the goal boxes were painted with a neutral gray enamel. The entire apparatus was placed on a $43\frac{1}{2}$ in. by $22\frac{1}{2}$ in. table which was 27 in. high.

The doors at the choice point were operated by a release mechanism that could drop them swiftly when the first part of the rat's body crossed the response criterion line. At this point the door from the straight alley and the door farthest from the end box that was entered were closed. As soon as the animal had completely entered an end box, the door nearest to the end box

(15)



.

entered was closed. The door at the starting box was equipped with a conventional string, pulley, and weight arrangement. All doors were padded with sponge rubber at the bottom so that their closing was as silent as possible.

Forced trials were performed by placing a false wall of $\frac{1}{2}$ in. plywood on the opposite side of the choice point from the maze arm into which the animal was to be forced. There was one of these for each side. They fitted flush against the rear wall on either side of the choice point and against the curved surface leading to the maze arm from the choice point in such a way that the maze arm was completely blocked off, forming a solid straight wall extending from the starting box to the back of the choice point on the desired side. The false wall was the same height as the rest of the maze. It was painted with the same gray enamel as the rest of the maze interior.

The end boxes were of distinctly different shape as illustrated and were the same height as the alleys. The negative end box was lined on the floors and sides with white posterboard. The posterboard formed a false wall with a $\frac{1}{4}$ in. space between the posterboard and wooden walls of the box. This space was filled with crushed food identical to that used in the positive goal

box in order to equalize the food odor. The positive goal box was lined with black posterboard on the walls and floor. The entire floor of the positive goal box was covered with Purina Dog Chow Checkers of the same type as the animals were normally fed. The posterboard was used rather than paint in the goal boxes in order to minimize any differential paint odor.

The goal box for preliminary training was placed at the end of the straight alley. It was constructed of plywood with inside dimensions of 11 5/8 in. wide, $7\frac{1}{2}$ in. deep, and $5\frac{1}{2}$ in. high with a rectangular floor plan. It was covered with a hinged cover of ordinary window screening. The entire interior of the box was painted with a neutral gray, flat paint. About midway through the experiment this box was replaced with another rectangular box with inside dimensions of 5 in. wide, 12 in. deep, and $11\frac{1}{4}$ in. high. It was covered with a cover made of $\frac{1}{4}$ in. mesh hardware cloth. The interior was painted with a neutral gray, flat paint.

The maze was lighted by a clear 200 watt bulb placed directly over the center of the straight alley. The bulb was 68 in. from the floor of the alley. There were also 60 watt frosted bulbs in semi-spherical white reflectors shining directly into each goal box and respective maze arm. The bulbs were placed 4 in. from the top of the walls of the maze. Both of these lamps

were placed above the first training goal box 6 in. from the top of the walls of the box and above the second training goal box 1 in. from the top of the walls of the box. The lighting procedure above reduced differential illumination, resulting from the reflection of the white and black walls of the respective end boxes on the maze arm walls that was visible from the choice point, so that it was just barely discriminable to the experimenter.

B. Subjects

The subjects were experimentally naive albino rats from the rat colony of the psychological laboratory of Michigan State College. The animals varied in age from 60 to 165 days at the beginning of the experiment. Their mean age was approximately 110 days. The animals remained in their typical home environment throughout the experimental period, except that they were fed separately in individual feeding ceges for part of the period. The animals were fed Purina Dog Chow Checkers both in the experimental apparatus and in the feeding cages. This was the same food that they were normally fed in their home environment.

C. Habituation Period

All animals except for four in the first group, who are designated in the data, were given a two day habituation period. This consisted of handling by the experimenter and being allowed to run free on a small table. On the first day of this period the animals were marked by a combination of cuts in their ears and ink markings on their tails. On the first day of this period they were given a large amount of food in their home cage. They were not fed on the second day. Water was continuously available both of these days and throughout the remainder of the experimental period.

D. Preliminary Training

On the day immediately following the habituation period a five day preliminary training period was begun. The animals were divided into two training groups at the beginning of this period which were labeled A and C. Each animal in both of these groups was fed in the individual feeding cages within between 5 and 10 minutes after the individual training session.

Animals in group A were handled for a short period by the experimenter and allowed to run free on a small table exactly as in the habituation period. These

animals were fed 7 grams of food the first day and $4\frac{1}{2}$ grams on the remaining days in the preliminary training period.

Group C animals were placed in the starting box of the experimental apparatus and allowed to run down the maze straight alley to the pre-training goal box. The animal was then allowed to eat for 30 seconds. The experimenter attempted to time the animals only during the period they were actually eating so as to compensate for interruptions in eating in so much as this was possible. If the animal did not eat within 5 minutes after entering the goal box, he was removed and this observation was recorded. This reward procedure was followed in both the pre-training goal box and the positive goal box for the remainder of the experiment. These animals were given two massed trials each day. They were fed $6\frac{1}{2}$ grams the first day and 4 grams on the remaining days during the preliminary training period. Consequently a very high drive level prevailed for all animals.

E. Training

The preliminary training period was followed on the next day by a four day training period. The training groups were subdivided into three different training

training condition groups. These were labeled groups I, II, and III which combined with the preliminary training conditions gave a total of six primary groups designated I-A, II-A, III-A, I-C, II-C, and III-C.

Groups I-A, II-A, I-C, and II-C were equalized for sex, initial position preference, and position of the positive goal box. This produced eight possible combinations in each of these primary groups. There were 16 animals in each group giving two animals under each possible combination of conditions.

Groups III-A and III-C were equalized for position preference only but the other variables were equalized as much as possible. There were 14 animals in each of these two groups.

Because of the equalization of position preference on the basis of the first training trial, most animals were not assigned to a group until after the first trial. This procedure made possible the important step of beginning all groups at a level of exactly 50% correct responses.

During the training period the animals were given four trials per day for a four day period. Two of each day's trials were to the negative end box and two were to the positive goal box so that each animal had two reinforced and two non-reinforced trials per day. The trials of the first day were alternately free and forced starting with a free trial. On succeeding days the first two trials were free and the second two were forced. The forced trials were arranged after the first day so that only the following combinations of right and left responses were possible: LLRR, RRLL, RLLR, and LRRL. This was done to prevent an alternation pattern of response.

The direction of each free response was recorded. The animals were fed in the positive goal box in exactly the same manner as in the preliminary training trials. On trials to the positive goal box the animals were timed from the moment the first part of their body touched the criterion line until they started eating. They remained in the negative end box for 30 seconds after entering it. On the first two days all trials were separated by a period of approximately 30 minutes. This time decreased somewhat on the second two days.

The animals were fed in the feeding cages between 5 and 10 minutes after their last trial for the day. They were allowed to remain until they had finished the day's ration. All animals were fed 4 grams of food on the first two days of this period and $5\frac{1}{2}$ grams on the second two days.

Group I animals were trained consistently to one side. For Group II animals the position of the positive

and negative end boxes was reversed after one free and one forced trial only one of which was a correct or reinforced response. For Group III animals the position of the positive and negative goal boxes was reversed after two free and two forced trials only two of which were correct or reinforced responses.

F. Extinction

On the day immediately following the last day of the learning trials all animals were given twelve massed extinction trials. The only change made in the apparatus for these extinction trials was to remove the food from the floor of the positive goal box. The animals were allowed to remain in the end boxes approximately 10 sec. for each extinction trial.

On all trials (preliminary training, training, and extinction) the animal was forced to run after the passage of a reasonable period of inactivity. This was done by poking the animal with a long stick. This was not recorded, if it was done in the starting box or straight alley. However, in the few instances that forcing was necessary at the choice point or immediately in front of the choice point on a free trial, it was recorded in the data. These few instances occurred primarily during extinction and <u>never</u> on the critical trials of the training period, i.e. trials 1, 2, or 3.

Animals were discarded from the data which refused to eat within the alloted 5 minute period on the first trial of the training period. They were also discarded if they refused to eat on both rewarded trials of any subsequent day in the period. Animals in preliminary training group C were discarded from the data, if during the preliminary training period they refused to eat on all trials of the first four days or if they refused to eat on either of the last day's trials.

V RESULTS AND DISCUSSION

The following measures were used for the statistical analyses of the results: (1) number of correct responses for all free trials for each animal during training, (2) number of correct responses for particular trials for all animals during training, (3) number of correct responses on all extinction trials for all animals, and (4) mean eating latency (time interval between the instrumental response and the eating response) per day per animal.

Analysis of variance of the total training period demonstrates no significant differences between preliminary training conditions or for the interaction between preliminary training and training (see Table I). Therefore, all further analyses of the training data combine preliminary training conditions and inspect differences only between training conditions. To assure further that there were no significant differences between preliminary training conditions a chi square was calculated on the largest single trial difference (Trial 2, Group II). This difference is not significant as shown in item 1 of Table III. The training data have not indicated any differences in learning due to differential preliminary training conditions designed to provide expectancies for one group and not for the other. Therefore, at this point hypothesis C is supported.

(26)

TABLE	Т
	_

Analysis of Variance of Training Data for Preliminary Training Groups and Training Groups in Terms of the Number of Correct Responses per Animal*

Source of Variance	Degree of Freedom	Mean Square	Significant F	P
Total	8 7	1.45		
Preliminary Training	g 1	1.08		
Training	2	15.35	11.29	•01
Position Preference Sex X Initial Place	X ment			
of End Boxes	42	•94		
Preliminary Training Training	g X 2	.15		
Erro r	40	1.36		

* These data meet the homogeneity requirements for analysis of variance.

Т	ABLE	II
_		

Group Comparisons Terms of the Mean Number of Correct Responses for the Total Training Period

	Correct	Responses	Actual	Require	d
N	Mean	S. E.	Diff.	Diff.	P#
32	7.12				
32	6.25	• 34	•8 7	•84	•02
32	7.12				
28	5.65	• 35	1.48	•97	•01
32	6.25				
28	5.65	•35	. 61	•60	•10
	N 32 32 32 28 32 28 32 28	Correct N Mean 32 7.12 32 6.25 32 7.12 28 5.65 32 6.25 32 7.12 28 5.65 32 6.25 32 5.65	Correct Responses N Mean S. E. 32 7.12 .34 32 7.12 .34 32 7.12 .35 32 7.12 .35 32 5.65 .35 32 5.65 .35	Correct Responses Actual Diff. N Mean S. E. Diff. 32 7.12 .34 .87 32 7.12 .34 .87 32 7.12 .35 1.48 32 6.25 .35 .61	Correct Responses Actual Diff. Require Diff. 32 7.12 .34 .87 .84 32 7.12 .34 .87 .84 32 7.12 .35 1.48 .97 32 6.25 .35 .61 .60

(

* P is based on a small sample t-ratio analysis.

TABLE III

Comparison of Preliminary Training Groups and Training Groups in Terms of the Number of Correct Responses on Pertinent Training Trials*

Compari	sons		Number Correct	Chi	
Groups	Trial	N	Responses	Square	P
	2 2	16 16	3 7	1.31	- 30
	-		•	2002	•00
To tal Total	1 2	92 92	46 67 x	9.18	.01
I A&C	2	32	26		
II A&C	2	32	10	14.27	•01
I A&C	3	32	28		
III A&C	3	28	6	24.09	•01
II A&C III A&C	3 3	32 28	19 6	7.59	.01

* A deduction of .5 from each discrepancy value has been made to allow for the small frequencies and emperical expected frequencies have been used in the calculation of all chi squares.

This figure represents the number of correct responses based on the position of the positive goal box for the first trial. All other figures for the number of correct responses are based on the position of the positive goal box for the indicated trial.

Analysis of variance of the total training period given in Table I demonstrated a significant difference among training conditions. Differences between the mean number of correct responses for the three training groups are all in the hypothesized direction and the significance of these differences is shown in Table II. The probability of a correct response for all 92 animals in terms of what was correct for the first trial (position of the positive goal box on the first trial) is well below the 1% level of confidence (see item 2, Table III). The tendency to make an incorrect response' on the second trial for Group II and on the third trial for Group III, just after reversal of end box position and in terms of this reversed end box position, as compared with Group I, which had no reversal, are both well below the 1% level of confidence (see items 3 & 4, Table III). The tendency to incorrect response for Group III animals on the third trial just after reversal of end box positions for that respective trial is significantly greater than the tendency to incorrect response for Group II animals on the third trial at well below the 1% level of confidence (see item 5, Table III).

The same relationships during training are shown graphically in Figure II. All the statistical analyses together indicate rather conclusively that all training groups differ significantly from each other in the





hypothesized direction for the entire training period and for the critical trials. In other words, reversal after one or after two reinforcements retards learning and <u>all animals learned on the basis of the first</u> <u>reinforcement</u>. At this point, therefore, expectancies do not appear to be necessary to mediate learning and hypotheses A and B are supported.

Analysis of variance of the total extinction data does not indicate any significant differences between the pertinent conditions nor is any significant interaction between the pertinent conditions indicated (see Table IV).

The eating latency measure did not meet the homogeneity requirements for analysis of variance when based on mean time per animal. Therefore, the individual within groups variance based on the individual variance per day is used as the error term. This allows comparison between preliminary training conditions and between training conditions, and analysis of the interaction between preliminary training and training conditions as shown in Table V. A further analysis of the differences between paired individual groups was not possible with a total number of subjects this small in view of the lack of homogeneity in the data.

Analysis of variance of the eating latency data in Table V indicates that the only significant difference

TABLE IV

Analysis of Variance of Extinction Data for Preliminary Training Groups and Training Groups in Terms of the Number of Correct Responses per Animal*

Source of Variance	Degree of Freedom	Mean Square	Significant F
Total	8 7	6.24	
Preliminary Training	1	2.78	
Training	2	2.54	(There are no
Position Preference Sex X Initial Placem	X lent		values.)
of End Boxes	42	3.87	
Preliminary Training Training	X 2	1.89	
Error	40	9.21	

* These data meet the homogeneity requirements for analysis of variance.

TABLE V

Analysis of Variance of Eating Latency Data for Preliminary Training Conditions and Training Conditions in Terms of the Mean Latency per Day per Animal

Source of Variance	Degree of Freedom	Mean Squ are	Significant F	P
Individual	91	5,377.76	;	
Preliminary Training	; 1	32,851.98	6.31	•02
Training	2	3,025.49	•	
Preliminary Training Training	; X 2	1,271.32	!	
Individual within Groups (error)	86	5,208.49	,	

.

.





Days

results from preliminary training conditions. The graph shown in Figure III illustrates this relationship. The preliminary training conditions significantly affect the time elapsing between the instrumental response and the eating response. That is, those animals given preliminary training in the straight alley eat significantly sooner than naive animals. However, this time difference appears in no way to affect learning. This lack of effect on learning is probably due to the immediate secondary reinforcement resulting from the observation of familiar food, because the delay in making the consummatory response is large for all animals (see Figure III).

The apparent reason for the exactly opposite results of no retardation of learning after reversal on the second trial obtained by Moore is apparently due to the lack of sufficiently immediate secondary reinforcement after the instrumental response in her experiment. (18)

(18) Moore, <u>op</u>. <u>cit</u>., p. 15.

A possible criticism of the obtained results and their interpretation could be drawn from an apparent artifact in the experimental procedure. The apparatus was designed in such a manner that the animal could experience a stimulus pattern while he was eating that would be the same as the stimulus pattern experienced

by the animal at the choice point before making the instrumental response. The lack of difference on the second trial in the amount of demonstrated learning between trained (C) and naive (A) groups, or even the increment on the second trial to the trained group might be attributed to expectancies mediated by this stimulus pattern. That is, on the second trial an animal upon reaching the choice point might approach the cues that have been associated with the eating of food. This stimulus pattern would consist mainly of floor cues and visual cues that would be different on either side of the choice point. However, such an explanation is a very remote possibility due to the very close similarity of stimulus patterns on either side of the choice point. These minimal stimulus differences would be so small that generalization of reinforcement and inhibition could completely nulify any differential value of the cues. Yet, the possibility still remains that an animal responded correctly on the basis of the close association of these cues with reinforcement rather than on the basis of reinforcement of the instrumental response. A sub-experiment designed and performed to test this possibility follows. The objective of the sub-experiment is to give just as great an opportunity for reinforcement of these minimal due

differences as in the main experiment without reinforcing the instrumental response, and to discover if an animal shows any tendency to respond appropriately on the basis of these cues alone.

VII SUB-EXPERIMENT

A. Procedure

Exactly the same apparatus was used in this experiment as was used in the main experiment. The subjects were 12 male and 12 female albino rats from the same rat colony. They varied in age from 80 to 105 days with a mean age of approximately 95 days. The habituation procedure was exactly the same as the habituation and preliminary training procedure for the naive (Group A) animals in the main experiment. Exactly the same procedure of food ration was employed in order to maintain the same high level of deprivation.

The training period was one day and it came immediately after the habituation period. All animals were placed in the starting box and allowed to run down the maze stem to the choice point which had the door closed. Then the animals were lifted out of the maze stem by the experimenter and placed in either the positive or negative end box and their respective maze arm. Twelve of the animals, 6 of which were female and 6 male, were each placed in the positive goal box and respective maze arm with all doors closed. Three of each group of males and females had the positive goal box on the right and the remaining three of each group had it on the left. All animals were then allowed to eat under the same

(39)

conditions as the animals in the main experiment. Approximately one-half hour later each animal was run down the maze stem to the closed choice point door and was then placed by the experimenter in the negative end box and allowed to remain there with all doors closed for the same time interval as animals in the main experiment. Exactly the same procedure was followed with the remaining 12 animals except that they were placed in the negative end box first and then in the positive goal box.

Approximately one-half hour following the last training trial each animal was given one test trial. This test trial consisted of allowing the animal to run the maze with all doors open. This was a free response trial with the end boxes placed in the same relative positions for each animal as they had been during the training trials. The direction of this free response was recorded.

B. Results and Discussion

The measure used was the number of responses in the direction of the positive goal box for all 24 animals on the test trial. The animals made 10 of these responses out of 24 possible. The chi square of the difference between this measure and the probability of

correct response on the second free training trial for all 92 animals in the main experiment (67 correct responses) is 43.29. This chi square with one degree of freedom is significant at far less than the 1% level of confidence.

The animals in the sub-experimental group, in other words, gave no evidence of correctly performing the instrumental response on the test trial even though the maze cues near the choice point were reinforced in much the same way as they were for animals in the main experiment on their first trial. <u>Therefore, we may</u> <u>conclude that the learning that took place on the first trial in the main experiment was due to direct</u> <u>reinforcement of the instrumental response and not to</u> the approach value of secondary reinforcing cues.

Having removed this objection we may conclude that all of the experimental hypotheses have been clearly supported by the data. In addition, the data have indicated at a high level of confidence that the consummatory response (act of eating food) has no effect on the learning of an instrumental response when the observed food has high secondary reinforcing value.

It then follows, that expectancies defined as fractional-anticipatory-goal responses would not need to be present in the organism for secondary reinforcement

to operate. The naive animals learn on the first trial as well as the animals with expectancy training. This would be in conflict with the assumptions of Denny and Moore (see pages 5 and 12).

It would appear that it is not necessary for the instrumental response to have secondary reinforcing value when all other forms of reinforcement are not immediately contiguous with the making of the response. This finding is exactly opposite to the postulations of Spence and the experimental conclusions of Grice (see page 7). It would appear, therefore, that the primary "gradient of reinforcement" originally postulated by Hull has a very real existence. The "backward action" of the reinforcing state of affairs on preceeding responses and stimuli is presumably mediated by the stimulus trace, and the stimulus trace is the basic unit of the "goal gradient" just as Hull postulates.(19)

(19) P. B., p. 158.

The viewpoint that behavior is specified in terms of a <u>particular learning situation</u>, solely by the action of Hull's goal gradient, is effectively supported. It follows, then, that particular need conditions have far greater specificity than has been heretofore supposed. This further suggests that the degree of

deprivation of a particular need may have a wide range of influence in providing optimum conditions for learning.

The mechanism of learning may be far less complex than has been previously supposed. It is suggested that all learning may be due to the temporal contiguity of a reinforcing situation with particular stimulus traces in the presence of a differentiated need condition, uncomplicated by expectancies. This is essentially the viewpoint presented in Hull's latest postulation concerning primary reinforcement. (20)

(20) Hull, Clark L., <u>Behavior Postulates and Corollaries</u> - <u>November</u>, 1949, p. 2.

VII SUMMARY AND CONCLUSIONS

The subjects were 92 experimentally naive albino rats. The apparatus was a modified T-maze designed to make the time interval between instrumental and eating response as short as possible. Half of these animals were given preliminary training in the maze stem in order to provide expectancies to the experimental apparatus. The other half were given no expectancy training. Each of these preliminary training groups was subdivided into three training groups. One training group had the relative position of the end boxes reversed after the first reinforcement, the second had the end boxes reversed after the second reinforcement, and the third group had no reversal. After eight training trials all animals were extinguished with 12 non-rewarded trials. A very high state of food deprivation was maintained throughout the entire experiment.

A sub-experiment was performed with 24 similar experimental animals to determine if the animals were learning on any other basis besides direct reinforcement of the instrumental response. No evidence of such a learning mechanism was indicated.

The following results were found: 1. The learning of both reversal groups was retarded as compared with the group having no reversal, and the

(44)

group reversed after the second reinforcement showed a greater retardation in learning than the group reversed after the first reinforcement.

2. Preliminary training (expectancy training) gave no better learning than no preliminary training.

3. Those animals having preliminary training ate significantly sooner than naive animals but this time difference does not seem to have any effect on learning when almost immediate secondary reinforcement through the perception of familiar food is available.

On the basis of these experimental findings we may state that there is a very large increment to response tendency after the first reinforcement in this experimental situation, and that <u>expectancies are not</u> <u>necessary to mediate learning</u>. Finally, it may be <u>postulated that all learning takes place as the result</u> <u>of the temporal contiguity of any afferent neural</u> <u>discharges resulting from environmental stimuli with a</u> <u>neural discharge representing reduction in a condition</u> <u>of deprivation</u>.

BIBLIOGRAPHY

1. Grice, G. Robert, "The Relation of Secondary Reinforcement to Delayed Reward in Visual Discrimination Learning", Journal of Experimental Psychology, vol. 38 (1948), p. 1-15.

2. Denny, M. Ray, Unpublished lecture series (untitled), 1950.

3. Hull, Clark L., <u>Behavior Postulates and Corollaries</u> - <u>November</u>, <u>1949</u>, p. 1- 10.

4. Hull, Clark L., "Goal Attraction and Directing Ideas as Habit Phenomenon", <u>Psychological</u> <u>Review</u>, vol. 38 (1931), p. 487-506.

5. Hull. Clark L., "Mind, Mechanism, and Adaptive Behavior", <u>Psychological</u> <u>Review</u>, vol. 44 (1937), p. 1-32.

6. Hull, Clark L., Principles of Behavior, New York: Appleton-Century-Crofts, Inc., 1945. 422 pp.

7. Moore, Dorothy Nell, 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 Dgree of M. A. (Michigan State College, 1949), p. 1-30.

8. Spence, K. W., "The Role of Secondary Reinforcement in Delayed Reward Learning", <u>Psychological Review</u>, vol. 54, p. 1-8, (1947).

9. Watson, John B., The Ways of Behaviorism, New York: Harper and Bros., 1928. 138 pp.

10. Wiener, Norbert, Cybernetics, New York: The Technology Press, 1948. 194 pp.

APPENDIX

Sex, position preference, and initial position of the positive goal box equalization layout for respective experimental groups; column 1 is sex; 2 is position preference; and 3 is the position of the positive goal box on the first training trial:

group:	I-A	II-A	III-A	I-C	II-C	III-C
animal #.	123	123	123	1.23	123	123
"12345678901123456 1123456	FFMMMFFFMMMMFFF	NRLRRLRLRLRL FFLRLLRLRLRL MMMMMFF	F FF MFMMMFFMMM MFFMMM	R L L R L L R R L R R L L R L L R R L L R R L L R R L L R R L L R R L R R L R L R R L R L R R L R R L R R L R R L R R L R R L R R L R R L R R L R R L R R R L R R R L R R R L R R R L R	L R R L R R L L R R L R R L R R L R R L R R L R R L R R L R R L R R L R R L R R L R R R L R R L R R L R R R L R R R L R	M A F A LLLLLLRRRR M F M M F F M F F F F M M

Summary of Animals Discarded from Data

Two animals refused to eat on all trials of the first four days of preliminary training.

Three animals refused to eat on both trials of the last day of preliminary training.

The following animals refused to eat on the first rewarded training trial:

Preliminary Training Gp.	Position of Positive Go	al Box H	Position Preference
A C C A C C A	R L L R R L		R L R R R L

The following animal refused to eat on both trials of the third day in the training period:

Group	Position of Positive Goal	Box	Position Preference	
I-A	R		L	





