LEARNING SET AND HYPOTHESIS THEORY

by

MARGARET JOOST
Department of Psychology

Submitted in Partial Fulfillment of the Requirements of the University Undergraduate Fellows Program

1978-1979

Approved by:

Bruce Bergum, Ph.D.

April 1979

ABSTRACT

Learning Set and Hypothesis Theory. (April 1979)

Margaret Joost, Texas A&M University

Advisor: Bruce Bergum, Ph.D.

The impact of set on problem-solving and summaries of the major theories of learning set are reviewed. One model, hypothesis theory, and relevant studies in the area are presented. The function of memory in the model is examined and a local consistency model is described that conforms with the research data. Two competing views of learning are compared and a test of the strength view versus the all-or-none view is discussed in terms of the blank trials law. In an ambiguous situation, trials with no feedback are not found to be equivalent to trials with positive feedback although equivalent performance is predicted by the blank trials law and all-or-none learning. Complete changes in hypothesis and significantly more errors appear in the no feedback condition as compared to a noncontingent positive feedback condition. This supports the prediction of the strength view that hypotheses of equal strength may compete for dominance in ambiguous circumstances.

DEDICATION

I would like to dedicate this thesis to my husband Victor and to Dyanne Fry and Mark Freeman, our former roommates and good friends.

ACKNOWLEDGEMENTS

My special thanks go to my advisor, Dr. Bruce Bergum, whose guidance has been invaluable. I would also like to extend my sincere thanks to everyone connected with the Undergraduate Fellows Program and to Dr. Glen Taylor, Mrs. Judy Bergum, and Dr. D.G. Barker for their assistance and suggestions. And thanks to my husband Victor for his patience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION AND LITERATURE REVIEW

The concept of set is a relatively old one in psychology and can be exhibited in many ways. Basically, however, the term refers to a tendency to respond to a given situation in a certain way. The phenomenon is not a behavior in itself but is rather a tendency that influences behavior. Thus, for example, problem-solving has often been studied in conjunction with set, the results suggesting that set can inhibit or facilitate problem-solving, depending on the set and the problem conditions. The more alike two problems are, the more a set for solving one problem will transfer to another problem situation. Whether set will facilitate or inhibit problem-solving depends than on how appropriate it is for that situation.

Set and Problem-Solving

An early study on the effect of set on human problem-solving was done by Maier (1931) in which he created a problem by suspending two strings from the ceiling of a room. The subject was instructed to tie the two strings together; however, the strings were too far apart for the subject to

---

The citations on the following pages follow the style of the *Publication Manual of the American Psychological Association*.

hold one and walk over to grasp the other. A pair of pliers on a bench was the only other object in the room. The solution was to tie the pliers onto the end of one of the strings to create a pendulum, swing the pendulum near the other string, and catch the pendulum when it came close enough to the other string. By removing the pliers and tying the strings together the problem is solved.

Only 39% of the subjects were able to solve the problem within ten minutes. Those who did not received a hint from the experimenter when he entered the room to check on the subject's progress: he brushed against one of the strings setting it in motion. Another 38% of the subjects were then able to solve the problem within an average time of less than one minute.

One explanation of this phenomenon involves the concept of functional fixity (Duncker, 1935, 1945) This concept implies that an object with a fixed function in the mind of the observer is less likely to be used creatively than one which has no fixed function. This is illustrated by Adamson's experiment (1952) in which the subject is presented with a box of candles, a box of matches, and a box of tacks and is told to mount the candles on the wall at eye level. The solution is to empty the boxes, tack them on the wall, and mount the candles on them with melted wax. Only 12 of 29 subjects were able to solve this problem, but when the boxes were presented empty along with the other supplies to another group, 24 of 28 solved the problem.

Presumably functional fixity, a type of set, is operating in this problem because the boxes were perceived as containers by the experimental group so that fewer subjects were able to perceive another use for them. The same explanation applies to the discussion of the Maier string problem. Apparently some subjects could not use the pliers as a weight for a pendulum because its customary use is so dissimilar.

The most extensive series of experiments on the conditions leading to and recovery from set was done by Luchins (1942) in his water jar problems. For these problems, the subjects are told they must measure a given quantity of water by using jars which have a fixed capacity. For example, the subjects are first given an example in which 20 quarts of water must be measured using a 29-quart jar and a 3-quart jar. The subjects are given a chance to solve this problem and then are shown the solution: filling the 29-quart jar with water and pouring off three 3-quart jars of water leaves the desired 20 quarts. The problems (see Table 1) are then presented one at a time, and 2.5 minutes are allowed for completing each problem.

Problems 2 through 6 are called the <u>Einstellung</u> (set) problems because they are all devised to be solved by the same method, and thus, to induce a set for that method. Problems 7 and 8 are called the <u>criticals</u> because they can

TABLE 1

WATER JAR PROBLEMS

PROBLEM	JARS			TOTAL
1	29	3	--	20
2	21	127	3	100
3	14	163	25	99
4	18	43	10	5
5	9	42	6	21
6	20	59	4	31
7	23	49	3	20
8	15	39	3	18
9	28	76	3	25
10	18	48	4	22
11	14	36	8	6

be solved by either the set method or by a simpler method. Problem 9 is the <u>extinction</u> problem which cannot be solved by the set method but only by a simpler method. Some subjects in the original experiment were so ingrained in the set method that they were unable to solve the extinction problem. Finally, problems 10 and 11 are also <u>criticals</u>, similar to problems 7 and 8, and the difference in solutions between these two sets of problems is used to measure the effect of the extinction problem on recovery from set.

Luchins concluded that individuals of all age, education, and intelligence groups become set to about the same degree. He also found that set effect can be increased by having the subjects work faster. Increasing the difficulty of the task also increases the set effect. However, minimizing the effect of set can be accomplished by lessening the importance of the method used. Thus, interspersing criticals early in the einstellung problems greatly reduces set.

In general, it has been found (Forgus, 1966, pp. 284-286) that there is no correlation between those who adopt a set in one situation and those who adopt a set in another situation unless the types of set involved are very similar. Finally, no correlation was found between susceptibility to and recovery from set. Typically people are more likely to adopt a set in those situations with which they are unfamiliar and feel less secure. Bergum (1975) has shown that more creative people are less susceptible to the effects of set on problem-solving.

These results suggest that set can result from both organismic and environmental factors. The concern in this paper, however, is mainly with the environmental factors related to set.

Learning Set

The concept of set in discrimination learning in monkeys was studied extensively by Harlow (1949). A Wisconsin General Test Apparatus was used, with the monkey in a cage on one side facing the test situation and the experimenter on the other side behind a one-way vision screen to minimize distraction and possible cue-giving by the experimenter. The first problems investigated were object discrimination problems. Here, the monkey is presented with two different objects, each covering one of the two wells in the test tray. The rewarded object always covered a well with food in it regardless of its randomly varied position. The other object always covered an empty well. After a number of trials with the pair of objects the monkey comes to choose the food-related object almost 100% of the time.

To discover how learning of this type would improve with practice, Harlow presented a series of 344 discrimination problems using 344 different pairs of objects to the same monkeys. He found that over the series of problems the ability to make object-discrimination choices improved dramatically. The percentage of correct choices (see Figure 1) increased from approximately 75% on the first block of problems to nearly 100% on the last block. Harlow designated this learning-to-learn phenomenon as learning set. The monkeys learned how to discriminate the correct choice on the first trial of each problem so that on the second trial they nearly always made the correct choice near the end of the problem series. This is in contrast to the near chance level of performance on trial 2 near the beginning of the problem series.

The form of the learning curve also changes across problems (see Figure 2). For the first few problems the curve appears to show much more trial and error learning than on later problems. For trial 2 performance on the final block of problems the learning curve has nearly reached asymptote and continues almost linearly. Following the formation of a discrimination learning set by the monkeys the problems are usually solved on the first trial and this solution is exhibited in performance on the second trial.

Harlow then trained the same eight monkeys on the more difficult task of discrimination reversal problems. A series of 112 problems of this type were given as before, each for several trials. In discrimination reversal, the animal must first learn a discrimination problem like those

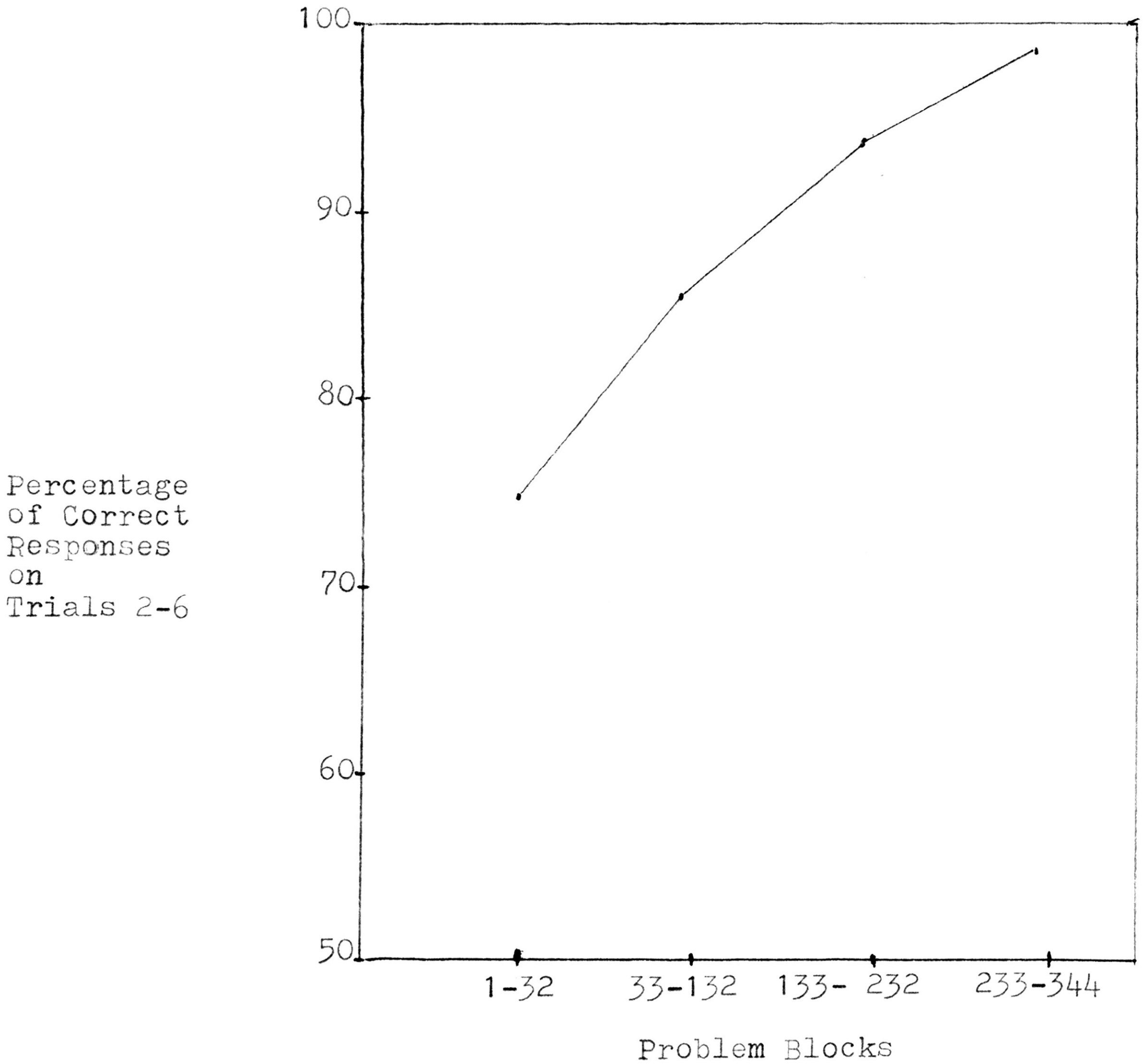

Figure 1. Performance on simple discrimination problems by monkeys across problem blocks (from Harlow, 1949)

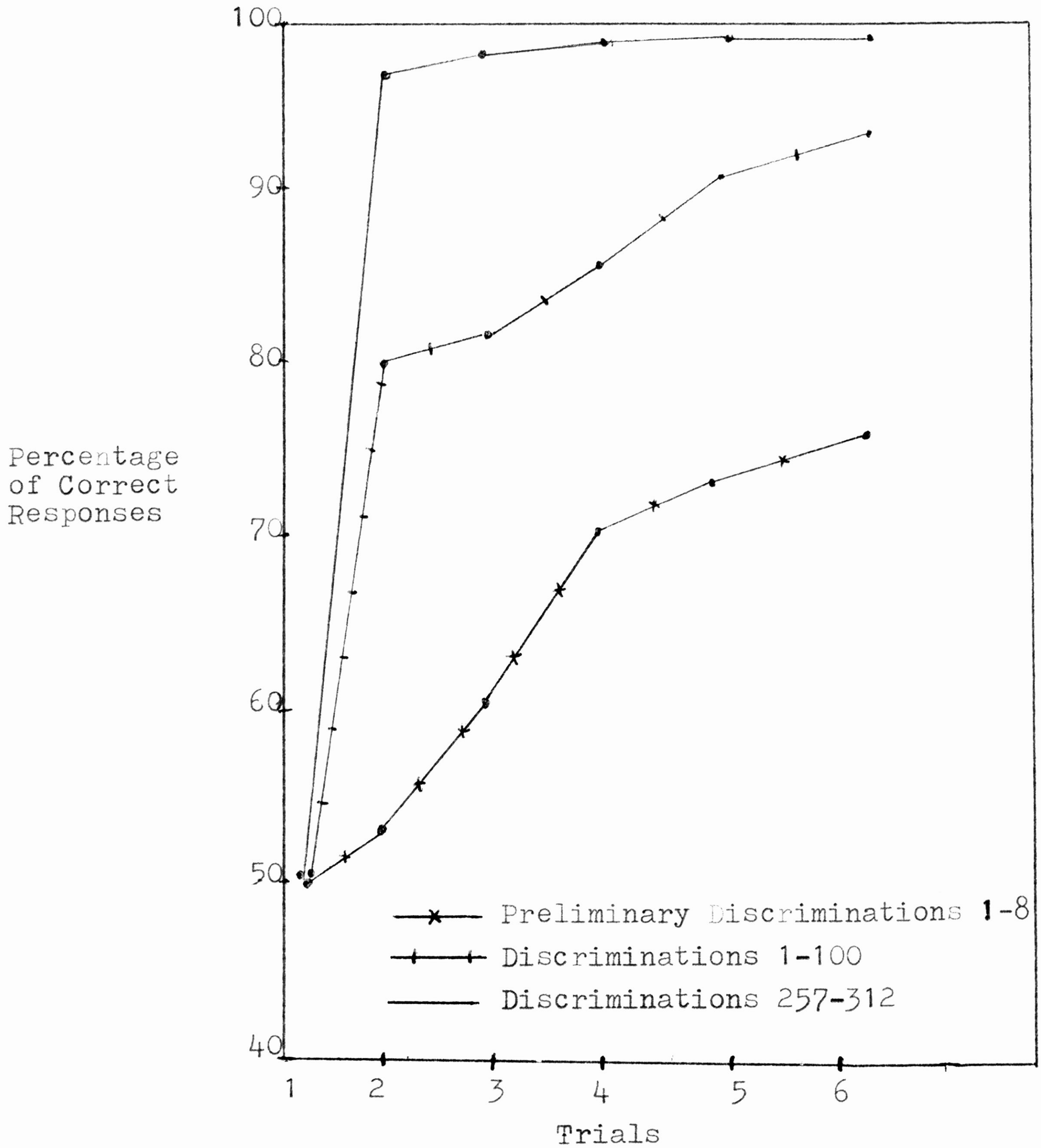

Figure 2. The discrimination learning curves for selected blocks of problems (from Harlow, 1949)

described earlier. But, at some point during training, a reversal of reward is made and the animal must learn to ignore the formerly rewarded object and choose the newly rewarded object. Once again, toward the end of the problem series performance approached 100% correct responses on the trial following the reversal. In fact, the monkeys learned the reversal faster than the original discrimination problem even though reversal is much more difficult. Transfer of training is assumed to account for this occurrence. The point is that the individuals do not learn 344 different discriminations or 112 different reversals but rather a generalized method of solving a discrimination problem. The monkey acquires an increased capacity to respond efficiently in a problem-solving setting. Presumably, learning sets acquired in the natural state would enable the animal to adapt better to those conditions.

Harlow also showed that an inappropriate learning set could inhibit learning. Six monkeys which had previously been trained on object discrimination problems were subsequently given a series of position reversal problems. For each of these problems, 25 trials of object discrimination were run, then, within the same problem, the trials were switched to position discrimination. These subjects, whose only previous experience had been with object discrimination, did well on the prereversal task but had much more difficulty learning the reversal to position discrimination (Figure 3).

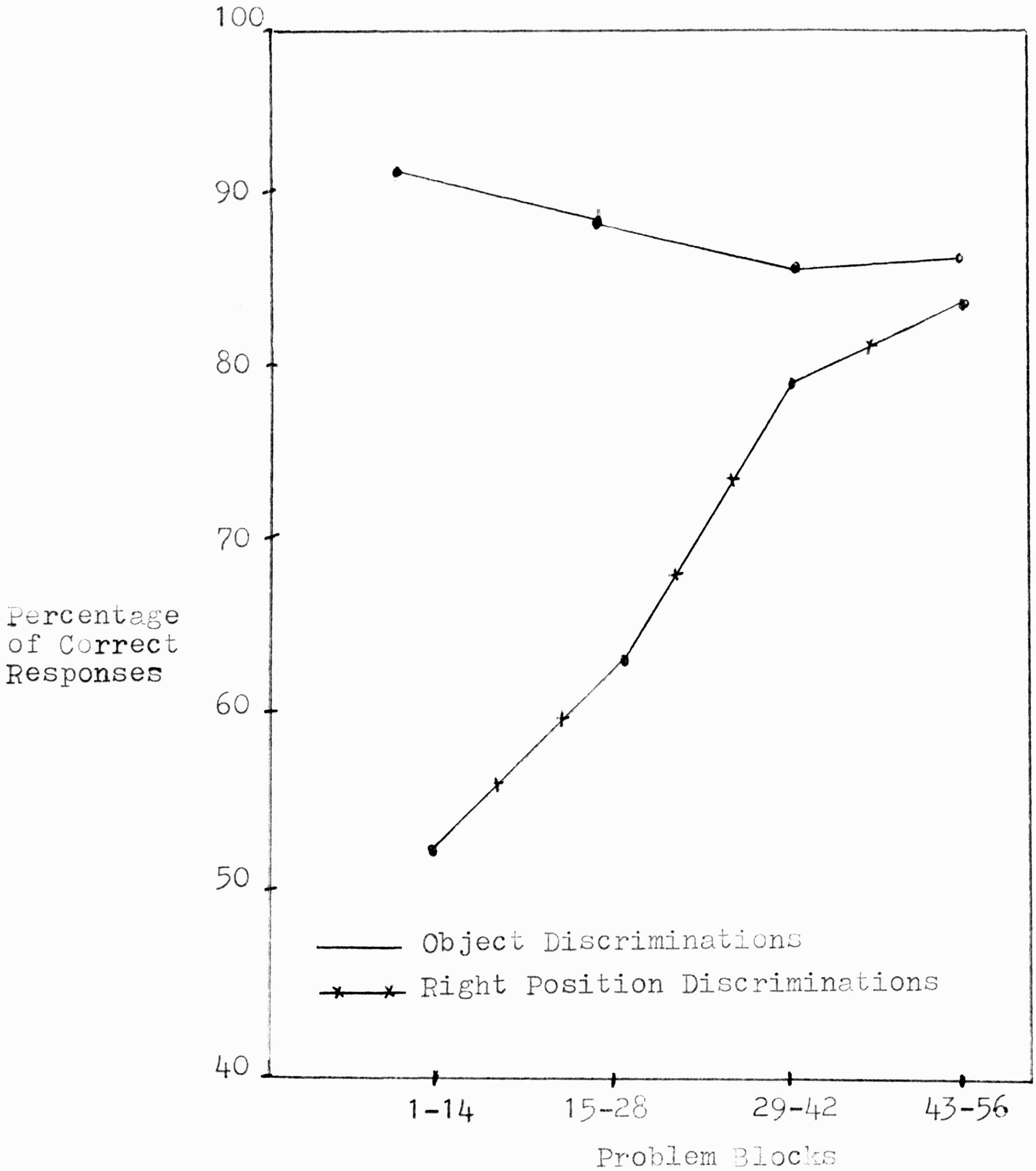

Figure 3. Contrasting learning curves from problems requiring a shift from object discrimination to position discrimination (from Harlow, 1949)

Thus, set can both inhibit and facilitate problem-solving. Facilitation occurs in more similar situations where the set is relevant and inhibition occurs where the transferred set is irrelevant.

Theoretical Considerations

The Hull-Spence theory of learning set, as suggested by Reese (1964), views learning set as a close approach to one-trial learning which occurs after practice on problems with a common solution. The learning curve is correctly predicted, since the large number of new cues in the environment are associated early in learning, while later there are fewer new cues. This predicts the characteristic shape of the typical learning curve. The theory also requires a minimum number of practice problems to bring the habit strengths and inhibitory strengths of all elements in the situation to approximate equivalence, a requirement of one-trial learning. Overlearning on some problems would change these strengths and tend to inhibit learning set until more training trials can equalize them again.

Transfer of learning sets is explained in terms of orienting responses, where each kind of learning set (position discrimination,for example) would require a different orienting response on the part of the animal. In this case, transfer of the response would facilitate learning of similar problems and inhibit learning of different problems.

Finally, in terms of reinforcement, the theory correctly predicts that early in training a single reward has a greater effect than a single nonreward, while later in training the opposite is true. This is due to large increments in habit for early rewards while later the increments have become smaller so that increases in inhibition have a greater effect then.

Reese (1964) speculates that if habit and inhibition ever become asymptotic in value, performance could not rise above chance level. No study has yet found this phenomenon even after extensive training. Furthermore, the direct strengthening of habit by reward can be questioned in that required avoidance of a rewarded object and approach of a nonrewarded object has been demonstrated (McDowell & Brown, 1965a, 1965b). Finally, Levine (1971) has shown that nonlearning of simple discriminations can occur under perfect reward conditions if the subjects are given pretraining on complex types of learning sets. The Spence-Hull theory, then, probably does not give the best explanation of learning set.

Harlow and Hicks (1957) described a theory of learning set called error-factor theory which assumes that there is only one learning process at work in learning set. In this theory, only inhibition is said to occur rather than both habit and inhibition as in the previous discussion. The process is assumed to follow a pattern of suppressing incorrect responses. The subject makes many wrong responses near the beginning of training, which compete with the correct response. Inhibition operates on these incorrect responses so that as learning set is formed only the correct response remains. The error-factors which are the source of the incorrect responses are said to have different initial strengths and some are never completely inhibited. This would apparently be the source of mistakes that occur after learning set is established. Intuitively, the theory seems sensible, particularly in the case of animal studies.

A third theory of learning set is feedback theory (Medin, 1972). According to this position, the individual learns what to expect from each available cue (feedback) and makes the response that will lead to the highest expected feedback. Thus, in this case, it is not learning-how-to-learn that is emphasized but rather learning-what-to-expect from different cues. Rewards and nonrewards do not affect learning directly but they do affect the feedback generated for different cues and thus they increase or decrease the expected value of each choice. Rather than inhibiting incorrect responses or transferring orienting responses, the organism learns to select the correct cues and respond to them.

Improved performance is assumed to be caused by the generalization of feedback to both correct and incorrect objects between problems. Transfer of feedback from rewards and nonrewards are not associated with the particular objects involved, but rather with the problem itself. When initial feedback from anticipated rewards is high, learning improves because the selection of cues becomes more efficient. When feedback for expected reward is low, other cues can compete more effectively for attention and performance is less efficient. For good performance to occur, the correct cue must have a relatively higher reward as an expected feedback. However, blank trials performance in humans, to be discussed later, would seem to discount this theory because no feedback is given and yet performance is maintained.

A mathematical model produced by Restle (1958) predicted changes in the learning curve within problems in learning set. The assumptions of the theory include the idea that the solving of a discrimination problem involves two processes: the adaptation of invalid cues, and conditioning of valid cues to the correct response. Valid cues are called "type-*a*" cues which are consistently reinforced and are contained in all problems of the experiment even when the discrimination objects are changed. These cues are somewhat similar to the feedback discussed above, described as "the property of having been reinforced, as distinguished from other properties such as size, color, spatial shape, arrangement of parts..." (Restle, 1958, p. 79). These latter concrete properties are type-*b* cues which are valid within a problem but change from one problem to another.

Finally, all other cues arising from the environment or from within the animal are type-c cues and are consistently invalid.

Type-c cues become adapted out early in the problem-solving process and learning proceeds quickly. Type-b cues are adapted to more slowly, but when this process is complete the individual responds only in terms of type-a cues. As soon as one of the objects becomes associated with type-a cues (gains the property of having been reinforced) the individual responds correctly for the remainder of that problem.

Restle then suggests equations with which he fits curves to previously reported data and makes some predictions about future research. The theory is not really meant to to give an explanation of underlying events in learning set behavior but rather to quantify the data found.

Hypothesis Theory

Finally, Levine described a model of hypothesis behavior for discrimination learning set. Hypotheses were first described by Krechevsky (1932). He observed position preferences, alternation, and light-going preferences in rats during discrimination problems which he called "hypotheses". A hypothesis is defined as a "specifiable pattern of response to a selected stimulus set" (Levine, 1959, p. 353). Other hypotheses were demonstrated in monkeys by

Harlow (1950), and humans have shown set response patterns in discrimination learning also (Goodnow & Pettigrew, 1955; Goodnow & Postman, 1955; Goodnow, Shanks, Rubinstein, & Lubin, 1957). Bruner, Goodnow, and Austin's (1956) analysis of hypothesis testing included the construction of hypotheses in the learning process, where later theories (Levine, 1959; Bower & Trabasso, 1964) assume that all hypotheses are already known to the individual. For these later theorists the important task in discrimination learning is hypothesis sampling, not hypothesis formation. However, there are cases where the subjects are not aware of all the possible hypotheses and this can lead to non-solution of problems (Glassman & Levine, 1972; Levine, 1971; Levine, Yoder, Kleinberg, & Rosenberg, 1968). It appears that in this case the important behavior is the covert hypothesis selection rather than the overt response or choice made by the individual, except in that it should reflect this internal behavior.

Hypotheses are manifested by different series' of responses across the trials of a problem. Levine (1959) originally postulated nine hypotheses for learning set behavior, each of which could be distinguished by a different series of responses: position preference, position alternation, stimulus preference, stimulus alternation, win-stay/lose-shift with respect to postion, lose-stay/win-shift with respect to position, problem-solving behavior (win-stay/lose-shift with respect to object), problem-solving behavior on a later trial (after trial two), and random response. It is assumed that these hypotheses are mutually exclusive and that two or more never operate on the same problem.

Through examination of the experimental data, it was then determined that position alternation, stimulus alternation, and win-stay/lose-shift and lose-stay/win-shift both with respect to position could be assumed to have zero strength. Position preference was found to have a stable strength (.18) for monkeys trained over a 60, 90, or 120 day period across all problem blocks, so that its only contribution to learning set was inhibition of the correct response. Both stimulus preference and random responding were less for the 120 day trained group and this seems to have produced the greater increase in learning by that group than that for the 60 or 90 day groups. Finally, the problem-solving behavior on trial two or later is much greater for the 120 day group than for the others. Learning set development then is the gradual strengthening of the win-stay/lose-shift strategy with respect to object, by 100% reinforcement of this hypothesis and extinction of most irrelevant hypotheses through random reinforcement. Not all hypotheses seem to extinguish at the same rate, as position preference did not seem to extinguish at all, even after four months of training. Yet, theoretically there is nothing special about the win-stay/lose-shift with respect to object hypothesis. Any hypothesis could be learned, including the seemingly improbable lose-stay/win-shift with respect to object (McDowell & Brown, 1965a).

Following his work on discrimination learning in monkeys, Levine (1966) examined human discrimination learning. In view of his hypothesis theory, Levine assumes that in discrimination learning individuals sample hypotheses from a set of hypotheses known to both the subject and the experimenter. Generally, this set of hypotheses is made known to the individual through the instructions and the pretrain-he or she receives for the experiment. In addition, the subjects are typically instructed to sample only simple hypotheses consisting of only one attribute and to sample only one hypothesis at a time. A typical experimental problem consists of four dimensions (e.g.; color, size, form, and position) with two attributes to a dimension (e.g.; black and white, large and small, triangle and square, left and right). Figure 4 is an example of one trial for a typical hypothesis sampling problem.

Levine originally assumed that memory for hypothesis sampling was perfect and that the subject always sampled from a set of hypotheses that could logically be correct given the information in previous trials. This process had been similarly described earlier by Bruner, Goodnow, and Austin (1956, pp. 129-134) as focussing, or eliminating

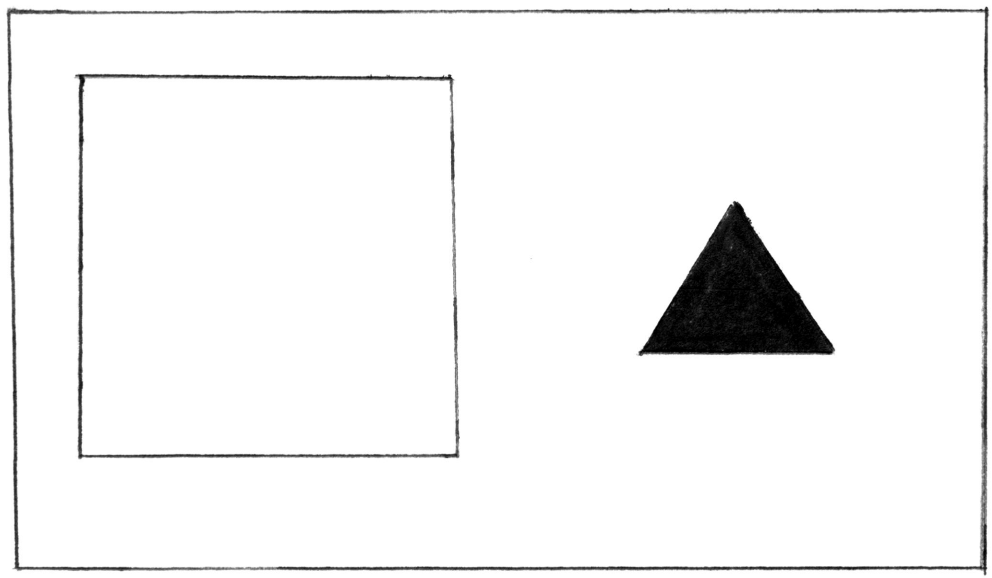

Figure 4. A typical stimulus card for a hypothesis-sampling experiment

all those hypotheses which could not be correct. However, human performance has consistently fallen below that predicted by focussing, and Levine (1966) modified his position.

Another different assumption about memory was made by Restle (1962), who assumed that there was no memory of previous trials or hypotheses when a hypothesis was sampled. A disconfirmed hypothesis was returned to the domain and sampling was always done using the entire hypothesis set. One inference that follows from this is that the probability of sampling a just-disconfirmed hypothesis will be equal to the probability of sampling any other hypothesis in the domain. In other words, if there are k hypotheses in the domain, then the probability of choosing any of them is always 1/k for every hypothesis. However, this has not been supported by findings that the probability of resampling a just-disconfirmed hypothesis is significantly lower than 1/k (Coltheart, 1971; Erickson, 1968; Levine, 1966). Also, the probability of choosing the correct hypothesis increases over trials rather than remaining constant (Erickson, 1968; Levine, 1966; Nahinsky, 1968).

If no memory is involved, it can be assumed that problems given singly or concurrently would produce equivalent performance to solution. However, giving problems concurrently produces a decrement in performance over giving them consecutively (Chumbley, 1969; Restle & Emmerich, 1966). Also, if there is no memory of prior trials, random reinforcement should not affect later performance on a problem, but performance decrements have also been found in this case (Holstein & Premack, 1965; Levine, 1962). Further investigation (Merryman, Kaufman, Brown, & Dames, 1968) has shown that random confirmation produced large decrements in later performance while random disconfirmation had little effect.

Bower and Trabasso (1964) assumed that learning occurs only after an error and that after a correct response the subject simply maintains his hypothesis. Various studies have shown that information processing is facilitated more by a correct trial than an error trial (Kenoyer, 1972; Levine, 1966; Nahinsky, 1968).

In light of these findings a partial memory model was introduced by Trabasso and Bower (1966) called a local consistency model in which the individual remembers the immediately preceding stimulus and response. This stored information is then compared to the current stimulus and response information following an error on the current trial, and if an attribute has consistent response assignments on both trials it remains in the sample set. Kenoyer (1972) predicted from this model that any series of feedback over three trials ending in an error should yield equivalent performances, but this notion was not supported by the data. Performance following the preprogrammed reinforcements was best following three "corrects" and decreased with increasing numbers of "wrongs" given. Frankel, Levine, and Karpf (1970) found that inconsistent patterns during blank trials occurred significantly more often following two initial "wrongs" than two "corrects".

In Levine's model (1966) this decrement in performance would be accounted for by coding errors. The subject codes information about each attribute of the stimulus that is picked and if a "correct" follows the choice then that information becomes the hypothesis set. If the choice is incorrect the complement of those coded attributes becomes the the hypothesis set; however in this case recoding is required which may be imperfect or incomplete and can lead to more errors in later performance. Since time is required for recoding, longer intertrial intervals should lead to better performance and this has been found to be the case (Chumbley, 1969). Coding demands might also account for the longer response latencies that occur after errors than after correct trials (Erickson & Zajkowski, 1967),as well as for the lower confidence ratings for hypotheses following errors than following correct trials (Coltheart, 1973b). Coding demands could also account for the research referred to earlier in which it was found that more information was processed and memory facilitated by a correct trial more by a correct trial than in the case of an error (Kenoyer, 1972; Levine, 1966; Nahinsky, 1968).

Another modification of the local consistency model was offered by Erickson (2968), who noted that a local consistency model with a memory for two hypotheses would best account for his data. It was assumed that the subject could recall only two tested hypotheses at a time and that when a new hypothesis was added, one of the others was lost and returned to the hypothesis domain. Some additional evidence for this exists in the fact that subjects exhibit good recall for at least two of the hypotheses they have tested, although memory for particular stimuli is poor (Coltheart, 1971).

With regard to learning in hypothesis sampling, two views are evident. One is that hypotheses in the set increase or decrease in strength as they are confirmed or disconfirmed and that this accounts for the different probabilities with which a hypothesis is chosen. They are assumed to continue to increase and decrease in strength even after solution. Another view is that hypotheses are sampled from a logical subset of hypotheses, all of equal strength. As hypotheses are disconfirmed, they are discarded from the subset. This is congruent with the idea of focussing and also follows from the coding explanation of what is remembered during problem solving.

Levine (1970) proposed the idea of subset sampling in which the subject narrows down a set of equally probable hypotheses across trials. If the stimuli are arranged orthogonally (each attribute occurs only once on a trial and all possible stimulus combinations are presented) the individual can eliminate half of the hypotheses on the first trial from the domain, half of those remaining on the next trial, and so on until only one hypothesis remains. The responses may be consistent with a single working hypothesis or with the majority of hypotheses in the subset, since it has been shown that subjects can evaluate several hypotheses per trial.

The trial of last error then, is not always the solution trial, although it may be. The subject's working hypothesis may still be contained in a subset of hypotheses being eliminated but if the subject samples only one hypothesis at a time and the correct hypothesis is the final one sampled when all others have been eliminated, then the trial of last error may also be the solution trial.

Levine (1969) reinterprets studies where increases in confidence ratings of hypotheses continue up to and past the trial of final error (Coltheart, 1973b; Falmange, 1968), as well as studies in which response latencies decrease in a similar pattern (Erickson & Zajkowski, 1967; Levine, 1969). These effects are both attributed to the decrease in size of the hypothesis subset which occurs as the subject nears solution. Thus, the trial of last error is not necessarily the solution trial according to Levine (1969). He had subjects ring a bell when they felt sure of their solution and found that response latencies decreased through the trial of last error, but remained constant beyond the solution trial. Fink (1972) found support for Levine's interpretation in that all of the latency decrease for his subjects, following the trial of final error, could be accounted for by those who selected more than one attribute. Those subjects who sampled only one hypothesis on the trial of final error did not show a latency decrease following this trial.

The none-to-all theorem of hypothesis learning (Levine, Miller, & Steinmeyer, 1967) states that the correct hypothesis never occurs before the last error and always occurs after it. This would hold, then, for those sampling the correct hypothesis alone for the first time on the trial of final error and for those who hold the correct hypothesis in a subset of logically possible hypotheses. This theorem also helps explain shy, for orthogonally arranged stimuli, presolution performance is suppressed below chance level to .3 or .4 probability. This is a fairly reliable phenomenon and is apparently due to the fact that not only the correct attribute, but the entire correct dimension, goes almost unsampled prior to solution (Coltheart, 1973a; Glassman & Levine, 1972; Levine, 1971; Levine et al., 1968). A possible explanation was offered by Coltheart (1973a) who suggested that when a hypothesis leads to an error, both the attribute and its entire dimension are rejected. When three values on a dimension were used rather than two (Gumer & Levine, 1971) to check for local consistency as an explanation for this suppression, subjects were found to use both dimension exclusion and local consistency to a degree. In this experiment, when an attribute choice led to an error, the third attribute (not appearing as a choice on that trial) was sampled more often than the other two attributes in the dimension, but not as often as hypotheses from other dimensions.

The second view of learning in hypothesis theory is that the strengths of hypotheses differ prior to the learning situation and that these strengths change due to reinforcement. Subjects' responses are made on the basis of the different strengths of hypotheses in the domain. The data on the decline in response latencies and the increase in confidence ratings of hypotheses already mentioned would support this view if it is assumed that the greater the strength of a hypothesis, the more confidence one would place in it as the solution and the faster one would respond on the basis of it.

Falmange (1968) found an increase in ratings of the correct rule and a decrease in ratings of irrelevant rules over trials. Further, it was found (Falmange, 1972) that the strength of a just-disconfirmed hypothesis falls to zero and never returns to the strength of an undisconfirmed hypothesis. The strength of a just-confirmed hypothesis increases with consecutive confirmations. Coltheart (1973b) found that confidence ratings increased on correct trials but decreased after a series of errors. The strength position attributes all-or-none learning to the artificial problem setting in which a choice must be made, but strengths cannot be indicated. Since it is the only way learning can be exhibited in this kind of situation, it is natural that it would be interpreted as all or none. However, the strength position argues for an underlying mechanism of interrelated strengths of hypotheses due to prior reinforcement and/or innate preferences.

The presolution response suppression discussed earlier appears to be similar to the behavior on unsolved and insoluble problems in that the correct hypothesis is not sampled Levine, et al., 1968). Nonlearning can be produced artificially by removing the relevant attribute from the stimulus problem, or by training the subjects to sample from a very large irrelevant set of hypotheses (Levine, 1971). In either case, the subject is unable to select the correct hypothesis because it is not in the domain from which he is sampling, so that regardless of feedback or the simplicity of the problem it will not be solved. Training on a large or complex set of hypotheses can produce nonlearning because the set is difficult to exhaust and so the behavior of sampling from it is difficult to extinguish once it has been reinforced during training. These data could present an explanation of the actions of set in problem-solving. The subject does not consider hypotheses not in the immediate domain and thus, problems with a simple solution may not be solved if the subject is sampling from a large complex domain of hypotheses irrelevant to the problem (Levine, 1971).

In the area of switching hypothesis domains, it has been found that subjects using more complex domains are more likely to switch to another domain of complex hypotheses than to a set of simpler ones (Lane, McDaniel, Bleichfeld, & Rabinowitz, 1976). The study found that instructions can aid in switching domains of hypotheses. Referring to the discussion of set in problem-solving it is apparent that this theory is helpful in explaining the inhibiting effects set can have on learning.

Another contribution to the possible inhibiting effect of set is evident in the finding that when a domain of hypotheses is exhausted it is usually resampled before a switch to a new domain is made (Erickson, 1968; Glassman & Levine, 1972; Levine, 1971). Levine (1971) also found that subjects learn nothing about hypotheses not in their domain. When given a multiple choice test following an unsolved simple discrimination problem which had been preceeded by problems with complex position sequences as solutions, most of the subjects chose a position sequence as the correct rule rather than the simple discrimination.

Random reinforcement has been shown to inhibit learning (Holstein & Premack, 1965; Levine, 1962) and partial reinforcement also has an effect. Levine proposed that solutions experienced on early problems suggest the class of solutions for later problems. Taddonio and Levine (1975) showed that subjects who experienced partial reinforcement began the next problem by searching for a hypothesis in a more complex domain. The research on how subjects switch hypothesis domains is helpful in explaining how set can inhibit or facilitate problem-solving.

While the effects of feedback on hypothesis sampling were discussed earlier, some final aspects remain to be included here. The no-feedback condition, or blank-trials as it is called, is used by hypothesis theory researchers to determine which hypothesis an individual is relying on at any time during the problem. Following one or more trials with feedback the subject responds on a few trials where no feedback is given. The pattern of responses on these trials is used to determine the hypothesis the subject was relying on at that time. If the pattern is inconsistent with any simple hypothesis it is called an error. Levine (1966) first formulated the "oops" error which he defined as a perceptual or motor slip on a trial and which had a probability of about .02. These errors can occur on blank-trials or other feedback conditions. Other researchers have found the probability of error to be somewhat greater than .02 (Chumbley, 1969; Coltheart, 1971).

Another form of response which was found to account for part of Levine's 1966 error data is called majority rule (Frankel, et al., 1970). In this case, two or more hypotheses are considered by the individual and he responds in a manner consistent with the majority of the hypotheses held. If both stimuli correspond to an equal number of the hypotheses the subject responds randomly. Majority rule was introduced to account for the fact that as blank-trials were extended to 30 consecutive trials, inconsistent response patterns also increased. Levine (1970) reanalyzed his 1966 data and found that 38% of what had been labeled "oops" errors could be accounted for by majority rule. However, when Harpur (1976) tested for majority rule it was not found and he concluded that earlier subjects were testing conjunctive hypotheses instead of responding with respect to the majority.

The behavior of blank trial and non-blank trial subjects has not been found to differ to any great extent (Aiken, Santa, & Ruskin, 1972; Coltheart, 1973a; Karpf & Levine, 1971; Kenoyer, 1972) although one study (Karpf & Levine, 1971) found blank-trials subjects to be more variable in their behavior. This study also investigated verbal reports of hypotheses versus inferring hypotheses from patterns of blank-trial behavior and no differences in the two techniques were found.

The basic assumptions behind the blank-trials procedure is that the subjects always respond on the basis of some hypothesis and that no outcome produces no resampling by the subject. If a subject does not resample following a blank trial, but retains his hypothesis, then the blank-trial is functionally equivalent to a correct trial. Levine has repeatedly shown that for a series of four blank-trials, 90% of the hypothesis patterns produced correspond to one of the eight simple hypothesis patterns for a four-dimensional problem (Andrews, Levinthal, & Fishbein, 1969; Levine, 1966; Levine, Miller, & Steinmeyer, 1967). When the number of blank trials is increased to 30 consecutive blank-trials, the number of perfectly consistent hypothesis patterns decreases to between 43% and 85% (Frankel, et. al., 1970). When the number of blank-trials is increased from four to seven performance again decreases (Chumbley, 1969).

Aiken, Santa, and Ruskin (1972) found support for the idea that blank-trials are equivalent to success trials in that the mean number of called errors was similar for a 50% called reinforcement group and for a 100% called reinforcement group. If subjects had resampled on some of the nonreinforced trials in the 50% group, a greater number of called errors would have been expected. Howevever, Coltheart (1973a) has found some evidence that blank-trials are not entirely equivalent to success trials. Longer latencies following an error just preceding blank-trials persists over all four blank trials whereas if a blank-trial had the same effect as a success trial, the latencies on blank-trials

2-4 should be shorter.

Current theories of learning in hypothesis theory group around two major positions. The strength view assumes that hypotheses exist with different interrelated strengths and that they are sampled from the domain in accordance with their various strengths. These strengths or probabilities of being selected can be changed through learning and these changes are reflected in the individual's responses. This view is supported by studies that show different probabilities for different hypotheses in humans and other organisms. The response latency and confidence rating data can also be interpreted to support this view. Performance decrements in the solutions of concurrent versus consecutive problems might also be attributed to differential changes in hypothesis strength in the two conditions. However, blank-trials performance is somewhat more difficult to explain from this theoretical position. It does not seem likely that a no-feedback trial would affect hypothesis strength in the same way as a positive feedback trial.

The all-or-none view of learning suggests that all hypotheses are at full and equal strength in the domain. Coding demands and subset sampling are introduced to explain hypothesis selection. Blank-trials performance is dealt with by defining no feedback as equivalent to positive feedback in that it does not cause the individual to resample hypotheses from the domain. All-or-none learning may be an adequate interpretation of the behavior of trained individuals whose hypotheses have been well learned, but for those subjects still in the training period strengths of hypotheses would seem to be more important.

Since the blank-trials law follows from all-or-none learning and would not be predicted by the strength view, a test of this law might be helpful in resolving the conflict between the two views. In the original work on blank-trials, identical stimuli were repeated for all blank trial tests of each problem. This repetition may have influenced the subjects' responses. In order to avoid repetition of the same stimulus materials and yet be able to give several sets of blank trials more complex problems should be used.

In a situation where two hypotheses are brought to equal strength and the subject is given no feedback as to which one is correct, these hypotheses should compete for dominance and the individual may switch from one to another. However, the all-or-none view of learning makes the prediction that hypothesis switches will not occur unless negative feedback is given. The blank-trials law sets forth the equivalence of positive feedback and no feedback in their effects on performance. The purpose of the present study was to test whether the blank-trials law will hold for an ambiguous situation where two hypotheses are brought to equal strength by equivalent feedback and then either no feedback or noncontingent positive feedback is given to differentiate them.

METHOD

Subjects

The subjects were 16 female and 16 male Texas A&M University undergraduates from introductory psychology classes. Half of the males and half of the females were randomly assigned to one of the two experimental conditions. One additional male did not complete the experiment due to experimenter error.

Apparatus

Ten complex problems were devised, each consisting of stimuli with six dimensions and two attributes to a dimension. Table 2 illustrates the dimensions, attributes, and the arrangement of trials such that, with certain exceptions, each hypothesis had a distinguishable pattern of responses every four trials. Patterns of feedback are also illustrated.

Each problem consisted of 32 trials and different problems were distinguished by different letters and colors. Other attributes perseverated across problems. Each trial consisted of two stimuli drawn approximately 3 cm. apart on a white unlined index card. The letters were either 1 inch (2.54 cm) or .5 inch (1.27 cm) tall. The underline was 2 cm long and the outline enclosing the letter and its underline was either a circle 4 cm in diameter or a

TABLE 2

REPRESENTATION FOR THE LEFT STIMULUS ONLY OF AN EIGHT TRIAL SEQUENCE DESIGNED SO THAT THE PATTERN OF RESPONSES INDICATES THE HYPOTHESIS IN USE

EXPERIMENTAL	CONTROL	TRIAL	LETTER	COLOR	SIZE	OUTLINE	UNDERLINE	POSITION
Correct Feedback	Correct Feedback	1	A	Red	Large	Circle	Solid	Left
		2	V	Blue	Large	Square	Solid	Left
		3	A	Red	Large	Square	Dotted	Left
		4	V	Blue	Small	Square	Dotted	Left
No Feedback	Noncontingent Positive Feedback	5	A	Blue	Large	Square	Solid	Left
		6	V	Red	Small	Square	Solid	Left
		7	V	Red	Large	Circle	Solid	Left
		8	A	Blue	Small	Circle	Dotted	Left

square 4 cm on a side. All dimensions were represented and each stimulus on a trial contained only one of the attributes of each dimension. No two trials were identical.

All trials of each problem were arranged in a particular way. Two attributes were randomly chosen to be possible answers for each problem, with the stipulation that both were not on the same dimension. The two always coincided on half of the trials of a problem and were separated on the other half. All problems were arranged beginning with a set of four trials where the two attributes coincided and alternating with sets of four trials where they were separated. Thus the patterns of the two chosen hypotheses for each problem was identical for alternate sets of four trials and exactly opposite on the remaining sets of four trials. Four of the problems were randomly chosen to be experimental and four to be control. The remaining two were used as practice problems.

Procedure

Each subject was tested individually and was seated across the table from the experimenter, facing the deck of practice problems. The individual received oral instructions from the experimenter that each problem had only one answer which remained unchanged throughout the problem and that feedback would usually, but not always, be given following the response to a trial. Instructions were also given on how to respond by indicating the left or right stimulus on a trial, and what the possible answers could be. The subject was then given an opportunity to solve two practice problems during which he also received some experience with blank trials.

The deck of four experimental and four control problems was then given to the individual and his response to each trial was recorded by the experimenter. For all trials, the subject was allowed to proceed at his own rate and, after giving a response to a card, he turned the card face down out of the way.

All subjects completed four experimental and four control problems in order to control for differences in ability. Experimental and control problems were alternated to control for practice effects. Finally, for half of the males and half of the females, the experimental problems were switched to control and the four control problems to experimental to account for possible differences in difficulty of problem.

For experimental problems, the subject was given the correct feedback on the first four trials and alternate sets of trials where the two chosen possible hypotheses coincided. On the remaining sets of four trials, where the hypotheses were separated, no feedback was given. For the control problems, the only difference was in feedback. The subject was given noncontingent positive feedback where the two possible hypotheses were separated instead of no feedback as was the case in the experimental problems. That is, the experimenter gave the subject positive feedback on those trials regardless of the subject's response.

RESULTS

The problems were scored using a criterion of solution of four consecutive correct trials with response-contingent feedback. When this criterion was reached on a problem, the number of the first correct trial following the trial of final error prior to solution was recorded. It was then determined when a correct hypothesis was first used consistently by examining the first complete set of blank-trials, or noncontingent positive feedback trials for control problems, following the trial of final error. This determined the "correct" hypothesis of the two possible hypotheses for each problem and any responses inconsistent with it were scored as errors. A total of 146 errors was observed for the experimental group, while a total of 16 errors was observed for the control group. To test whether the greater number of errors for the experimental group was significant, a Wilcoxon's Signed Rank Test was performed on the data. Significantly more errors were found for the experimental problems ($p < .005$).

In addition, there were 10 complete changes in hypotheses on the blank trials of the experimental problems and no consistent changes on the control problems. These changes were measured by comparing the first completely consistent set of blank-trials following the trial of final error to any remaining sets of blank-trials and a hypothesis switch was defined as a set of four trials where the subject relied on the other possible hypothesis for that problem.

DISCUSSION

The results supported the prediction of the strength view of learning that no feedback and positive feedback do not have an equivalent effect on performance. In that there were significantly more errors in the no-feedback condition and also 10 complete hypothesis switches compared to no switches in the noncontingent positive feedback condition, the blank-trials law prediction of no differences was contradicted. Since the blank-trials law follows from all-or-none learning, this view has not been supported. The strength view of learning would predict competition between two hypotheses brought to equal strength by equivalent feedback. Most of the errors on the experimental problems were due to the complete hypothesis switches on the blank trials. Yet there were only 10 switches on 128 total experimental problems given. This could be the result of greater differences in the strengths of hypotheses than could be equalized by the feedback from one problem.

The results of the study are not consistent with most of the work on blank-trials however this is probably due to the differences in methodology. One of the arguments of the strength view of learning is that in choice situations the relationships of strengths between the hypotheses are not easily represented. In these problems, the methodology was such that various hypotheses were allowed to compete for dominance and the difference in performance on the experimental and control problems showed that changes in strengths in hypotheses can be seen even in this limited situation.

The significance of hypothesis theory does not lie in merely whether hypothesis strengths are changeable. The model is important in explaining how a learning set can be produced and how sets inhibit and facilitate learning. Problems in learning that can occur in sampling hypotheses and switching domains should be considered when creative thinking and problem-solving is required.

REFERENCES

Adamson, R. E. Functional fixedness as related to problem-solving: A repetitition of three experiments. *Journal of Experimental Psychology*, 1952, *44*, 288-291.

Aiken, L. S., Santa, J. L., & Ruskin, A. B. Nonreinforced trials in concept identification: Pre-solution statistics and local consistency. *Journal of Experimental Psychology*, 1972, *93*, 100-104.

Andrews, O. E., Levinthal, C. F., & Fishbein, H. D. The organization of hypothesis-testing behavior in concept-identification tasks. *American Journal of Psychology*, 1969, *82*, 523-530.

Bergum, B. O. Self-perceptions of creativity among academic inventors and non-inventors. *Perceptual and Motor Skills*, 1975, *40*, 78.

Bower, G., & Trabasso, T. Concept Identification. In R. C. Atkinson (Ed.) *Studies in mathematical psychology*. Palo Alto, California: Stanford University Press, 1964.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking*. New York: Wiley, 1956.

Chumbley, J. Hypothesis memory in concept learning. *Journal of Mathematical Psychology*, 1969, *6*, 528-540.

Coltheart, V. Memory for stimuli and memory for hypotheses in concept identification. *Quarterly Journal of Expermental Psychology*, 1971, *89*, 102-108.

Coltheart, V. The effects of acoustic and semantic similarity on concept identification. *Quarterly Journal of Experimental Psychology*, 1972, *24*, 55-65.

Coltheart, V. Concept identification with and without blank-trials. *Quarterly Journal of Experimental Psychology*, 1973, *25*, 1-9. (a)

Coltheart, V. Confidence ratings as a response index in concept identification. *Journal of Experimental Psychology*, 1973, *97*, 46-50.

Duncker, K. *Zur Psychologie des Produktiven Denken*. Berlin: Springer, 1935.

Duncker, K. On problem-solving. (Tr. L. S. Lees) *Psychological Monographs*, 1945, *58*, No. 270.

Erickson, J. R. Hypothesis sampling in concept identification. *Journal of Experimental Psychology*, 1968, *6*, 12-18.

Erickson, J. R., & Zajkowski, M. M. Learning several concept identification problems concurrently: A test of the sampling with replacement assumption. *Journal of Experimental Psychology*, 1967, *74*, 212-218.

Falmange, R. A direct investigation of hypothesis-making behaviors in concept identification. *Psychonomic Science*, 1968, *13*, 335-336.

Falmange, R. Memory process in concept identification. *Journal of Experimental Psychology*, 1972, *92*, 33-42.

Fink, R. T. Response latencies as a function of hypothesis-testing strategies in concept identification. Journal of Experimental Psychology, 1972, 95, 337-342.

Forgus, R.H. Perception. New York: McGraw-Hill Boook Company, 1966.

Frankel, F., Levine, M., & Karpf, D. Human discrimination learning: A test of the blank-trials assumption. Journal of Experimental Psychology, 1970, 85, 342-348.

Glassman, W. E., & Levine, M. Unsolved- and insoluble-problem behaviors. Journal of Experimental Psychology, 1972, 92, 146-148.

Goodnow, J. J., & Pettigrew, T. F. Effect of prior patterns of experience upon strategies and learning sets. Journal of Experimental Psychology, 1955, 49, 381-389.

Goodnow, J. J., & Postman, L. Learning in a two-choice probability situation with a problem-solving setting. Journal of Experimental Psychology, 1955, 49, 16-22.

Goodnow, J., Shanks, B., Rubinstein, I., & Lubin, A. What is the human subject responding to in a two-choice task? WRAIR Problem-solving memo., 1957, No. 3.

Gumer, E., & Levine, M. The missing dimension in concept learning: Dimensionality or local consistency? Journal of Experimental Psychology, 1971, 90, 39-44.

Harlow, H.F. The formation of learning sets. Psychological Review, 1949, 56, 51-65.

Harlow, H.F. Analysis of discrimination learning by monkeys. Journal of Experimental Psychology, 1950, 40, 26-39.

Harlow, H. F., & Hicks, L. H. Discrimination learning theory: Uniprocess vs. duoprocess. Psychological Review, 1957, 64, 104-109.

Harpur, J. G. Hypothesis testing in concept identification within a modified selection paradigm. Psychological Reports, 1976, 39, 1063-1066.

Holstein, S. B., & Premack, D. On the different effects of random reinforcement and pre-solution reversal on human concept identification. Journal of Experimental Psychology, 1965, 70, 335-337.

Karpf, D, & Levine, M. Blank-trial probes and introtacts in human discrimination learning. Journal of Experimental Psychology, 1971, 90, 51-55.

Kenoyer, C. E. Experimental tests of models that assume local consistency and resampling only on error trials. Journal of Experimental Psychology, 1972, 92, 215-219.

Krechevsky, I. "Hypothesis" versus "chance" in the pre-solution period in sensory discrimination learning. University of California Publications Psychology, 1932, 6, 27-44.

Lane, D. M., McDaniel, J. R., Bleichfeld, B. E., & Rabinowitz, F.M. Toward the specifications of hypothesis domains and einstellung. Journal of Experimental Psychology, 1976, 2, 489-496.

Levine, M. A model of hypothesis behavior in discrimination learning set. Psychological Review, 1959, 66, 353-356.

Levine, M. Cue neutralization: The effects of random reinforcements upon discrimination learning. Journal of Experimental Psychology, 1962, 63, 438-443.

Levine, M. Hypothesis behavior by humans during discrimination learning. Journal of Experimental Psychology, 1966, 71, 331-336.

Levine, M. The latency-choice discrepancy in concept learning. Journal of Experimental Psychology, 1969, 82, 1-3.

Levine, M. Human discrimination learning: The subset-sampling assumption. Psychological Bulletin, 1970, 74, 397-404.

Levine, M. Hypothesis theory and nonlearning despite ideal S-R reinforcement contingencies. Psychological Review, 1971, 78, 130-140.

Levine, M., Leitenberg, H., & Richter, M. The blank-trials law: The equivalence of positive reinforcement and nonreinforcement. Psychological Review, 1964, 71, 94 -103.

Levine, M., Miller, P., & Steinmeyer, C. H. The none-to-all theorem of human discrimination learning. _Journal of Experimental Psychology_, 1967, _73_, 568-573.

Levine, M., Yoder, R. M., Kleinberg, J., & Rosenberg, J. The presolution paradox in discrimination learning. _Journal of Experimental Psychology_, 1968, _77_, 602-608.

Luchins, A. S. Mechanization in problem-solving: The effect of Einstellung. _Psychological Monographs_, 1942, _54_ (6, whole No. 248).

Maier, N. R. F. Reasoning in humans. II. The solution of a problem and its appearance in conciousness. _Journal of Comparative Psychology_, 1931, _12_, 181-194.

McDowell, A. A., & Brown, W. L. Further evidence for the excitatory properties of nonrewarded cues. _Journal of Genetic Psychology_, 1965, _106_, 137-140. (a)

McDowell, A, A., & Brown, W. L. Learning mechanisms in response shift learning set to non-rewarded cues. _Journal of Genetic Psychology_, 1965, _106_, 173-176. (b)

Medin, D. L. Role of reinforcement in discrimination learning set in monkeys. _Psychological Bulletin_, 1972, _77_, 305-318.

Merryman, C., Kaufman, B., Brown, E., & Dames, J. The effects of "rights" and "wrongs" on concept identification. _Journal of Experimental Psychology_, 1968, _76_, 116-119.

Nahinsky, I. D. A test of axioms of all-or-none concept identification models. _Journal of Verbal Learning and Verbal Behavior_, 1968, _7_, 593-601.

Reese, H. W. Discrimination learning set in rhesus monkeys. _Psychological Bulletin_, 1964, _61_, 321-340.

Restle, F. Toward a quantitative description of learning set data. _Psychological Review_, 1958, _65_, 77-91.

Restle, F. The selection of strategies in cue learning. _Psychological Review_, 1962, _69_, 329-343.

Restle, F., & Emmerich, D. Memory in concept attainment: Effects of giving several problems concurrently. _Journal of Experimental Psychology_, 1966, _71_, 794-799.

Taddonio, J., & Levine, M. Hypothesis theory and the PREE. _Bulletin of the Psychonomic Society_, 1975, _5_, 74-76.

Trabasso, T., & Bower, G. H. Presolution and dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. _Journal of Mathematical Psychology_, 1966, _3_, 163-173.