

AD No. 23 524
ASTIA FILE COPY

EXPERIMENTAL ANALYSIS OF PROBLEM-SOLVING BEHAVIOR

Howard H. Kendler

Technical Report I

Office of Naval Research Project NR 150-064

Contract Nmr 187 (00)

Department of Psychology

University College of Arts & Science

New York University

New York, N. Y.

November 15, 1953.

BEST AVAILABLE COPY

Preface

The author of this report wishes to express his appreciation to May D'Amato, Sanford Katz, Sol Klier, Stanley Pliskoff and Robert Vineberg, all of whom served at various times as graduate research assistants on this project.

Chapter I

Introduction

At the time this research program was initiated, its aim was essentially exploratory in nature. It sought to determine the influence of some variables upon the development and expression of a specific 'mental set' (9). This program of research was dependent to some extent on a given research strategy, a discussion of which is necessary before reporting the experimental findings.

In such an undeveloped and unstructured field as the psychology of thinking, the selection of a hypothesis to test, as well as the experimental situation in which to test it, is for all practical considerations a matter of taste. And, as is the case with matters of taste, one is not justified in defending his preference as being correct. Rather, one is forced to recognize, assuming the necessary degree of insight is present, the arbitrary nature of one's predilection.

Scientific tastes, like other personal tastes, have important behavioral consequences and therefore this writer feels that it would be helpful for all research workers to make explicit the nature of the preferences which dictated their work. Perhaps if this were done it would enable psychologists to recognize the relationship between a particular research strategy and the fruitfulness, or lack of it, of the research stemming from it.

With such an introduction, there appears to be no alternative but to try to set down the decisions which determined the execution of the experiments included in this report. These decisions are not logically interrelated and hence they are presented in a casual order.

1. In a field such as thinking, which possesses so many personal qualities, the research worker cannot avoid the temptation to tailor his research program so that it will reflect "thinking" as it "actually occurs," whether it be in such scholarly areas of mathematical reasoning and scientific theorizing or more common situations such as parlor room problems or mechanical puzzles. The result of yielding to such temptation has usually been the imposition upon one's research techniques, as well as the theoretical constructs associated with them, preconceived ideas which stem from non-investigative sources such as one's own introspective experience, cultural pressures, philosophical orientations, etc. It should be remembered that whether or not a scientist succumbs, consciously or unconsciously, to such temptations, is not in any sense right or wrong. Nevertheless, we should attempt to recognize the consequences of such a decision, because of the impact such decisions have upon the techniques used to gain knowledge, as well as the knowledge itself. It is this writer's feeling, in spite of his own history of submitting to the above described temptation, that the most strategic research program in the field of human problem solving behavior is one that resists such temptations by restricting itself to investigating empirical relationships in well defined and effectively controlled situations, and from the data obtained to cautiously abstract experimentally meaningful concepts and theories. Constructions derived from experimental events, it is felt, are likely to be more fruitful both for theory and application, whereas those imposed upon events by cultural and personal factors will tend to be misleading and sterile. It will be seen that this decision influenced the collection of most of the data included in this report.

2. The present research program was motivated by an attempt to extend conditioning theory to the realm of problem solving behavior. This decision seems natural considering the writer's addiction to conditioning principles in his work in the field of learning. It should be recognized that this decision has important implications (i.e., "entailed decisions"). By failing to recognize any hiatus between simple associative learning and problem solving behavior, one's explanations must not only be consistent with the results of the specific thinking experiment but also with the mass of data collected in the field of conditioning and learning. This is obviously a more difficult task. Secondly, the decision of working within a conditioning framework raises many problems, both specific and general, associated with the extension of principles abstracted from one empirical area to another. There are many fine nuances to this problem, a full discussion of which would require space beyond the limits of the present report. It may be sufficient at present to merely mention two points; the first being that an attempt to extend conditioning principles of thinking behavior does not imply necessarily that each relationship discovered in conditioning will be simply reflected in a mirror-like fashion in problem solving behavior. For example, the facilitating effect increases in the drive variable have upon performance in the conditioning situation would not in most cases be expected to occur in the problem solving situation. The correct response, at the beginning of the thinking experiment, is low in the response hierarchy. If the assumption that behavior is some multiplicative function of learning and motivation (4,5) is valid, increasing drive while other factors remain constant would actually retard the appearance of the correct response since it would increase the difference between the response strength of the dominant incorrect response tendency and the correct response tendency. The second point, in connection with the extension of conditioning principles to problem solving behavior, is that this sort of an attempt does not necessarily mean that the same empirical variables which have been demonstrated to be important in conditioning will also be relevant to the problem solving situation. For example, the amount of work (4,6) is highly correlated with rate of extinction. The question is whether increased work would also operate to weaken "neutral sets." Some of our research sought to answer this question. The results were mostly negative but, as we shall discuss later, such findings do not mean the application of conditioning theory to problem solving behavior is destined to fail.

3. The third decision undertaken in the course of this research program was the adoption of a "hit and run" research tactic. Unless a given study or a given series of interrelated studies provided unequivocal results, the research technique was abandoned. It was felt that in such an undeveloped area as the psychology of thinking, it would be more strategic to deal with experimental designs, the variables of which have obvious effects upon behavior. There is always the danger when one abandons an experimental hypothesis and the techniques associated with the testing of it that a little more persistence and imagination might have uncovered significant and important relationships. Perhaps more patience should have been demonstrated. The writer, however, had the uncomfortable feeling that many of the specific problems that concerned him were leading to blind alleys.

Before the results of our research program are reported, it is appropriate to describe briefly the plan for their presentation. This report will consist of 5 chapters, of which 4 will report the studies of a particular problem area. These "empirical" chapters will consist of a general introduction followed by reports of the individual experiments and will conclude with a general discussion. Because the experiments are essentially exploratory in nature, they will for the most part be described briefly and informally. A full technical account of all the research would make the report overly long and cumbersome. Those readers who might be interested in a more detailed description of the procedure or the results of any specific experiment can receive the necessary information upon request.

Chapter II

Tests of the Extinction Hypothesis

General Introduction

The experiments reported in this chapter stemmed directly from the extinction hypothesis offered by Kendler, Greenberg and Richman (6) to account for their result that massed practice was superior to distributed practice in establishing a mental set. According to these authors the learning of a mental set mainly reflects that portion of the learning process in which distributed practice is "inferior" to massed practice; viz. experimental extinction (14, 15). The stronger mental set found in their massed group would be attributed to the assumption that the non-set responses of the massed group were weakened more rapidly and completely than those of the distributed group. Consequently, the massed group would have a stronger mental set.

The extinction hypothesis states essentially that factors which facilitate experimental extinction can also facilitate the development of mental sets. This formulation does not deny the influence of other variables (e.g., personality, perceptual orientation, etc.) on human thinking; it merely states that "extinction variables are among the numerous variables functionally related to thinking behavior." The particular extinction variable for investigation in problem solving behavior was the work variable.

Experiment 1

The purpose of this investigation was to observe the influence of work upon the development and expression of a mental set. It has been demonstrated that increased work facilitates the extinction process (16) and consequently it should be expected, according to the extinction hypothesis, that increased work should increase or decrease the strength of a mental set, depending upon the context within which the work is presented.

In the present study, the work variable was controlled by having the Ss use either mercury or water in a Luchins' type problem. It was predicted that with brief practice, the S using mercury (and hence having greater work) would evidence a stronger set because it would be expected that the competing non-set responses would be more quickly and effectively weakened by the greater work inhibition developed. With increased practice, however, it would be predicted that the Ss using mercury would evidence a weaker mental set as compared to those Ss using water. This would follow because after a brief amount of practice the set response would become dominant and consequently the additional work would result in greater reactive inhibition for the set response.

Experimental Procedure.

The experimental procedure utilized in this study was similar to the procedure used by Kendler, Greenberg and Richman. Simple problems involving three glass containers were individually presented to Ss who were instructed to obtain a specific quantity by the appropriate manipulations of the container. A set solution to all problems was developed during the training series by insuring its successful utilization. The strength of this set response was measured by a test problem which could be solved either in the set manner, involving use of all three containers, or in a more direct way, involving the use of only two of the containers.

The Ss were 128 male undergraduate students, mainly from the elementary psychology classes of the University College of Arts and Science of New York University.

The arithmetical problems used were as follows:

Table 1
The Series of Arithmetical Problems

Problem	Containers Given			To Get
	a	b	c	
1.	29	3		20
2.	9	59	5	40
3.	10	59	4	41
4.	18	43	10	5
5.	20	59	4	31
6.	9	25	6	4
7.	3	59	18	20
8.	4	49	14	17
9.	8	43	10	15
10.	23	49	3	20

The problems were presented in a concrete manner. A set of 10 pyrex tubular containers, each with its capacity in milliliters stenciled above the capacity line, was used. A large, unmarked pyrex container held 170 ml. of fluid and served as the storage container. Each problem was presented by having the appropriate containers placed on an aluminum cafeteria tray, 12 by 16 inches.

The Ss were divided into two main experimental groups; group M used mercury while the fluid used by Group W was water. In order to equate the apparent monetary value of the fluids used by Group M and Group W, the water was colored with ferric chloride and the Ss were told that the large container held a colloidal gold suspension.

Both groups M and W received two degrees of training and two degrees of spacing between successive training problems. Half of each group were given the first nine training trials prior to their receiving the test series while the remaining half were given only the first four training problems. Half of each group received the training problems under massed conditions while the remaining half had three minute period intervening between successive problems. In order to control the effects of a three minute interval prior to the test problem, half of each group had Problem 10 immediately following Problem 9, while the remaining half had a three minute interval between problems.

The above experimental manipulations resulted in the formation of 16 experimental sub-groups of eight Ss each. These sub-groups varied in terms of the liquid used (mercury or water), the number of training problems (4 or 9), the condition of spacing during training (0 and 3 min. interval) and the time interval between the last training problem and the test problem (0 or 3 min. interval). The sub-groups are designated by a letter (M or W) and number (4 or 9) preceding a hyphen which refers to the liquid used and the number of training problems and two numbers (0 or 3) following the hyphen referring firstly to the degree of spacing during the training series and secondly to the time interval between the last training problem and the test problem. Thus Group M 7-30 means that the Ss comprising this group used mercury, had nine training trials, had a three minute interval between successive training problems and had a zero interval between the last training problem and the first test problem.

The following instructions were read to each S in order to provide a rationale for the use of such "valuable" materials as mercury and "colloidal gold."

"An important problem in psychology is the effect of the value of a material upon the acquisition of motor skills in working with it. For instance, a brick-layer splitting a brick and a diamond-cutter splitting a diamond perform essentially the same operations, but the diamond-cutter is much more careful and many more years of experience and practice are required before he becomes skillful at it.

That is the problem we are interested in; what is the effect of the value of a material upon the acquisition of motor skills. For obvious reasons we cannot use diamond, but we are going to use mercury (or "colloidal gold") which is expensive. As a matter of fact, this much of it (experimenter points to the storage container) is worth ten dollars. For this reason we want you to be careful not to waste any when you use it."

Then the Ss were given the following instructions:

"You will receive a number of numerical problems. Each involves measurement of mercury by means of all or any of the containers. In all these problems you will be presented with a number of containers. The numbers on these containers represent the capacity of the container in milliliters. You will also be given a card stating the amount of mercury you are to obtain.

Here is the first problem. There is an empty 29 ml. container and an empty 3 ml container as well as a large quantity of mercury in this large container. You are to obtain 20 ml. of mercury. Here is the card stating the amount of mercury you are to obtain. The mercury is obtained in this manner: Fill the 29 ml container. Pour 3 ml. from it into the smaller container leaving 26 ml. remaining in the large one. Again fill the 3 ml. container. Now 23 ml. remain in the large container. Again fill the 3 ml container and the required 20 ml. are left in the large container. When you are finished, return the mercury to the storage vessel.

This problem can be solved in the following manner. Fill the 29 ml. container and then empty a sufficient quantity in the 9 ml container leaving 50 ml. in the large container. Then fill the 5 ml. container leaving 45 ml. Doing this once will give you the solution, that is, you will have 40 ml. of mercury in the large container. This problem

can also be solved by filling the 5 ml container and emptying it into the 59 ml. container. If you do this seven more times you will have 40 ml. Do you understand? If so here is the next problem."

The Ss who had a three minute interval between problems had their time occupied by rating on a five point scale how humorous certain cartoons were.

A time limit of 5 min. was imposed on each problem. If the S had not achieved the correct solution by then, the set solution was suggested. All the Ss who had the 9 training problems completed at least 5 of the 7 training problems (problem 3 to 9) without any assistance in the set manner.

Results and Discussion

The results of the 16 sub-groups are presented in Table 2. There are several possible comparisons that can be made. It will be recalled that our prediction was that with brief practice (four training problems) the Ss using mercury would evidence a stronger set. The results lend some support to this hypothesis. In all of the four comparisons between mercury and water groups with brief practice, the mercury group demonstrated a larger per cent of set solutions. By chance, the greater per cent of set responses in the mercury group for all four comparisons would occur .06 of the time. The difference between the total percentages of both the M4 and W4 groups was significant at the .15 level, in the direction obtained. The difference between the mercury and water groups with the greater amount of practice (9 training problems) was also in the direction predicted, i.e., the water groups should demonstrate a greater per cent of set responses. The difference between total percentages was significant, in the direction obtained, at the .05 level. In three of the four comparisons the water groups exceeded the mercury groups in terms of per cent of set responses.

Table 2

The N of Ss in Each Group Solving Problem 9 in the Set Manner

Group	Set Solutions		Group	Set Solutions	
	N	%		N	%
M4-00	5	62.5	W4-00	4	50.00
M4-33	5	62.5	W4-33	4	50.00
M4-03	2	25.0	W4-03	1	12.5
M4-30	3	37.5	W4-30	2	25.0
Total	15	46.9	Total	11	35.4
M9-00	3	37.5	W9-00	5	62.5
M9-33	4	50.00	W9-33	6	75.0
M9-03	6	75.0	W9-03	4	50.0
M9-30	1	12.5	W9-30	6	75.0
Total	14	43.8		21	65.6

At first glance the results tend to be quite consistent with the a priori predictions made by the extinction hypothesis. A closer examination of the data, however, reveals certain inconsistencies within the data as well as inadequacies in the formulation of the extinction hypothesis.

It will be recalled that Kendler, Greenberg and Richman investigated the influence of massed and distributed practice on the development of a manual set, and following the finding that the manual set was learned to a stronger degree under conditions of massed practice an extinction hypothesis was proposed to explain the results and integrate them with the facts of distribution of practice. When it was decided to investigate the work variable in order to test the extinction hypothesis it was thought wise to modify the experimental procedure used by Kendler, Greenberg and Richman. These investigators used galvanized iron containers with capacities measured in cubic inches and the commodity, the amount of which was manipulated, was dried white beans. It was thought that the work variable could be more easily manipulated by the use of glass cylinders in which liquids of different weights could be introduced.

It is difficult to know how this modification in experimental design effected the behavior of the Ss. Whereas Kendler, Greenberg and Richman, whose Ss also had 9 training problems, report that their massed and spaced Ss evidenced 74% and 52% set responses respectively on the test trial, the equivalent figures for the present experiment were 53% and 52%. If however, the results of the water and mercury groups are separated, we discover that in the water groups 75% of the massed Ss had a set response as compared with only 56% of the spaced Ss, results which are practically identical with those of Kendler, Greenberg and Richman. The relationship was different for the mercury groups; 56% of the spaced Ss solved the test problem indirectly while only 31% of the massed group did likewise. In general it can be said that the use of mercury depressed the amount of set responses.

Why the relationship reported by Kendler, Greenberg and Richman should be reflected in the Ss who used water but not with those Ss who used mercury, is a difficult, if not impossible, question to answer. One possible answer, consistent with the extinction hypothesis, is that in the massed mercury groups the non-set responses were extinguished rather early and consequently the work inhibition produced by the manipulation of mercury was mainly localized in the set response, thus weakening its effective reaction potential. The same line of analysis would imply that in the massed water group, because of the relatively small amount of work, the set response was not weakened very rapidly and hence still maintained a relatively high place in the response hierarchy of the Ss at the time of the test problem.

There are several objections to this analysis. Firstly, if the analysis were correct we would expect that the massed mercury group with brief practice should demonstrate less set responses than the massed water group. The percentages are 44% for the massed mercury group and 31% for the massed water group, a difference which is not significant but is in a direction opposite to that predicted. Secondly, if the specific hypothesis which instigated this experiment is to be tested adequately, it would be necessary to know more precisely the changes in the per cent of set responses for the massed water group as the number of training trials increased. With a precise knowledge of this relationship it would be then possible to discover whether the set responses of a massed mercury group would initially fall significantly below that of the massed water group and later significantly rise above it.

This writer believes that with additional knowledge the precision of the extinction hypothesis in relation to the present experimental problem can be improved considerably. Whether it would be strategic at the present time to expend such a large amount of research energy to attain this degree of precision is a question that at the present time this writer must answer in the negative because of reasons which will become apparent in the later part of this report.

This writer believes that a more reasonable explanation of the results stems from the effects of the differences between the experimental procedure used in this study as compared with that of Kendler, Greenberg and Richman. In their study, beans were used as the material to be measured and the various containers were filled to the top. In order to avoid the large amount of "spillage" which occurred in their study, the cylinders in the present study were not filled to the top; the capacity line was appreciably below the top of the cylinder. Since accuracy was demanded, the S had to exercise care and in many cases was motivated to pour the liquid back and forth until the precise amount was achieved. This required not only time but also a good deal of concentration that would probably obscure the conceptualization of the set method, i.e., the S would become so engrossed in filling the containers accurately that he would forget the method by which he had solved the problem.

The verbal reports of the Ss during and after the experiment tended to support this analysis. Frequently, during the experiment, a S would, after filling a particular container, stop and request the E to tell him which containers he had filled because he had forgotten. At the end of the experiment each S was asked whether he had discovered a common method in the solution of the problem. Of the 64 Ss who had 9 problems and would therefore be expected to remember the set method which they used to solve the training problems, seven Ss said they had not discovered any method and only one could verbalize precisely the "b-a-2c" method. Only 50% were able to verbalize any idea of a set involving all containers. This would suggest that the Ss not possessing the verbal conception of the set method would be more apt to solve the test problem directly. This would lead one to expect that the per cent of direct responses in the present study would be greater than found in the Kendler, Greenberg and Richman study. As previously reported, this difference was obtained; the groups with 9 training problems had 63% set responses in the Kendler, Greenberg and Richman study and only 55% in the present study. Other evidence which is consistent with the interfering effects precise measurements had upon the establishing of a set is the fact that Tresselt, in a personal communication, reports that only one of her 40 Ss was unable to verbalize the "set" solution after but a few training problems. In her experiment, although water was used, the containers were filled to the top by merely dipping them into a large receptacle of water. It is also interesting to note that in Tresselt's study 70% of her Ss solved the test problem in the set manner.

Experiment 2

Although certain doubts were cast upon the adequacy of the experimental procedure used in Experiment 1, these doubts were not sufficiently strong to result in the abandonment of the experimental technique. The writer tried to devise new experimental problems utilizing the same technique to evaluate the influence of work upon the development and weakening of a mental set. This experiment reports one such attempt.

The aim of this experiment was essentially similar to the previous one, i.e., to demonstrate that increased work would facilitate the extinction of a mental set. It was thought that a more direct attack upon the problem involving actual "extinction" trials might produce positive results.

Experimental procedure.

The basic experimental technique was similar to the one used in the first experiment. Fourteen Luchins type problems were used. An attempt was made to make the groups of Ss who were given the test problem as homogeneous as possible.

Problem 1 could be solved either directly (a-c) or indirectly (b-a-2c). Those Ss who solved it in a set manner were discarded. Problem 2 - 7 were designed to produce the desired set response. These problems were best solved in the set manner. Only those Ss who solved all of these problems in the set manner were permitted to continue the experiment. Problem 8, like the first problem, had two solutions. The function of this problem was to screen out those Ss who had not succumbed to the set. Accordingly, those Ss who solved this problem directly were discarded. Problems 9 - 12 were, for all practical purposes, insoluble. These problems constituted the extinction phase of the experiment. Problem 13, as did Problems 1 and 8, had two solutions and served as the test problem.

Two groups of 10Ss each were used. Both groups used water for the first six problems. On the seventh problem and thereafter, Group M used mercury while Group W continued to use water. To equate for the possibility that Ss using mercury might, because of its monetary value, be more careful than those using water, the water used for the seventh and all succeeding problems was colored with ferric chloride and Ss were told that this was a solution of colloidal gold. The Ss were told that the monetary value of the gold in the reservoir cylinder was estimated to be worth \$25., which was the value also assigned to the mercury.

Results and discussion.

Table 3 reports the results of the experiment.

Table 3

N and Per cent of Set Solutions for Problem 13

<u>Group</u>	<u>N</u>	<u>%</u>
M	6	60
W	3	30

Although this difference was not significant it was in a direction opposite to the one predicted. In order to discover whether Group M actually engaged in more work a record was made of how many times the Ss in each group lifted a filled container unnecessarily during the extinction problems (9 - 12). It will be recalled that these problems were insoluble and consequently any overt manipulations of the filled containers were unnecessary. The mean number of liftings for Group M was 6.4, while Group W had a mean of 9.9. This difference was significant at the .03 level.

It might be argued, by one who is inclined to defend the extinction hypothesis, that the direction of such findings is consistent with the extinction hypothesis. Although it was thought initially that mercury would produce more work, the findings suggest the reverse might be the case. But such arguments smack more of debate than of theorizing. They point to the lack of precision the implications of the extinction hypothesis have for the influence of work upon the weakening of a mental set.

Experiment 3

Introduction

In Experiment 2, it was thought that the series of extinction problems in which no solution was possible might not have provided a sufficiently sensitive test for the effects of the use of mercury or water. During the extinction series

there was no possible measure of the strength of the set tendency for Groups W and M. Instead of having a series of extinction problems, perhaps a series of ~~problems~~ which could be solved either in a set or direct manner could provide a more sensitive test of the strength of the set tendency, as it was undergoing weakening. The present experiment utilized such a procedure.

Experimental procedure.

The procedure of the present experiment was identical to that of Experiment 2, except for one modification. Instead of having a series of insoluble extinction problems, problems 9 - 12 in this study were capable of being solved either in the set or the direct manner. Each group consisted of 13 Ss recruited from the elementary psychology classes of the University College of New York University.

Results and discussion.

Table 4 reports the number and per cent of set solutions for problems 9 - 12.

Table 4

Number and Per Cent of Set Solutions for Problems 9 - 12

Problem	Group M		Group W	
	N	%	N	%
9	10	56	11	61
10	10	56	8	44
11	7	39	5	28
12	6	33	3	17
Total	33	46	27	38

Again, the direction of differences between groups is opposite to that predicted by the extinction hypothesis. Table 4 indicates that 67% of Group M extinguished the mental set by the twelfth problem while 63% of Group W did likewise. This difference was not significant ($p = .24$).

It is of interest to note the results obtained on Problem 8, which was the same for the Ss in both Experiments 2 and 3. It will be recalled that all Ss did the initial seven problems with water and on the eighth problem they were divided into either the mercury or water (colloidal gold) group. Only those Ss who solved Problem 8 in the set manner were permitted to remain in the experiment. On the eighth problem, 28 of the 38 Ss in the mercury group, and 28 of the 46 Ss in the water group solved it in the set manner. The difference is not significant ($p = .20$), but the finding is interesting in that it suggests that perhaps mercury has a more distracting effect on a mental set than does a gold colored solution. Since Groups M and W had exactly the same experience prior to problem 8, there should be no appreciable difference between the per cent of set responses on problem 8. If, however, we assume that mercury has a more distracting effect upon the mental set, then we would expect the per cent of set solutions to be less for the mercury groups. This is, of course, what occurred in Experiments 2 and 3, and it would appear reasonable to assume that the fascinating physical quality of mercury might have distracted Ss sufficiently to make them forget the set under which they were operating. Of course, if this was true, it would mean that the variable of work in our experimental technique is badly confounded.

Experiment 4

Introduction

This study was designed to investigate the implication of the extinction hypothesis that mental sets can be more effectively weakened under conditions of massed practice. The first three experiments failed to provide any significant evidence that increased work facilitates the extinction of a mental set. It was decided to do some pilot studies to determine whether an experimental situation other than one involving Luchins type problems would also provide evidence that massed practice would facilitate the weakening of a mental set. The experimental technique selected was a modification of one used by Maltzman (10) involving anagrams.

Experimental procedure.

Twenty Ss from the elementary psychology class at University College of New York University were given a series of 15 anagrams, all possessing the same letter order (54123) and all being food words. The words used were sugar, bacon, candy, cocoa, apple, cakes, prune, onion, sodas, berry, syrup, fudge, salad, fruit and grape. After the completion of the training series all Ss were given 15 additional anagrams having various letter orders, none of which were 54123. The words in this test series were all "nature words;" daisy, ponds, creek, swamp, tulip, brook, stern, ferns, beach, plant, pansy, stems, lilac, river and ranch.

During training the time interval for all Ss between successive anagrams was 15 sec. During the test series Group O (N=10) was presented with an anagram as soon as they had completed the preceding one while Group 30 (N=10) had a time interval of 30 sec. between anagrams. During the test series, the Ss were given a time limit of 30 sec. to solve the test anagram. If they failed to do this, they were given the next anagram after the appropriate time.

The purpose of this experimental design was to create a mental set during the training series and then weaken it, by the process of non-reinforcement, during the test series.

Results and discussion.

Group O during the test series had a mean of 6.2 failures, while Group 30 had a mean of 7.2. This difference was not significant ($p = .10$) in the direction obtained. It is possible, however, that the influence of the difference in spacing would dissipate as the test series progressed. It was, therefore, decided to test the significance of difference between the number of failures from test trials 2 - 4. The means for Groups O and 30 were .7 and 1.6, the difference of which was significant beyond the .01 level. The means for trials 2 - 5 were 1.1 and 1.8, while the means for trials 2 - 6 were 1.3 and 2.1. Both of these differences, with the use of a one-tail hypothesis, were significant well beyond the .05 level.

The results appeared to be promising. It was felt, however, that some defects in our experimental design existed. It would probably be better to have only one possible set for the training anagrams. In order to encourage the development of a set during the training series, two possible sets were included; a letter set (54123) and a meaning set (food words). Upon questioning at the conclusion of the experiment, it was found that the Ss fell into three groups; those who were aware of both sets and those who were aware of only one set. The numbers are too small to suggest any difference in the relative strengths of these sets but it is most likely that some differences do exist, the result of

which would be to increase the within group's variability.¹

General Discussion

The four experiments reported in this chapter were all designed to test some implications of the extinction hypothesis. There is no doubt that these results are not particularly favorable to the extinction hypothesis. It should be realized, however, that hypotheses and theories are not created in final form but rather are developed slowly, progressing from crude formulations to precise propositions. The rate of such development is largely dependent upon the contributions empirical studies can make to both the sharpening of the formulation and the development of experimental situations capable of producing systematic relationships. In this sense, these studies have been complete failures and the extinction hypothesis remains the same crude formulation as it was at the time it was proposed to explain one set of experimental data. Perhaps a more imaginative experimental approach will succeed in giving more substance to the extinction hypothesis.

There are other points to consider. In simple classical conditioning the extinction process requires, after a relatively high degree of conditioning has taken place, a substantial number of trials. In thinking behavior there are so many other response tendencies present with relatively strong associative values that the amount of weakening, required for the dominant response tendency to lose its position in the response hierarchy, is relatively little. Added to this factor is the ability of human organisms with verbal capacities to essentially change their environmental situation by shifting the verbal cues to which they are responding. The combination of these two factors might restrict the extinction process to such a narrow range that it would be exceedingly difficult to empirically observe its consequences.

The problem for this writer was to decide whether to pursue the extinction hypothesis or become preoccupied with another area. As is the case in many conflicts, the resolution depends upon the attractiveness of the various alternatives. As the next chapter will indicate, an area of thinking behavior appeared which this writer thought was more promising, and consequently the attention of this research program was shifted from the extinction hypothesis to the verbal control of problem solving behavior.

In closing, a few remarks may be made about the implications the findings reported in this chapter have for an attempt to extend conditioning principles to the realm of problem solving behavior. For some, the content of this chapter will increase their conviction that such an attempt is doomed to failure. From my point of view such a view would be horribly premature. These findings suggest that it will probably not be possible to simply select variables which operate in the conditioning situation and expect them to be as potent in the problem solving situation. An extension of conditioning principles to problem solving behavior does not mean the "extension" of conditioning variables, but rather the extension of relationships obtained in the conditioning situation.

¹An attempt was made to complete a more extensive study using the modification suggested. Since the experiment was begun near the end of the academic year, it was only possible to obtain a total of 18 Ss. The results of these Ss revealed absolutely no difference between Groups 0 and 30 during the test series. The experiment was not continued at the beginning of this academic year because of the priority given to the experiments reported in the third chapter.

Such an extension cannot be done in a simple minded fashion but rather requires a tremendous amount of work and thought in coordinating the independent and dependent variables in the conditioning situation with independent and dependent variables in the problem solving situation, and then testing whether the conditioning principles are applicable to thinking behavior.

Verbal Factors in Concept Formation Behavior

The initial experiment reported in this chapter was designed to determine the influence the learning of simple concepts (concepts based on one stimulus attribute) had upon the learning of a compound concept (concepts based on two stimulus attributes) which involved combinations of simple concepts. It was found that some Ss could not learn a compound concept even after they had learned the simple concepts of which it was composed. An explanation of this phenomenon was offered which emphasized the importance of verbal cues in concept formation behavior. The remaining experiments in the chapter sought to test the adequacy of the proposed hypothesis.

Experiment 5

Introduction

The empirical problem investigated was a specific one related to concept formation behavior in a conventional card sorting situation. It sought to determine the influence the learning of simple concepts had upon the learning of compound concepts which involved combinations of simple concepts. An example of a simple concept would require sorting a series of cards by the color of the figures on them. Another simple concept would require the Ss to sort the cards according to the size of the figures. A compound concept would involve sorting on the basis of both size and color.

Our experimental design involved three groups, each required to learn two simple concepts successively. Following this, each group was required to learn the test concept, which was a compound concept. One group learned both of the simple concepts of which the test concept was composed, the second group learned only one, while the third group learned neither of these concepts. The expectation was that the rate of learning the test concept would be directly related to the number of simple concepts learned which were involved in the test concept. Since the learning of the test concept depended upon the development of two specific response tendencies, it would be expected that previous training involving these response tendencies would facilitate the acquisition of the test concept.

Experimental procedure.

Male students from the elementary psychology course at University College of New York University were used. They were naive as to the nature of the experiment. The final results included the data from 62 Ss.

The conventional card sorting test procedure was used. The Ss were required to sort the cards with respect to a group of stimulus cards. Correct responses were followed by the E saying "Right," and incorrect responses were called "Wrong." The initial form of the N.Y.U. Card Sorting Test was used. This test was patterned after the Wisconsin Card Sorting Test (A). It consisted of 32 response cards which varied in terms of four forms (circle, crescent, square and triangle), and four "colors" (black, gray, yellow and orange), and two sizes. The designs appeared on a 3 in. square card with the designs themselves being either 1 in. or 2 in. in height.

Two studies with the same experimental design were conducted. In the first study, three groups were required to learn two simple concepts successively. The Ss were seated individually in front of a table on which was a tray with two rows of four compartments in which the cards could be placed. In the two center

compartments in the row farthest from the Ss were placed two stimulus cards; a large yellow crescent and a small black square. The response cards could be sorted into two categories represented by these two stimulus cards in four different ways. The basis of such sorting would involve four simple concepts; size, color, shape and "part-whole." Acquisition of the size concept would be indicated by the S placing the cards having large designs under the large yellow crescent and small designs under the small black square. The color concept required the S to sort chromatic cards (yellow and orange) under the yellow stimulus card, and achromatic cards (black and gray) under the black stimulus card. The shape concept involved placing figures with curved lines (crescent and circle) under the stimulus card with the crescent and placing figures with straight lines (square and triangle) under the stimulus card containing the square. This "part-whole" concept was based upon the idea that a triangle is part of a square. Consequently, when the S was learning this concept, he was required to sort part figures (triangle and crescent) below the crescent and whole figures (square and circle) below the square.

At the beginning of the experiment, the S was given the following instructions:

"I want you to put each of these cards (pointing to the response deck) in the slot below the card (pointing to the stimulus cards) with which you think it belongs. After you place each card, I will tell you whether you are right or wrong. Your object is to get as many as possible right. Do you understand? Go ahead and try the first one."

After the S reached a criterion of 10 successive sorting responses, he received the following instructions:

"That was very good. I would now like you to try a different problem. This one has a different solution."

In the first study (represented by the letter A), Group 1A learned the shape concept and then the size concept; Group 2A learned the shape concept and then the color concept; while Group 3A learned initially the color concept and then the "part-whole" concept.

Following the learning of the two simple concepts, the S was told:

"That was fine. Now here is a third problem for you to try. Now we will have four categories."

Two additional stimulus cards were added; a large gray triangle appearing in the first compartment of the tray and a small orange circle being placed in the fourth compartment. Ss in all three groups were required to learn the same test concept; a compound concept involving both the shape and size concept (for a schematic representation of the procedure of the first study refer to the first three columns of Table 5). Examples of correct responses would be as follows; a response card with a small black crescent would be placed under the small orange circle because both are small and have the same shape (the figures are composed of curved lines). A large orange square should be placed under the large gray triangle, because both are large and have straight lines. The criterion of learning for the test concept was also 10 successive correct responses. The Ss had a total of six decks (192 cards) in which to achieve the criterion of learning. This limit was required because of the hour duration of the experimental session.

The second study (represented by the letter B) had a similar design and procedure (refer to the first three columns of Table 5). Three groups of Ss were required to learn two simple concepts successively. Group 1B learned the shape and size concepts, while Group 2B learned the shape and color concepts, and Group 3B learned the color and "part-whole" concepts. The major difference between the first and second study was the nature of the test concept. In the second study the test concept, which was also a compound concept, involved both the color and "part-whole" concepts. According to such a concept, a correct sorting of a response card with a small black crescent would involve placing that card under the stimulus card having a large gray triangle, since both these cards have achromatic "coloring," and both are "part" figures.

The inclusion of the second study in our experimental design was to control for the possibility that the rate of acquisition in the first study would be a function of some general effect induced by the learning of the simple concepts rather than due to the specific interaction between the simple concepts and the test concept. For example, if Groups 1A and 1B learned the test concept more rapidly, then we would attribute such a finding to some general facilitating effect the learning of the simple concepts of size and shape had upon the learning of all compound concepts. If, on the other hand, Groups 1A and 3B learned the test concept most rapidly, we would ascribe such results to the interaction between the successive learning of the simple concept and the test concept.

Results and discussion.

There were no significant differences between Groups 1A and 1B, 2A and 2B, 3A and 3B in the rate of learning the training concepts.

The results of learning the test concept in the first study are presented in Table 1. Two of the Ss in Group 3A failed to learn the test concept within the allotted 192 test trials. These Ss were assigned scores of 192, and the calculation of both the mean and the standard deviation of Group 3A included these two unsuccessful Ss. It was felt that such a procedure was defensible since it did not exaggerate the differences between the groups.

Table 5

Mean Number of Trials Required To Attain Test Concept in the First Study

Group	Training Concepts	Test Concept	N	N Solving Test Concept	Mean Trials	σ	Range
1A	shape, size	shape size	10	10	23.3	21.2	5-65
2A	shape, color	shape size	10	10	42.4	22.1	12-81
3A	color, "part-whole"	shape size	11	9	104.5	65.0	17-192

It was deemed advisable, because of the problem raised by the two unsuccessful Ss in Group 3A as well as the large differences between the variance of Group 3A and the other two groups, to use a simple non-parametric test to evaluate the significance of the differences between the various experimental groups. U tests (11) revealed that the differences between Groups 1A and 2A, 2A and 3A, and 1A and 3A were significant, on the basis of a one-tail hypothesis, at the .02, .05, and .01 levels. The probability of three comparisons, producing by chance three significant differences beyond the .05 level, is .0025 (.05).

Table 6 reports the data of the second study. Only eight of the 31 Ss in this study were capable of solving the test concept within the allotted number of trials. The results, however, do suggest a trend; the per cent of Ss solving the test concept, as well as the number of trials required by these successful Ss to reach a criterion of learning, appears to be related to the number of simple concepts appropriate to the test concept which had previously been learned.

Table 6

Mean Number of Trials Required to Attain Test Concept -- Second Study

Group	Training Concepts	Test Concept	N	N Solv- ing Test Concept	Mean Trials of Ss Solving Test Concept	Range
1A	shape, size	color + "part-whole"	10	1	167.0	--
2B	shape, color	color + "part-whole"	11	3	89.0	67-130
3B	color, "part-whole"	color + "part-whole"	10	4	77.8	25-186

The results of both studies appear to be, in a general way, consistent with the prediction made that the rate of learning the test concept would be directly related to the number of simple concepts appropriate to the test concept, which had been learned. In the first study a very definite and significant relationship was obtained; the group 1A, which had learned both of the simple concepts of which the test concept was composed learned the test concept more rapidly than did the group 2A, which had learned only one of the significant simple concepts and this latter group 2A, in turn, acquired the test concept more quickly than did the group 3A, which learned neither of the appropriate simple concepts. The results of the second study failed to show the same significant differences, seemingly because such a large per cent of the Ss failed to acquire the test concept. The trend of the results is consistent, however, with the original hypothesis. It was found that the rate of learning of these Ss who did solve the test concept appeared to be directly related to the number of appropriate simple concepts that had been learned. The data of the second study also suggested that the per cent of Ss solving the test concept within the allotted trials was related to the number of appropriate simple concepts that had been learned.

There are several interrelated problems raised by the data, the implications of which are particularly relevant to the relationship between S-R theory and concept formation behavior. The first of these is the discrepancy between the results of Groups 1A and 3B. It will be recalled that these two groups were treated similarly; they were initially trained on two simple concepts and then were required to learn a compound concept composed of the two simple concepts on which they had been trained. The results indicate that, whereas all the Ss in Group 1A were able to learn their compound concept, only four of the 10 Ss in Group 3B were successful. An analysis of this discrepancy points to one of the basic mechanisms in human concept formation behavior.

The results of Group 3B demonstrate that "mastery" of the simple concepts does not necessarily insure the solution of the compound concept. If Ss can learn the simple concepts, why cannot they learn the compound concept which is a mere combination of the two simple concepts? The answer to this question lies in the understanding of what responses are learned when an S masters a simple concept. It

should be remembered that when a S was learning a simple concept, he was required to sort the response cards into only two categories. It would be possible for a S to sort the cards appropriately without discovering the rationale for the correct sorting procedure. For example, in solving the "part-whole" concept, a S by rote memory could learn that he should place all crescents and triangles below the large yellow crescent stimulus card and all squares and circles below the small black square stimulus card. Although such behavior results in the successful completion of the initial phase of the experiment, it fails to provide those verbal cues (part figures vs. whole figures) which greatly facilitate the acquisition of the test concept.

The essential implication of the above point is that the associations underlying the process of concept formation in a card sorting test are not between the cards and the sorting response but rather between the cards and some verbal response to them which, in turn, serves as a cue for the sorting response. Our contention is that when the Ss of Group 1A learned the simple concepts of shape and size, they were able to verbalize these conceptual differences, and these verbalizations served as cues which facilitated the acquisition of the test concept. Although it was probably easy for the Ss in Group 3B to verbalize the conceptual differences between the response cards when learning the simple concept of color, it was extremely difficult for these Ss to conceptualize the differences between the response cards when sorted on the basis of the "part-whole" concept. It would seem that a triangle, especially the equilateral triangle which appeared on the cards used in this experiment, is unable, for most Ss, to instigate the idea that it is part of a square.²

The emphasis on the importance of verbal cues is not, of course, all speculative. The Ss were questioned following the termination of the experimental session. As is so often the case, many of the Ss were found to be amazingly inarticulate; this can undoubtedly be attributed in part to defects in our questions. Their replies, however, did suggest some interesting relationships. The Ss in Group 1A were able to verbalize the rationale of their card sorting after having learned the test concept. Most of them were able to express themselves in terms of the abstract principles, i.e., the differences between the four categories was on the basis of the combination of both the size and shape factors. The remaining Ss expressed themselves in a somewhat more concrete fashion, describing specifically the underlying features of their sorting procedure,

² This inability to respond to the triangle as a part of the square was probably the primary reason why the learning of the "part-whole" concept during the initial phase of the experiment required a large number of trials. The mean number of trials required to achieve the criterion of learning the simple "part-whole" concept was 43.9 as compared to 5.7, 6.6, and 20.6 for the simple concepts of color, shape and size. These means, it should be noted, are based upon the results of all three groups in each study. The shape concept was always learned initially while the learning of the size and part-whole concept always followed the acquisition of a previous simple concept. The color concept was learned first in Groups 2A and 2B, and was the second simple concept to be learned in Groups 2A and 3B.

e.g., "small figures with small figures, round figures with round figures, large figures with large figures, sharp edge figures with sharp edge figures." Contrasted with this were the verbal reports of the successful Ss in Group 3B. Two of the Ss were practically incoherent, being unable to spontaneously provide any rational basis for their correct sorting behavior. Upon further questioning, it appeared that these Ss were sorting the cards correctly by responding on an abstract level to the differences in color but specifically to the differences in form, i.e., these Ss learned to place achromatic crescents and triangles in one category, chromatic crescents and triangles in another, achromatic squares and circles in another, and chromatic squares and circles in a fourth category. Although being able to conceptualize the differences between the achromatic and chromatic cards, they were unable to verbalize the differences between the cards containing part figures and whole figures. This type of solution, which is based upon a combination of both abstract and specific features was the most common one reported throughout the second study. Only one of the eight successful Ss (a S in Group 3A) was able to give a clear statement of the abstract principles which guided his sorting behavior.

The importance of verbal cues in concept formation is nicely illustrated although not elaborated, by Oseas and Underwood (13). They used as stimuli three different geometrical forms of three sizes and shades. Size and shade were relevant dimensions, while the form was irrelevant. This resulted in a task involving nine concepts, each being a different combination of size and shade. It was found that concepts based on small and large geometrical form were learned more rapidly than those based on medium size figures. They attributed these results to interference produced by stimulus generalization.

"In learning concepts of medium size, generalizing tendencies from both the small and large figures could interfere. But, in learning the concepts based on small figures, generalizing tendencies would be maximal from the medium-size figures, but minimal as regards the large because of their greater "distance" along the size dimension. Likewise, in learning the concepts based on large forms, interference again should be maximal from the medium-size figures but minimal from the small." (p. 116).

The same line of reasoning should apply to concepts involving the three shades (white, gray and black) along the brightness dimension. There wasn't, however, any difference in the speed of learning these concepts. The discrepancy between the learning of concepts based on size as compared with brightness according to our analysis would be due to the differences in the verbal responses produced by the stimuli. In the case of size, one does not initially respond to the size dimension with the verbal responses of large, medium and small. One has to learn that there are three sizes before these descriptive terms can be attached. It is during this learning of the existence of three sizes that the interfering effects of stimulus generalization are produced. In the case of shade, however, the S has available from the very beginning three verbal responses (white, gray and black) which differentiate the three points on the brightness dimension. Consequently, the interfering effects on the gray concepts arising from the stimulus generalization do not have any opportunity to operate. It would appear that the three distinctive words inhibit the expression of stimulus generalization based upon physical similarity, as was found to be the case in the study by Kuene (). Our point can be summed up by saying that if Oseas and Underwood had used three shades of gray rather than white, gray and black, they would also have found generalization effects during the learning of concepts based upon differences in shade.

Returning to the results of our own study, it might be argued that our analysis of the discrepancy between the results of Groups 1A and 3B says no more than that the Ss in Group 1A were able to perceive the necessary relationships while most of the Ss in Group 3B were unable to do this. We believe that such an argument is not entirely justified. Our emphasis has been on response produced cues which serve to mediate formation of the appropriate card sorting responses. Such an analysis points to what we believe are basic mechanisms as well as integrating thinking behavior and simple associative learning, rather than creating a hiatus between the two, as is customary by advocating a "perception-of-relationship" orientation. In addition to this, our analysis does not require that the appropriate relationships be perceived for the concept to be attained. For example, it would be our guess that the per cent of Ss solving the test concept would increase by conditioning the triangle and crescent to a common nonsense syllable. Such a procedure would provide a common verbal cue which should facilitate in the acquisition of the test concept, even though it would not emphasize the idea that a triangle is part of a square.

Let us now consider the significant differences found between the three groups in the first study. Our analysis would suggest that the advantage derived from the acquisition of the appropriate simple concepts did not stem from the mere repetition of correct sorting responses, but rather from the opportunity the initial training provided for the appropriate verbal responses to become dominant and thereby facilitate the acquisition of the test concept. Essentially the preliminary training, involving the appropriate simple concepts, modifies the cue situation rather than strengthening markedly any particular response tendency. Unlike a simple conditioning situation, in which the conditioning trials increase the habit strength between the conditioned stimulus and the conditioned response, the major effect of the training trials in the present study is to determine what verbal cues shall be dominant at the beginning of the test trials.

This point was borne out to some extent by the comments of the Ss following the completion of their experimental session. The most common introspective report of the Ss in Group 1A when describing the rationale of their correct sorting procedure for the test concept was the simple statement of the abstract relationships involved (size and shape). The most common report of the Ss in Group 2A, which had only learned one (size) of the appropriate simple concepts, was a more specific description even though some abstract features were mentioned. An example of this type of report is one in which the S categorized the groups as large squares and large triangles, large circles and large crescents, etc. The most common report of the Ss in the third group was vague in that it had failed to specify in any clear or simple manner the verbal cues which dictated their sorting response.

A basic problem raised by the present studies, but unfortunately not answered by them, is related to the question of whether these factors which influence the rapidity of acquisition of compound concepts (such as training on simple concepts) also influence the per cent of Ss solving the compound concept. The data from our two studies would suggest this is so since in the initial study the only two Ss who failed to solve the test concept were in the group that had no training on the appropriate simple concepts while in the second study the per cent of solvers in each group was directly related to the number of appropriate simple concepts learned. One might argue, however, that the training on appropriate simple concepts has the effect of hastening the appearance of the appropriate verbal responses but if unlimited time was afforded to all Ss the per cent of solvers would be unrelated to the type of preliminary training. Our findings give some support to this point since two of the Ss in Group 1B (no training on appropriate simple concepts) who were unable to solve the test concept within the allotted trials were able to solve the test concept with additional trials. The

theoretical problem involved in this question is whether the training on appropriate simple concepts merely hastens the appearance of the appropriate verbal responses, assuming that it is within the Ss repertoire, or whether this training "selects out" the appropriate verbal response which under other training conditions might not occur.

Our analysis of the present data, with its major emphasis on the cue situation being modified as a function of training, has certain similarities to Guthrie's (3) treatment of conditioning phenomena. This does not mean necessarily that only a simple contiguity conditioning theory, such as Guthrie's, is applicable to human thinking behavior. The stimulus situation is an important variable in all S-R theories, whether they be committed to a single trial associative principle or to a habit growth principle. Within a theoretical framework involving the habit growth principle it can be said that the new associations involved in the present study are not new associations with zero or relatively weak habit strengths but instead are established associations with their habit strengths approaching or having approached their asymptote values. The essential condition for the acquisition of the test concepts is that the relevant associations become dominant in their competition with the irrelevant associations. This comes about by rapid shifting in the dominant position in response hierarchy because of the almost equal strength of the competing associations involving the various verbal responses.

Experiment 6

In the previous experiment an analysis of human concept formation behavior was suggested involving a sequence of two S-R associations. In the first connection the stimulus would represent the test cards while the response would refer to the verbal (usually subvocal) responses made to them. The stimulus of the second would represent the verbal cue produced by the preceding response while the response would refer to the overt card sorting behavior.

The present experiment attempts to evaluate the adequacy of this type of conceptual analysis when applied to a comparison between a reversal and non-reversal shift in a concept formation test. Our specific assumption is that learning of a concept will be facilitated when the appropriate verbal cues are present. This hypothesis would predict that reversal shift would be more rapid than nonreversal shift. This would follow because at the beginning of the reversal shift the appropriate verbal cues are dominant; they are merely connected to the wrong sorting responses. This condition of having the appropriate verbal cues dominant would not prevail for the condition of nonreversal shift. At the beginning of the nonreversal shift the verbal cues dominant are those related to the initial concept which is, of necessity, different from the second to-be-learned concept.

Experimental procedure.

The experimental design is represented in the following 2 x 2 table.

Table 7

Schematic Representation of Experimental Design
Second Concept

		Reverse A	Reverse B
First Concept	A	Group 1 (N = 10)	Group 2 (N = 10)
	B	Group 3 (N = 10)	Group 4 (N = 10)

The N.Y.U. Card Sorting Test, Form 3, was used. It is a rather complex test and need not be described here. The reason for its complexity is that our supply of Ss was exhausted and we were forced to use Ss who had experience with the initial form of the N.Y.U. Card Sorting Test. Consequently a card sorting test, as unlike Form 1 as possible, was designed. The two concepts used can be represented by the letters A and B.

Two stimulus cards were used. Half of the Ss (a total of 40) initially learned Concept A while the other half learned Concept B. Half of each of these groups then learned the reverse of Concept A, while the remainder learned the reverse of Concept B. The Ss were not informed that a shift from the first concept to another concept would be made.

Results and Discussion.

According to our analysis, the combined Groups 1 and 4 (reversal shifts) should learn the second concepts more rapidly than the combined Groups 2 and 3. The results obtained were in that direction and also significant ($p = .01$). Unfortunately, there was a difference between the groups in the rate of acquisition of the first concept in favor of Group 1 over Group 2, and Group 4 over Group 3. The Ss were randomly assigned and consequently this difference in the rate of learning the first concept was purely a function of chance. The result, however, is that such a chance difference in the learning of the first concept invalidates the significant difference obtained in the learning of the second concept.

Experiment 7

During the time the data from the preceding experiment was collected, Buss reported (1) that in human concept formation learning a reversal shift occurred at a more rapid rate than did a nonreversal shift. Such a finding was consistent with our prediction. Buss, however, accounted for his obtained results in terms of partial reinforcement and a reinforcement-inhibition theory of discrimination learning.

The present experiment was similar to Experiment 5 in its attempt to compare the effects of reversal shift with nonreversal shift in concept formation learning. An experimental modification was introduced which had as its aim an evaluation of the relative merits of the explanations offered by Buss, as compared to ours, to account for the superiority of a reversal shift over a nonreversal shift.

Experimental procedure.

This experiment, at the time of writing, has not been completed, but the results are so definite that it was decided to report the data of the initial 46 Ss. These Ss were from the elementary psychology course of University College of New York University.

The N.Y.U. Card Sorting Test, Form 2, was used. This test is similar to Form 1 (a description of which appears in the report of Experiment 5), except that a  - shaped form was substituted for the triangle, and red was substituted for the orange color. The two stimulus cards used were not identical with any of the response cards. They were a small dark gray ellipse and a large orange pentagon.

The 46 Ss were divided into two groups of 24 Ss each. Group S learned initially the shape concept; all  - shaped and square figures had to be placed

under the large orange pentagon, and all circles and crescent figures were required to be sorted below the small dark gray ellipse. Group C initially learned the color concept; all gray and black figures had to be placed below the small dark gray ellipse while all red and yellow figures were required to be sorted below the large orange pentagon.

After reaching a criterion of 15 successive correct responses on the first concept, half of the Ss in each group were required to learn the reverse of the shape concept ("curved shapes" under the pentagon and "straight shapes" under the ellipse), while the remainder had to learn the reverse of the color concept (chromatic figures under the gray ellipse and achromatic figures below the orange pentagon). The groups were equated in terms of their performance on the first concept. The learning of the reverse concepts will be referred to as the second concept. Table 8 represents the experimental design.

Table 8

Schematic Representation of Experimental Design

		Second Concept	
First Concept	Shape	Reverse Shape	Reverse Color
		Group S-RS (N = 12)	Group S-RC (N = 12)
First Concept	Color	Group C-RS (N = 12)	Group C-RC (N = 12)

Both Buss and this writer would expect Groups C-RS and S-RC to demonstrate slower learning than Groups S-RS and C-RC. Buss would expect this event to occur on the basis of a reinforcement-inhibition theory with a principle of partial reinforcement added while this writer would predict it on the assumption that the appropriate verbal cues are dominant for the latter two groups at the beginning of the learning of the second concept.

According to Buss' analysis, the Ss in the nonreversal shifts (Groups C-RS and S-RC) would receive partial reinforcement when responding during the learning of the second concept in the manner which had been correct for the first concept. For example, placing a yellow crescent under the large orange pentagon stimulus card would be a correct response for a S in Group C-RS during the learning of the first concept. It would also be a correct response for the same S during the learning of the second concept. This partial reinforcement, according to Buss, would serve to maintain the color concept, and thus retard the learning of the reverse shape concept. For the reversal shift groups (Groups C-RC and S-RS) there would be no reinforcement of the discrimination learning previously, i.e., responses which were correct during the learning of the first concept would be 100% non-reinforced during the learning of the second concept.

In order to test the relative merits of Buss' formulation and that of the present writer it was necessary to devise some experimental technique capable of controlling the effects of partial reinforcement. This was accomplished by

using a 16 response card deck at the beginning of the learning of the second concept. All response cards, the correct sorting of which would be the same for the Ss in the nonreversal shift groups for both the first and second concepts, were removed. Consequently, during the initial learning of the second concept, the Ss of both the reversal and nonreversal shift groups were undergoing 100% non-reinforcement of the concept learned previously.

During the learning of the second concept, the Ss in all groups used this 16 card deck until they reached a criterion of 10 successive correct responses. With this modified deck the correct responses for both the reverse shape and reverse size concepts were the same and hence it was necessary to re-introduce the previously discarded 16 cards in order to compare the effects of reversal shift with the nonreversal shift. The criterion of learning with the 32 card deck was 15 successive correct responses.

Following the learning of the second concept, half of each group learned the shape concept while the other half learned the color concept. We will restrict our analysis to four major groups differing in terms of the second and third concepts they learned. The four groups will be designated as follows; Groups RC-C, RC-S, RS-S, RS-C. The letters preceding the hyphen refer to the second concept, while the letter following the hyphen refers to the third concept. In the learning of the third concept, the entire 32 cards were used.

Results and discussion.

Since both the reversal and nonreversal groups had suffered 100% non-reinforcement of the correct response tendency for Concept 1, Buss would be forced to predict no difference between these two groups in the rate they learned the second concept. Although this writer would not deny the possible effects of partial reinforcement in a situation such as Buss used, the elimination of this variable in our experiment should not effect our hypothesis, i.e., a reversal shift should be more rapid than a nonreversal shift. Table 9 reports the results of the learning of the second concept in terms of the performance of the various groups with both the 16 and 32 card deck.

Table 9

Median and Range of Number of Trials to Reach Criterion of Learning for Second Concept.

Group	16 Card Deck		32 Card Deck	
	Median	Range	Median	Range
S-RS	6.0	2-16	0.0	0-106
C-RS	5.0	1-16	60.0	3-160
C-RC	3.0	1-26	0.0	0-33
S-RC	2.5	1-18	30.0	0-160

With the 16 card deck the correct responses for both the reverse shape and reverse color are the same. There was no significant difference in the performance of the four groups with the 16 card deck. With the 32 card deck, however, there were very large and significant differences. In both comparisons (S-RS vs. C-RS, C-RC vs. S-RC) the reversal groups learned the second concept faster than the nonreversal groups. The U-test was used to evaluate these differences and the level of significance obtained in both comparisons was at the .01 level.

Table 10 reports the results of the learning of the third concept. These results are consistent with those obtained in relationship to the acquisition of the second concept (both of the differences were significant at the .01 level).

Table 10

Median and Range of Number of Trials to Reach Criterion of Learning
for Third Concept

<u>Group</u>	<u>Median</u>	<u>Range</u>
RS-S	1.0	1-25
RC-S	4.5	1-160
RC-C	2.0	1-14
RS-C	8.0	1-36

They once again demonstrate that a reversal shift is simpler to accomplish than a nonreversal shift.

Experiment 8

Our analysis of human concept formation behavior into a sequence of two S-R associations has an implication for that age-old controversy between the continuity and non-continuity interpretation of discrimination behavior. Since we have emphasized the importance of the appropriate verbal cue in human concept formation learning it would appear that a reversal shift in concept formation learning prior to the appearance of the appropriate verbal cues should not retard learning, i.e., the non-continuity hypothesis should apply to the human concept formation behavior. This pilot study was designed to test this possibility.

Experimental procedure.

The N.Y.U. Card Sorting Test, Form 1, was used. Two stimulus cards were used; one containing a large figure, the other a small figure. For the first ten trials, the experimental group was required to place response cards with small figures below the stimulus card with the large figure, and response cards with large figures below the stimulus card with the small figure. The correct response for the control group was to place the same size response card under the same size stimulus card. From the eleventh trial on, the correct response for both groups was similar to the situation which prevailed for the control group during the first 10 trials.

There were 15 Ss in each group from the elementary psychology course at University College of New York University.

Results and discussion.

The experimental group required a mean of 76.2 trials for the emergence of the correct response, while the control group required a mean of 68.8 trials. This difference was not significant.

Although the direction of the results is consistent with the implication of the continuity hypothesis, it is obvious that the data do not warrant any extensive analysis or speculations. What the pilot study does suggest is that the card sorting technique is applicable to the problem under consideration. It is this writer's intention to initiate some systematic investigations in this area.

General Discussion

The results reported in this chapter are suggestive that verbal responses play an important role in problem solving behavior. Such a conclusion is neither startling nor profound. But it is felt that such a conclusion is helpful in pointing to a problem area, a clarification of which can contribute much to our understanding of human problem solving behavior.

One of the major implications of our orientation is that there is a basic theme common to the understanding of all behavior ranging from animal behavior in simple learning situations to human problem solving behavior, i.e., in order to predict behavior it is necessary to control the stimulus situation. This task is relatively easy when one is concerned with predicting behavior of animals in a simple situation because the experimenter has direct control of the physical environment to which the animal is responding. This condition of direct control of the stimulus situation does not prevail in most cases of human problem solving behavior, since the human subjects, when attempting to solve a problem, are usually responding to their own sub-vocal verbalizations. The task of the experimental psychologist then becomes the one of controlling indirectly these sub-vocal verbalizations.

How can this be done? This is not a novel question. It's the sort of question which motivated the gestalt oriented psychologists like Lewin (8) to insist that behavior could not be related directly to the objective physicalistic environment of the subject but rather had to be correlated with the subject's psychological environment. The important point of view is how is the "psychological environment" (or, what we prefer calling when dealing with verbal organisms, "the verbal stimulus situation") to be defined? The gestalt oriented psychologists failed to give a systematic answer to this question. Because of their concern and experiences with problems of perception, they frequently used introspective reports to describe the state of the subject's psychological environment. To the modern behaviorist the approach to the problem of sub-vocal verbalization through the medium of introspection provides only half of the picture required for a full understanding of behavior. Prediction of behavior from the observable or manipulable stimulus situation requires that the stimulus situation arising from internal responses should be defined in terms of past and/or present states of the organism and its environment.

We have attempted to do this in our work. The second and third experiments in this chapter have defined the verbal cue in terms of past states of the organism and its environment. It would appear from those results that one of the major factors transferred from one concept formation problem to another with the same card sorting test is the Ss' verbal responses, and the cues derived from them.

Among the problems remaining is how do these verbal responses gain ascendancy; gradually or suddenly, consciously or unconsciously? Perhaps systematic work involving the continuity - non-continuity orientation with humans will be capable of clarifying this problem.

Chapter IV

The Acquisition of Flexibility in Problem Solving Situations

General Introduction

Training people to be flexible in their problem solving behavior has always been an appealing thought to an educator. How this was to be accomplished has been for the most part a problem which has concerned social philosophers rather than experimental psychologists. In recent years, however, the renewed interest in the field of problem solving behavior has once again brought this important educational problem to the attention of research psychologists.

The experiments reported in this chapter were for the most part empirically oriented studies, their major concern being to discover techniques capable of influencing the degree of flexibility in a problem solving situation.

Experiment 9

This experiment was designed to determine whether the development of a mental set or "Einstellung" could be influenced by various types of training immediately prior to the presentation of the set material. This training involved the presentation of a group or groups of problems which was followed by a number of set-inducing problems of the type used by Luchins (9). It was thought that Ss who were given a number of similar problems all of which should be solved by the same principle would succumb more easily to the set inducing problems. Conversely, if Ss experienced a number of problems, all of which were different from each other, then these Ss should be expected to resist more effectively the set-inducing problems.

Experimental procedure.

This experiment was conducted on a group basis. A total of 862 Ss was used from various colleges within New York University, Pace College and City College of New York. The Ss were each given a booklet of one of 14 experimental forms.

Initially, they were given a series of training problems designed to train the Ss to be either rigid or flexible. Then they were presented with two or three Luchins' type problems which could be solved only in a set (b-a-2c) manner. The strength of this set was measured by a test problem which could be solved in the set manner or by a more direct method (a-c).

The experimental groups differed from each other in terms of several variables. These variables were:

1. Rigid or Flexible Training: Rigid training always involved a series of problems which could be solved by a simple formula, while flexible training involved a series of problems involving different principles.
2. Type of Training Problem: There were two main types of training problems used; those involving numerical manipulations and those not requiring numerical manipulations.
3. Number of Training Problems: Some Ss received six training problems while the remainder received 16 training problems.

4. Number of Set Inducing Problems: Some Ss received two set-inducing problems while the remainder received three.
5. Set Breaking Training: While the experiment was in progress another group was added which differed in terms of the training received. The Ss who received this type of training had actual experience in breaking mental sets, other than the one induced in the test series.

Results and discussion.

Much to the surprise of this writer, the experimental variables investigated failed to provide any consistent findings. The only consistent finding was that a control group that received the set-inducing problems with no previous training, flexible or rigid, had the highest per centage of set solutions on the test problem. This would seem to imply that any training prior to set-inducing problems produces some degree of flexibility. Perhaps the important factor is the shift from one type of problems to the set-inducing problems. It is difficult to understand, if this one shift was important, why the additional shifts to which the flexible Ss were exposed did not increase their flexibility.

During the course of the experiment, when it became apparent that our experimental techniques were not going to influence the "flexibility" of our Ss, it was hypothesized that Ss cannot learn to be flexible by passively experiencing a number of different problems, but rather require active set-breaking experiences. This is why the "set breaking training group" was introduced but as has already been indicated the effects of such training was not noticable.

Another possible explanation of our negative results is that the rigid training was so simple (all problems were alike) that the Ss became suspicious and actively looked for "trick" questions. If this were so, and comments of some Ss imply this, then the rigid training, by its very nature, was defeating the purpose for which it had been devised.

Another possibility is that the Luchins' problems reflect certain basic thinking and personality variables incapable of being influenced by brief pre-training sessions.

The negative results of this study raise interesting problems in their relationship to the positive findings reported by Schroder and Rotter (16). They report:

"Conditions in this experiment leading to rigid behavior consist of repetition of the same or similar situations with reinforcement always following the same type of solution. The more trials the greater the rigidity. Non-rigid behavior was learned by introducing trials where an alternative solution was possible and sufficiently easy that it could readily be grasped when S discovered that the previously correct method would not work. The more times S is forced to adopt another solution the greater is the potential that he will expect and look for change in the situation." (16, p. 148).

It is this writer's hypothesis that the key difference in Schroder and Rotter's procedure as compared to ours is that in their study the test problems were similar to the training problems while in our situation they were not. This would suggest that flexibility, like other responses, might be weakened by generalization effects.

Experiment 10

In order to test the hypothesis that the differences between the results obtained by Schroder and Rotter and those of Experiment 9 might be due to generalization effects, the present pilot study was initiated. The problems used in this study involved anagrams during both the training and test series. It would be expected, therefore, that in such a situation flexible problem solving behavior would be acquired.

A 2 x 2 experimental design was used, varying in terms of degree of flexibility and the amount of pre-training. Four sets of 10 anagrams served as experimental material. The anagrams in any one group had a common letter order.

The 44 Ss were divided into two major groups. In the rigid group (Group R), the anagrams in each set were presented successively, while in the flexible group (Group F), the anagrams from the various sets were presented randomly. During the training series half the Ss were presented with a total of only two sets (Sets C and D) while the remaining half received all four sets (Sets A, B, C and D).

The test series consisted of 11 anagrams. The letter orders of these 11 test anagrams in order of their presentation, were D D C X X X D D X C C. D represents an anagram with a letter order similar to Set D, while C represents an anagram with a letter order similar to Set C. X represents a letter different from those of any of the four training sets. The X anagrams were also different from each other in the test series.

Our prediction was that the performance on the first, second, eighth and eleventh test anagram would be superior for Group R. This would follow because it was assumed that one of the important stimulus components becoming conditioned to the anagram solving response was the response produced cues from the previous solution. Since the Ss in Group R had received 10 successive D type anagrams at the end of their training series it would be expected that the cues from a D solution would be strongly conditioned to another D response. The same line of reasoning would also predict that whenever a C anagram followed another C anagram Group R should do better.

Our expectations were that on all X anagrams Group F would do better. In the flexible group the cues from one solution had not been strongly conditioned to any one type response.

It was felt that it was not possible to predict the results of the third, seventh and tenth test anagram. The C or D response was more strongly conditioned in the rigid group but it was conditioned to a particular cue situation (the solution of a previous similar anagram). In the above mentioned test anagrams, the C and D anagrams followed anagrams of a different letter order. The initial factor mentioned would favor Group R, while the second factor would favor Group F.

Results and Discussion.

There were no consistent findings with those groups who had only two sets of anagrams. For the groups having the four sets of anagrams, Group R tended to solve C and D test anagrams, while Group F tended to perform in a superior fashion on the X anagrams. Only the difference on the first test anagram was significant at the .01 level, but the general results suggest that this might be a promising technique, to investigate rigid and flexible problem solving behavior.

General Discussion

The results of these two studies suggest that at the present time training to be flexible in specific tasks might be more feasible than to provide training in "general flexible" thinking behavior. But what is most obvious is the necessity to develop experimental techniques capable of providing clear cut answers to our theoretical questions. This writer's feeling is that the responses to the Luchins' type problems are a function of so many unrelated variables that the technique is not promising for those who are concerned with systematic answers to theoretical questions. Perhaps the anagram technique will be more fruitful. Certainly research psychologists in this area are in great need of additional research techniques.

Chapter V

Frequency of Reinforcement in Concept Formation Behavior

General Introduction

The experiments reported in this chapter were conducted with the hopes of explaining the results reported by Grant and Berg (). Their experiment attempted a behavioral analysis of certain aspects of the reinforced process in the Wisconsin Card Sorting Test. This test consisted of a pack of four stimulus cards and 64 response cards which were devised so that each card contained from one to four identical figures of a single color. Four types of figures (star, cross, triangle and circle) and four different colors (red, yellow, green and blue) were used. A single card might have three green stars, or one red circle, or any of the 64 possible combinations of colors, numbers and forms. The four stimulus cards were one red triangle, two green stars, three yellow crosses and four blue circles. Each response card could then be sorted according to the color, the number, or the form of the figure.

The initial correct sorting category was color. As soon as the S made a certain number of consecutive correct responses (reinforcing trials), the E shifted the problem, with no explanation to S, and began to call the number category "right," and all others including color "wrong." A total of five shifts took place; from color to number which has already been mentioned, from number to form, from form to number, then to color, and finally to form.

Seven groups of Ss were run with a criterion of 3, 4, 5, 6, 7, 8 and 10 consecutive reinforcements required before each shift was made. In the treatment of the data the first three groups were combined to form a "low reinforced" group and the remaining groups were combined into a "high reinforced" group. The results indicated that the high reinforced group shifted more easily than the low reinforced group. Such a finding would appear at first glance to be inconsistent with the principle of conditioning that increasing the number of reinforcements creates greater resistance to extinction.

A closer examination of Grant and Berg's data indicates that the greater flexibility demonstrated by the high reinforced group was based upon the last two shifts (from number to color and from color to form). Since these concepts had been learned previously such differences between the high and low reinforced groups might be attributed to relearning. Since Grant and Berg's experimental procedure confounds the effects of extinction and relearning it was decided to experimentally separate these two variables and discover whether increases in reinforcements would produce greater degrees of rigidity.

Experiment 11

This experiment was somewhat similar to Grant and Berg's, except that it sought to separate extinction effects from relearning effects.

Experimental procedure.

Three groups of students from the elementary psychology class at University College, New York University, were given the N.Y.U. Card Sorting Test, Form 1. Four stimulus cards were used, with the appropriate concept being a compound concept involving both size and shape. The groups varied in terms of the criterion of learning they were required to achieve (5, 10 and 25 successive reinforcements). Following learning, all of the three groups received 25 "extinction trials" during which time all responses of the S were called "wrong."

Results and discussion.

The main results are presented in Table 11.

Table 11

Showing Means of Scores Obtained by Groups 5, 10 and 25 during Various Stages of Experiment 11

<u>Stage of Experiment</u>	<u>Score</u>	<u>Group Means</u>			<u>"One-tail" Significances</u>
		<u>5</u>	<u>10</u>	<u>25</u>	
Learning	N of Trials to Learn	39.8	62.3	57.1	None
Extinction	N of Perseverative Responses	12.1	11.3	9.0	5 and 25 (p .03)
Relearning	N of Trials to Relearn	50.5	18.3	8.0	5 and 25 (p .02) 10 and 25 (p .05)

Although different criteria of learning were used, there was not a significant difference between the number of trials required by the various groups to reach their criterion of learning. The results of the extinction series are of interest. One would expect by the simple application of conditioning principles to this experimental situation that there should be a positive relationship between number of reinforcements during learning and number of perseverative responses, i.e., responses which would have been correct during the learning stage. The results indicated a negative relationship. This result is not surprising if we consider the total pattern of stimulation. The greater the number of reinforcements the more strongly is the sorting response conditioned to the "right" response of the E given to the preceding sorting response. When this cue element is withdrawn, according to our analysis, the groups with the greater amount of reinforcements will suffer the greater "generalization" loss.

The results of the relearning series are consistent with our expectations since they demonstrate a positive relationship between number of reinforcements and speed of learning.

Experiment 12

The results of Experiment 11 suggested that the greater the number of reinforcements in concept formation learning, the faster would relearning of the concept occur following a series of "extinction trials." The results also suggested that the greater the number of reinforcements the fewer would be the number of perseverative responses occurring during the extinction series. The purpose of this study was to obtain additional data about the above described phenomena as well as to determine whether these effects would be duplicated in an experiment involving simple concepts, similar to the type used by Grant and Berg.

Experimental procedure.

The Wisconsin Card Sorting Test was used with form serving as the correct concept. Thirty-six Ss were divided into three groups, each being required to reach a different criterion of learning (4, 8 and 16 successive reinforcements). Following this, each S received a series of 20 extinction trials and then was required to relearn the original concept.

Results and discussion.

There were no significant differences or, for that matter, any suggestive trends obtained in any of the three stages of the experiment; learning, extinction or relearning. One possible explanation is that the learning of the simple concepts in the Wisconsin Card Sorting Test involves mainly the acquisition of the correct verbal response. Once this is achieved additional reinforcements have little or no effects.

Experiment 13

This pilot study was conducted with the specific purpose of demonstrating that increasing reinforcements tend to reduce flexibility, i.e., the ability to shift from one concept to another. One reason why this may be difficult to demonstrate in a typical card sorting test, involving successive shifts, is that the Ss not only learn specific concepts but also learn to respond to the major features of the stimulus cards, i.e., the reinforcements they receive strengthens response tendencies associated with specific concepts as well as certain general responses, oriented towards certain features of the cards. Consequently, increasing the number of reinforcements will strengthen both response tendencies, the first of which will interfere with shifting while the second will facilitate it.

Experimental procedure.

Thirty-one Ss were divided among four groups. These groups were trained until they achieved a criterion of 0, 5, 25 or 50 successive correct responses on the form concept with the Wisconsin Card Sorting Test. Following this learning, the correct concept was shifted. The second "correct" concept was one involving position, i.e., a correct response placed the response card to the right of the previous response card regardless of the content of either card.

Results and discussion.

Only 10 of the 31 Ss learned the second concept within 200 trials; 50% of Group 0, 30% of Group 5, 13% of Group 25, and 42% of Group 50. The means of the number of trials to learn the second concept by the successful Ss in each group were; 49 for Group 0, and 148, 176 and 107 for Groups 5, 25 and 50 respectively.

Although the results do not provide any definite evidence of greater number of reinforcements producing greater rigidity, they do suggest that having no reinforcements makes a S more "flexible" than those Ss having some reinforcement on the form concept.

Experiment 14

The purpose of this experiment was the same as the preceding experiment. It sought to accomplish this aim by modifying the experimental procedure so that a larger proportion of the Ss would solve the test problem.

Experimental procedure.

The Ss in this experiment were given the same two problems as the Ss in Experiment 13, except that the procedure was modified so that the S was required to make a series of five successive sorting responses while stating explicitly the hypothesis which dictated his responses. Only after this was done was the S informed by the E whether his sorting responses, as a group, were right or wrong.

It was thought that such a procedure would result in the elimination of reinforcements of sorting responses which were initiated by incorrect hypotheses.

The Ss were divided into four groups of 10 Ss each. They varied in terms of number of reinforcements they had to achieve in the learning of the first concept (0, 1, 5 and 10 successive groups of correct sorting responses).

Results and discussion.

Table 12 indicates the results of the experiment.

Table 12

Percent of Ss Solving Second Concept and Their Mean and Median Number of Trials to Achieve Learning

Group	% Solving Second Concept	N of Trials Required To Learn Second Concept	
		<u>M</u>	<u>Med.</u>
0	70	6.6	4
1	60	17.5	17
5	80	17.0	16
10	50	20.8	22

Although the results of percent of Ss solving the second concept do not suggest any trend the performance of the Ss who learned the second concept indicates that 0 reinforcement facilitated the rate by which the second concept was acquired. The difference between the mean of Group 0 and the combined mean of Groups 1, 5 and 10 in terms of number of trials required by the successful Ss to learn the concept was significant at the .01 level.

It should be noted that the modification introduced in the procedure increased the per cent of Ss solving the second problem from 32% of the previous experiment to 65% in this experiment.

The failure to obtain a clear cut relationship between number of reinforcements and rigidity might be due to the fact that the initial concept is so simple that the maximum of learning is obtained very rapidly, and consequently variations in number of reinforcements above 0 do not have any effect. Perhaps a more complex initial concept, such as used in the N.Y.U. Card Sorting Test would reveal the "long sought-after" relationship between number of reinforcements and rigidity of problem solving behavior.

Experiment 15

Previous experimentation has indicated that the number of trials is not highly related to the ability of Ss to shift from one concept to another. One explanation of this is that the procedure used in the simple card sorting experiments confounds the factors of learning to respond specifically to one concept (e.g., color or form) with the learning to respond generally to one feature of the cards. That is, the stronger the number of reinforcements on the initial concept the stronger will be the tendency for the S to persist in both tendencies. The tendency to persist in responding to the specific feature of the initial concept will retard the learning of the second concept; the tendency to persist in responding to only one feature of the cards will facilitate the acquisition of the second concept only if the second concept is based upon one feature of the card. Our prediction would be that if Ss initially were required to learn a concept based on one stimulus attribute (color) and then were required to learn a second concept based on two stimulus attributes (size and shape), the number of trials to learn the second concept would be positively related to the number of reinforcements received on the initial concept. This would follow because both the specific and general response tendencies acquired when learning the initial concept will interfere with the learning of the second concept in the above-described situation.

Experimental procedure.

Fifty-four students from the elementary psychology course served as Ss. The N.Y.U. Card Sorting Test, Form B, was used. With the first concept to be learned being a simple concept based on color, while the second concept to be learned was a compound concept based on both size and shape. The Ss were divided into four groups varying in terms of the criterion of learning (0, 5, 15 and 30 successive correct responses), they were required to achieve. The shifting from the first to the second concept was immediate and without the Ss' knowledge.

Results and discussion.

Fifteen Ss were unable to solve the second concept (6 from Group 0, and 2, 3 and 4 from Groups 5, 15 and 30 respectively).

Table 13 indicates the mean and median number of trials required to reach the criterion of learning for the second concept.

Table 13

Mean and Median Number of Trials to Reach Criterion of Learning for Second Concept.

<u>Group</u>	<u>Mean</u>	<u>Med.</u>
0	39.7	27.0
5	51.9	38.0
15	59.2	52.0
30	65.2	68.0

Although the trend of results is in the predicted direction, the analysis of variance failed to result in a significant F. The high variability of scores in each group suggests that large numbers of Ss will be required to establish statistical significance.

General Discussion

The results of the experiments in this chapter were, in a certain sense, rather distressing. An attempt was made to demonstrate a rather simple relationship between number of reinforcements and inability to shift from one concept to another. Although this writer feels that the experiments reported suggest a technique by which this relationship can be demonstrated, such demonstration will require large number of Ss.

Why cannot this relationship be shown in as simple a fashion as in the conditioning situation? An obvious answer to this question is that greater variability exists with the human Ss. But why is there greater variability? This question emphasizes the point made throughout this report, viz., that the stimulus situation, in terms of verbal cues, varies tremendously not only between individuals but also within the same individual from time to time. Unless techniques are developed to control (as was done in Experiments 6 and 7) and hold constant these verbal cues, research in this area must be conducted with extremely large groups of Ss.

The experiments reported in this chapter also suggest that the individual trial in conditioning will not simply be coordinated to the single trial in concept formation tests. Perhaps a simpler relationship will appear if we consider the acquisition of a specific concept as one single trial. For example, in learning the shape concept in the N.Y.U. Concept Formation Test, Form 1, four figures are used; a square and a triangle which are "straight-lined" shapes and a crescent and a circle which are "curve-lined" shapes. If Ss after learning this discrimination were required to sort pentagons, hexagons, and two different type ellipses, this second discrimination might be considered the second trial. The number of trials, defined in this manner, would be limited by the imagination of the card sorting test designer. Perhaps with this definition of a single trial the relationship between number of reinforcements and rigidity may be more easily demonstrated as well as the numerous other relationships in conditioning involving frequency.

References

1. Buss, A.H. Rigidity as a function of reversal and nonreversal shifts in the learning of successive discriminations. J. exp. Psychol., 1953, 45, 75-81.
2. Grant, D.A. and Berg, E.A. A behavioral analysis of degree 32 of reinforcement and ease of shifting to new responses in a Weigle-type Card sorting problem. J. exp. Psychol., 1948, 38, 404-411.
3. Guthrie, E.H. The psychology of learning. New York: Harper, 1952.
4. Hull, C.L. Principles of behavior. New York: Appleton-Century, 1943.
5. Kendler, H.H. Drive interaction: II. Experimental analysis of the role of drive in learning theory. J. exp. Psychol., 1945, 35, 188-198.
6. Kendler, H.H., Greenberg, A., and Richman, H. The influence of massed and distributed practice on the development of mental set. J. exp. Psychol., 1952, 43, 21-25.
7. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. J. exp. Psychol., 1946, 36, 271-290.
8. Lewin, K. A dynamic theory of personality. (Trans. by D.K. Adams and K.E. Zener). New York: McGraw-Hill, 1935.
9. Luchins, A.S. Mechanization in problem solving. Psychol. Monogr., 1942, 54, 1-95.
10. Malozuan, I. and Morrisett, L. Different strengths of set in the solution of anagrams. J. exp. Psychol., 1951, 44, 242-246.
11. Mann, H.B. and Whitney, D.R. On a test of whether one or two random variables is statistically larger than the other. Ann. math. Statist., 1947, 18, 50-60.
12. Mowrer, O.H. and Jones, H.M. Extinction and behavior variability as functions of effortfulness of task. J. exp. Psychol., 1943, 33, 369-386.
13. Osgood, L. and Underwood, B.J. Studies of distributed practice: V. Learning and retention of concepts. J. exp. Psychol., 1952, 43, 143-148.
14. Reynolds, B. The acquisition of a trace conditioned response as a function of the magnitude of the stimulus trace. J. exp. Psychol., 1945, 35, 81-95.
15. Rohrer, J. Experimental extinction of a function of the distribution of extinction trials and response strength. J. exp. Psychol., 1947, 37, 473-493.
16. Schroder, H.M. and Rotter, J.B. Rigidity as learned behavior. J. exp. Psychol., 1952, 44, 141-150.
17. Sheffield, Virginia F. Resistance to extinction as a function of the distribution of extinction trials. J. exp. Psychol., 1950, 40, 305-314.
18. Solomon, R.L. Effort and extinction rate: A confirmation. J. comp. Psychol., 1948, 41, 93-101.
19. Wilcoxon, F. A statistical consideration in psychological research. Psychol. Bull., 1951, 48, 150-152.
20. Wertheimer, M. Productive thinking. New York: Harper, 1945.