1971

Learning and transfer in hierarchically structured concept attainment tasks.

Sergio V. De Luna
University of Massachusetts Amherst

LEARNING AND TRANSFER IN HIERARCHICALLY STRUCTURED CONCEPT ATTAINMENT TASKS

A Master's Thesis Presented

By

SERGIO V. de LUNA

Submitted to the Graduate School of the

University of Massachusetts in

partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

September, 1971

Department of Psychology

LEARNING AND TRANSFER IN HIERARCHICALLY

STRUCTURED CONCEPT ATTAINMENT TASKS

A Master's Thesis presented

By

SERGIO V. de LUNA

Approved as to style and content by :

Chairman of Committee

Member

Member

September, 1971

Introduction

The research program described in this report was conducted in conjunction with a research project directed by Dr. John A. Emrick. This project was designed to investigate the validity of a test model for individualized instruction and utilized concept learning tasks and examined acquisition trends across levels of conceptual difficulty. The rationale for this test model followed from a detailed examination of decision rule logic in single skill testing. Furthermore, this testing is considered a necessary and essential feature of nearly all individualized instruction programs. A more detailed account of the specifics of this test model is described by Emrick and Adams (1969) and also by Emrick (1971).

The Concept Identification Task

The subject's task in a concept identification problem traditionally has been described as involving at least two components: the identification of the relevant dimensions and the identification of the rule or rules which bring the attributes together in a particular fashion (Bruner, Goodnow and Austin, 1956; Haygood and Bourne, 1965; Bourne, 1968).

Given a set of dimensions _a_, _b_, _c_..._x_ each with _n_ values or attributes (a1, a2, a3...an; b1, b2, etc.) the subject's task is to learn which of these attributes satisfy the conditions defining the concept. These attributes which satisfy the concept definition are said to be relevant and the dimensions to which they belong are called relevant dimensions. In a concept made by the attributes a1, b2 and c3, these attributes would then be the relevant attributes; all other attributes (including the remainders of dimensions _a_, _b_ and _c_) and all other dimensions are said to be irrelevant to this particular concept.

However, the learning or identification of a concept goes beyond simple identification of relevant attributes. Two or more quite different concepts might consist of identical attributes but differ in the way the latter are arranged. For example, attributes a1 and b2 might be the relevant attributes of two different concepts, namely "both a1 and b2 must be present" vs. "a1 or b2 must be present but not both".

The particular arrangement of the relevant attributes of a concept is known as the conceptual rule. Neisser and Weene (1962) have shown that when the number of relevant attributes is restricted to two, there are 10 such conceptual rules. Table 1 describes these 10 possible conceptual rules.

Insert Table 1 about here

A substantial interest exists concerning the relative influence of each of these two components -- attribute learning vs. rule learning -- in the process of concept identification.

For the most part, research has concentrated either on the identification of the relevant attributes (Hovland and Weiss, 1953; Archer, Bourne and Brown, 1955; Bourne and Haygood, 1959) or on the relative difficulty of different rules (Bruner, Goodnow and Austin, 1956; Neisser and Weene, 1962).

Research on attribute learning has demonstrated the effects of such variables as the number of relevant and irrelevant dimensions (Walker and Bourne, 1961) and the amount of intra- and inter-dimensional variability (Battig and Bourne, 1961). For example, Battig and Bourne's (1961) investigation on the effects on error-rate of changes in the number of dimensions and changes in the number of values within each dimension revealed that college students made more errors following both inter- and intra-dimensional variations. Further, this relationship between error-rate and intra-dimensional variability was found to correspond very closely to a straight line function.

The amount of irrelevant and relevant information has also been shown to contribute to task complexity. Although it would seem on an intuitive basis that increased relevant information should increase the difficulty of the conceptual task, it is not so obvious that increased irrelevant information should do so. Actually, the amount of irrelevant information affects only the complexity of the stimulus pattern, since the number and type of categories into which the patterns must be sorted will remain the same. Further, Walker and Bourne's (1961) study indicated an interaction between the amount of both relevant and irrelevant information and problem difficulty. Errors increased at a positively accelerated rate with increases in relevant information, but this effect depended on the level of irrelevant information employed in a problem.

Of principal interest to this proposal are the results of research concerned with the learning of conceptual rules. Neisser and Weene (1962), using the rules described in Table I, demonstrated that rules are not of equal learning difficulty even though they refer to the same set of attributes. Neisser and Weene further showed that the different rules fall logically into three categories or levels based on the number of component elements. Table II shows the three levels and the rules comprising each level. Their results indicated that the degree of difficulty of each rule increased from level to level, level I containing the easiest rules.

Insert Table 2 about here

Although the results seem hardly surprising, it is not immediately clear why the different rules would distribute themselves along this continuum of difficulty.

Several explanations have been offered in an attempt to explain why certain rules are more difficult to attain than others. One possibility suggested by Haygood and Bourne (1965) is that subjects are forming and testing various rule-hypotheses until the correct one is discovered. Thus, as concept increases in complexity, more rules become available, reducing the probability of an early solution. This explanation is very similar to - if not the same as - the decision-tree model suggested by Hunt (1962, cited by Haygood and Bourne, 1965). It easily can be shown that the rules of each level in Table II are built upon the rules of the previous levels. Therefore, as one moves from one level to another, larger samples of stimuli would be required to eliminate competing rules, which is consistent with Hunt's hypothesis. However, even though this may be true in some cases, this explanation does not seem entirely satisfactory. For one thing, it requires that subjects have all the rules available, which may not necessarily be the case. For example, Haygood and Bourne (1965) and Bourne (1968) reported that when naive subjects were given the relevant attributes of a concept and were asked to identify it, subjects showed a strong preference for conjunctive rules even though there were no pre-established rules (i.e., any rule would define the concept). Moreover, subjects seemed surprised at the possibility that other rules might have been involved. Another study reported by Neisser and Weene (1962) evaluated this assumption of availability of rules. A computer was programmed to identify concepts of varying difficulty (see Table II) using a logical elimination strategy. The results indicated that the time (number of steps) required for the computer to identify each concept was inversely related to the structural simplicity of the rule. These results strongly imply that something other than - or in addition to - simple logical elimination is involved in human concept identification strategy.

Another hypothesis suggested by Neisser and Weene (1962) concerns differences in difficulty of comprehension or verbal expression within the various classes of rules. For example, the authors found that rules within the second classification were not of equal difficulty, two being very difficult. Also, Haygood and Bourne (1965), working only with rules from the second classification, showed that the rules differed in relative difficulty. Yet this hypothesis of differential rule difficulty, even though plausible, still does not seem to account for all the data. For example, Neisser and Weene (1962) reported that subjects seemed to have better verbal understanding of complex rules such as "either/or" than of the more rapidly learned (i.e., "easier") category-2 rules.

In view of all these arguments, Neisser and Weene (1962) suggest that their data can be better explained in terms of a hierarchical organization. The hierarchy can be seen by looking at the problems in Table II. For example, compare three rules shown at the Table: (A), (A. -B) and (A. -B) v (-A. B). According to Neisser and Weene (1962) the facilitative effect of learning lower-level concepts before learning more complex concepts lies in the fact that to solve rule (A. -B) subjects must learn what (A) and (-B) mean; following the same reasoning, learning (A. -B) will facilitate learning of (A. -B) v (-A. B). It thus appears important to turn to the issue of hierarchical conceptual learning.

The Hierarchical Organization of Concepts

Neisser and Weene's data tend to provide experimental support for a theoretical model being developed by Gagne and which is fully described in Gagne, 1965. Specifically, Gagne's model describes learning as increasing in stages of complexity and difficulty in hierarchical terms. The differential difficulty of concept learning for ostensibly similar concepts -- as reported by Neisser and Weene --

corresponds well to his theoretical interpretation.

The basic working principle of Gagne's model is the description of learning as a cumulative process. More specifically, he states that "within limitations imposed by growth, behavioral developments result from the cumulative effects of learning" (Gagne, 1968, p. 178). This statement is better understood if one considers the area of concentration of Gagne's work and, within it, the type of learning which constitutes his primary concern.

It is perhaps not an understatement to say that Gagne has been basically concerned with applied research and that, for the most part, his work deals with instructional procedures for the teaching of mathematical concepts (Gagne, 1962a, 1962b, 1963, 1965). In these works he has consistently shown that a complex task can be broken down into its components such that performance in each step of this sequence is dependent upon mastery of the previous steps (for example, Gagne, 1962).

Another characteristic of Gagne's model - and perhaps a consequence of the ones above mentioned - is that it involves mostly what he calls "rule" or "principle" learning. A rule or principle is basically a concept but is distinguished from the latter in that:

1. While attainment of a concept can be shown by means of an identificatory response (concrete concept or concept by observation) the rule or principle has to be demonstrated (abstract concept or concept by definition)

(Gagne, 1966).

2. A rule or principle is composed by associations, motor and verbal chains, multiple discriminations, concepts, and simple rules (in the case of complex rules) (Gagne, 1965, 1968).

Gagne himself provides a clear example to illustrate the distinction above. Suppose one is looking for a criterion to judge whether or not a person has the concept "radius of a circle". He might ask the person to draw the radius in a circle or to identify which of several lines drawn in a circle is the radius. But what criterion should one use to be sure that a learner knows what "work" represents in physics? There seems to be just one and that is to require the subject to demonstrate that the work produced by a force exerted on a body is the product of the force times the distance the body is moved. Further, this implies that the learner must appreciate or already knows such things as "force", "displacement", "multiplication" and the like.

One of the implications of a rule or principle (as opposed to a concept by observation) is that it is not "learned" but has to be taught (Gagne, 1966). The distinction here seems to relate to the level of abstraction involved in each of these two kinds of concepts. For example, one might expect a subject to learn to identify the radius of a circle even though he is not able to define what the radius of a circle is. The relevant attributes of the concept are all physically contained in the instance and can be isolated, for example, simply by differential reinforcement. A rule or principle, on the other hand, requires relational operations that go far beyond the observable properties of the stimuli (as, for example, in the principle of "work"). According to Gagne, even in the case where the subject has mastered all the discriminations and concepts involved in the rule, he is not likely to demonstrate the rule if he has not been taught it. Therefore, rule learning as defined by Gagne seems to differ considerably from the process usually studied in psychological research, which deals with what he calls concept by observation.

Moreover, when one thinks of the concepts that constitute a mathematical rule, it is apparent that the hierarchical organization of information becomes an end rather than a means. In the learning of the rule 2N-1, to learn what means is not a facilitatory device but rather a prerequisite (unless, of course, the rule is changed). This notion of hierarchies comprised by prerequisites is recognized by Gagne.

> The hypothesis is proposed that specific transfer from one learning set to another standing above it in the hierarchy will be zero if the lower one cannot be recalled and will range up to 100% if it can be. (Gagne, 1962, p. 358)

A close examination of Table II shows the difference between rule learning as defined by Gagne and the process described by Neisser and Weene (1962). In the latter, subjects could learn rules from any levels independently of previous mastery of rules from lower levels. For example, subjects in Haygood and Bourne's (1965) study commenced with level-II concepts and were quite able to learn them. It might be argued, however, that level-I concepts in Neisser and Weene's experiment consisted of the simple presence (or absence) of an attribute and, therefore, were not different from attribute identification tasks. Since one of the conditions of Haygood and Bourne (1965) was learning of the rules given the attributes, the task in both studies becomes comparable.

There is enough evidence, however, that rules can be learned at any level independently of learning rules from presumably subordinate levels. Haygood and Bourne (1965) and Bourne (1968) have consistently shown that if subjects are given training in discovering rules there is an improvement from problem to problem much like the phenomenon of learning set described by Harlow (1959). Moreover, Haygood and Bourne's (1965) study also included a condition in which subjects had to learn both a rule and the attributes. Although the performance of this group was considerably poorer than that of the other two groups (rule learning with attributes given and attribute identification with rule given) there is no doubt that subjects did learn the task.

Therefore, although the relationship between Neisser and Weene's results and the work developed by Gagne seems to compliment each other, more basic research is needed in order to clarify some of the problems involved in hierarchical organization of concepts. For example, no satisfactory definition of complexity has been given. While Gagne seems to imply that complexity is a function of increasing the number of subordinate rules and concepts assumed prerequisite to the learning of some higher-level rule, Neisser and Weene apparently show complexity as the addition of subsets of rules.

Rationale for the Experiment

This project has been designed in order to further investigate some of the variables involved in the learning of hierarchically structured concepts.

By setting up several levels of concept learning tasks and by imposing a subordinate - superordinate relationship within the relevant dimensions from level to level, it is possible to make the conceptual tasks to correspond roughly to the notion of a hierarchy as described by Gagne (1965) and Neisser and Weene (1962). Thus, given a certain number of dimensions, complexity across levels can be manipulated by systematically varying the number of changing dimensions and/or by increasing the number of such dimensions selected as relevant to the problem. Table III shows three levels constructed on the basis of the criteria of complexity given above. There are several advantages in this procedure.

Insert Table 3 about here

First, it allows for elimination of the effects of the relative difficulty of different rules. As it can be seen in Table III, level-I concepts contain 3 varying dimensions, one being relevant to the problem. Each subsequent level builds on previous levels by adding new relevant dimensions to the problems (level II adds one more dimension, level III adds two more dimensions). Therefore, the transfer of learned prerequisites from one level to another can be assessed independently of the effects of different rules.

Another advantage of this approach is that it allows for evaluation of altering solution attributes within dimensions across subsequent problem levels. Neisser and Weene's study, for instance, failed to show what was the precise nature of the transfer effects from level to level. The effects of learning the rule (A) should be more dramatically shown in the learning of such a rule as (A. -B) than a rule as (-A. B). In fact, Neisser and Weene (1962) were not able to account for the fact that (-A) was more easily learned than (A). Likewise, learning the rule (-A. B) should yield more positive transfer to learning (A. -B) v (-A. B) than should learning the rule (-A v B).

As it can be seen in Table III, the attributes of a particular dimension may be kept constant or may be varied across levels. Thus, dimension A is represented by values 2, 3 and 4, respectively, while dimension B remains constant. In order to gain information about the effects of shifting the attributes of the relevant dimensions from level to level, each level can be subdivided into two groups: for group 1, attributes remain constant whereas for group 2, attributes vary across levels. Since both groups 1 and 2 have to solve problems based on the same rules, transfer effects from level to level should be higher for group 1 than for group 2.

As stated earlier, Gagne predicts that, given a hierarchy of tasks, transfer will be zero if previously learned prerequisites are not recalled, and will range up to 100% if they are. Therefore, transfer from level to level depends upon mastery of subordinate concepts. This prediction was not tested in Neisser and Weene's experiment primarily due to the nature of their design (repeated measure design). That is, since relatively few subjects received all treatments in unspecified order, there was no basis (indeed no intention) for assessment of hierarchical transfer effects.

To test Gagne's assumption of mastery of prerequisites in hierarchical learning tasks, three groups of subjects were run according to three different criteria: 1. low level of training, 2. moderate training level, and 3. high level of training. The theoretical number of trials required to learn in each level was computed following a model suggested by Trabasso and Bower (1968). Specifically, given Trabasso's and Bower's (1968) assumption of equisaliency of dimensions, and their calculations of expected error rates in conjunctive problems of this type, average required trials to mastery was estimated to be 20. The number thus estimated was used as the criterion for Group 2. Making this number of trials equal to n, Groups 1 and 3 received n/2 and 2n trials, respectively. Table 4 lists the expected number of trials for each group.

The specific hypothesis tested by this procedure predicted error rates on subsequent problems to be a decreasing linear function of prior training criteria. Since Group 1 was advanced to new problems prior to the predicted solution state (see Table 4) it was predicted that the performance of subjects in this group would be considerably poorer than that of subjects in Groups 2 or 3.

Criterion tests were administered to half of the subjects immediately following completion of the training phase for each concept. Thus, after completing the 10 trials for level 1, half of the subjects in Group 1 received a criterion test, half did not. A similar split was made for the criterion groups (Groups 2 and 3). Following completion of the third concept (level 3 problems) all subjects received a criterion test. This testing procedure was intended to control for learning effects in testing.

The experiment was conducted in two phases: a short "warm-up" session and the training phase itself. The goals of the "warm-up" session were as follows:

1. Since feedback would be presented via magnetic tape it was decided that confirmation of the correct response would be given in terms of the position occupied by the positive stimulus. Consequently, one of the goals of the "warm-up" session was to test and/or teach the understanding of ordinal position, and the ability to match the position of the stimuli with the corresponding spaces provided in the answer sheets.
2. The second objective of the "warm-up" session was to acquaint subjects with the ultimate goal of the problem, namely to "guess" which of the stimuli was correct, and to discover the conceptual rule.

METHOD

Experimental design.

The experiment involved a 4-way repeated measure analysis of variance design. The repeated measure variable was represented by the three levels of problems (concepts 1, 2 and 3). The independent group factors were: training condition, with three levels (low, moderate and high levels of training), transfer conditions, with 2 lines (shift of attributes and no-shift of attributes), and one test factor (continuous testing *vs*. final test only).

Subjects.

There were 96 *Ss* from third grade, randomly assigned to each of the experimental conditions. Subjects were recruited from elementary schools in Greenfield, Whately and W. Springfield. Also, girls and boys were evenly distributed across treatments.

Material.

Stimuli were composed of combinations of attributes of six dimensions. Table 5 lists the six dimensions and their respective attributes. One particular stimulus could be a large, green triangle filled in with slashes and with two Xs at the bottom (borders), or a small, blue diamond filled in with Os and with four Xs at the top and so on.

These stimuli were mounted on 35 mm slides such that each slide was composed of an array of four stimuli. The relevant

----

Insert Table 5 about here

----

dimensions for each level and their respective attributes are described in Table 6, and Figure 2 exhibits one instance of stimulus-array for each level, as they appeared in the experiment. As can be seen in Figure 2, the method of pre-

----

Insert Table 6 about here

Insert Figure 1 about here

----

senting the variations in the dimensions differed from concept to concept. The three dimensions varied in the first concept were presented in such a way that all four attributes were presented in each problem. For the second and third concepts taught, only two attributes within each dimension were varied in a given problem.

Except for dimension shape which was made irrelevant for all problems, all dimensions and attributes were orthogonally manipulated with the only restriction being that each slide should contain one and only one positive instance.

Apparatus.

The training and test materials were projected via a 35 mm carousel slide projector on a screen in full view of the Ss. Also aside from the introduction and pre-training which was presented verbally by the E, all subsequent training and testing instructions were presented via a magnetic tape recorder. The Ss responded individually to each of the training and testing problems by marking their choice in a response booklet. This response booklet for the training problems contained one page for each problem. On the page there were four boxes, one box corresponding to each stimulus position on the projected screen. When instructed by E via the tape, the S selected his choice and indicated same in his response booklet. He was then informed by E what the correct stimulus choice was for that problem. The S was then directed to observe whether he had chosen correctly.

Procedure.

Each step of this experiment was conducted in two phases:

training and testing. The training phase consisted of a number of trials with feedback appropriate to the concept being taught. Testing consisted of blank trials. The experiment was conducted such that eight children were escorted from their classroom to the experimental room and seated. They were then instructed to fill out certain information on the training booklet in front of them. This information included their name, their age, sex and set number from 1 to 8. They were then given some preliminary instruction concerning the nature of the task in which they were to engage. This pre-training included a brief instructional unit in which they were taught how to make responses for specific choices on the screen and also an introduction as to the nature of the specific problems that they would be attempting to solve. Specifically, the children were told that they would be playing a learning game. Instructions went as follows:

> "The nature of the game will be for you to choose the correct picture when I show you several pictures on the screen like this (the slide projector was then turned on and four stimulus figures appeared on the screen). Here we see four pictures. This is the first picture (the E points to the leftmost picture); this is the second picture (he points to the second picture); this is the third picture and this is the fourth picture (he points to the rightmost picture). Now look at the first page of your booklet. Do you see those four boxes? (E waits for Ss to acknowledge) Each one of those boxes goes with a picture you see on the screen. The first box would go with the first picture; the second box would go with the second picture; the third box would go with the third picture and the fourth box would go with the fourth picture. Now, everybody look at the pictures again. Do you see the circle? (Pause) All right,

now suppose that you wanted to choose the circle. Mark an X on your answer sheet that shows that you are choosing the circle. (Pause) How many people chose the third box? Raise your hand if you chose the third box. (Pause) All right, let's try another one. Turn over to the next page. (E then projects a new slide on the screen in which the circle moves to position 2.) All right, now let's see if you remember how to play this game. Suppose that you wanted to choose the circle again. Mark the box that would show that you are choosing the circle. How many chose the second box? (Pause) Very good. All right, let's try once again. (E advances to a new slide.) Turn to the third page. Now mark the box for the circle. How many marked the first box? (Pause) Very good. From now on I'll be talking to you over the tape recorder but I want you to keep in mind a few things that are very important. First, this is a learning game so you want to try to do your best but you also want to be sure that you do your own work. Don't be concerned with what your neighbor is doing because he may be doing things wrong. Second, we'll have a lot of problems to do and each problem goes on a different page. I'll tell you which page it goes on so you be sure you look to see that you are on the correct page. It is very easy to skip a page and be on the wrong one, so look very carefully. Third, once you've made a mark for your choice, don't change it. If you have a problem, simply raise your hand and we'll help you. All right? Very good. I'll be talking to you on the tape recorder from now on. Remember, if you have a problem, just raise your hand".

The rest of the experiment was presented automatically via the magnetic tape recorder and slide projector. Two Es participated in this training, and occasionally a third was added to assist in the training. For the first five or six problems, the second E stood at the front of the room and when the correct choice was announced via the tape recorder he also indicated the correct choice by pointing a marker on the projector screen. The instructions presented on the magnetic tape recorder initially introduced the Ss to the specific problems."

Initial problems were presented at a fairly stable rate of 15 seconds observation time and 10 seconds per problem. For later problems this rate was advanced to roughly 10 seconds observation time and five seconds response and feedback time (such that four problems per minute were presented for the later slides in the sequence). The test items were presented at a fairly stable rate of 15 seconds per item; there was no feedback interval.

The eight Ss who served simultaneously at each session of the experiment actually constituted four subgroups of two Ss each. One subgroup of Ss remained throughout all activities for a given training condition. That is, they received all training and all test items. The second group received only the first five of each ten-item test. They were excused from the room and waited outside after they completed the first five items for each of the three tests. The third and fourth groups were excused from the experiment immediately following training for the first and second concepts. They were reintroduced after the tests were completed. All children received the first five items of the terminal test. However, only the first and third groups of children received the last five items of the terminal test. This procedure did not produce any noticeable negative side effects, particularly with the children who remained throughout nature of the problems they would be solving as follows:

> "All right boys and girls, we're now ready to begin. Now, as we explained to you, the purpose of this game is to choose the correct picture. Now, when I show you a problem on the screen, look carefully at each picture. Then, when I tell you, choose one of the pictures by making a mark in your booklet. After everybody has had time to choose the picture, I'll tell you which picture was right so you can see if you chose the correct one. Now, theres a reason why certain pictures are correct and others are not. When you discover this reason, you'll be able to get all of the problems right. So this means that at first you'll get some of the problems wrong. Don't feel bad but try to find the secret so you'll get the rest correct. Work quickly but carefully. Do your own work and don't change any answers once you've made them. I'll say which page each problem goes with so you'll be sure that you're not on the wrong page. All right, let's begin.
>
> Here is the problem for page one. You all should be on the first page of your booklet. See each picture carefully. Now mark the one that you think is correct. (Pause) If you marked the third picture, you were correct. The third picture."

This procedure was repeated for each of the training problems. The number of problems presented was determined by the learning condition and the concept level of the particular training sequence.

A sample of the instructions given to the children receiving continuous testing (in this case, the first concept tested) is given below:

> "All right, let's continue with the game only we're going to play it a little differently than before. Each of you has a sheet of paper on which you have written your name. Now I'll show you some problems just like before and for each problem you are to choose the picture that you think is correct. However, I'm not going to tell you which one is correct for these problems. All right, now I'll tell you which line you should be on for each of these

the experiment (that is, received all training and testing). Moreover, the children who did not receive continuous testing (i.e., were excused from the experiment during the first and second tests) appeared somewhat upset that they were not able to participate in everything.

RESULTS

Separate analyses of variance were performed on training and blank trial (test) data. Also, a conditional analysis of initial response tendencies was performed on the initial training trials. The results of these analyses are as follows:

Training data. Responses obtained for the training problems were tabulated from the response booklets and summarized into ten-trial blocks for analysis. Differences in performance as a function of training paradigm (shift *vs.* no-shift) and level of training were assessed by means of analysis of variance on the last ten-item block score. This analysis was performed across concepts (i.e., the concept was treated as a repeated measure variable) and the results are summarized in Table 7.

Insert Table 7 about here

These results reveal significant differences attributable to the level of training ($F = 7.24$, $p < .01$), problem complexity ($F = 10.80$, $p < .01$) and a differential effect of training paradigm on problem complexity ($F = 5.01$, $p < .01$). Specifically, average performances in terms of the three training levels were 5.28, 7.06 and 7.02, respectively, indicating that the principal effect for training level occurs between underlearning and criterion learning levels.

The mean performance scores as a function of problem complexity were 7.00 for concept 1, 6.71 for concept 2 and 5.66 for concept 3. Hence, the assumption that the three concepts are ordered in terms of complexity is supported.

The differential effect of training paradigm on performance as a function of problem complexity is displayed in Figure 2. In this figure, a systematic response decrement occurs for

----

Insert Figure 2 about here

----

mean shift performance across concepts. On the other hand, the no-shift groups display nearly identical mean scores across concepts.

No other source of variance attained significance although the shift groups averaged slightly more correct responses than the no-shift groups (6.73 vs. 6.18). Recalling that this difference was greatest on the concept 1 problems (7.75 vs. 6.25), the data were re-examined for evidence of initial stimulus bias. Specifically, the number of responses to "dotted" (positive attribute for the no-shift groups) and "wavy" (positive attribute for the shift groups) given by both groups in trial 1 of concept 1 was examined. An equal saliency of cues assumption leads to the prediction that responses to trial 1 of concept 1 would be equally distributed among stimuli. Table 8 displays the actual response pattern to the trial 1 stimuli for each of the two training groups. A chi-square

---

Insert Table 8 about here

"goodness-of-fit" (to the equal saliency hypothesis) was performed on these response patterns. The results display a departure from this distribution at the .001 level for the shift group ($X^2 = 100.60$) *vs.* the .20 level for the no-shift group ($X^2 = 6.16$).

One possible explanation for these non-uniform response patterns would be sampling bias. This hypothesis, however, is unlikely given random assignment to groups (Exact probability test: $p < .001$). Another possibility is that pre-training tended to sensitize the two groups differentially to the stimuli, yielding the observed patterns. However, this explanation requires the assumption of a within-subject interaction to this pre-training, since the latter was uniform across *S*s. A third possibility is that, due to the methodology employed (groups of 8 children presented with the stimuli and recording their own responses in booklets), some *S*s delayed responding until--or corrected their responses following--feedback. The assumption that one child out of 5 so behaved tends to redistribute the first-item responses into parallel patterns as shown below:

Group	Dotted Triangle	Wavy Circle	Starred Diamond	Slashed Square
No-shift	.15	.67	.08	.10
Shift	.17	.49	.14	.20

However, the circle stimulus still receives excessive proportion of responses, which can be accounted for in terms of pre-training sensitization (or carryover). This explanation receives support upon inspection of trial 2 response data, as presented in Table 9. Specifically, the unadjusted response

---

Insert Table 9 about here

---

patterns of trial 1 are compared with those of trial 2 for each group. An approximate decrement of 25% in the proportion of responses to the circle stimulus occurs on trial 2 for both groups. The proportion of Ss in the shift group which chose the circle on trial 2 can be used as an estimate of the initial tendency to adjust choices on the basis of feedback.

It was subsequently decided to block the training groups in terms of early trial performance patterns. This was done by blocking the Ss in terms of score (number of correct responses) on the first ten training trials. Two blocks (score split =

5 correct responses) were established and the analysis of last 10 training trials were recomputed separately for each concept. The method of unweighted means for unequal cells was used, as suggested by Winer (1962, p. 242).

The results of the concept 1 blocked analysis are summarized below (Table 10). This analysis displays a significant effect for training level ($F = 7.43$, $p < .01$), block ($F = 26.95$, $p < .01$)

Insert Table 10 about here

---

and a trend for the training level by block effect ($F = 2.42$, $p = .08$). The main effect for paradigm failed to reach significance ($F = 2.93$, $p < .10$) although the shift group had a higher mean score than the no-shift group (6.18 vs. 7.07). The average by training level were 5.38, 6.71 and 7.82, respectively.

Tables 11 and 12 summarize the results of the blocked

Insert Tables 11 and 12 about here

---

analysis for concepts 2 and 3, respectively. Concept 2 analysis reveals no significant effect although level of training shows a very strong trend towards significance ($F = 2.97$, $p < .06$). The mean scores in terms of level of training were 5.64, 7.47 and 6.90, respectively, pointing to a possible superiority of the criterion learning groups over underlearning groups.

The results of the analysis of concept 3 show a significant effect for training level ($F = 8.46$, $p < .001$). The specific mean scores for each respective training level were 3.89, 6.53, and 6.27. All other factors failed to reach significance.

Conditional analysis of inter-concept transfer. Another analysis was performed on the data to test for specific transfer from concept to concept, using responses to the first trial of concepts 2 and 3 as dependent variables. Since the positive instances for the no-shift groups in concepts 2 and 3 always contained the same attributes as concept 1 (plus one or two new attributes), it was hypothesized that *S*s from those groups would make more correct first-trial responses to subsequent problems than would *S*s in the shift groups. A *Z* test for differences between proportions was utilized to evaluate this hypothesis (Walker and Lev, 1953, p. 77). The proportion of *S*s in the no-shift groups responding correctly in trial 1 of concepts 2 and 3 was compared to that of *S*s in the shift groups. Table 13 exhibits both the proportions and the results

---

Insert Table 13 about here

---

of the analysis. The obtained *Z* for each comparison is shown below:

Concept 2	Concept 3
NS *vs.* SH: $Z = 2.81^{*}$	NS *vs.* SH: $Z = .37$
L1 *vs.* L2: $Z = -1.62$	L1 *vs.* L2: $Z = 1.19$

$^{*}p < .05$

The analysis indicates a significant difference between the proportion of Ss responding correctly on trial 1 of concept 2, the difference favoring the no-shift group. All other comparisons were non-significant.

Blank trial analysis. The data obtained from the blank trial inserted at the end of each concept training phase were grouped and analyzed in terms of the design variables. Two such analyses were performed: one for the 48 Ss who were tested across all concepts, and one for the 96 Ss who received the final test. The analysis of tests across concepts is summarized in Table 14. These results show a significant effect for

----

Insert Table 14 about here

----

concept ($F = 4.68$, $p < .05$), the average scores for each particular concept being 6.04, 6.93 and 5.46, respectively. A significant interaction was also found for paradigm by concept ($F = 8.82$, $p < .001$) which is shown graphically in Figure 3. As the conceptual complexity increases there tends to be an

----

Insert Figure 3 about here

----

increment in the performance of the no-shift groups and a decrement for the shift groups.

Paradigm and training level did not reach significance (both $F = 2.82$, $p < .10$) but the average score for the shift group is slightly higher than the no-shift group (6.78 vs. 5.51) and the averages for training level increase from level to level (4.98, 6.31 and 7.15).

The analysis of the final test (concept 3) is presented in Table 15. These results show a significant main effect for

---

Insert Table 15 about here

---

training level ($F = 4.21$, $p < .05$) such that the average scores were 3.88 for level 1, 4.62 for level 2, and 6.15 for level 3. Although no other effect reached significance, an interesting interacting trend was observed for paradigm by level. This trend is presented in Figure 4, and as can be seen, a strong linear increment as a function of training level occurs for the no-shift groups, whereas the shift groups display an irregular pattern.

Insert Figure 4 about here

DISCUSSION

Two major hypotheses were evaluated in the experiment. The first was related to the possible facilitative effect of a hierarchical organization of concepts. That is, it was expected that the no-shift group would perform better than the shift group in concepts 2 and 3 because of the hierarchical nature of the problems. For the no-shift group, specific transfer from concept to concept would yield at least a partial solution (thus decreasing the likelihood of an error) because the relevant attribute of each concept would continue to be relevant in subsequent concepts. Conversely, specific transfer from concept to concept for the shift group would likely lead to errors since the relevant attributes were changed from concept to concept.

The specific transfer hypothesis provided the basis for predicting superior performance for the no-shift group on concepts 2 and 3. Since no transfer effects should be apparent in concept 1, the expectation was that the two paradigm groups would not differ in performance on this problem. When the performance of the two paradigm groups are analyzed separately for each concept (Tables 10 and 11) it is apparent that the predictions concerning the facilitative effects of a hierarchical organization of concepts did not receive support. The predicted superiority of the no-shift groups over the shift groups is not present in those results. However, the analysis of variance for performance across concepts did indicate a significant interaction between paradigm and concept, both for the training data (Table 7) and blank trials (Table 14). Therefore, some differential effect for treatment did occur across concepts. The failure to obtain significant main effects for paradigm may be best examined through the results of the separate analysis of each concept.

When performances in concept 1 are compared, a significant difference is found between the no-shift and shift groups, in favor of the latter. However, as shown in the Results section, these differences may be explained through a "response corrector" hypothesis.

The non-significant difference between the two paradigm groups in concepts 2 and 3 may be related to the nature of the problems rather than to paradigm itself. Table 13 shows that the no-shift groups did show some specific transfer, at least from concept 1 to concept 2, as attested by a significant _Z_. However, Tables 11 and 12 show no main effect for paradigm and the test for differences of proportions (Table 13) is not significant for concept 3. The problems for concept 2 presented three new features in relation to concept 1: (1) the solution rule shifted from simple presence to conjunction; (2) each dimension contributed only two attributes in each slide; and (3) the position of the borders became a varying, yet irrelevant, dimension. Feature (2) may explain some of the results obtained.

It was said above that for concept-2 problems each dimension contributed only two attributes per slide (see Figure 2). Therefore, the no-shift groups entered level-2 problems with a 50% chance of making a correct response (assuming that transfer from level 1 occurred) since their previous positive attribute would always be present and reinforced 50% of the time. Subjects in the shift groups, on the other hand, would respond correctly to trial 1 of concept 2 only by chance (their previous reinforced attribute was not present on this particular trial). This prediction was supported as indicated by the analysis in Table 13. The initial advantage of the no-shift groups over the shift groups, however, was not carried on since no significant difference was found for paradigm in concept 2. Given that KCR was being provided Ss in the shift group would be cued to the correct solution in early trials. Once one of the elements of the conjunction had been identified, those Ss, too, would have a 50% chance of making a correct response. Moreover, if concept-2 problems did not present great difficulty, the initial advantage of the no-shift groups would not be strong enough to act as a factor (especially in the last two levels of training conditions where only the last ten-item block was used for purposes of analysis).

The lack of significant differences between paradigms for concept 3 also seems to be due to the nature of the problems. The results in Table 12 indicate no difference between the no-shift and shift groups as to their responses to trial 1 of concept 3. Contrary to what happened in concept-2 problems, the no-shift groups had a 75% chance of making an incorrect response against a 25% chance of being correct, on trial 1 of concept 3. This was so because in concept-3 problems the two values contributed by each dimension were unevenly distributed (3 "green" vs. 1 "red"; 3 "circles" vs.

1 "square", etc.). Furthermore, concept-3 problems seemed to have involved a double complexity. That is, besides requiring the conjunction of four attributes, the problems involved border and position as relevant dimensions. Incidental verbal reports from Ss after the experiment was over tended to indicate that position of the borders was very seldom attended to or attended to only in later trials. If that is actually the case, it is conceivable that the predicted superiority of the no-shift group towards the shift group may have been overcome by a very complex task. This interpretation seems to be supported by the fact that the only strong difference between the two criterion learning groups (L2 and L3) is found in concept 3 (Table 15).

The second major hypothesis evaluated in the experiment was related to the effects of the different levels of training conditions for each concept. As it may be recalled, the theoretical number of trials required for mastery of each concept was computed and three groups were formed according to the number of trials allowed during the training phase. The three groups corresponded to a low level of training (L1), a medium level (L2), and a high level of training (L3). It was predicted that the last two groups would outperform the low level groups in all concepts.

These predictions seem to have received support from the analyses of the training data (Tables 7, 10, 11 and 12) as well as test data (Table 15). The analyses shown in Table 14 (test data) did not reveal a significant effect for training condition but the trend observed ($p < .08$) points in the predicted direction.

Implications and Conclusions

A thorough evaluation of the hypotheses concerning the effects of a hierarchical organization of concepts has to take into consideration variables that bear on methodological aspects of the research. These variables have already been discussed in previous sections but their importance to further developments on the issue calls for a general review.

As may be recalled, the problems given to each paradigm group across the levels of conceptual complexity were comparable in the sense that (1) both groups received, at each level, problems which were defined by the same conceptual rule; and (2) at each particular level both groups had the same dimensions as relevant, the attributes being different. However, a more powerful test of the effects of the hierarchical organization of concepts would be obtained if the two groups received the same problems at level 3. Such factors as stimulus preference--if existent--might be eliminated in a design where the three levels of complexity aimed to a final task which was common to both no-shift and shift groups.

The reliability of the training data was somewhat decreased due to the methodology used. In fact, there is no guarantee that Ss did not change their response following feedback.

Two aspects of the procedure employed account for the unreliability of the training data. First, Ss recorded their own responses in the booklet. Since Ss were run in groups, it became quite difficult to prevent the changing of responses in spite of the fact that Ss were instructed not to do so.

Second, after all Ss had (supposedly) made their choice and recorded their responses, knowledge of correct results (KCR) was provided. Therefore, Ss were not only told whether or not their response was correct but also which response was correct. The precision of response recording might be increased either by running Ss individually (thus increasing the control of E over S) or by transferring the actual recording of responses from Ss to either E or some mechanical device. Also, the use of a simple "correct-incorrect" feedback system might make it less tempting to Ss to change incorrect responses.

A final remark should be made as to the actual stimuli used. The inferences drawn from the pattern of responses obtained from no-shift and shift groups at concept 1 might have been clearer if the problems used during the pre-training phase were unrelated to the experimental stimuli. As it may be recalled, one of the hypotheses suggested to explain the discrepancies found between the two groups on concept 1 was the fact that Ss might have been sensitized to respond to the circle. Furthermore, it might be helpful to know Ss' behavior towards the experimental dimensions before engaging in the actual research. In the present research, the hypothesis of a stimulus preference cannot be ruled out completely, even though it seems unlikely that that has been the case.

References

Archer, E. J., Bourne, L. E., Jr. & Brown, F. G. Concept identification as a function of irrelevant information as instructions. Journal of Experimental Psychology, 1955, 49, 153-154.

Battig, W. F. & Bourne, L. E., Jr. Concept identification as a function of intra- and inter-dimensional variability. Journal of Experimental Psychology, 1961, 61, 4, 329-333.

Bourne, L. E., Jr. & Haygood, R. C. The role of stimulus redundancy in concept identification. Journal of Experimental Psychology, 1959, 58, 232-238.

Bourne, L. E., Jr. Concept attainment. In Theodore R. Dixon & David L. Horton (Eds.), Verbal behavior and general behavior theory. New York: Prentice-Hall, Inc., 1968, chap. 10.

Bourne, L. E., Jr. & Guy, D. E. Learning conceptual rules: I. Some interrule transfer effects. Journal of Experimental Psychology, 1968, 70, 3, 423-429.

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

Emrick, J. A. & Adams, E. N. An evaluation model for individualized instruction. Research Report R.C. 2674 -- Thomas Watson Research Center, Yorktown,Hts., N.Y., October, 1969.

Emrick, J. A. The validation of an evaluation model for individualized instruction. USOE Research Grant, DEG 1-71-0002 (509), September, 1970.

Gagne, R. M. The acquisition of knowledge. <u>Psychological Review</u>, 1962a, 69, 4, 355-365.

Gagne, R. M., Mayor, J. R., Garsterns, H. L. & Paradise, N. E. Factors in acquiring knowledge of a mathematical task. <u>Psychological Monographs</u>, 1962b, 76 (7, Whole No. 526).

Gagne, R. M. <u>The conditions of learning</u>. New York: Holt, Rinehart & Winston, 1965.

Gagne, R. M. The learning of principles. In Herbert Klausmeier & Chester W. Harris (Eds.), <u>Analysis of concept learning</u>. New York: Academic Press, 1966.

Gagne, R. M. Contributions of learning to human development. <u>Psychological Review</u>, 1968, 75, 3, 177-191.

Harlow, H. Learning set and error factor theory. In S. Koch (Ed.), <u>Psychology, A study of a science</u>, vol. 2., New York: McGraw-Hill, 1959.

Haygood, R. C. & Bourne, L. E., Jr. Attribute- and rule-learning aspects of conceptual behavior. <u>Psychological Review</u>, 1965, 72, 3, 175-195.

Hovland, C. I. & Weiss, W. Transmission of information concerning concepts through positive and negative instances. <u>Journal of Experimental Psychology</u>, 1953, 45, 175-182.

Myers, J. L. <u>Fundamentals of experimental design</u>. Boston: Allyn & Bacon, 1966.

Neisser, U. & Weene, P. Hierarchies in concept attainment. <u>Journal of Experimental Psychology</u>, 1962, 64, 6, 640-645.

Trabasso, T. & Bower, G. H. Attention in learning: theory and research. New York: John Wiley & Sons, Inc., 1968.

Walker, C. M. & Bourne, L. E., Jr. The identification of concepts as a function of amounts of irrelevant and relevant information. American Journal of Psychology, 1961, 74, 410-417.

Walker, H. M. & Lev, J. Elementary statistical methods. New York: Holt, Rinehart & Winston, 1958.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

TABLE 1

CONCEPTUAL RULES DESCRIBING PARTITIONS OF A POPULATION WITH TWO FOCAL ATTRIBUTES (RED AND SQUARE)

Name	Verbal description
1. Affirmation	All red patterns are examples of the concept.
2. Conjunction	All red <u>and</u> square patterns are examples.
3. Inclusive disjunction	All patterns which are red <u>and</u> square <u>or both</u> are examples.
4. Conditional	<u>If</u> a pattern is red <u>then</u> it must be square to be an example.
5. Biconditional	Red patterns are examples <u>if and only if</u> they are square.
6. Negation	All patterns which are <u>not</u> red are examples of the concept.
7. Alternative denial	All patterns which are <u>either not</u> red <u>or not</u> square are examples.
8. Joint denial	All patterns which are <u>neither</u> red <u>nor</u> square are examples
9. Exclusion	All patterns which are red <u>and not</u> square are examples.
10. Exclusive disjunction	All patterns which are red <u>or</u> square <u>but not</u> both are examples

Modified from Haygood and Bourne, 1965.

TABLE 2

TYPES OF ATTRIBUTES WHICH CAN BE DEFINED BY PRESENCE OR ABSENCE OF TWO FEATURES

Level	Name and symbolic designation	Description of positive instance
I	Presence (A)	A must be present
	Absence (-A)	A must not be present
II	Conjunction (A.B)	Both A and B must be present.
	Disjunction (AvB)	Either A or B or both must be present.
	Exclusion (A.-B)	A must be present and B not present
	Disjunctive absence (-Av-B)	Either A or B or both must be absent
	Conjunctive absence (-A.-B)	A and B must both be absent.
	Implication (-AvB)	A may be absent but if A is present then B must be also; thus A implies B.
III	Either/or (A.-B)v(-A.B)	Either A or B must be present but not both together.
	Both/neither (A.B) v (-A.B)	Both A and B must be present, unless neither is.

Modified from Neisser and Weene, 1962.

In fact, there are five pairs of complimentary rules rather than ten different rules.

TABLE 3

LEVELS OF COMPLEXITY AS DEFINED BY THE NUMBER OF RELEVANT DIMENSIONS AND VARYING DIMENSIONS IN EACH LEVEL

Level	Name and symbolic designation	Description of positive instance	Number of varying dimensions	Number of irrelevant dimensions
I	Presence of A (A)	A2 must be present.	3	2
II	Conjunction (A.B)	Both A3 and B1 must be present.	4	2
III	Conjunction (A.B.C.D.)	A4, B1, C2, and D1 must all be present.	6	2

TABLE 4

NUMBER OF TRIALS FOR EACH LEARNING CONDITION ON EACH LEVEL

	Level		
Learning condition	I	II	III
1. Underlearning	10	10	20
2. Learning to criterion	20	20	40
3. Overlearning	40	40	80

TABLE 5

DIMENSIONS AND RESPECTIVE ATTRIBUTES

Shape (A)	Color (B)	Size (C)	Pattern (D)	Borders (E)	Position (F)
1. Square	1. Blue	1. Large (25 sq. cm.)	1. (((wavy	1. X	1. Top
2. Circle	2. Red		2. 000 dots	2. XX	
3. Diamond	3. Green	2. Small (12.5 sq. cm.)	3. /// slashes	3. XXX	2. Bottom
4. Triangle	4. Yellow		4. *** stars	4. XXXX	

TABLE 6

DESCRIPTION OF STIMULUS VARIATIONS COMPRISING

THE THREE EXPERIMENTAL CONCEPTS

Training paradigm	Stimulus variables	Level 1	Level 2	Level 3
Intradimensional shift	Relevant	Pattern(wavy)	Pattern (starred) Color (yellow)	Pattern (slashed) Color (red) Border (XXX) Position (bottom)
	Irrelevant	Shape Size	Shape Position	Shape Size
	Relevant	Pattern (dotted)	Pattern (dotted) Color (green)	Pattern (dotted) Color (green) Border (XXXX) Position (top)
	Irrelevant	Shape Size	Shape Position	Shape Size

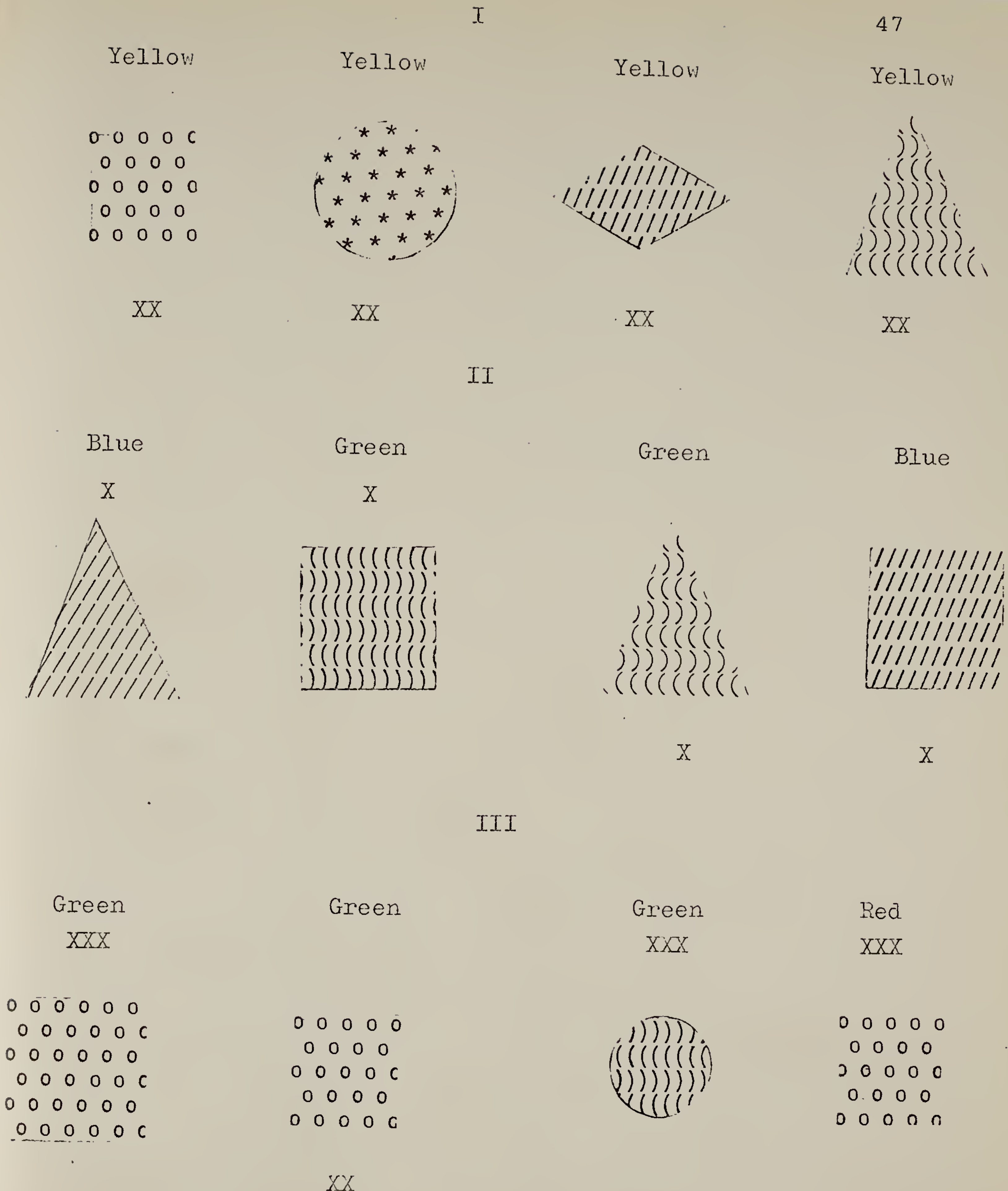

FIGURE 1 Sample problems for the different conceptual rules .

TABLE 7

SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE LAST 10 TRAINING TRIALS OF EACH CONCEPT

SOURCE	DF	SS	MS	F
Between				
P	1	21.67	21.67	1.5
L	2	198.42	99.21	7.2**
PL	2	4.67	2.34	.2
S/PL	90	1234.65	13.72	
Within				
C	2	95.92	47.96	10.8**
PC	2	45.09	27.54	5.0**
LC	4	31.59	7.89	1.8
PLC	4	11.85	2.96	.7
SC/PL	180	799.54	4.44	

**$p < .001$

P = paradigm

L = level of training

C = concept

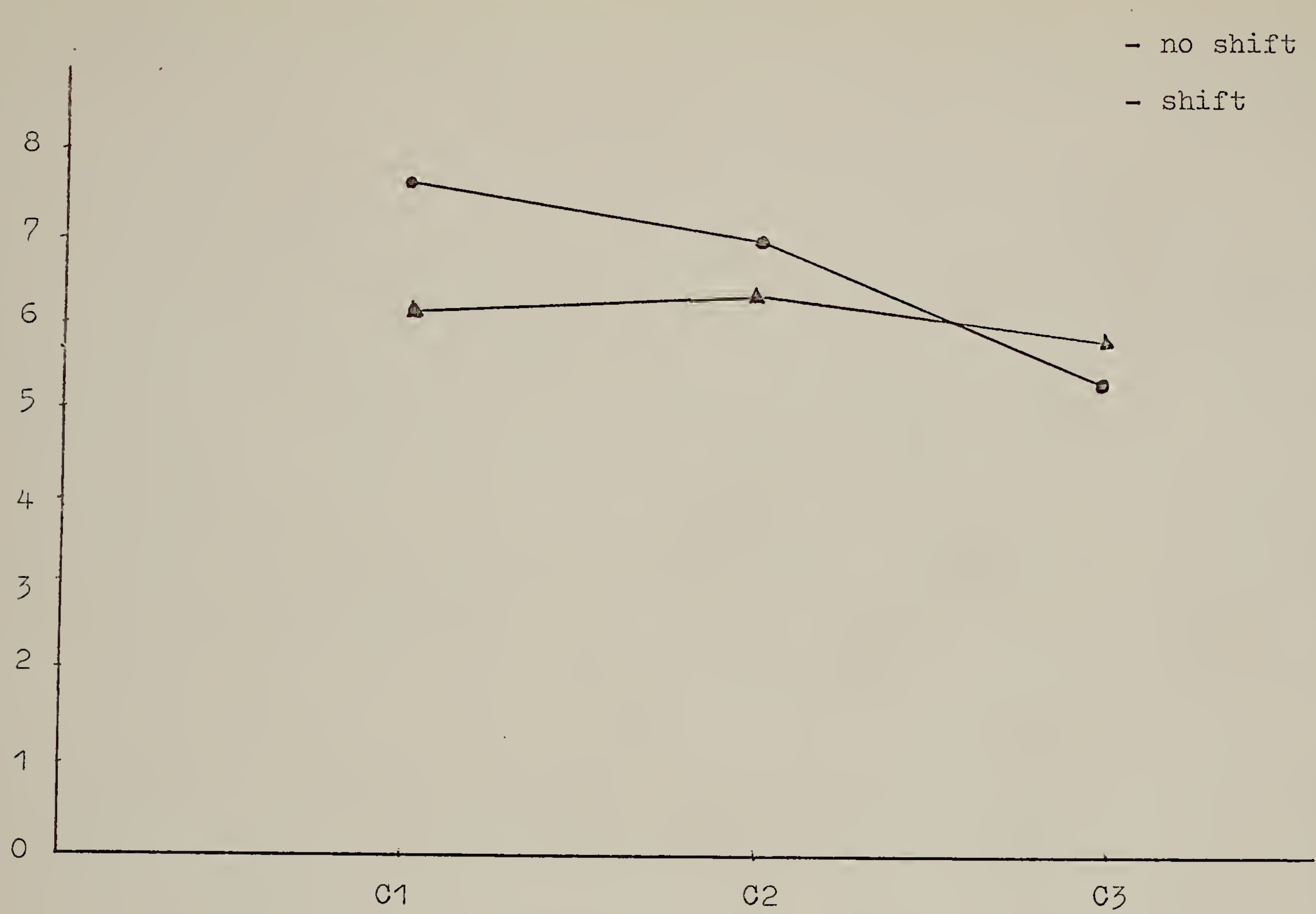

FIGURE 2 Plot of mean training scores (last ten items) for differential effect of training paradigm by problem complexity (Table 7).

TABLE 8

RESPONSE PATTERNS OF THE SHIFT AND NO-SHIFT GROUPS TO THE TRIAL 1 TRAINING STIMULI

Group	Varying dimension	1	2	3	4
	Pattern	Dotted	Wavy	Starred	Slashed
	Shape	Triangle	Circle	Diamond	Square
	Size	Large	Small	Large	Small
NO-SHIFT		.37	.29	.16	.18
SHIFT		.00	.87	.05	.08
PREDICTED*		.25	.25	.25	.25

*(Assuming equal saliency of cues)

TABLE 9

RESPONSE PATTERNS OF THE SHIFT AND NO-SHIFT GROUPS

TO TRIALS 1 AND 2 of CONCEPT 1

Trial 1	Group	Stimulus description (varying dimensions)			
		1	2	3	4
		Dotted	Wavy	Starred	Slashed
		Triangle	Circle	Diamond	Square
		Large	Small	Large	Small
	NS	.37	.29	.16	.18
	SH	.00	.87	.05	.08
Trial 2	Group	Stimulus description (varying dimensions)			
		1	2	3	4
		Starred	Dotted	Slashed	Wavy
		Circle	Diamond	Triangle	Square
		Large	Small	Small	Large
	NS	.08	.45	.33	.08
	SH	.50	.07	.08	.35

TABLE 10

SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE LAST TEN TRAINING TRIALS OF CONCEPT 1

SOURCE	DF	SS	MS	F
L	2	84.61	42.30	7.43*
P	1	16.72	16.72	2.94
B	1	153.45	153.45	26.95*
LP	2	3.83	1.92	-
LB	2	27.62	13.81	2.42
PB	1	6.51	6.51	1.14
LPB	2	3.29	1.64	-
S/LPB	84	478.27	5.69	

*$p < .001$

TABLE 11

SUMMARY OF THE ANALYSIS OF VARIANCE OF THE LAST TEN TRAINING TRIALS OF CONCEPT 2

Source	DF	SS	MS	F
L	2	49.27	24.63	2.97
P	1	5.78	5.78	-
B	1	12.61	12.61	1.52
LB	2	2.64	1.32	-
LP	2	2.71	1.35	-
PB	1	1.18	1.18	-
LPB	2	4.80	2.40	-
S/LPB	84	696.79	8.29	

TABLE 12

SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE LAST TEN TRAINING TRIALS OF CONCEPT 3

Source	DF	SS	MS	F
L	2	118.35	59.17	8.46*
P	1	5.36	5.36	-
B	1	10.66	10.66	1.52
LP	2	3.69	1.84	-
LB	2	18.88	9.44	1.35
PB	1	1.18	1.18	-
LPB	2	32.61	16.30	2.33
S/LPB	84	587.59	6.99	

*$p < .001$

TABLE 13

PROPORTION OF CORRECT RESPONSES ON TRIAL

1 OF CONCEPTS 2 AND 3

	C 2			C 3		
	NS	SH	TOT.	NS	SH	TOT.
L1	.87	.31	.59	.25	.12	.18
L2	.87	.75	.81	.31	.62	.46
L3	.81	.56	.68	.43	.43	.43
TOT.	.85	.54		.33	.39	

TABLE 14

SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE BLANK TRIALS OF EACH CONCEPT

Source	DF	SS	MS	F
Between				
P	1	57.51	57.51	2.82
L	2	114.67	57.33	2.82
PL	2	70.39	35.19	1.73
S/PL	42	854.04	20.33	
Within				
C	2	53.29	26.65	4.68*
PC	2	100.26	50.13	8.82**
LC	4	37.54	9.38	1.64
PLC	4	30.15	7.54	1.32
CS/PL	84	478.08	5.69	

*$p < .05$

**$p < .001$

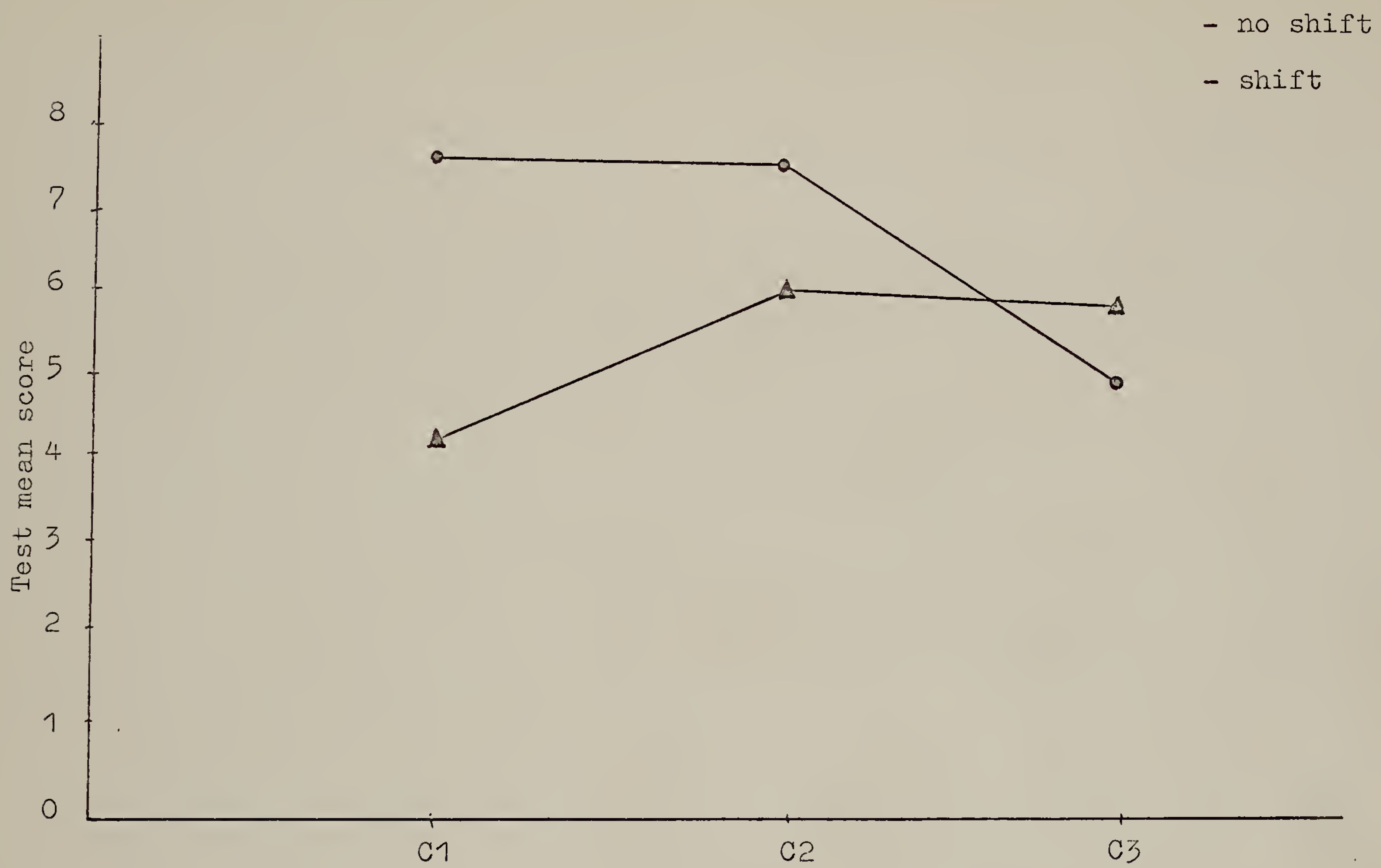

FIGURE 3. Plot of mean blank trial scores for differential effect of training paradigm by problem complexity (Table 14).

TABLE 15

SUMMARY OF THE ANALYSIS OF VARIANCE FOR

BLANK TRIALS (CONCEPT 3)

Source	DF	SS	MS	F
P	1	10.01	10.01	.97
L	2	86.52	43.26	4.21*
T	1	38.76	38.76	3.77
PL	2	38.02	19.01	1.85
PT	1	2.34	2.34	-
LT	2	12.27	6.13	-
PLT	2	18.19	9.09	-
S/LPT	84	861.62	10.26	

*$p < .05$

T = test condition: continuous (TL) or final (T2).