ABSTRACT

INDUCING STAGE III SERIATION CAPABILITIES
IN KINDERGARTEN CHILDREN THROUGH
CUE FADING AND REINFORCEMENT

By

Larry Eugene Schafer

The purpose of this study was to investigate the effectiveness (acquisition, retention, and transfer) of using cue fading and reinforcement to instruct children who were in Piaget's seriation stage II (the child orders with some difficulty but fails to insert a disarranged set of objects into an ordered set) for performance at Piaget's seriation stage III (the child both orders with ease and inserts objects into an ordered set).

Of the 95 kindergarten children who were given a seriation pretest, 34 were found to be in seriation stage II. The group of 34, stage II children was divided into a control group, which received no training, and an experimental group, which received training for seriation stage III.

Each experimental subject was individually given approximately 30 minutes of training on each of three consecutive days. The primary objective of the training

was to induce the ability to insert objects into an ordered set of objects. The same basic procedure was used during each 30 minute training session. Materials were set up at a number of training stations (45 training stations for session 1, 30 for session 2, and 24 for session 3), and each experimental subject was individually guided from one station to the next. The number of objects in the individual tasks was increased in stages throughout each session. At the beginning and whenever the number of task objects increased during a session either the ease of object discrimination was high and then gradually decreased in levels, or cues were introduced and then gradually faded in levels. Three slightly different practice tasks were used for each cue or discrimination level. The subjects were required to meet a performance criterion for each cue or discrimination level before they were allowed to progress to the next level.

Posttests were given approximately one, eight, and 132 days after training. Each posttest consisted of both near and far transfer measures. The materials used in the near transfer measure were similar to those materials used in the training, whereas the materials used in the far transfer measure were unlike those materials used in the training.

The results of repeated measures, and multivariate analyses revealed that the experimental subjects acquired and retained the specific target capabilities of the training (near transfer data) but failed, in general, to transfer those capabilities to tasks involving new materials (far transfer data). Although no massive, overall transfer effect was observed, the experimental group did outperform the control group on the far transfer measure of the second posttest. Two explanations (test-retest and novelty) for the experimental group's unexpected far transfer means were contrived.

The training method used in this study was found to be reasonably successful. However, because such large amounts of time, space, and material would be required, the unaltered use of the training method in the classroom would be prohibitive. Although this study does not seem to prescribe any specific seriation training techniques for immediate, direct use in the classroom, it does offer a tested method of cueing and cue fading that might be subsequently used to improve the seriation lessons found in the modern elementary science and mathematics programs.

INDUCING STAGE III SERIATION CAPABILITIES
IN KINDERGARTEN CHILDREN THROUGH
CUE FADING AND REINFORCEMENT

By

Larry Eugene Schafer

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

College of Education

1971

To Penny

ACKNOWLEDGMENTS

The author extends a genuine expression of gratitude to Professors Joe Byers, Glenn Berkheimer, Julian Brandou, and Gordon Wood for the valuable guidance and instruction which they gave not only during the preparation of this thesis but during the other phases of doctoral study as well.

To Professor Byers, the author's thesis advisor, goes a special note of thanks; for it was through interaction with Dr. Byers that the author gained a much deeper sense of what it means to be a scholar and a scientist.

The dedication of this thesis to Penny, the author's wife, only begins to express the appreciation which she deserves. Her sacrifices and patient, loving support made this thesis possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE PROBLEM

During the past fifteen years American developmental psychologists have increasingly turned their attention toward the study of cognitive development in children (Sigel & Hooper, 1968, p. 2). Undoubtedly the work of Swiss psychologist, Jean Piaget, has played a major role in causing that change of attention to occur. In response to Piaget's prolific output of research and theory, American psychologists have followed at least two avenues of research. One avenue reflects efforts to replicate or validate Piaget's findings, while the other reflects attempts to modify cognitive growth through training (Sigel & Hooper, 1968, p. 4).

In both validation and training research, Piaget's conservation capabilities (area, number, substance, weight, and length) have been prime targets of study. Although the conservation capabilities are central to Piaget's theory of intellectual development and therefore, justifiably draw upon the efforts of researchers, parallel lines of investigation must be launched into other important Piagetian capabilities. The research reported here was an attempt to extend a line of investigation into another of Piaget's capabilities, namely, the capability of serial ordering. Elkind (1964) has already taken the validation avenue to the study of serial ordering. The research reported here, on the other hand, was along the training avenue and thus, was a study of the effect of training on the development of serial ordering.

The purpose of Chapter I is to provide a basic introduction to the research reported. The objectives of the research will be discussed, and a rationale for the study will be described. Chapter I will be concluded with an overview of the experimental procedure and a statement of the general research hypothesis.

The Objective of the Study

In a concrete form, serial ordering consists of arranging material objects in order according to a particular attribute. For example, wooden sticks could be serial ordered according to length from the shortest to the longest to form a stairsteps-like figure. Children usually acquire the ability to serial order material objects before the age of seven years (Piaget, 1965, p. 133; Elkind, 1964).

Piaget (1965, pp. 122-134) has identified three seriation stages. In the first stage, children approximately four years of age have the ability to make pairwise discriminations but are unable to use that ability to serial order four or more objects. Given a set of four sticks which are easily distinguishable according to length, the child in stage I is able to correctly identify the shortest and longest sticks of any pair presented to him; however, he is unable to serial order the sticks according to length.

By the time children reach approximately five years of age, their entrance into stage II is indexed by an ability to serial order by trial and error. Children in seriation stage II rarely serial order a set of sticks by successively choosing the shortest stick from the sticks yet to be ordered. In addition to lacking spontaneity in serial ordering, the seriation stage II child fails to correctly insert a disarranged set of objects into an already serial ordered set.

Children six or seven years of age generally exhibit characteristics of the third and final seriation stage. Children in seriation stage III serial order with few errors and are able to correctly insert a disarranged set of objects into a serial ordered set of objects.

Piaget observed the seriation stages as they occurred in absence of any known systematic instruction in serial ordering. The question arises: Can specific instruction alter the rate at which young children progress through the seriation stages? This study, at least in part, was an initial attempt to answer that question. Specifically, this study was designed to investigate the effectiveness of using cue fading and reinforcement to train kindergarten children at seriation stage II (order with difficulty) for performance at seriation stage III (order objects with ease and insert a disarranged set into an ordered set). The effectiveness of the training was determined by repeatedly measuring the retention and transfer of the induced capabilities.

In addition to the retention and transfer measures there were the four individual difference (I.D.) measures of chronological age, reflectivity, impulse control and field independence. These I.D. measures were included for the following reasons: 1) to increase the precision of the comparisons among the dependent measures by covarying on the I.D. measures and 2) to provide the opportunity for uncovering potential relationships between I.D. measures and training effects.

The two reasons for including the I.D. measures imply suspected relationships between the various I.D. measures and the seriation abilities indexed by the dependent measures. Both Piaget (1965, p. 124) and Elkind

(1964) have found that as children grow older they become better able to perform seriation tasks. Consequently, chronological age and performance of the posttests are expected to be positively correlated. Kagan and Moss (1963) have found that reflective children demonstrate higher standards of mastery on intellectual tasks, greater persistence with such tasks, choose more difficult tasks, and work longer on the items than do impulsive children. In addition, they have found that analytic response styles are used more by reflective subjects than impulsive subjects. Consequently, it is suggested that the child who is more reflective will learn more from the seriation training and will do better on the seriation posttests. Banta (1968) has defined impulse control as the ability to control motor response. It is suggested, therefore, that children who can control their impulse to respond will likely have at least the time to evaluate an anticipated response and thus will make fewer errors in performing the seriation postests. According to Witkin (1950), a field-independent person tends to experience his surroundings analytically with object experiences being discrete from their backgrounds and tend to impose structure on a field which lacks it. Therefore, since the seriation posttest tasks require analysis and the ability to impose structure on a field, it is suggested that there will be a positive correlation between field independence scores and seriation posttest scores.

Rationale for the Study

The development of the periodic chart of chemical elements is evidence of the effective use of serial ordering in scientific inquiry. Mendeleef observed that when the then known 65 elements were serial ordered according to the atomic weights, similar physical and chemical properties periodically occurred. From this observation evolved the periodic chart of elements and the subsequent prediction of the chemical and physical properties of elements yet to be discovered.

The use of seriation is by no means reserved for the mature, competent scientists. Serial ordering can be used by the young elementary school child to investigate his environment. By ordering material objects according to one property, the child may more easily discover other properties related to the property used in ordering.

In a similar fashion, the young child may use seriation to study the relationship between experimental variables. Suppose, for example, that a child performed an experiment designed to study the relationship between the amount of incandescent lighting and the growth of a particular kind of plant. To study the results of the experiment, the child could serial order his plants according to the amount of incandescent light they received and then observe the height of plants as a function of the amount of lighting. Once the plants had been ordered with respect to the amount of light received, other plant characteristics, such as the shade of green, the stem thickness, and the size of leaves, may be discovered to be related to the amount of incandescent lighting.

Examples have been given to show how the elementary school child could use the serial ordering capability to facilitate his study of science. The authors of two major, contemporary elementary science programs, Science - A Process Approach and the Science Curriculum Improvement Study, apparently believe that seriation can facilitate the learning of science since they have included serial ordering lessons in their programs. Assuming that seriation is an important aspect of elementary science education, the research reported here is relevant to the improvement of science education because it is an investigation of techniques used in helping children acquire seriation capabilities.

Not only is serial ordering an important aspect of elementary science education, but it is also relevant to the child's development of the number concept. The importance of the relationship between seriation and the number concept has been emphasized by Piaget (1965):

> Number is at the same time a class and an asymmetrical relation, the units of which it is composed being simultaneously added because they are equivalent, and seriated because they are different one from another [p. 184].

To find the cardinal value of a class of objects (a ball, block, and stick) all objects must be considered alike to the degree that they each contribute one unit to that cardinal value. While all objects are considered the same in one respect, they are simultaneously considered different. Each object must be counted once and only once and therefore must be distinguished from the other objects. In order to avoid counting an object more than once, the objects are enumerated in a particular order which can be purely arbitrary. The objects arranged in the order in which they were counted form a series. The series is formed not on the basis of some physical attribute but on the basis of ordinal position (the first object counted, the second object counted, etc.). In addition, as the objects of a group are enumerated, the number of objects which have already been counted gradually increase (1 object counted, 2 objects counted, 3 objects counted, etc.) just as the lengths of sticks gradually increase in a serial ordered set of sticks.

Piaget does not suggest that the concept of number is merely reducible to the concepts of classification and seriation. He does argue, however, that without the concept of seriation, the concept of number cannot exist. Our technical society is requiring an increasing emphasis on the use of numbers and numerical relationships. Therefore, more attention needs to be given to the factors influencing the development and use of the number concept. This research provided some of that attention by focusing on one method of influencing the development of the seriation concept, a concept most important in acquiring the idea of number.

Thus far, argument has been presented to support the contention that serial ordering is relevant to elementary science education and to the development of the number concept. This research was designed to study an attempt at training kindergarten children for the final stage of seriation. Children generally reach the final stage of seriation by the age of six or seven years. What rationale, then, is there for training children to do that which they will eventually be able to do without specific training?

Benjamin S. Bloom (1964) has found support for the proposition that ". . . a characteristic can be more drastically affected by the environment in its most rapid period of growth than in its least rapid period of growth [p. 210]." Therefore, the extent to which the concept of seriation is relevant to processing information and developing number concepts may very well depend upon the kinds and amount of stimulation that occurred during the development of seriation. This research did not attempt to relate training during development with performance after development. This research did, however, provide the first step since it focused on ways to induce change during development.

In his book, Intelligence and Experience, J. McV. Hunt (1961) reviews experimental work and theories related to intellectual development. The evidence, he maintains, suggests that intelligence is a hierarchically arranged set of central processes, which develops as a result of child-environment interaction.

With this view of intelligence, Hunt (1961) writes:

> . . . it is no longer unreasonable to consider that it might be feasible to discover ways to govern the encounters that children have with their environments, especially during the early years of their development, to achieve a substantially faster rate of intellectual development and a substantially higher adult level of intellectual capacity. . . . The fact that it is reasonable to hope to find ways of raising the level of intellectual capacity in a majority of the population makes it a challenge to do the necessary research [p. 363].

This thesis, therefore, was partially justified because it was a study of the effectiveness of training procedures used in an attempt to accelerate children through one phase of cognitive development.

According to Kohlberg (1968), Piaget's interactional view of cognitive development suggests that massive general types of experience play a vital role in cognitive development and that in general these broad types of experiences cannot be replaced by limited specific training. Similarly, Sigel and Hooper (1968, p. 259) point out that while Piaget does not deny that learning processes are involved in cognitive development, he does not believe that American learning theory can adequately explain the development of logical reasoning.

This thesis was contrary to the Piagetian viewpoint in at least two respects. First, the training was specific, and second, techniques borrowed from behavior theory were used in an attempt to accelerate a child through one of Piaget's stages. Therefore, in addition to being relevant to elementary science education and to Hunt's quest for the means of raising intellectual capacity, this thesis has implications for testing the effectiveness of specific behavioral techniques, as well as Piaget's contentions about the necessary role that massive generalized experience plays in cognitive development.

Cue fading was chosen as the substantive element of the seriation training primarily because it has been found to be a particularly successful technique for training young children to perform rather difficult tasks. Of those studies showing successful uses of cue fading in the training of young children (Hively, 1962, 1965; Moore and Goldiamond, 1964; Sidman and Stoddard, 1967; and Bijou, 1965), Bijou's study provides the most impressive example. By using a cue fading procedure, Bijou was able to successfully train young, normal and retarded children to identify object images which showed particular angular rotations with respect to sample stimulus images.

For the most part, the target capabilities of the cue fading studies have been complex discriminations. The target behavior of this thesis required that the subjects make discriminations. For example, to correctly insert objects into an ordered set, the subject had to discriminate between the correct and incorrect positions in the seriated set. Since the evidence shows that cue fading can be used as an effective means of producing discrimination learning, and since the acquisition of the target behavior in this thesis involved discrimination learning, the cue fading method, as used in this thesis, was expected to contribute to the successful induction of the insertion capability in kindergarten children.

Overview of Procedure and General Research Hypothesis

A seriation test was given to those children attending kindergarten at Scott Elementary School, DeWitt, Michigan. The results of the test were used to identify those students in each of Piaget's three stages of seriation development. All stage II subjects were given the covariable measures to determine impulse control, reflectiveness, and field independence. Stage II subjects were then randomly divided into experimental and control groups. The experimental subjects received approximately 30 minutes of individual training during each of three consecutive days. The control subjects received no training. Three posttests were given to both control and experimental subjects after training was completed. The first posttest was given one day after training. The second and third posttests were given approximately one week and nineteen weeks after the training. All tests contained retention and transfer items. The results of the posttests were analyzed by a repeated measures form of multivariate analysis.

The general research hypothesis was that the experimental subjects will perform significantly better than the control subjects on the posttests. The basis for this hypothesis will be presented in Chapter II.

CHAPTER II

REVIEW OF LITERATURE

Seriation studies and studies related to training methodology will be reviewed in this chapter. Seriation studies will be reviewed in order to relate that which is known about seriation to this thesis and to show how the techniques and findings of those seriation studies have been used in the thesis research. To provide support for the research hypotheses, Piagetian training studies and studies from instructional psychology will be reviewed and related to those methods used in seriation training.

Seriation Studies

Although acquiring the ability to serial order is an important aspect in the cognitive development of children, the amount of research directly related to serial ordering along one dimension is surprisingly small. Much of the seriation research has seemingly been stimulated by the work of Piaget.

To study the seriation capabilities of young children Piaget (1965, p. 123) used the following technique.

Ten sticks, differing in length by 0.8 cm. were presented to a child who was instructed to form a series from the shortest (A=9 cm.) to the longest (J=16.2 cm.). Once the sticks (A-J) had been ordered, the child was given nine more sticks (a-i), one at a time and in any order, and was asked to insert these sticks in the right places. If the child had ordered and inserted the sticks correctly, the final series would have been A a B b .... I i J.

By using the technique described above, Piaget (1965, p. 124) has identified three distinct stages in the seriation of the sticks. The first stage is characterized by the child's inability to make a complete series with sticks A through J. The child in this first stage may make several short series which are placed side by side, or he may construct a staircase by considering only the tops of the sticks while disregarding the bottoms. In stage two, the child is able to order by trial and error, but fails to insert the additional sticks (a .... i). Children in the third stage of seriation are able to order without hesitation and are able to correctly insert the additional set of sticks into the ordered set.

Piaget (1965, p. 124) did not report the mean ages of his subjects in the three seriation stages. He did, however, present sample protocols typical of those from subjects of the three different stages. These protocols showed the ages of subjects in seriation stages one, two, and three to be about four, five, and six years, respectively.

Elkind (1964) replicated Piaget's experiments on discrimination, seriation, and numeration and then, unlike Piaget, applied a statistical analysis. In addition to seeing whether or not Piaget's results were verifiable, Elkind attempted to test Piaget's intimation that the perceptibility of size differences might influence the age at which the stages appear, but not the order of their appearance.

Assuming that the dimensionality of materials effects the perceptibility of size difference, Elkind used sets of one, two, and three dimensional items to test Piaget's intimation. The sets of one, two, and three dimensional items were respectively, sticks (one-fourth inch diameter dowels of various lengths), slats (one and one-half inch by one-fourth inch rectangular pieces of wood of various lengths), and blocks (three-fourths inch square pieces of wood of various lengths). Both ordering and inserting tests were administered with each of the three different sets of materials.

For the seriation test, Elkind found that the ease of task performance increased as the dimensionality of the materials increased. That is, the subjects found seriating sticks (one dimension) the most difficult, seriating slats (two dimensions) of intermediate difficulty, and seriating blocks (three dimensions) the least difficult. Elkind, therefore, claimed support for Piaget's contention that the perceptibility of size difference does influence task performance.

Elkind's analysis further revealed that the inserting problem was more difficult than the ordering problem. Moreover, a significant Test X Material interaction disclosed that difficulty was more pronounced when a relatively difficult task (insertion) was paired with a material of low dimensionality (sticks).

The relationship between age and seriation capability revealed in Elkind's study was consistent with Piaget's findings. Children about four years of age could neither order nor insert, children about five years of age could order but not insert, and children about six years of age could both order and insert.

Shantz (1967) has studied the effects of additional relevant and irrelevant information on children's ability to perform double seriation tasks. Contrary to expectation, she found that additional, relevant information did not significantly increase the number of correct responses. However, as was expected, the added irrelevant information did significantly reduce the number of correct responses.

In addition, no support was found for the hypotheses that the amount of added, relevant information required decreases with age and that the amount of added, irrelevant information that can be tolerated without affecting the response increases with age.

Elkind and Shantz's results will be compared with the understanding that the comparison may be jeopardized by the fact that Elkind and Shantz used different seriation tasks (single vs. double seriation) and different critical attributes (length vs. color brightness, amount of border, etc.). Two different interpretations will be made with respect to Elkind's results, and each interpretation will be shown to be inconsistent with the results of Shantz.

In Elkind's study, the slats (two dimensional materials) were constructed by adding the dimension of width to the sticks (one dimensional materials). Similarly, the blocks (three dimensional materials) were constructed by adding both dimensions of width and thickness to the sticks. On one hand, since neither width nor thickness varied within any set of Elkind's materials, these dimensions must be considered irrelevant to the task of seriating with respect to length. On the other hand, these added irrelevant dimensions of width and thickness may have increased the dominance of surface area and volume, two dimensions that are redundant to the relevant dimension of length. Thus, Elkind's increase in dimensionality may be interpreted as an increase in the number of irrelevant dimensions (width and thickness) or as an increase in the number of redundant, relevant dimensions (surface area and volume).

Elkind and Shantz's findings are inconsistent regardless of the way Elkind's increases in dimensionality are interpreted. If, on one hand, Elkind's increase in dimensionality is interpreted as an increase in the number of irrelevant dimensions, then Elkind found that increases in the number of irrelevant dimensions facilitated seriation performance. To the contrary, Shantz found that increases in the number of irrelevant dimensions retarded seriation performance. On the other hand, if Elkind's increase in dimensionality is interpreted as an increase in the number of redundant, relevant dimensions, then Elkind found that increases in the number of redundant, relevant dimensions facilitated seriation performance. Again the results were inconsistent, since, contrary to expectation, Shantz found that increases in redundant, relevant information did not facilitate seriation performance. From this comparison of two seriation studies, it is obvious that further research is needed to resolve the inconsistencies which exist between the findings of Elkind and Shantz.

Prentice (1963) administered various seriation tasks to 200 children in nursery school, kindergarten, and second grade. The tasks varied according to the number of elements (five, ten, or fifteen), the increment between elements (small or large), the materials (sticks of different lengths, pictures of objects that move at different speeds, or pictures of sticks of different lengths), and the instructions (to seriate, to insert, and to successively choose the smallest element of a group). From the data Prentice made the following observations:

1. There was a significant tendency, especially strong for nursery school children, for series of five elements to be easier than series of ten or fifteen elements and for series of fifteen elements to be easier than series of ten elements.

2. Older children were found to perform with significantly greater ease when the increment between elements was large rather than small.

3. Children who correctly serial ordered by a logical analysis of the problem tended to correctly insert elements into partially ordered sets.

4. For kindergarten and nursery school children, but not for second graders, it was significantly easier to make successive choices of the smallest element in a group than it was to form a series.

5. The children found the order of difficulty in using the materials to be, from easy to difficult, sticks, pictures of sticks, and pictures of moving objects.

The seriation literature reviewed above will now be related to the training research reported here. Both Piaget (1965) and Elkind (1964) have found evidence for the existence of three qualitatively different stages in the development of the seriation capability. It should be made clear that these stages do not merely reflect a smooth, continuous growth in the development of a single seriation capability but instead reflect changes in the kinds of seriation capabilities which can be performed.

The concept of stage development and the existence of the particular capabilities at each stage of development have been accepted for use in the thesis research. This acceptance is reflected in the stated purpose of this thesis: that stated purpose being to investigate a method of inducing seriation stage III capabilities in children possessing those capabilities characteristic of seriation stage II.

Sticks were seriated and inserted according to length in the studies of Piaget (1965), Elkind (1964), and Prentice (1963). Likewise, sticks were used in this study in both testing and training. Piaget (1965) presented his subjects with 10 sticks to order and then another nine to be inserted into the ordered set of 10. Elkind (1964) on the other hand, used a minimum of four and a maximum of nine sticks in ordering tasks. The insertion tasks in Elkind's study required that five sticks be inserted into an ordered series of nine. In this thesis the minimum and maximum number of sticks used in the insertion tasks consisted, respectively, of two sticks inserted into an ordered set of four and six sticks inserted into an ordered set of twelve. Thus, in this study the numbers of sticks used in the tasks are more consistent with the numbers used by Elkind than with the numbers used by Piaget.

Prentice's (1963) results have shown, as expected, that the increment between adjacent elements in an ordered set is a relevant factor in determining the ease of seriation. The increment used in the training study was relatively large. Piaget used an increment of 0.30 inch, Elkind used a 0.50 inch increment, and in this research a 0.75 inch increment was used. The purpose of using a comparatively large increment in this study was to reduce the discrimination problem and, hopefully, in doing so to release the subjects to concentrate on the seriation problem.

In the seriation studies reviewed above, two different methods were used to present the insertion items to the subjects. Piaget presented the sticks to be inserted

one at a time, while Elkind, Shantz, and Prentice presented the insertion sticks all at one time. To give the subjects more flexibility in responding, the method of presentation used by Elkind, Shantz, and Prentice was also used in this training study.

In the seriation studies discussed thus far, none of the researchers employed a partial credit method of scoring performance on the individual task items. The subject received either a score of one if, at the end of the task, all elements were in serial order or a score of zero if the elements were not in serial order. This training study used three different methods of scoring. One method was exactly like the one used in the seriation studies reported above, and, therefore, no partial credit was given. The other two scoring methods used in this training study were two different ways of giving partial credit to task performance.

One of Prentice's results was considered in designing the training procedure used in the research. Prentice found that the number of elements used in a task item was inversely related to the ease of task performance. Therefore, in the initial stages of training small numbers of elements were used to make the insertion tasks relatively easy. As the training progressed and the subjects became more capable of inserting, the numbers of elements in the training tasks were increased.

Only certain aspects of the seriation studies reviewed above were relatable to the thesis reported here. This is somewhat understandable because the seriation studies and this thesis were not designed for the same or similar purposes. Seemingly, seriation training studies would be more relatable to this thesis. Unfortunately, only one such seriation study has been found, and it will be reviewed subsequently.

Coxford (1964) studied the effects of instruction on the stage placement of children in some of Piaget's seriation experiments. Rather than attempt to induce serial ordering and insertion capabilities, Coxford attempted to induce: 1) the capability to construct serial correspondence between two sets of materials (stage II) and 2) the capability to conserve serial and ordinal correspondence between two sets of materials (stage III).

Balloons and sticks were used by Coxford to determine stage placement. To test for the ability to construct serial correspondence (stage II), the subjects were given disarranged sets of sticks and balloons and were asked to order the balloons and sticks in two separate, parallel arrangements so that the biggest stick went with the biggest balloon, etc. To test for the conservation of serial correspondence (stage III), the subjects were presented with an ordered set of sticks and an ordered set of balloons. The two sets were parallel and ordered in the same direction but one set was spread out relative to the other set. The subject's task was to find the object in one set which corresponded to a specified object in the other set. In testing for the conservation of ordinal correspondence (stage III), each subject was presented with one ordered set of materials and one disarranged set. The task was to find the object in the disarranged set which corresponded to a specified object in the ordered set.

The results of Coxford's study revealed that the training used to induce the conservation of serial and ordinal correspondence was successful but that the training used to induce the ability to construct serial correspondence was unsuccessful. In other words, Coxford was able to induce stage III capabilities in stage II subjects but he was unable to induce stage II capabilities in stage I subjects.

Coxford's study and this thesis research will be compared with respect to the purposes of training, the training materials and methods, and the testing procedures. Although both studies investigated the effectiveness of seriation training, little direct correspondence is found between them.

The training procedures used in the two studies were designed for different purposes. On one hand, Coxford's training was designed to help children acquire the ability to construct a serial correspondence between two sequences of objects, to conserve a serial correspondence when it is no longer perceptible, and to conserve an ordinal correspondence. On the other hand, the training used in the thesis was designed to help children acquire the ability to insert a disarranged set of objects into an ordered set of objects.

With respect to training procedures, Coxford used four, ten to fifteen minute sessions, spaced one week apart. Essentially, his training involved practice of the target behaviors. In the thesis reported here more of a massed training procedure was employed since three, thirty minute sessions were held over three consecutive days. Rather than have subjects practice the target behavior during the three sessions, cue fading was used to help the children gradually acquire the insertion capability with little difficulty.

Piaget (1964, pp. 17-18) has provided the following criteria for cognitive reorganization: a) stability over time, b) broad transfer across tasks, and c) acquisition of new, more complex cognitive operations. Coxford's study fails to account for these criteria in measuring the training effects since no retention or transfer measures were used and since no attempt was made to see if new, more complex operations had been acquired subsequent to the induction of the target capability. In contrast to Coxford's study, retention and transfer measures were used in the thesis research. Thus, the thesis research accounted for two of Piaget's three criteria, whereas Coxford's study accounted for none.

The literature reviewed above revealed that very little systematic research has been performed to investigate the area of seriation. Most of the existing seriation studies, have been directed toward determining the effects of different materials on the child's performance of various seriation tasks. The area of research devoted to the study of seriation training was found to be nearly void. Only one seriation training study was found, and the rigor of that study was questioned. In general, this thesis study finds relevance and justification in attempting to partially fill that void which exists in seriation training research.

Training Methodology: Support for the Research Hypotheses

Review of Piagetian Training Studies

Piaget has proposed that limited, specific training will produce no significant changes in the cognitive development of children (Kohlberg, 1968). Furthermore, even though Piaget believes that learning processes are involved in cognitive development, he does not believe that American learning theory can adequately explain the development of cognitive capabilities (Sigel and Hooper, 1968, p. 259).

Piaget's proposals and beliefs regarding the ineffectiveness of limited, specific training and the inadequacies of American learning theory have not gone unexamined. In most of the many attempts to induce Piagetian capabilities, specific, short-term training methods have been used. These specific short-term training methods tend to reflect two general positions regarding cognitive development. One position is consistent with Piaget's thinking, and the other is more representative of an American learning theory viewpoint. The Piagetian position is characterized by emphasis upon internal cognitive structures, logical operations, and equilibration mechanisms. The learning theory approach as considered in this review is characterized by emphasis upon such factors as learning set, reinforcement, corrective feedback, the influence of irrelevant cues, verbal rule learning, cue fading, and practice.

The seriation research reported in this thesis was an attempt to investigate the effectiveness of using cue fading and reinforcement to induce the Piagetian capability of seriation. Therefore, the nature of the seriation training obviously reflected the American learning theory position. Since this thesis followed a learning theory approach and since there exists a number of Piagetian training studies reflecting both Piagetian and learning theory approaches, an examination of the relative effectiveness of the two approaches might reveal some expectations regarding the success or failure of the seriation training methods.

In the Piagetian training studies, researchers have focused their attentions almost exclusively on the training of conservation capabilities (Sigel and Hooper, 1968, pp. 258-434; Brainerd and Allen, 1971). Therefore, little will be lost by limiting the review to those training studies concerned with the induction of conservation. Since stick <u>length</u> and the <u>number</u> of lines per card were the critical attributes used in seriation training, the review will be further limited to those training studies concerned with the induction of <u>length</u> or <u>number</u> conservation.

Flavel (1963, p. 245) has defined conservation as "the cognition that certain properties (quantity, number, length, etc.) remain invariant (are conserved) in the face of certain transformations (displacing objects or object parts in space, sectioning an object into pieces, changing shape, etc.)." Hence, a child who conserves number will maintain, without counting, that the number of objects before him remains the same regardless of any changes in their spatial orientation. Similarly, a child who conserves length will insist that the lengths of two sticks remain the same regardless of their relative positions.

Brainerd and Allen (1971) have reviewed training studies in the conservation of "first-order" quantitative invariants (number, length, substance, weight, and area). They conclude from their review that the common element among the successful methods (defined by statistical significance) is training in reversibility. Reversibility refers to the idea that every direct operation (action or transformation) has an inverse which cancels or negates it. For example, the lengthening of a row of objects in number conservation tests can be cancelled or reversed by returning the row to its original length.

According to Piaget (1964), "reversibility of thought" is a cognitive capability which must be present in the intellectual repertoire of the child in order for him to conserve. The child who is to conserve number, for example, must realize that since a rearranged set of objects can always be returned to the arrangement observed at the time of counting, the number of objects remains the same. Since Piaget stresses the importance of reversibility and since Brainerd and Allen conclude from their review that reversibility training is commonly found in the successful training studies but not in the unsuccessful studies, it would appear that the evidence supports the Piagetian position more than it supports the learning theory position.

The apparent support for the Piagetian position rests on the validity of Brainerd and Allen's analysis of the length and number conservation studies. A careful reanalysis of the same studies, however, brings that validity into question. Apparently, Brainerd and Allen attended primarily to the reversibility element and failed to regard other training conditions which may have accounted for the success or failure of training. Therefore, the conditions used to induce length and number conservation will be surveyed, the analysis will be broadened, and the various conclusions of Brainerd and Allen will be refuted or reaffirmed.

Before beginning the analysis, perhaps typical tests for number and length conservation should be described. A test for number conservation usually begins with the examiner placing two numerically equivalent rows of objects parallel to each other with the objects in exact one-to-one correspondence. Either by having the child count the objects or by having him observe the one-to-one correspondence, the numerical equivalence of the two rows is established. One row is then lengthened or shortened, and the child is asked if the numbers of objects in the rows are the same or different. Usually the child is requested to explain his answer. The child who fails to give a reasonable explanation is often judged to be a nonconserver, even though he maintains that the numbers are the same. A conserver might reasonably explain that the numbers of objects must still be equal since objects were neither added nor taken away. Testing for conservation of length would proceed in a similar manner. Two sticks of equal length would be placed parallel with the ends aligned. Equivalence of length would be established, and then one stick would be moved so the ends of the sticks would no longer be in line. The question of equivalence would be asked and then followed by the request for an explanation.

The training methods used to induce conservation capabilities tend to reflect two general positions regarding the development of conservation. One position is consistent with Piaget's thinking and the other is more representative of an American learning theory viewpoint. The Piagetian position is characterized by emphasis upon internal cognitive structures, logical operations, and equilibration mechanisms. According to Piaget (1952), before conservation can appear, the child must acquire the following operations: multiple classification, multiple relationality, atomism, reversibility, and seriation. The researchers favoring the Piagetian position construct their training procedures in accordance with Piaget's proposed prerequisite operations and his general mechanism for cognitive development (equilibration). The learning theory approach, on the other hand, is characterized by an emphasis upon such factors as learning set, reinforcement, corrective feedback, the influence of irrelevant cues, verbal rule learning, cue fading, and practice. Some of the researchers (Gelman, 1969; and Kingsley and Hall, 1967) who favor this approach would contend that a young child may in some way be able to conserve but may fail to conserve because of inattention to the relevant quantitative attribute or because of the strong tendency to attend to changes in irrelevant attributes such as shape, position, or color.

Unsuccessful Attempts to Induce Length and Number Conservation

Relatively unsuccessful attempts to train for conservation of length and number have been reported by Smedslund (1963), Wohlwill and Lowe (1962), and Mermelstein and Meyer (1969).

Smedslund (1963) hypothesized that training methods would be successful if they were consistent with Piaget's concept of equilibration. Piaget (1950) asserts that the logical structures of the child develop as a function of the internal process called equilibration. Through interaction with his environment the child has experiences which are sometimes inconsistent with his way of thinking. This inconsistency puts the child in a state of disequilibrium. To make experience and thinking match and to return to a state of equilibrium, the child goes through a process of reorganizing his thoughts. This process of reorganization is called equilibration. According to Smedslund, conservation can be induced by stimulating the equilibration process through the presentation of conflict situations.

Smedslund studied the relative effectiveness that five different conflict producing procedures had upon the induction of length conservation in young children. In all of the procedures, Muller-Lyer arrowheads (<---> >-----<) were used to create a perceptual change of length.

Although Smedslund did not report a statistical analysis of his data, such an analysis has been performed and reported by Brainerd and Allen. The results show that only one of Smedslund's five experimental treatments (the anticipation condition) produced significant increments in length conservation with respect to the control condition.

Although the anticipation method of training was found to be better than no training (control), it was not found to be superior to the other four training methods.

In the anticipation method, subjects were asked to anticipate and judge the relative lengths of two equally long sticks before and after movements back and forth between an optimal comparison position (sticks side-by-side without Muller-Lyer arrowheads) and a perceptually distorting position (sticks spread apart with Muller-Lyer arrowheads). The procedure for one unit of the anticipation method of training is more explicitly explained as follows:

1. Sticks with the angles ( ) were presented to subjects. Subjects were then asked, "If the sticks are moved together from under the angles, will this one be longer, will they be equally long, or will this one be longer?" Following subject's response, the sticks were placed together and shown to be equal.

2. The subject was then asked to anticipate the relative lengths of the sticks if they were moved out to the angles again.

3. The sticks were moved out under the angles, and the subject was asked again about the relative lengths of the sticks.

4. In the final step of the instructional unit, the subject was requested to anticipate the relative lengths of the sticks if they were moved in from under the angles.

Each subject in the anticipation group received four instructional units similar to the one just described. The only difference between the units was the color of the sticks used. It should be pointed out that the experimenter extended neither reinforcements nor corrective feedback to the subjects during any of the training.

Brainerd and Allen suggest that Smedslund's anticipation method of training was successful because the subjects were given reversibility training. A careful examination of the instructional steps, however, reveals that the last step in the usual method of reversibility training was omitted. Usually, an equality - apparent inequality - equality sequence is used in reversibility training. In Smedslund's anticipation method, the final state of equality was not demonstrated. The subject was merely asked to anticipate the relative stick lengths _if_ the transformation was reversed (i.e., if the sticks were moved from under the arrowheads back to the side-by-side position). The reverse transformation was not actually carried out. Thus, the success of the anticipation method cannot be attributed to reversibility training.

The failure of the other four training methods used by Smedslund is not surprising. The maximum number of times a subject responded under any of the unsuccessful training conditions was 16. Furthermore, no feedback was given to the subject even when he did respond.

Wohlwill and Lowe (1962) studied the effects that four different methods had on the induction of number conservation in young children. The four training conditions were: Reinforced Practice, Addition and Subtraction, Dissociation, and Control. The training series for each condition consisted of nine trials administered on each of two successive days.

Wohlwill and Lowe's results revealed that all three experimental conditions as well as the Control condition produced significant effects on a nonverbal measure of conservation, but virtually no demonstrable effects on a verbal test of conservation. Moreover, the experimenters found no reliable differences in effectiveness among the four training conditions.

In view of what they consider to be a predominantly negative experimental outcome, Wohlwill and Lowe offer a general conclusion. Since the children in their experiment rather consistently responded on the basis of length differences in making numerical comparisons between two collections, the experimenters maintain that there exists support for interpreting the lack of conservation as a failure to differentiate number from irrelevant perceptual cues. Admittedly, Wohlwill and Lowe's conclusion is based on somewhat scanty evidence. Nevertheless, their conclusion has been supported by the well designed and rigorous research of Gelman (1969), which will be reported subsequently.

Merelstein and Meyer (1969) studied the effects that five treatment conditions (four experimental and one control) had on the induction of the number conservation in young children. The subjects were tested three weeks, two and one-half months, and five months after the last training session, and the results indicated conservation was not induced by any of the training techniques.

The four experimental training conditions were similar to Smedslund's Cognitive Conflict technique, Beilin's Verbal Rule Instruction, Bruner's Language Activation technique, and Sigel's Multiple Classification training. Eight training trials were given under each experimental condition.

The amount of training given under each experimental condition was indeed minimal (eight trials). Of the eight cognitive conflict trials, only the last three were designed to actually produce a conflict situation. Thus, subjects in the Cognitive Conflict group, essentially received only three trials of training. Under the Multiple Classification condition an attempt was made to give subjects training in multiple labeling (a poker chip can be called a checker or a toy coin), multiple classification (a poker chip can be classified according to color, shape, texture), multiple relations (a poker chip can have color and shape), and reversibility. Since eight trials were used to train the subjects for all five of the target capabilities, the Multiple Classification subjects probably never acquired any of the capabilities to a respectable degree.

Some questions can be raised as to whether or not certain training situations were properly designed. For example, in the reversibility training, transformations were never reversed, and in Language Activation training the question was phrased in such a way that it would tend to minimize rather than maximize language activation.

Considering the inadequacies of the training and testing techniques, Mermelstein and Meyer's study could hardly be considered a fair test of the relative effectiveness of the four training conditions studied. In the face of such inadequacies, it becomes difficult to accept Mermelstein and Meyer's general conclusion that specific training, regardless of procedure, is ineffective in inducing conservation concepts.

Brainerd and Allen have observed that unsuccessful attempts to induce number and length conservation are characterized by the absence of adequate reversibility training. They use this observation to support the contention that the presence or absence of reversibility training determines the success or failure of training attempts.

Although Brainerd and Allen's observation is correct, the use of that observation in support of their contention is questionable since the failures can be explained by other inadequacies. In all the unsuccessful attempts, the number of training trials was very small, none of the subjects were trained to criterion performance level, and feedback given to the subjects was almost nonexistent. Therefore, it is quite possible that failures could be attributable to insufficient amounts of training rather than to the method of training, as suggested by Brainerd and Allen.

Successful Attempts to Induce Length and Number Conservation

Relatively successful attempts to induce number and/or length conservation have been reported by Goldschmid (1968), Wallach and Sprott (1964), Wallach, Wall, and Anderson (1967), Gelman (1969), and Kingsley and Hall (1967). Goldschmid (1968) used six experimental groups and one control group to study the relative effectiveness of using three different training procedures to induce various conservation capabilities. Each subject in half the experimental groups was trained to conserve discontinuous quantity (<u>amount</u> of beads poured into containers of different shapes), two-dimensional space, and substance (<u>amount</u> of clay). Each subject in the other half of the experimental groups was trained to conserve continuous quantity (<u>amount</u> of liquid poured into containers of different shapes), number, and weight.

One of three different training procedures was used with each of the three groups receiving training on the same concepts. The different procedures consisted of reversibility training, compensation training, and a combination of reversibility and compensation training. The reversibility procedure involved repeated demonstrations showing that any transformation can be reversed. The child was to infer from reversibility training that since an action can be reversed, the original quantity remains unchanged regardless of the transformation. Compensation training involved step-by-step transformations of one of two quantitatively equal objects (or sets of objects). This training procedure was used to show the child that change in one dimension (tall to short) was compensated for by change in another dimension (skinny to fat). The combination of reversibility and compensation training involved aspects of both procedures described above. One of the three training procedures was used with each subject during three one-half hour sessions. Each session was devoted to training on one conservation capability. Training characteristics such as the number of training trials, the use of a training criterion, and the use of feedback were not reported by Goldschmid.

Posttests were administered three and six weeks after training. During the two posttests every subject was tested on six conservation tasks, the three he received training on (acquisition), and the three he received no training on (specific transfer). Only during the second posttest were the subjects tested on conservation of length, and area (nonspecific transfer).

The results from the posttests revealed that all experimental groups performed significantly better than the control group on all test items of both posttests, and that reversibility training was more effective than either compensation or combination training. Goldschmid concluded from the data analysis that the training procedures were effective means of inducing conservation concepts, that the acquired conservation capabilities were transferable, and that the effects were durable over a six week time period. In addition, Goldschmid claimed support for the idea proposed by Piaget that reversibility and compensation are thought processes underlying conservation.

Wallach and Sprott (1964) successfully induced number conservation in children by using a combination of reversibility and addition/subtraction training. The training materials consisted of six dolls and six beds. At the beginning of each trial the beds with the dolls in them were placed side-by-side in a row before the subject. The dolls were then taken out and placed closer together or farther apart in a row in front of the beds, so that there was a bed without a doll or a doll without a bed. In half the trials either a doll or bed was added or removed so that the two rows became the same length. After the transformations, the subject was asked the following series of questions: "Do you think we can put a doll in every bed now? Will there be any beds left over? Any dolls left over?" After answering the question the subject was asked to put a doll in each bed.

The subjects were presented the training situations in the following order: 1) dolls together; 2) dolls together, bed removed; 3) dolls apart; 4) dolls apart, bed added; 5) dolls together; 6) dolls together, doll added; 7) dolls apart; 8) dolls apart, doll removed. Each situation was repeated until the subject made the correct prediction and then confirmed it. All experimental subjects were trained to a criterion of a correct prediction on the first trial of four situations in succession.

Two posttests were given, one immediately after training and the other between 14 and 23 days after training. Dolls and beds as well as checkers and cards were used in the first posttest. In the second posttest bowls and spoons, checkers and cards, and dolls and beds were used. The questions asked during the posttests were not the same as those used in training. Training questions asked nothing about the number of objects, whereas testing questions asked if there were the same number of objects.

The analysis of the posttest results indicates that the experimental group performed significantly better than the control group on both posttests. The conclusion was made that the treatment was an effective means of inducing sustained number conservation.

These results seemingly provide more support for the Piagetian contention that reversibility is a fundamental prerequisite for conservation. In spite of the results, Wallach and Sprott were not willing to claim unquestionable support for the Piagetian contention. They observed that over half the experimental subjects responded correctly in all training trials to the question, "Do you think we can put a doll in every bed now?" From this observation, they concluded that many subjects already knew that transformations were reversible prior to both the training and the acquisition of number conservation. Thus, Wallach and Sprott reasoned that the training did more than simply supply the subjects with the knowledge of reversibility.

Wallach and Sprott (1964) gave reversibility training the credit for successful induction of number conservation. However, since the training procedure was actually a combination of reversibility and addition/subtraction training, no reasonable conclusion can be made about the individual effectiveness of the two kinds of training.

Wallach, Wall, and Anderson (1967) followed up the Wallach and Sprott study with an attempt to investigate the separate effects that reversibility and addition/subtraction training had upon the acquisition of number conservation. All experimental subjects were trained to a performance criterion of four successful trials in succession.

The results of the study showed that reversibility training was a successful means of inducing sustained number conservation, but that the addition/subtraction training was not. Furthermore, the results indicated that number conservation induced by reversibility training did not transfer to the conservation of liquid amount.

The expected conclusion made from the results would seemingly be that children acquire number conservation once they become able to recognize reversibility. Wallach,

Wall, and Anderson, however, did not make such a conclusion. Instead, they maintained that reversibility was successful, not because it led the subjects to recognize reversibility, but because it led them to stop relying on misleading cues.

The evidence for the experimenters' conclusion came from subjects' explanations of the conserving responses. Rarely did the conserving subject use the idea of reversibility to explain that number was conserved. Most often the conserving subject gave an explanation which referred to the spacing of the objects (closer together or farther apart) rather than to the overall length of the row. The implication was that a shorter row did not necessarily mean that there were fewer objects but that the objects were just closer together. Wallach, Wall, and Anderson interpreted this most frequent explanation as evidence for suggesting that the subjects had learned to ignore the misleading cue of row length.

The study by Wallach, Wall, and Anderson generates the question: Can number conservation be induced by a method which trains the subjects to disregard the misleading cues while at the same time minimizes any suggestion of reversibility? Gelman (1969) used such a method and found the answer to be affirmative.

Gelman (1969) hypothesized that a young child may in some way be able to conserve but may fail to conserve because of inattention to the relevant quantitative attribute or because of the strong tendency to attend to changes in irrelevant attributes such as shape, position, or color. She investigated the supportability of her hypothesis by using discrimination learning set training on length and number tasks. In learning set training procedures, the subjects receive training with a large number of problems containing many different stimuli; although the stimuli differ, across problems there is one common relationship, and the attentional responses to this common cue are reinforced. Gelman trained her subjects to attend to the relevant quantitative attributes of length and number while disregarding irrelevant attributes of color, shape, and spatial arrangement. The training consisted of 32 six-trial problems. In half the problems the relevant attribute was length, and in the other half the relevant attribute was number. The number problems were alternated with the length problems. Three stimulus objects (sticks or rows of chips) were used in each problem. Two of the objects (sticks or rows of chips) were identical while the third was different. For example, two rows of five chips were used with one row of three chips and two 6-inch sticks were used with one 10-inch stick. In the training problems, the subject's task was to point to two sticks that had the same (or different) length or to point to the two rows that had the same (or different) number of items in them. After a subject made a correct response he was told that his response was correct and he was given a prize (trinket). When an incorrect choice was made, the subject was told that his response was wrong. Contrary to Brainerd and Allen's report, a noncorrection procedure was followed.

For the six trials of each problem the stimulus objects remained the same. The variation between trials within a problem took the form of changes in spatial orientation. Gelman used Figure 1 to illustrate the between trial variation within a problem.

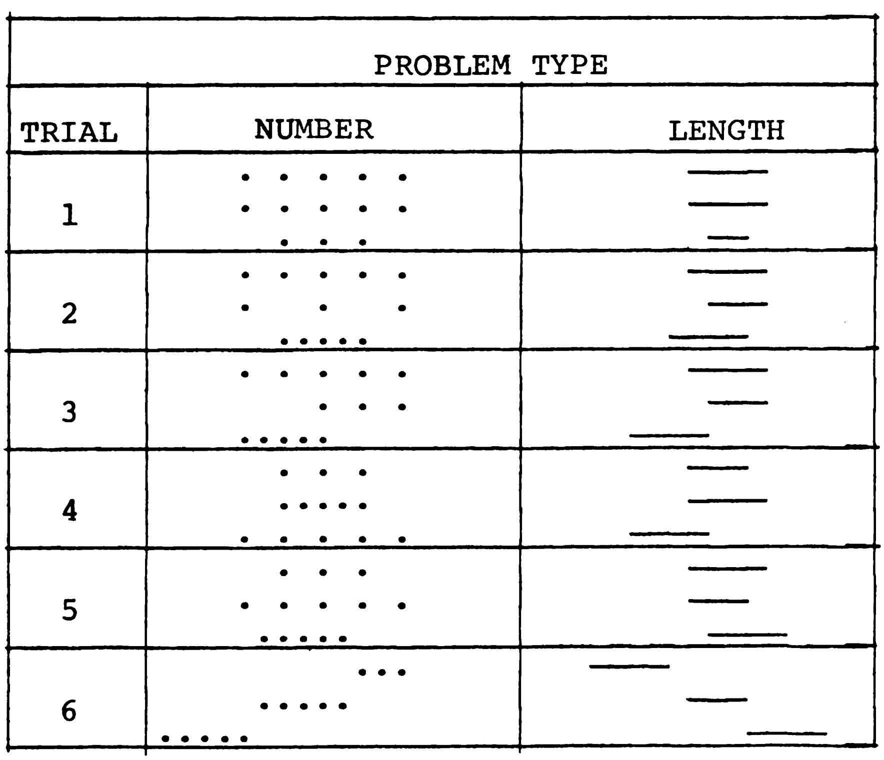


Fig. 1. Gelman's (1969) illustration of between trial variation within a training problem.

There is no indication in this figure that reversibility training did take place, as suggested by Brainerd and Allen.

Between problem variations consisted of changes along the following irrelevant attributes: 1) color of chips and sticks, 2) size and shape of chips and sticks, 3) starting arrangements (horizontal, vertical, etc.), and 4) quality combination (i.e., whether the odd stick or set of chips was small or large compared to the other two sticks or sets).

Gelman employed two control groups. One control group received oddity training with toys. The other group received the same training as the experimental group except that no feedback was given.

Conservation posttests were given to all groups the day after training and then 2-3 weeks later. The posttests contained items designed to test conservation of length and number (specific transfer) and conservation of mass and liquid (nonspecific transfer). The results showed near perfect specific (length and number), and approximately 60% nonspecific (mass and liquid amount) transfer of training. Over the 2-3 week retention interval these effects were found to be durable. Gelman concluded that support was found for the hypothesis that young children fail to conserve because of inattention to relevant quantitative relationships and attention to irrelevant features in classical conservation tests.

Gelman's study does not stand alone. Kingsley and Hall (1967) also successfully trained subjects to conserve length and weight with learning set procedures. Rather than specifically train the subjects to ignore irrelevant and attend to relevant aspects of the conservation problem, Kingsley and Hall used Gagné's learning set analysis and trained the subjects on a graded series of subtasks related to conservation (e.g., the appropriate use of scales and measuring instruments, the effects of addition and subtraction, the relation between spatial orientation and length). The subject's experiences during the sub-training most likely led them to ignore the irrelevant cues. Thus Kingsley and Hall's subtask training indirectly served the same purpose as the learning set training used by Gelman. Contrary to Brainerd and Allen's review, Kingsley and Hall reported no use of reversibility training.

Discussion

The unsuccessful studies were characterized by very few training trials (maximum number of trials was 18), no training performance criteria, and almost no means of providing subjects with feedback. In comparison, the successful studies used more training (Goldschmid, 1.5

hours; and Gelman, 192 trials), trained to criteria (Wallach and Sprott; Wallach, Wall, and Anderson; and Kingsley and Hall), and often provided the subjects with feedback. This comparison suggests that future training studies use adequate amounts of training and provide the subjects with feedback, and that possibly those methodological comparisons made in the unsuccessful studies should be repeated with more emphasis given to the extensiveness of the training and to the use of feedback.

Contrary to Brainerd and Allen's conclusion, the successful and unsuccessful studies are not necessarily distinguished on the basis of whether or not reversibility training was given. It is true that the unsuccessful studies did not employ reversibility training. It is not true, however, that all successful studies employed reversibility training. In fact, the most successful attempt (Gelman) did not include reversibility training.

Gelman's study and the study by Kingsley and Hall lend strong support to the contention that children fail to conserve not because they lack certain cognitive capabilities (reversibility) but because they are unable to distinguish the relevant from the irrelevant cues. The successful studies which employed reversibility training do not challenge this contention, but instead support it, since reversibility training has the potential for leading subjects to disregard irrelevant cues and attend to relevant cues. In fact, reversibility training might be considered one kind of learning set training.

With respect to the theoretical position, the work reviewed here seemingly provides more support for the learning theory position than for the Piagetian position. First, somewhat contrary to Piaget's notion, young children can acquire at least specific conservation concepts through specific experience. Second, the learning theory position with emphasis on attention, stimulus factors, and learning set can explain the success of both the studies using learning set procedures (Gelman, and Kingsley and Hall) and the studies using reversibility training (Goldschmid; Wallach and Sprott; and Wallach, Wall, and Anderson). The Piagetian position with emphasis on requisite cognitive structures and reversibility training, however, fails to explain the success of the learning set procedures.

As a final note of clarification, it should be stated that this review has cited no evidence which suggests that conservers need not recognize reversibility. In fact, it seems reasonable to propose that all "real" conservers should be able to recognize reversibility. The evidence merely suggests that either children already recognize reversibility before they are taught to attend to the relevant cues, or that they acquire the ability to recognize reversibility as a consequence of learning to attend to the relevant cues.

Although the conclusions resulting from this review of Piagetian training studies cannot be directly used in formulating hypotheses for the thesis study, the conclusions do reveal some hint of expectations regarding the success of the seriation training methods. In general, the evidence shows that not all Piagetian capabilities are resistant to training and furthermore, that the learning theory position can be used as an adequate bases for training. Therefore, since learning theory is the basis for the seriation training, and since the seriation capability is seemingly not as complex as the conservation capability, the seriation training methods are expected to be successful.

In general, the Piagetian training studies previously reviewed were attempts to answer academic or theoretical questions rather than educational questions. Like many "academic" studies, the training studies were characterized by relatively short training periods, retention intervals which were of little more than minimal length, and measurements of transfer which, for the most part, were limited to the use of tasks very similar to those used in training.

According to Piaget, the design characteristics of the training studies did not permit adequate tests of his theoretical proposition concerning cognitive reorganization. In addition to maintaining that cognitive change cannot be induced by limited, short-term training, Piaget (1964, pp. 17-18) contends that broad transfer, stability over time, and the acquisition of new, more complex capabilities are the three criteria necessary for inferring cognitive reorganization. If these criteria are interpreted in a broad sense, then the review of conservation training studies has shown that those studies rarely even approached a consideration of Piaget's three criteria.

If Piaget is correct in his selection of criteria, then certainly more massive, long-term studies are needed. It would seem that these studies while answering theoretical questions could at the same time answer educational questions. In fact, studying the effects of induced intellectual capabilities on school learning might be interpreted as an attempt to account for Piaget's criteria of broad transfer and acquisition of new, more complex capabilities.

Although the training studies have not produced convincing answers to the theoretical questions and have certainly been of little immediate educational worth, they have indeed established the basis necessary for the development of those massive, long-term studies which have both theoretical and educational relevance. For example, without the conservation training studies, no methods would have been developed for inducing conservation capabilities. As a result, no rationale would exist for attempting the large scale investigations in the schools to see if extended training in conservation would have a significant impact on: 1) the nature (quality, rate, and sequence) of intellectual development observed over a considerable time period, and 2) the nature of school learning observed, as well, over a considerable time period.

The more thorough conservation training studies have produced a foundation upon which the more massive, theoretically and educationally relevant studies can be built. A similar foundation for the seriation capabilities has not yet been constructed. The seriation training study reported in this thesis is, therefore, particularly relevant since it contributes to the sound foundation from which will eventually spring those longitudinal studies in which intellectual development and school learning will be observed as a function of induced seriation.

Review of Instructional Psychology

Studies in instructional psychology will be reviewed to provide support for the general research hypothesis. That hypothesis states that the instruction given to the experimental subjects will be successful in helping them acquire the ability to correctly insert objects into a seriated set of objects.

The organization of the review will be structured according to Gagné's set of instructional conditions. According to Gagné (1965, pp. 205-236), the manager of instruction may manipulate the following set of conditions: a) Conditions for the Acquisition of Capabilities - 1. the techniques to gain and maintain the attention of the learner, 2. the establishment of certain preconditions for learning, 3. the presentation of the stimuli directly involved in learning, 4. the use of prompting and guiding to facilitate the learning process, 5. the specification of the conditions of responding, 6. the employment of feedback methods; b) Conditions for the Promotion of Retention; c) Conditions to Enhance the Transfer of Learning. The manipulation of Gagné's conditions in the thesis research will be discussed and literature relevant to the employed conditions will be reviewed.

Conditions for the Acquisition of Capabilities

The techniques to gain and maintain the attention of the learner. In summarizing a review of recent research on methods of gaining and maintaining attention, Gagné and Rohwer (1969) concluded that the application of techniques using reinforcement contingencies seemingly holds much promise for maintaining attentional sets. In support of their conclusion Gagné and Rohwer cited a study by Parker and Nunnally (1966), who found that penny rewards increased children's "time of looking" at nonsense syllables presented in a spin-wheel game. In a similar manner, Staats and associates (1962, 1964a, 1964b) found that token reinforcers served well as effective reinforcers in teaching beginning reading skills to four-year-olds. The token reinforcers were accumulated by the children and eventually exchanged for small trinkets and toys. Staats and associates concluded that the reinforcement system solved the major problem in teaching young children, namely, to keep them at the task over long periods of time. Since the effectiveness of token reinforcement systems, as shown by Staats and associates, has been replicated a number of times (Howard and Tracy Kendler and associates, 1959, 1960; and Whitlock and Bushell, 1967), the token system, therefore, seems to be a reliable technique which can be used to teach young children various skills.

The material reinforcers used in this thesis consisted of marbles which the subjects stacked in "marble banks." Unfortunately, at the request of the subject's teachers, no "back-up" reinforcers (candy or trinkets) could be exchanged for the marbles. Although the desired token reinforcement system was not employed in the thesis research, it was assumed that acquiring and stacking marbles in columns would provide the subjects with a significant increment of reinforcement to keep them at the tasks.

Non-material, social reinforcement in the form of verbal praises (Gerwirtz and Baer, 1958a, 1958b) and teacher attention (Harris, Wolf, and Baer, 1964; and Harris, Johnston, Kelley and Wolf, 1964) has been found to facilitate the training of young children. Zigler and Kanzer (1962) found that "praise" ("good," "fine") reinforcers were more effective with lower socioeconomic class children than with middle-class children and that "correct" ("right," "correct") reinforcers were more effective with middle-class than with lower-class children. Since there existed the possibility of having both lower and middle-class children in the seriation study, both kinds of social reinforcers described by Sigler and Kanzer were used. Whenever a subject responded correctly to a seriation training task, he was told that his response was correct ("Right!," "Correct!"), and he was given praise ("Good job!," "Very nice!"). With the research on social reinforcement showing that social stimuli can be used as reinforcers in training young children, the social stimuli used in seriation training were assumed to provide a measure of reinforcement sufficient enough to keep the subjects interested in acquiring the target seriation capabilities.

<u>The establishment of certain preconditions for learning</u>. According to the Gagné model, another instructional event which should be considered by the instructor is the establishment of preconditions for learning. To facilitate learning, students may be given verbal directions and pretraining and may be asked to recall past learning which is relevant to the subsequent instruction.

In the thesis study, an orientation session was given before each of the three training sessions to prepare the subjects for the instruction which followed. The orientation sessions were used to introduce the tasks and materials, point out the cues, illustrate correct and incorrect responses, show how reinforcement would be administered, and allow the subjects to demonstrate their understanding of the instructions. It was assumed that these sessions would prepare the subjects for instruction and thus facilitate their acquisition of the target capabilities.

<u>The presentation of the stimuli directly involved in the learning</u>. Instructional designers often have the option of presenting materials in the form of objects, pictures, or words. Gagné and Rohwer (1969) have suggested

that when such an option exists in preparing materials for young learners, the stimulus type chosen should be either objects or pictures rather than words. With respect to the use of various stimulus types in seriation research, Prentice (1963) has found that children seriate sticks more easily than they seriate pictures of sticks. Furthermore, Elkind's (1964) seriation study revealed that the ease of performing seriation tasks increased as the dimensionality of the objects increased.

The findings of Elkind and Prentice and the suggestion of Gagné and Rohwer seemingly imply that, for young learners, task difficulty increases as the abstractness of the stimulus material increases. This implication was used in this thesis to determine the order of presenting two different training materials. To help the children acquire the ability to insert, the material which made the insertion task the easiest for the children was used first. This material, according to the above implication, would have a concrete nature. Thus, sticks, rather than lined cards, were used in the initial stages of the training sequence since stick length was assumed to be a more concrete quality than line density.

The use of prompting and guiding to facilitate the learning process. The degree of guidance given to the learner during the learning process must be considered in any thorough design of an instructional procedure. In some procedures very little guidance is given, and as a consequence the frequency of error responses is relatively high at the start of the training and gradually decreases as learning takes place. In other procedures much guidance is given, and the subjects acquire the target behavior having emitted very few, if any, error responses.

B. F. Skinner (1961, pp. 59-66) has argued that errors are not essential to the successful acquisition of target behaviors. He maintains that a cue fading process can be used to minimize error and thus accelerate learning. In the cue fading process cueing or prompting stimuli are used to supplement the stimuli to which the learner is to eventually respond without aid. The cueing stimuli are chosen to increase the probability of correct responding. As the learning process proceeds, the supplementary cueing stimuli are gradually withdrawn or faded at a rate which maintains a minimum frequency of error responses. Eventually no cues are needed and the terminal behavior is acquired.

Seemingly, Terrace's pioneering work (1963a, 1963b) in errorless discrimination stimulated the production of relevant research on fading. Terrace (1963a) found that if cue fading techniques were used, pigeons could learn a red-green discrimination without producing error responses.

In a second study, Terrace (1963b) extended the use of fading and obtained errorless transfer of training from the red-green discrimination to horizontal-vertical discrimination.

Studies by Hively (1962, 1965), Moore and Goldiamond (1964), Sidman and Stoddard (1967), and Bijou (1965) have shown that cue fading can be efficiently and effectively used in training young children to make relatively complex discriminations. For example, Bijou (1965) revealed an impressive demonstration of the successful use of errorless training techniques and fading when he induced right-left form concepts in young normal and retarded children. On a terminal task item, the subject was presented with five stimuli and was directed to choose the one stimulus which was the same as a given sample stimulus except for an angular rotation. Bijou cued for the correct response by deforming the correct stimulus choice. Then as training progressed, the deformed choice was gradually altered so that it showed progressively greater resemblance to a rotation of the sample stimulus.

The target behavior of this thesis required kindergarten subjects to make discriminations. Given an object to be inserted into a seriated set of objects, the subjects had to discriminate between the correct position in the seriated set and the incorrect positions. To induce this discrimination capability, cues were introduced and then gradually faded.

Since the evidence shows that cue fading can be an effective means of producing discrimination learning, and since the acquisition of the target behavior in the thesis study involves discrimination learning, it is hypothesized that the cue fading method, as used in the thesis study, will contribute to the successful induction of the insertion capability in kindergarten children.

The specification of the conditions of responding. Anderson (1967) claims that enough is known to make a reasonably safe conclusion about the relative learning effects of overt and covert responding. Anderson concludes that learning is facilitated when subjects are required to make overt, constructed responses.

In the thesis study, the subjects who received the training were required to make overt, constructed responses. For each training trial, the experimental subjects responded by attempting to insert a disarranged set of objects into a serial ordered set of objects. According to Anderson's conclusion the response mode used in the training should facilitate learning.

The employment of feedback methods. Stimuli presented to a subject following his emitted response can serve informative as well as reinforcing functions. The completeness of the information provided by the feedback may vary considerably. For example, the subject may be merely told whether his response was correct or incorrect, or he may be shown the correct response and then given a full explanation of why the response was correct or incorrect.

Bourne and Pendleton (1968) compared two conditions of feedback completeness in concept identification problems. In one condition the subject was told whether his response was correct or incorrect. In the other condition, the subject was, in addition, shown the correct response whenever a mistake was made. The results indicated that the latter, more complete, feedback condition was superior to the former, less complete, feedback condition. In a similar study, Travers, Van Wagen, Haggood, and McCormick (1964) studied the relative effectiveness of four different feedback systems used to teach elementary school children the English equivalents of German words. They found that the two conditions in which the subjects' incorrect responses were corrected were significantly more effective than the two conditions in which no correction procedures were used.

In this thesis study, informative feedback was given to the subjects during seriation training. Whenever a subject responded incorrectly, he was told why his response was incorrect, and he was shown the correct response. Thus, at the end of every training trial, the subject was given the opportunity to see the correct behavioral product (a seriated set of objects) regardless of whether he or the experimenter was responsible for its construction. Since the kind of informative feedback used in this thesis has been found to facilitate learning (Bourne and Pendleton, 1968; Travers et al., 1964), it was assumed that such feedback would facilitate the acquisition of seriation capabilities.

The studies reviewed from instructional psychology have provided support for the methods used in the seriation training. The conservation studies have suggested that Piagetian capabilities can be induced and that learning theory can provide a sound basis for constructing the training methods used in successfully inducing Piagetian capabilities. Furthermore, Coxford's results have suggested that some seriation capabilities can be induced. Therefore, as a consequence of this accumulated support, it is hypothesized that the experimental subjects in this thesis will acquire the insertion capability characteristic of seriation stage III.

Conditions for the Promotion of Retention

There is reason to believe that retention of knowledge is more a function of the degree of original learning than a function of other variables which may, for instance, influence the efficiency of learning. For example, it has been known that when the degree of original learning is held constant, learning speed (Underwood, 1954), intralist similarity (Underwood and Richardson, 1958), meaningfulness (Underwood and Richardson, 1956), and associative strength (Underwood and Keppel, 1963) are of no significant consequence for retention.

This phenomenon has received additional support from a recent study by Olton (1969). Olton examined the effects of grammatical context on the retention of paired associates. He found that when original learning was equated, the amount retained by the group which learned paired associates in a grammatical context was no different than the amount retained by the group which learned paired associates in absence of grammatical aid. As expected, however, learning was more rapid when the pairs were embedded within the grammatical context.

The review of literature up to retention has been focused on methods which increase the degree of acquisition. The review led to hypothesizing that the methods used in seriation training would produce significant increments in seriation ability. Therefore, if retention is positively related to original learning, and if this relationship can be extrapolated from adult verbal retention to the retention of induced nonverbal capabilities in young children, then the subjects given what has been hypothesized to be successful seriation training will be expected to show a significant amount of retention.

The pattern of reinforcement used in conditioning a response determines to a large degree the extent to which the response is resistant to extinction. Deese and Hulse (1967, p. 152) report that there exists literally hundreds of experiments showing that partial patterns of reinforcement produce greater resistance to extinction than a continuous pattern of reinforcement. For example, Jenkins, McFann, and Clayton (1950) found that a variable ratio schedule, in which different numbers of responses were required for successive reinforcements, produced five times as many responses during extinction as did a continuous reinforcement schedule, in which reinforcement was given after every response.

The pattern of reinforcement used in the seriation training was continuous in nature. Therefore, the pattern used in seriation training is not expected to contribute much to the long term retention of the acquired seriation capabilities.

Distribution of practice is another factor which influences retention. Although spaced practice, in certain situations, impedes learning, it may, in those situations, facilitate retention (Underwood, 1964). For example, Underwood (1964) reported that spaced practice in which interference is present (learning to make different responses to the same stimuli) seriously impedes learning in the initial stages but greatly facilitates retention after eight days.

In a study having nearly direct applicability to instruction, Rothkopf and Coke (1966) found that distributed practice was superior to massed practice in promoting retention of prose material. Each of eight sentences were rephrased and repeated once, either in immediate succession or after other intervening material. Retention was facilitated when repetition of the sentences was delayed.

In the seriation training, "practice" was massed for the same material and task, but distributed between different materials and tasks. Approximately 30 minutes of training were given to every experimental subject on each of three successive days. The materials used in the first and second training sessions were the same, but the tasks were slightly different. During the second and third training sessions the tasks were the same but the materials were different.

The 24 hour spacing between the three training sessions was expected to contribute to the retention of the acquired seriation capability. However, since only three sessions were used as compared to Underwood's eight, and since the condition in which spaced practice best facilitates retention (namely, a condition where responses interfere) was not present, the extent to which the spacing of practice in seriation training facilitates retention is expected to be somewhat limited.

In addition to the degree of original learning, the pattern of reinforcement, and the distribution of practice, delay of feedback has been found to influence retention. Lintz and Brackbill (1966) found that bigrams learned by a paired-associate method were recalled better after one week when feedback was delayed. Similar effects were found by Sassenrath and Yonge (1968).

During the last two sessions of seriation training, feedback was delayed for most of the responses. The training trials required the subjects to insert a number of objects into a serial ordered set. The feedback for each object was delayed until all objects used in a trial had been inserted. According to the research on delayed feedback, the feedback delay used in seriation training would be expected to promote retention.

The amount of original learning, pattern of reinforcement, distribution of practice, and delay of feedback have been found to be factors which influence retention. Research relevant to these factors has been reviewed and the findings have been related to the methods used in seriation training. In general, the research suggested that certain characteristics of the seriation training (degree of original learning, delay of feedback, and, possibly, distribution of practice) should promote the retention of the acquired seriation capabilities. Therefore, it is hypothesized that the experimental subjects will retain the capabilities acquired in the training.

Conditions to Enhance the Transfer of Learning

Transfer of training may be described as the influence of prior learning or experience in one task on performance in another. Depending on the transfer tasks and the characteristics of the original learning, the carry-over effects may be positive (facilitating), negative (inhibiting) or negligible.

Gagné (1965, p. 231) has distinguished between lateral and vertical transfer. Vertical transfer refers to the degree to which learning in one task influences *learning* in a different task. Lateral transfer refers to the degree to which the learning of a task influences the *performance* of the same general class of task in a different stimulus setting. In this thesis, lateral transfer of the acquired seriation capabilities was the main concern. More specifically, the transfer tasks used in this thesis study were designed for the purpose of identifying whether or not the subjects could transfer the acquired insertion capabilities to different stimulus objects.

Transfer has been studied as a function of the degree of similarity between the stimuli in two tasks and as a function of the degree of similarity between the responses in the two tasks. Some time ago Osgood (1949) presented a transfer surface which illustrated the direction (positive or negative) and magnitude of transfer as a function of the stimulus and response similarities existing between the learning task and the transfer task. In part, the surface showed that positive transfer increased as stimulus similarity increased, provided that the responses were the same or similar. In several studies (Bugelski and Cadwallader, 1956; Dallett, 1962; and Wimer, 1964) most of the possible relations have been examined. All three studies cited above revealed that for the same or very similar responses, positive transfer increases as the stimulus similarity between the learning and transfer tasks increases. Therefore, if this finding is generalizeable, a consideration of the stimulus similarity between the training and transfer tasks used in this thesis should give an indication of the degree of transfer to be expected, since the responses required in both training and transfer tasks were nearly the same.

Three seriation posttests were administered. The first two posttests consisted of two different measures and the third posttest consisted of three different measures. The two measures of the first two posttests and the first two measures of the third posttest were called near and far transfer measures. The third measure of the third posttest was called a far-far transfer measure.

The stimulus materials used in the training tasks consisted of sticks and lined cards. The sticks were placed side-by-side to form what looked like stairsteps. The lined cards were ordered side-by-side from the card with the fewest black lines to the card with the most black lines.

Sticks and lined cards were the materials used in the near transfer tasks of the first two posttests. Only sticks were used in the near transfer tasks of the third posttest. Since the materials used in the near transfer tasks were the same as those materials used in training, it may be more appropriate to call the near transfer measures retention measures. However, since the number of objects used in the training and testing tasks were not the same, the testing tasks will be called near transfer items.

The far transfer trials of the first and second posttest were the same. The materials used in these far transfer trials consisted of "cars" and colored blocks. The "cars" (sticks with wooden wheels attached) were ordered end-to-end according to length along a track to form a train. The colored blocks were painted various shades of blue and were ordered from light blue to dark blue. Only "cars" were used in the far transfer measures of the third posttest.

The far-far transfer trials of the third posttest consisted of "happies" and story cards. The "happies" were rectangular pieces of cardboard with smiling faces drawn on them. The faces were used so that the girth or "fatness" of a "happy" could be distinguished from its height. The girth of the "happies" was the relevant ordering dimension used in the testing trials. The irrelevant dimension of height was not correlated with width.

The storycards were rectangular pictures showing a stick-man, a diving board, and water. When the cards were sequenced properly, they showed the story of the man climbing up the diving board and diving into the water.

The materials used in training and testing have been briefly described. An examination of the stimulus similarities between the training and testing materials will now be made to predict whether or not transfer will be expected.

Since the "cars" were simply sticks with wooden wheels, and since length was the relevant ordering attribute for both "cars" and sticks, the stimulus characteristics for the two materials are considered to be similar. Therefore, the training with sticks is expected to transfer to performance with "cars."

Because of the stimulus similarity between the lined cards used in training and the colored blocks used in testing, more transfer is expected. When the lined cards were ordered according to increasing line density, the amount of light reflected from each card in sequence gradually decreased. The cards then appeared to be ordered from light to dark; thus, training with lined cards is expected to transfer to performance with colored blocks.

The materials used in training were shown to be similar to the materials used in the far transfer trials of the three posttests. The far-far transfer materials of the third posttest, as the adjective "far-far" implies, were less similar to the training materials than the far transfer materials. For example, there existed little similarity between the stimulus characteristics of the story cards (far-far transfer) and sticks (training). Therefore, stimulus similarity cannot be used as a basis for predicting transfer of training from sticks and lined cards to "happies" and story cards.

Stimulus-response similarity is not the only basis for expecting transfer. In addition to stimulus-response similarity, certain training conditions determine the extent of transfer. Two such conditions have been found to be the degree of original learning and the variety of training tasks.

Grant and co-workers (Grant and Berg, 1948; Grant and Cost, 1954) reported two investigations in which transfer was studied as a function of the degree of original learning. The results of both studies revealed that transfer to new problems increased as the amount of learning on previous problems increased.

A series of studies (Adams, 1954; Callatine and Warren, 1955; and Morrisett and Hovland, 1959) has been performed to determine whether the extent (depth) of learning or the variety (breadth) of learning is the dominating factor affecting the transfer of learning. The results of those studies led to the final conclusion that transfer improves with the number of different training problems provided that a high degree of learning occurs with each problem.

Research concerning the relationship between transfer and the nature of training (variety and extensiveness) has been reviewed. The conclusions from that review will be used in conjunction with an examination of the seriation training to determine whether or not transfer from that training can be expected.

Seriation training was divided into three sessions. Since a different combination of task and material was used during each session, the three sessions may be considered as training on three different problems. The subjects receiving seriation training were exposed, on the average, to a minimum of 43 trials for each of the three different problems. Morrisett and Hovland (1959) exposed their subjects to 64 trials on each of three different problems and found maximum transfer. Although the number of trials per seriation problem was somewhat less than the number used by Morrisett and Hovland, the level of learning was maintained by requiring the subjects to meet criterion performance levels on the average of 11 times per seriation problem.

The seriation training conditions were similar to those conditions which Morrisett and Hovland found to promote transfer. Furthermore, high levels of learning were maintained on a variety of seriation training problems. Thus, in accordance with the conclusions of the transfer studies, the variety and extensiveness of the seriation training is expected to promote transfer of training.

Stimulus similarity has already been proposed as a basis for expecting the experimental subjects to transfer their acquired capabilities to the tasks of the far transfer measures. In addition, the extensiveness and variety of seriation training are expected to further increase transfer to those tasks.

Stimulus similarity does not serve as an adequate basis for predicting transfer to the "happies" and story-cards used in the far-far transfer measures of the third posttest. The basis for predicting such a transfer lies, instead, in the extensiveness and variety of the seriation training.

In summary, the capabilities acquired during seriation training are expected to transfer to the far and far-far transfer measures. The basis for the expected transfer was found in the variety and extensiveness of the seriation training and in the stimulus and response similarities existing between the training tasks and the far and far-far transfer tasks.

Research Hypothesis

Since not all levels of the research design were completely crossed with all other levels, the analysis of the data was necessarily performed in two parts. One part, the Experimental versus Control Group Analysis, was an analysis of the experimental and control groups' means from the near and far transfer measures of the three posttests. The other part, the Experimental versus Control versus Special Control Group Analysis, was an analysis of the experimental, control, and special control groups means from the near, far, and far-far transfer measures of the third posttest. In the following sections the research hypotheses for each of the two parts will be constructed from the bases established in the review of literature.

Hypotheses for the Experimental versus Control Group Analysis

As a result of the training, the experimental subjects will acquire the insertion capability. Furthermore, since the experimental subjects received what is expected to be successful training, and the control subjects received no training at all, the experimental group should outperform the control group at least on some of the posttest measures. The expected result stated as a research hypothesis is:

> H_1: There will be a Treatment main effect with the direction of the effect favoring the experimental group.

When sticks were used in ordering tasks, the correct arrangement of sticks usually looks like stairsteps. The stairsteps configuration can be used by the subject to monitor his performance. When the configuration looks like stairsteps, performance is correct, but when the configuration does not look like stairsteps errors have been made and need correcting.

Sticks were used in the near transfer tasks. Consequently, the configuration was available and could be used to monitor performance. None of the far transfer tasks offered any configural aid. Therefore, the near transfer tasks are expected to be less difficult than the far transfer tasks.

Since the control subjects had already shown some ability to seriate with sticks on the seriation pretest and since configural aids were present in near transfer tasks and not in far transfer tasks, the control subjects are expected to show better performance with the near transfer tasks than with the far transfer tasks. Likewise, the experimental subjects should show better performance on the near than on the far transfer tasks since the experimental subjects were trained with the near transfer materials but not with the far transfer materials.

With both experimental and control subjects expected to show better performances on the near than on the far transfer tasks, the following hypothesis regarding test type effects (near versus far transfer) can be made:

> H_2: There will be a Test Type main effect with that effect favoring the near transfer test type.

The materials and tasks used in the training were nearly the same as those used in the near transfer measures. Therefore, the experimental group's performances of the near transfer measures should provide information concerning the degree to which the induced capabilities were retained. Since the experimental subjects are expected to retain those capabilities acquired in training, and since the control subjects received no training at all, the research hypothesis associated with the comparison of the experimental and control groups' near transfer performances becomes:

> H_3: The experimental group's mean for the near transfer measures will be greater than the control group's mean for the near transfer measures.

The far transfer tests basically provided measures of transfer since the materials used in those tests were unlike the materials used in the training. According to the literature review, the experimental subjects are expected to transfer the capabilities acquired in training.

This expected result should be confirmed by the experimental group's acquisition of an overall far transfer mean that is greater than the control group's overall far transfer mean. Stated in the form of a research hypothesis, the expected result regarding transfer effects becomes:

> H_4: The experimental group's overall far transfer mean will be greater than the control group's overall far transfer mean.

<u>Hypotheses for the Experimental versus Control versus Special Control Group Analysis</u>

For this second part of the overall analysis, three groups (experimental, control, and special control) will be compared on three different measures (near, far, and far-far transfer measures) of the third posttest. The experimental and control groups were the same as those used in the first part of the analysis. The subjects of both the experimental and control groups were in stage II at the beginning of the study, whereas the subjects of the special control group were in stage III.

The purpose of the training was to induce stage III capabilities in the experimental group. In other words, the training was designed to provide the experimental group with those capabilities already possessed by the special control group. In accordance with the review of literature, the experimental group is expected to acquire, retain, and transfer those stage III capabilities possessed by the special control group. Therefore, the experimental and special control groups should perform equally well on the posttest measures. Since the control group entered the study with stage III capabilities and was given no training, the special control group with its stage III capabilities should perform better than the control group on the posttest measures. Put in the form of research hypotheses the propositions contrived above become:

> There will be no difference between the experimental and special control group means for each of the following transfer measures of the third posttest:
>
> H_5: near transfer
>
> H_6: far transfer
>
> H_7: far-far transfer
>
> The special control group's mean will be greater than the control group's mean for each of the following measures of the third posttest:

H_8: near transfer

H_9: far transfer

H_{10}: far-far transfer

The research hypotheses of major interest have been presented. These hypotheses, in general predicted that the experimental group will outperform the control group on the posttest measures, that the near transfer measures will be easier for all subjects than the far transfer measures, and that the special control group's performance being no different from the experimental group's performance will be superior to the performance shown by the control group.

CHAPTER III

THE RESEARCH PROCEDURE

Pretest Materials and Tasks

Orange painted sticks were the objects used in the seriation pretest. The sticks were three-quarters inch thick, three-quarters inch wide, and varied in length from one and one-half to nine inches.

Two kinds of tasks were presented. One task required the child to serial order a set of disarranged sticks from the shortest to the tallest to make stair-steps. The other task required correct insertion of a set of three disarranged sticks into a serial ordered set of sticks. There were three ordering tasks and each was followed by a corresponding insertion task. This order of task presentation was chosen so that the ordered set obtained from an ordering task could be used in the insertion task which followed. The number of objects in the ordering tasks was gradually increased as the test continued, but the number of sticks to be inserted was always three. The pretest tasks were presented as follows: 4 sticks were ordered and 3 were inserted, 6 were ordered and 3 inserted, and 8 were ordered and 3 inserted. Appendix

A provides a detailed description of the materials used in each pretest task.

To help each child understand what was to be expected of him on the test, the experimenter provided examples of how to perform the ordering and inserting tasks. The experimenter's monologue used during the illustration and the verbal instructions given to the subject during the testing are found in Appendix B.

The subjects were assigned to stages according to the following criteria. To be considered in stage III, the subject had to correctly perform all three of the ordering tasks and at least two of the insertion tasks. For the stage II subjects two or three ordering tasks had to be performed correctly and two or three insertion tasks had to be performed incorrectly. All subjects not fitting stage II or stage III criterion were considered to be in stage I.

The Sample

The subjects for this study were selected from the three kindergarten classrooms of Scott Elementary School, Dewitt, Michigan. To identify stage II children for inclusion in this study, the seriation pretest was individually administered to each available kindergarten child. Of the 95 children pretested, 32 were found to be in stage II as defined by the pretest. The mean age of these stage II children at the time of pretesting was approximately 5.44 years.

Training and Posttest Materials and Tasks

The following six different kinds of materials were used in the study: sticks of various lengths were ordered side-by-side according to length, cards with various numbers of parallel black lines were ordered according to line density, wooden "cars" of various lengths were ordered end-to-end according to length, blocks painted various shades of blue were ordered according to the shade of blue, rectangular pieces of poster board (happies) which varied in width and height were ordered according to width, and story cards which showed frames of a stickman diving into water were ordered according to the sequence of events.

Below is a more detailed description of the specific tasks and materials used in the different portions of the study. Pretest materials and tasks have already been described; consequently, their description will be omitted in the description which follows.

Training Materials and Tasks - First Training Session

Forty-five different tasks were used during the first day of training. A set of materials was made for each of the tasks. The forty-five sets of materials were placed at the different training stations which were located in one large room.

The same kind of task was presented at all of the stations. At each station, a disarranged set of sticks and two ordered sticks appeared before the subject. Only one of the sticks in the disarranged set would fit between the two ordered sticks to make a serial ordered set of three. The subject's task was to find that one stick and place it between the two ordered sticks. An example of how the sticks were arranged at each station is illustrated in Figure 2.

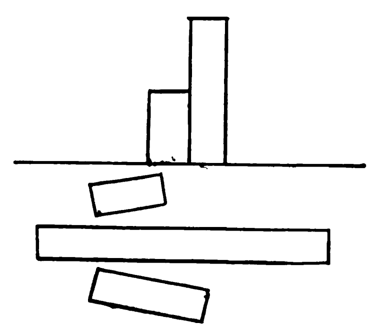


Fig. 2. An illustration of the arrangement of sticks at each station during the first day of training.

The forty-five stations consisted of three groups of fifteen stations. The three groups of fifteen stations differed with respect to the number of sticks in the disarranged sets of sticks. The first fifteen stations had two sticks in the disarranged sets, the second fifteen stations had three sticks in the disarranged sets, and the third fifteen stations had four sticks in the disarranged sets.

Each group of fifteen stations consisted of five subgroups of three stations each. The five subgroups within each group of fifteen stations differed with respect to the fineness of discriminations required in task performance. In other words, each subgroup corresponded to a discrimination level; hence, the three stations within each subgroup provided practice at one discrimination level.

In the progression from one subgroup to the next, finer and finer discriminations had to be made. This was true only for progression within one of the three large groups. Progression from one large group to the next resulted in an abrupt drop in discrimination level since all of the large groups started with approximately the same low discrimination level. Appendix C gives the exact lengths of the sticks used at each station.

Training Materials and Tasks - Second Training Session

For the second day of training thirty training stations were used and each station was supplied with a set of orange sticks. Six sticks were used at each of the first fifteen stations and nine sticks were used at each of the remaining fifteen stations. The sticks in each set were arranged so that part of the set was serial ordered and part was disarranged. At each of the first fifteen stations, four of the six sticks were serial ordered along a line to make stairsteps, and the remaining two were disarranged. At each of the second fifteen stations, six of the nine sticks were ordered and the remaining three were disarranged.

The nature of the task was the same at all thirty training stations. The child was instructed to insert the disarranged sticks into the serial ordered set so that all of the sticks were used in making stairs. The tasks presented at the first fifteen stations required that two sticks be correctly inserted into a set of four sticks. The task for the second group of fifteen stations required that three sticks be correctly inserted into a set of six sticks.

For each of the two groups of fifteen stations, relatively strong cues were given in the beginning stations and then these cues were faded in five steps as movement through the fifteen stations progressed. Within each large group of fifteen stations there were five subgroups of three stations each. Each of these five subgroups corresponded to a cue level and all three stations comprising a subgroup provided practice at one particular cue level.

The cues appeared in the ordered set of sticks. The cues were designed to aid the child in finding the positions in the set where insertions would take place. Although the cues had the potential for helping the child locate the positions for insertion, they could not have helped the child choose the correct stick for those positions.

An ordered set of sticks which contained cues looked as if some of the sticks were missing. In other words, the lengths of sticks increased in regular increments except for those places where insertion would take place. At those places where a stick would be inserted, the increment was comparatively large. Figure 3 illustrates an ordered set which cues for the insertion of two sticks.

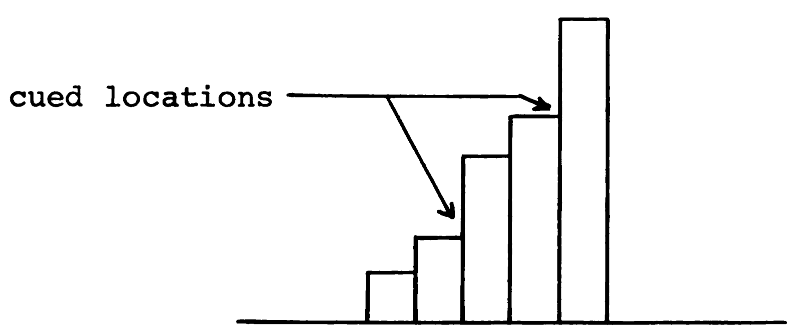


Fig. 3. An illustration of an ordered set of sticks displaying two cues.

In the progression from one cue level to the next, the large increments used in cueing were gradually reduced to regular increments in five stages. The last set of materials in each of the two groups of fifteen stations offered no cues at all. In those ordered sets containing no cues, the lengths of the sticks increased in regular increments and the regular increment was the same as the regular increment in the ordered sets containing cues. Figure 4 illustrates the five stages in the reduction of cues. The lengths of the sticks used at each training station are given in Appendix D.

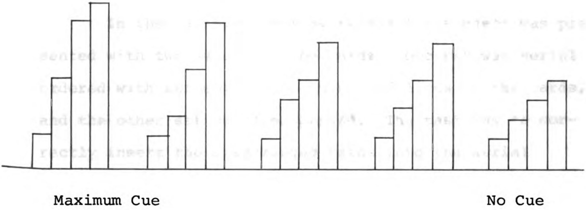


Fig. 4. An illustration of cue reduction during the second day of training.

Training Materials and Tasks - Third Training Session

During the first two days of training, sticks were used. To help the subjects generalize the process of serial ordering, a new material was used in the third day of training.

The materials consisted of rectangular pieces of posterboard upon which were cemented photographs of parallel, evenly spaced, black lines. These cards measured one and three-quarters by four and one-half inches, and the lines on them were parallel to the short side. There were twelve different numbers of lines used on the cards. These numbers ranged from three to thirty-two in such a way that if X were the ordinal position of a card (the card with the Xth most lines), then Y, the number of lines, could be calculated from the following formula: $Y \doteq e^{X}$

In the various training trials the subject was presented with two sets of lined cards. One set was serial ordered with respect to the number of lines on the cards, and the other set was disarranged. The task was to correctly insert the disarranged cards into the serial ordered set.

As in the second day of training, cues were used to aid the subjects in making the correct responses. Twice the cues were introduced and faded in the training sequence of twenty-four stations. Strong cues were used at the beginning of the sequence and these were faded in four stages until no cues were used for those trials appearing in the middle of the sequence. Just after the cues had been faded for the first half of the sequence, more cards were introduced into the tasks. To help the student overcome the added difficulty of having to deal with more cards, the cues were introduced again and then faded in four stages until the last three trials contained no cues.

The width of the lines served as the cue. For those sets of cards in which the cues were present, the number of lines on a card and the width of lines were directly related.

As mentioned above, the cues were faded in four stages. The variation in line widths of a strongly cued set of cards was comparatively large. In the progression from one cue level to the next, the variance in line widths was reduced toward a mean line width. The sets of cards in which no cues were used had no variation in line width. A detailed description of the cards used in the various training trials is found in Appendix E.

Posttest Materials and Tasks - Posttests One and Two

The four different kinds of materials used in each of the first two posttests were: sticks, lined cards, wooden "cars," and blue painted blocks. Both ordering and inserting tasks were performed with all of the materials except sticks. Only inserting tasks were performed with sticks, since all subjects in the study had demonstrated the ability to order sticks on the pretest.

All of the training was focused on the insertion capability. Consequently, there existed the possibility that the training would be very specific and would result in the subjects being able to insert but not being able to order. Since stage III performance with any set of materials requires both serial ordering and inserting, the posttest included both serial ordering and insertion tasks for all materials except sticks.

The testing materials were set up at stations and the subjects were guided individually from one station to the next in the testing sequence. The materials were presented in the following order: sticks, lined cards, "cars," and colored blocks. Whenever both ordering and inserting tasks were performed with a given material, the ordering task always preceded the corresponding inserting task. The tasks and materials are described below in the order in which they were presented during the posttest. Since the materials and tasks were the same in the first two posttests, the description applies to either of those posttests.

Sticks like those used in the training were used to test the child's ability to insert a set of disarranged sticks into a serial ordered set. Three different trials were presented and the number of sticks in both the disarranged set and the ordered set were increased from one trial to the next. In the first trial, two sticks were to be inserted into a serial ordered set of four; in the second trial, three were to be inserted into six; and in the final trial, four sticks were to be inserted into a serial ordered set of eight.

Lined cards, like those used in the training, were used in a second series of trials. The first task required that four lined cards be serial ordered from the card with the fewest lines all the way to the card with the most black lines. The second task required that two additional lined cards be inserted into the four cards just serial ordered. The ordering and inserting sequence was repeated again with six cards to be ordered and three to be inserted into the ordered six.

Wooden "cars" were used in the next trials. A wooden "car" was a stick with short dowels glued to one side to give the impression of wheels. The "cars" looked like boxcars of a train and were to be serial ordered with respect to length. Rather than order the "cars" side-by-side to form a staircase-like figure, the "cars" had to be ordered from end-to-end along a track to form a "train."

There was one ordering and one insertion task performed with "cars." Six "cars" were presented disarranged and the subject was instructed to order them from the shortest to the longest along a track drawn on a sheet of paper. If the subject could not order the cars, the tester arranged the cars in serial order for the insertion task which followed. Once an ordered set of cars had been formed, either by the subject or the tester, the subject was given three additional "cars" to insert into the ordered set of six.

Colored blocks were used in the last two trials of the test. A block was constructed of wood three-quarters inch wide, three-quarters inch thick, and four inches long. A piece of posterboard was cemented to one side of each block, and each piece of posterboard was painted a different shade of blue.

In the first test trial with colored blocks, the subject was instructed to place eight blocks side-by-side in serial order from the lightest to the darkest blue. If the child failed to order the eight colored blocks, the tester ordered the blocks for use with the insertion task which followed. In the final task, the subject was presented with three disarranged colored blocks and the serial ordered set from the previous trial. The subjects were instructed to insert the three disarranged blocks into the ordered set so that all blocks would be arranged from the lightest to the darkest blue.

For a major portion of the analysis of the data, scores on near and far transfer items will be used. The score obtained from the trials involving sticks and lined cards will be considered a near transfer score and the total score from the trials involving "cars" and colored blocks will be considered far transfer scores.

The test trials in which sticks and lined cards were used are considered near transfer trials because the materials used in the trials were the same kind as those used in the training. Since the "cars" and colored blocks were not used in training and since the results of a previous study (Schafer, 1969) suggest that these materials were more difficult to order and insert than sticks, the trials with "cars" and colored blocks are considered far transfer trials. A detailed description of the test materials used in the first two posttests is presented in Appendix F.

Posttest Materials and Tasks - Posttest Three

The third posttest consisted of near, far, and far-far transfer tasks. There were two near transfer tasks, and in each, a disarranged set of sticks was to be inserted into an ordered set. In the first near transfer task, five sticks were to be inserted into an ordered set of 10. In the second near transfer task, six sticks were to be inserted into an ordered set of 12.

There were two far transfer tasks, each of which required that a disarranged set of "cars" be inserted into an ordered set of "cars." In the first task three "cars" were inserted into an ordered set of six, and in the second task, four "cars" were inserted into an ordered set of eight.

The near and far transfer tasks of the third posttest were not the same as those tasks used in the near and far transfer tasks of the first two posttests. In the near transfer tasks of posttests one and two, sticks and lined cards were used; whereas in the near transfer tasks of posttest three only sticks were used. Both "cars" and colored blocks were used in the far transfer tasks of posttests one and two; whereas only "cars were used in the far transfer tasks of posttest three. The reason for eliminating the tasks with lined cards and colored blocks from posttest three was to provide the time for administering the far-far transfer measures.

"Happies" and storycards were the two different materials used in the far-far transfer tasks of the third posttest. "Happies" were rectangular pieces of white posterboard on which were drawn smiling faces. The faces were used only to give each card a particular orientation so that height and girth of the card could be determined. The "happies" were ordered according to girth or "fatness" (i.e., the width of the card). The height of the "happies" varied randomly with respect to the girth; hence, height was considered an irrelevant dimension. Evidence from a pilot study revealed that ordering the "happies" with respect to girth could be challenging for some first grade children. Considering the evidence from the pilot study and Shantz's (1967) results, which showed that irrelevant information (height, in the case of "happies") retarded performance of seriation tasks, the tasks with "happies" were considered to be relatively difficult for the kindergarten subjects of the seriation study. Hence, the tasks with "happies" were called far-far transfer tasks.

Two specific tasks were performed with the "happies." After a brief orientation session in which the subjects were taught both to distinguish between fat and skinny "happies" and to disregard the height of the "happies," the subjects were given six "happies" which they were to order side-by-side from skinniest to fattest. If the task was performed correctly the next task was begun; but if the task was not performed correctly, the tester arranged the six "happies" in the proper serial order for use in the next task. The second task with "happies" challenged the subjects to insert two additional "happies" into the ordered set of six.

Storycards were used in the last two tasks of the far-far transfer measures. The storycards were rectangular pieces of posterboard, each showing a picture of a stick man, the ground, a diving board, and water. The picture on each card showed the man at a different stage in the process of climbing up the diving board and diving into the water.

To perform the first task with storycards, the subject had to put nine storycards in order showing the sequence of the stick man climbing the ladder and diving into the water. Each subject was given the first card in the sequence. Before going onto the second task with the storycards, the tester made sure that all cards were in the proper order. With the correct sequence of cards present, the subject's second task with cards was to correctly insert two additional cards into the ordered set. A detailed description of the materials used in posttest three is presented in Appendix G.

Verbal Instructions Given During the Posttests

To help standardize the administration of the posttests, the same verbal instructions were given to each subject. The general format of instructions used with each material was the same. First, an orientation session was presented to make sure the subject understood the nature of the task. Next the subject was instructed to order or insert according to a specified attribute. Each time a subject finished a task he was asked to recheck his work and correct any mistakes which may have been made. No reinforcements were given during the actual testing portions of the posttests. The verbal instructions given during posttests one and two are presented in Appendix H. The verbal instructions for posttest three are presented in Appendix I.

Training Procedures

One group of experimental subjects received training during three consecutive days of the first week. The other group of experimental subjects received the same training during three consecutive days of the second week. Since the training was the same for both groups, the training procedures will be described for the three training sessions of one week.

The following topics will be discussed in this section: the two kinds of reinforcement used, the methods of administering the reinforcements during the three days of training, and the performance criteria for the progression through the training trials of each day.

Reinforcements and the Orientation Sessions Offered During Training

Two kinds of reinforcement were assumed to be working during training. One kind of reinforcement was the praise given to the child by the experimenter following a set of target responses. The praises took the following forms: Good boy (girl), Very nice job, Good job, Very good, Very nice. The praises preceded and followed a statement indicating what behavioral product was being reinforced. For example, after a child had correctly inserted sticks into an ordered set, the experimenter would have said, "You are correct, Very nice job. See how all of the sticks go from the shortest to the tallest to make stairs. Good job."

Another kind of reinforcement was offered in the form of marbles. At the beginning of a training sequence, a child was given a box made of clear plastic. This box was divided into columns so that marbles could be stacked one on top of the other. The child was told that the plastic box was a "marble bank" and that he would be given marbles for the bank when he performed correctly. The subject was shown how the marbles stacked up, and he was asked to try to get as many stacks of marbles as he could.

The marble reinforcers were given out in fundamentally two different ways. Each task in the first training sequence challenged the subject to insert one stick between two ordered sticks. A marble was given for each task performed correctly. The tasks performed in the second and third training sequences challenged the subject to insert a number of objects into an ordered set of objects. For each of these tasks, the subject was given one marble for each object correctly inserted and one additional marble for inserting all the objects correctly.

After the "marble bank" was introduced and just prior to beginning each training sequence, each subject was given a short orientation session. During each of the three orientation sessions, the sequence of events occurred as follows: 1) the materials were introduced and a general description of the task was given, 2) the cues and the methods of using the cues were described (only for the second and third training sessions), 3) the correct performance of the task was demonstrated, 4) the method of reinforcement was shown, 5) the task was performed incorrectly and reasons were given to explain why the task performance was considered incorrect, and 6) the subject was given opportunity to perform the task and reinforcement was administered accordingly.

Immediately after each orientation session, the corresponding training sequence was begun. The sequence of events for each training trail was generally the same regardless of the training sequence. Each training trial was begun by telling the subject what he was to do with the materials (for example, "Put the disarranged sticks in with the ordered set of sticks so that all of the sticks go from the shortest to the tallest to make stairs."). When the subject had completed the task, verbal praises and marbles were given where they were due, and any mistakes were pointed out and corrected. By correcting mistakes, the subject was given the opportunity to see the correct configuration at the end of each trial.

The Performance Criteria for Progression Through the Training Trials

The materials were arranged in discrimination or cue levels, and each of these levels consisted of three different practice trials. A subject's progression through the training sequence depended on his ability to meet a performance criterion at each discrimination or cue level. The use of performance criteria was intended to insure that the subject had mastered the trials at one level before moving on to the next more difficult level. It was assumed that by having the subjects master the trials at one level, the probability of success at the next level would be increased, and thus the subjects would pass through the training with little difficulty and much reinforcement.

The criterion for passing from one cue level to the next was the same for the second and third days of training. This criterion will be discussed below.

At each cue level, the subject received three different trials in a particular order. If no mistakes were made on these trials, the subject passed directly to the next cue level. On the other hand, if any mistakes were made on the first attempts of any of the three trials, the subject was recycled through the same three trials in the same order. During any one recycle, the subject was immediately passed to the next cue level when any two trials of the recycle were performed correctly. Therefore, if the first two trials of any recycle were performed correctly, the third trial was omitted and the subject immediately began training at the next cue level. At any cue level, no subject ever required more than three recycles.

The criterion used for the first day of training was different from that criterion used during the second and third days of training. At each discrimination level of the first day of training the subject received three different practice trials. Since there was a relatively high probability that these three trials could have been correctly performed by chance, the subject was recycled through the trials regardless of his initial performance on the trials. The recycling was continued until any two trials of any recycle were correctly performed. Whenever any two trials of a recycle were correctly performed, the subject immediately progressed to the next discrimination level. At any one discrimination level, no subject ever required more than three recycles.

Sequence of Events and Design

The following topics will be discussed in this section: 1) the rationale for the use of two groups, 2) the rationale for the sequence of treatment, 3) the sequence of events, and 4) the design. The discussions of rationale are intended to lend meaning to the detailed description of the sequence of events. The topic of design will mainly be a discussion of the control or lack of control of certain sources of invalidity.

Rationale for the Use of Two Groups (Groups I and II)

Two groups (groups I and II), each with an experimental and control subgroup, and a special control group were used in this study. The subjects in groups I and II were identified by their performances on the seriation pretest as being in stage II with respect to seriation. The subjects in the special control group were identified as stage III subjects with respect to seriation.

The focus of the study centered primarily on the subjects in groups I and II. These subjects were observed on all three posttest measures. The special control group subjects, however, were only observed on the third and final posttest.

All stage II subjects were randomly divided into two groups, group I and group II, and each of these groups was randomly divided into an experimental and control group. Both subgroups of experimental subjects received the same training but at different times. Basically, the experiment with group II was a replication of the experiment with group I.

Why run the experiment twice, once each with two relatively small groups, when it might have been run once with a relatively large group? With one large group and the use of two trainers, the time needed for training half of all available subjects would have been six days, two days for each set of materials. With one large experimental group, not all of the experimental subjects could have received the same treatment at approximately the same time. Half of the experimental subjects would have received the same treatment on each set of materials a day later than the other half. Furthermore, with six days of training, a weekend would have interrupted a portion of the training sequence.

With two groups and two trainers, the experimental subjects of each group received all of the training in three consecutive days. Thus, the design rendered the experimental subjects of each group more homogeneous with respect to the time of treatment and as a consequence possibly reduced the magnitude of extraexperimental interference. With a reduction of interference between training sessions, more of what was learned was likely to be retained from session to session. More retention between sessions may have led to more transfer of training and thus to a greater total impact on learning.

Rationale for the Sequence of Treatment

It seems obvious that a capability is most efficiently acquired when training progresses from relatively easy tasks to the more difficult tasks. In the training, sticks were used prior to lined cards because it was assumed that the training tasks with sticks would be easier than the training tasks with cards. The basis for this assumption lies in the findings of an unpublished study (Schafer, 1969). The data from that study suggested that "cars" and colored blocks, such as those used in the transfer tasks of this study, were more difficult to use in insertion tasks than were sticks.

The ordered sets of "cars" and colored blocks differed in one fundamental way from the ordered set of sticks. When the sticks were ordered, they formed a geometric figure which was indicative of the correctness of the order. The degree to which the set looked like a triangular shaped staircase was directly related to the degree of ordering.

On the other hand, when "cars" and colored blocks were ordered they did not form geometric patterns which indicated the correctness of the order. Regardless of order, the "cars" formed a line and the colored blocks formed a rectangle

Since the geometric pattern formed by the set of lined cards is not related to the correctness of ordering, the lined cards would fall into the same class as the "cars" and colored blocks. Hence, the assumption would be that the lined cards, like "cars" and colored blocks, present a more difficult problem for insertion than sticks, and therefore, should be used after sticks in the training sequence.

Prior to the training with lined cards, sticks were used in two somewhat different training sessions. The two sessions in which sticks were used varied with respect to the kind of task performed. During the first session, the position for insertion (between the two ordered sticks) was designated and the subject was to find the appropriate stick for that position. During the second session, the task essentially required the subject to find the appropriate position in an ordered set for each of a given number of sticks.

The training given during the second day was designed to develop the target capability of inserting a disarranged set of sticks into an already ordered set which contains no cues. The first day of training was designed to develop a capability which could be used in the second day of training. During the second session the positions for insertion in the ordered set were cued and then these cues were faded as training progressed. If the positions for insertion were cued but the subject could not find the appropriate stick for those positions, the cues would be essentially ineffective in helping to shape the target behavior. Therefore, so that the cues could be effectively used in the second session, the subjects were trained in the first session to find the appropriate stick for a given position.

To summarize, it can be said that the order in which the three training sessions were presented was arranged so that learning which occurred during one session would facilitate learning which took place during the next session.

Sequence of Events

The sequence of events is outlined below. The outline shows when the various groups (groups I and II and their respective experimental and control subgroups, and the special control group) were formed, when the training was administered to the experimental groups, and when testing was carried out. The numbers refer to the days of the experiment numbered from the administration of the seriation pretest to the administration of the third and final posttest.

Days	Events				
1	Seriation pretest				
2	Seriation pretest				
3	Seriation pretest				
4					
5	Stage I, II, and III subjects were identified. Stage II subjects were randomly divided into two groups, groups I and II.				
	Group I		Group II		
6					
7					
8	Group I received the CATB and was randomly divided into experimental (E) and control (C) subgroups				
	E	C			
9	1st day of training				
10	2nd day of training				
11	3rd day of training				
12	Posttest 1	Posttest 1			
13					
14					
15			Group II received the CATB and was randomly divided into experimental (E) and control (C) subgroups		
			E	C	
16			1st day of training		
17			2nd day of training		
18			3rd day of training		

Sequence of Events (continued):

Days	Events				
	Group I		Group II		
	E	C	E	C	
19			Posttest 1	Posttest 1	
20					
21					
22	Posttest 2	Posttest 2			
23					
24					
25			Posttest 2	Posttest 2	
26					
. . .					The special control group was formed from a random selection of stage III subjects
					Special Control
148			Posttest 3	Posttest 3	
149					
150	Posttest 3	Posttest 3			
151					
152					Posttest 3

The two groups (I and II) essentially differed only with respect to the times of testing and training. Therefore, to make clear the differences between groups I and II the following description of particular time intervals is necessary:

a) Group II experimental subjects received their training approximately seven days after group II experimental subjects received their training.

b) All the subjects in each group received the first posttest on the day after the corresponding experimental group finished training.

c) Group I was given the second posttest 10 days after the experimental subjects of group I finished training. Group II was given the second posttest six days after the experimental subjects of group II finished training.

d) Posttest three was given to group I approximately 139 days (four months, 2.5 weeks) after the training of group I experimentals, and posttest three was given to group II approximately 130 days (four months, 1.2 weeks) after the training of group II experimentals.

Covariance and Dependent Measures

The covariance and dependent measures will be discussed in this section. Particular concern will be given to the derivation of the dependent variable scores.

Covariance Measures

Since the number of subjects used in the study was somewhat small, the probability of obtaining randomly equivalent groups was lessened. To help combat the effects of having groups that are not randomly equivalent in all respects and to increase the precision of the analysis, the following four covariables were used: chronological age, field independence, reflectivity, and impulse control.

The values for the chronological age variable were the ages of subjects in months at the time the first post-test was taken. The values for the other three covariables consisted of scores obtained on the Cincinnati Autonomy Test Battery (CATB) (Banta, 1968). The field independence score (Early Childhood Embedded Figures Test) was the total number of trials, out of 14, in which the subject was able to distinguish field from ground. According to Witkin (1950), a person who scores high on field independence generally experiences his surroundings analytically and tends to impose structure on a field which lacks it. Reflectivity (Early Childhood Matching Familiar Figures

Test) was measured by the total number of trials, out of 12, in which the subject was able to match familiar figures. Kagan and Moss (1963) have found that reflective children, children who take more time to respond but who get more items correct, usually demonstrate higher standards of mastery on intellectual tasks, choose more difficult tasks, and work longer on the items than do impulsive children. The average rate of line drawing (Draw-a-Line Slowly Test) was considered to be a measure of impulse control, and this average rate was calculated in inches per hundredths of a minute. Impulse control, according to Banta (1968) is a measure of the child's ability to control motor responses. A child with a high impulse control score is less likely to control his motor responses. More thorough descriptions, including reliability measures, of the Early Childhood Embedded Figures Test, the Early Childhood Matching Familiar Figures Test, and the Draw-a-Line Slowly Test are found in Appendix J.

Dependent Measures

In the study of seriation, researchers have exclusively used a stringent method of scoring. With the stringent method no partial credit was given. The task performance was either perfectly correct (no objects out of order) and a score of "one" was awarded, or the performance was judged incorrect (at least one object out of place), in which case a score of zero was awarded.

Since a set of objects can be more or less ordered, in this thesis two lenient or partial credit methods of scoring (L1 and L2) were developed and used in addition to the stringent method (S). In other words, three different scoring methods (S, L1 and L2) were used to derive scores from the same performance data.

A preliminary data analysis (see page 122) revealed that the three different scoring methods produced indistinguishable results. Consequently only one scoring method was needed. Since the stringent method provided the most meaningful scores, only stringent scores were used in the final analysis and only the stringent scoring method will be described here. The two lenient scoring methods are described in Appendix K.

The first two posttests included near and far transfer tasks. The third posttest included near, far, and far-far transfer tasks. For every subject, a stringent score was calculated for each different set of tasks (near, far, and far-far) of each posttest. Thus, every experimental and control subject had a near and far transfer score from each of the three posttests and a far-far transfer score from posttest three.

The stringent score calculated for each different set of tasks of each posttest was the percent attained of the maximum attainable score. Thus, with a score of one given for each task performed without error and a score of zero given for an other than perfect performance, the derived stringent score for a set of tasks (near, far, and far-far) of a particular posttest was the percent of tasks performed without error. For example, a subject's score of 57 on the near transfer tasks of posttest one would mean that the subject performed 57 percent of the near transfer tasks of posttest one without error.

For the final analysis, each experimental and control subject contributed seven scores derived in the manner described above. It was this set of scores which constituted the dependent measures.

Design and Methods of Statistical Analysis

Since the experiment with one group was essentially a replication of the experiment with the other group, the general research design will be discussed with respect to one group. After the members of a group had been administered three parts of the Cincinnati Autonomy Test Battery, they were randomly divided into experimental and control subgroups. Each experimental subject then was given approximately thirty minutes of training on each of the three consecutive days, while the control subjects were given no training. Following the training all subjects of a group were posttested three times. Group I subjects were given posttests one, ten, and approximately 139 days after training. Group II subjects were given posttests one, six, and approximately 130 days after training.

The basic research design, being similar to Campbell and Stanley's (1963) posttest-only control group design, controls for most of the critical sources of internal validity. Assuming that randomization renders the experimental and control groups essentially equivalent on extraexperimental dimensions, the design may still have a few relatively minor sources of internal invalidity. First, since the experimental subjects received all of the attention during the three training sessions while the control subjects received none, the Hawthorne effect may produce some difference between the posttest scores of the experimental and control subjects. Second, since the experimental subjects had the opportunity to become familiar with the experimenters during the training, they may feel more at ease during the posttesting and consequently do better than the control subjects.

Although the design used in this study may not control for all sources of internal invalidity, those critical sources considered by Stanley and Campbell are seemingly controlled. Consequently, the number of alternative explanations for the effects observed may be considerably reduced.

The matrix in Figure 5 shows the independent variable factors of Group (groups I and II) and Treatment (experimental and control). The matrix also shows the dependent variable factors of Posttest (posttest 1, 2 and 3), Test Type (near, far, and far-far transfer), and Scoring Method (S, L1, and L2). Since there were three different types of tests and three different ways of scoring each type, the design revealed a multiple variable aspect. Furthermore, since the performance of the different test types was repeatedly measured across the posttests, there was a repeated measures aspect of the design. As a consequence of the multiple variable and repeated measures aspects of the design, the major portion of the data analysis was accomplished with a repeated measures, multivariate analysis of variance.

The repeated measures, multivariate model was considered more appropriate than the univariate repeated measures model for a number of reasons. Since the dependent measures were taken on each subject, those dependent measures could not be treated as if they were independent from one another. Consequently, the univariate model would have been inappropriate since a separate, independent

		Posttest 1						Posttest 2						Posttest 3								
		NT			FT			NT			FT			NT			FT			FFT		
		S	L1	L2	S	L1	L2	S	L1	L2	S	L1	L2	S	L1	L2	S	L1	L2	S	L1	L2
Group I	E																					
	C																					
Group II	E																					
	C																					
Special Control		No Data																				

Legend: E = experimental subgroups
C = control subgroups
NT = near transfer
FT = far transfer
FFT = far-far transfer
S = stringent scores
L1 = lenient scores (sequence)
L2 = lenient scores (correlation)

Fig. 5. Matrix showing independent and dependent variables and the respective levels for each.

univariate analysis would have been necessary for each different dependent variable. With the multivariate model, on the other hand, the dependent variables can be considered simultaneously and a single probability statement can be applied to all dependent measures taken together. In addition to being more suitable for analyzing multiple dimensional designs, the multivariate model, unlike the univariate model, avoids the precarious assumption that the off diagonal elements of variance-covariance matrix are equal.[1]

The computer program used in the data analysis was a modified version of the program developed by Jeremy D. Finn (1968). Neither this program nor the model which it serves could be used to analyze the entire set of data as illustrated in the data matrix of Figure 5. Figure 5 shows that the far-far transfer measure and the control group factor were not completely crossed with the other levels of the design. The far-far transfer measure was administered only during the third posttest, and the special group was tested only during the third posttest.

As a consequence of the incompleteness of the

---

[1]Information concerning the appropriateness of multivariate analysis for the design of this study was acquired through personal communication with Dr. William Schmidt, Educational Psychology, Michigan State University, 1970.

design, the data was initially analyzed in two parts. One part, the Experimental versus Control Group Analysis, consisted of a repeated measures, multivariate analysis of variance of the following factors and their respective levels: Groups (groups I and II), Treatment (experimental and control), Posttest (posttests 1, 2, and 3), Test Type (near and far transfer), and Scoring Methods (S, L1, L2). The other part, the Experimental versus Control versus Special Control Group Analysis, was a multivariate analysis of the performance of all subjects, including subjects of the special control group, on all measures of posttest 3, including the far-far transfer measure.

Preliminary analyses considerably simplified the analyses reported in Chapter IV. To determine whether or not the three different methods of scoring (S, L1, and L2) were providing unique bits of information a regression analysis was performed. The results of that analysis revealed that knowledge of one score allows near perfect prediction of the other two ($\chi^2 = 454.55$, df = 98, $p < .00001$). Since the stringent method of scoring was the method used in other seriation studies and since the stringent score (percent of tasks performed without error) was considerably more meaningful than the two lenient scores, stringent (S) scores were used in the final analysis.

Two further simplifications were made as a result of preliminary analysis. First, another regression analysis revealed that the four covariates (age, field independence, reflectivity, and impulse control) were not significantly associated with the dependent variables (chi square for the test of the hypothesis of no association between dependent variables and covariables = 30.4855, df = 28, $p < 0.3404$). Therefore, in the final analysis the covariates were eliminated. Second, the Experimental versus Control Group Analysis, was carried out with the two groups (groups I and II) separated. No main effect, nor any interaction effects associated with the Group variable were found. As a consequence, the Group variable, being of no particular interest, was eliminated in the final analysis, and the results were reported in terms of only Experimental versus Control.

Since the Groups factor and two of the three different scoring methods could be eliminated, the matrices for the two parts of the analysis reported in Chapter IV became as shown in Figures 6 and 7. The data from the matrix in Figure 6 was analyzed with a repeated measures, multivariate analysis of variance. The data from the matrix in Figure 7 was analyzed with a multivariate analysis of variance. The results of both analysis are described in the next chapter.

	Posttest 1		Posttest 2		Posttest 3	
	NT	FT	NT	FT	NT	FT
Experimental (N=15) (pooled from groups I and II)						
Control (N=17) (pooled from groups I and II)						

Fig. 6. Matrix for the Experimental versus Control Group Analysis. The stringent scores (S) were the numbers in the cells. Note, there was an attrition of two experimental subjects because of illness.

	Posttest 3		
	NT	FT	FFT
Experimental (N = 15)			
Control (N = 17)			
Special Control (N = 13)			

Fig. 7. Matrix for the Experimental versus Control versus Special Control Group Analysis. Only stringent scores (S) were used. FFT refers to the far-far transfer measure.

CHAPTER IV

ANALYSIS OF DATA

Because not all levels of the design were completely crossed with all other levels and because the special control group was administered only the third and final posttest, the analysis of the data, was necessarily performed in two parts. The matrices corresponding to those two parts have already been shown in Chapter III (Figures 6 and 7). In this chapter, the analysis of the data within each of those matrices will be described and discussed. See Appendix L for the compilation of all the data used in this study.

Experimental versus Control Group Analysis

The data considered in this part of the analysis consisted of the experimental and control subjects' near and far transfer scores from each of the three posttests. Each score was the per cent of seriation tasks performed without error. The group means (mean per cents correct) and standard deviations calculated from the near and far transfer scores appear in Tables 1 (near transfer and 2 (far transfer).

Table 1

Near Transfer Means and Standard Deviations
(Percentages)

	Experimental		Control		Overall	
	Mean	SD	Mean	SD	Mean	SD
Posttest 1	83.9%	18.7%	38.7%	24.4%	59.9%	31.9%
Posttest 2	80.0	22.6	52.4	21.4	65.3	26.4
Posttest 3	76.7	40.3	52.9	43.6	64.1	44.4
Overall	80.2	28.9	48.0	32.1		

Table 2

Far Transfer Means and Standard Deviations
(Percentages)

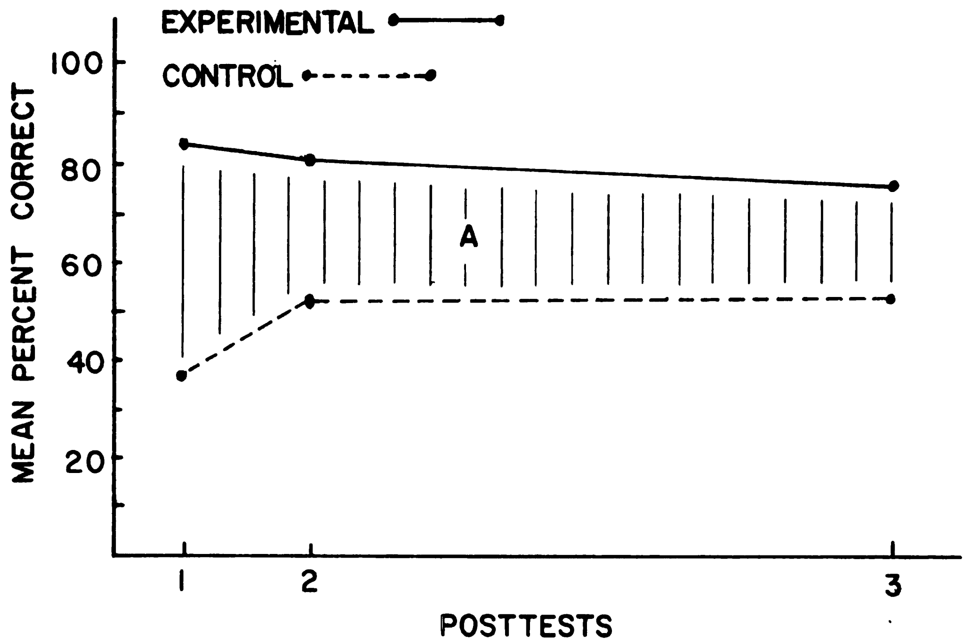
	Experimental		Control		Overall	
	Mean	SD	Mean	SD	Mean	SD
Posttest 1	41.7%	19.7%	35.3%	22.8%	38.3%	21.9%
Posttest 2	68.3	33.5	39.7	24.3	53.1	32.8
Posttest 3	46.7	42.7	41.2	30.8	43.8	37.6
Overall	52.2	35.3	38.7	26.3		

The means from Tables 1 and 2 are shown in Figure 8 in the form of two graphs, one (A) showing a comparison of the experimental and control groups' near transfer means from the three posttests and the other (B) showing a similar comparison with the far transfer means. The graph of near transfer means reveals that the experimental group's near transfer capabilities remained relatively unchanged across the posttests and were, at the same time, seemingly superior to the control group's corresponding capabilities. The graph of far transfer means, on the other hand, does not allow a set of conclusions which are parallel to those derived from the near transfer graph. Contrary to what might be initially expected, the experimental group's superior far transfer capability was seemingly shown on the second posttest rather than on the first. In any respect, there clearly was no massive, substantial far transfer superiority revealed by the experimental group.

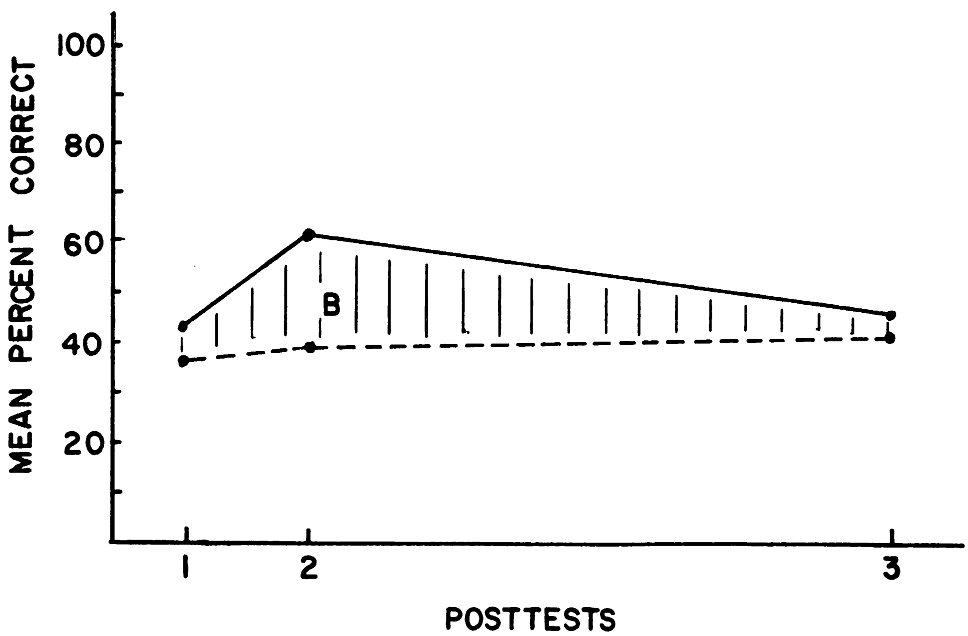
Since there were repeated measures (posttests 1, 2, and 3) on multiple variables (near and far transfer measures), a repeated measures, multivariate analysis of variance (Bock, 1963) was used to statistically analyze the data described above. For this particular kind of analysis the scores were assumed to be distributed in a multivariate normal form. See Appendix M for a thorough

Fig. 8. Graphs of the experimental and control groups' near and far transfer means. To give at least some indication of the time span between posttests, the distances from the absissa to the posttest numbers were scaled according to the function $N'=.5\sqrt{N}$, where N' was the number scaled and N was the number of days between the end of training and the particular posttest. Posttests were given approximately 1, 8, and 132 days after training. Note, the dependent measure used to calculate the means was the per cent of seriation tasks performed without error.

NEAR TRANSFER MEANS



FAR TRANSFER MEANS



display of information (sample correlation matrices, variances, hypothesis mean products) relevant to the multivariate analysis reported here.

The repeated measures, multivariate analysis tested null hypotheses for the following sources of variation: 1) Treatment main effect, 2) Posttest main effect, 3) Test Type main effect, 4) Treatment x Posttest interaction, 5) Treatment x Test Type interaction, 6) Posttest x Test Type interaction, and 7) Treatment x Posttest x Test Type interaction. Of those sources of variation, the Treatment main effect and the Test Type main effect were of prime interest since they were the sources associated with the following two research hypotheses developed in Chapter II:

> H_1: There will be a Treatment main effect with the direction of the effect favoring the experimental group.

> H_2: There will be a Test Type main effect with the direction of the effect favoring the near transfer test type.

Rather than begin the discussion of results with the simple main effects, the complex Treatment x Posttest x Test Type interaction will be treated first. The reason for considering this interaction first stems from the need to discuss all the other significant results in light of this significant three-way interaction. For each test to be described in this analysis, the probability of falsely rejecting a true null hypothesis will be 0.05 (i.e., $\alpha = 0.05$).

Because of the particular nature of the analysis, there existed no single test for the Treatment x Posttest x Test Type interaction. Instead, the test was performed in two parts. If significance was found in the test of either part, then a significant Treatment x Posttest x Test Type interaction was considered present.[2] Table 3 shows the multivariate F ratio for part 1 of the three-way interaction to be significant ($p<0.0235$). Hence, the null hypothesis associated with the Treatment x Posttest x Test Type interaction was rejected.

Since the three-way interaction is not easy to interpret in any simple numeric way, the graphs shown in Figure 8 will be used to facilitate an interpretation. Whereas a two-way interaction is determined by the degree to which the shapes of two <u>lines</u> correspond, a three-way interaction is determined by the degree to which the shapes of two <u>surfaces</u> correspond. The two surfaces

---

[2]Dr. William Schmidt, Educational Psychology, Michigan State University, personal communication, February, 1971.

Table 3

Multivariate Test for the Treatment x Posttest x Test Type Interaction

Source	df	Multivariate F Ratio	p less than
Treatment x Post-test x Test Type - part 1	1,30	5.7668	0.0235
Treatment x Post-test x Test Type - part 2	1,30	3.2314	0.0839

referred to in consideration of the Treatment x Posttest x Test Type interaction are surfaces A and B in Figure 8. Since the surfaces, A and B, are not geometrically similar, a three-way interaction is implied.

The bottom sides of the two surfaces (i.e., the performances of the control group) seem to be "parallel." Hence, the difference in the shapes of the top sides (i.e., the performances of the experimental group) appears to hold within it the locus of the three-way interaction. Since the ability to perform a task often tends to gradually decline in time from the end of instruction, the slight gradual decrease in the experimental group's performance of the near transfer tests (top line, surface A,

Figure 8) was not unexpected. In a reasonable manner, the experimental group's performance of the far transfer tests (top line, surface B, Figure 8) might be expected to follow a similar decline, thus showing a decrease from maximum transfer on the first posttest to minimum transfer on the third posttest. However, contrary to expectation, maximum transfer was found on the second posttest rather than on the first. A t test (Winer, 1962, pp. 36-38) revealed that the experimental group outperformed the control group on the second far transfer measure ($t=2.736$, $df=25$, $p<0.01$). Although there is no simple analytic way to unequivocally identify the locus of the three-way interaction, it would appear that the locus stems from the experimental group's unexpected low performance on the first far transfer test (posttest 1) coupled with its relatively high performance on the second far transfer test (posttest 2).

There was a significant Treatment main effect (Multivariate $F= 8.1623$, $df= 1/3$, $p<0.0077$). Since the graphs in Figure 8 indicate the performance superiority of the experimental group and since the experimental group's overall mean per cent correct (66.2%) was greater than the control group's overall mean per cent correct (43.4%), the significant Treatment main effect favored the experimental group. Consequently, the null hypothesis associated with the Treatment main effect must be rejected in favor of the research hypothesis, H_1, which predicted a Treatment main effect showing higher performance for the experimental group.

Differences between the experimental and control groups with respect to seriation ability undoubtedly contributed to the Treatment main effect. However, since a three-way interaction (Treatment x Posttest x Test Type) was observed, the difference in seriation ability was not the only factor which was involved in producing the Treatment main effect; the Posttest and Test Type factors must have been involved as well. Seemingly, the locus of interaction which contributed to the Treatment main effect was the experimental group's second fair transfer mean. That mean was by far numerically greater than any other far transfer mean and set a pattern among the far transfer means which was unexpected and inconsistent with the pattern of near transfer means. Apparently, then, the experimental group's capabilities (Treatment factor) at the time of the second posttest (Posttest factor) interacted with the materials and tasks of the far transfer test (Test Type factor) to produce an elevated mean which contributed to the experimental group's superior overall performance as indicated by the Treatment main effect.

The repeated measures, multivariate analysis revealed a significant Test Type main effect (Multivariate $F = 16.1776$, $df = 1/30$, $p < 0.0004$). Since for both groups (experimental and control) the near transfer score was numerically greater than the far transfer score on each posttest and since the overall mean for the near transfer measures (63.1%) was numerically greater than the overall mean for the far transfer measures (45.1%), the direction of the observed Test Type main effect favored the near transfer test type. Thus, support was found for the research hypothesis H_2, which predicted a Test Type main effect favoring the near transfer test type.

Since there existed a three-way interaction, the data must be examined to determine whether or not that interaction contributed to the observed Test Type main effect. The experimental group's elevated second far transfer mean has been suggested as the locus of the three-way interaction. Since this elevated mean would contribute to a smaller rather than to a larger, difference between Test Type means, the three-way interaction likely played no part in producing the Test Type main effect.

Tests of those hypothesis which were of prime interest and which were the most soundly based have been described. The multivariate analysis tested the null hypotheses associated with four other sources of variation. Because these four remaining hypothesis tests are somewhat peripheral to the central focus of this study, the results of those tests will be described and discussed in brief form.

The multivariate test for a Posttest main effect was necessarily performed in two parts. The first part compared the mean associated with posttests 1 (49.1%) and 2 (59.2%) and revealed that the mean per cent correct on posttest 2 was significantly greater than the mean per cent correct on posttest 1 (Multivariate $F=4.4951$, $df=1/30$, $p < 0.0424$). The second part of the Posttest main effect test compared means for posttests 2 (59.2%) and 3 (53.9%) and revealed that the mean per cent correct on posttest 2 was significantly greater than the mean per cent on posttest 3 (Multivariate $F=14.0421$, $df=1/30$, $p < 0.0008$). The observed superiority of the overall mean associated with posttest 2 most likely was enhanced by the experimental group's elevated second far transfer mean, which has been suggested as the locus of the significant three-way interaction.

The Posttest x Test Type interaction was tested in two parts and neither test revealed significance (Multivariate $F=0.3834$, $df=1/30$, $p < 0.5410$; and Multivariate $F=0.0964$, $df=1/30$, $p < 0.7588$). Failure to index a significant Posttest x Test Type interaction implied that the difference between near and far transfer means (calculated across treatment groups) did not change significantly across the three posttests.

The multivariate test of the null hypothesis associated with the Treatment x Posttest interaction was performed in two parts. Since the analysis revealed no significance for either part (Multivariate $F=0.3963$, $df=1/30$, $p < 0.5338$; and Multivariate $F=2.0586$, $df=1/30$, $p < 0.1621$), the difference between the experimental and control groups' means (across test types) did not change significantly across the three posttests.

A significant Treatment x Test Type interaction did occur (Multivariate $F=6.0730$, $df=1/30$, $p < 0.0202$). On the near transfer measures (across posttests) the experimental group's mean (80.2%) was 32.2 mean percentage points greater than the control group's mean (48.0). On the far transfer measures (across posttests) the experimental group's mean (52.3%) was only 13.6 mean percentage points greater than the control group's mean (38.7%). This change (18.6 percentage points) in treatment group differences in going from near to far transfer measures probably accounts for the significant Treatment x Test Type interaction. The significant Treatment x Test Type interaction is interpreted to mean that the differences between the experimental and control groups' near transfer means was greater than the difference between the experimental and control groups' far transfer means.

The significant Treatment x Test Type interaction was not likely enhanced by the observed three-way interaction since the experimental group's elevated, second far transfer mean, thought to be the locus of the three-way interaction, contributed to widening the difference between the two groups' performances of the far transfer tasks. Had the experimental group's elevated, second far transfer mean not occurred, the differences between differences associated with the Treatment x Test Type interaction would have been even greater than they were.

No post hoc, multivariate procedures were available to examine the specific experimental-control differences which comprised the significant Treatment main effect. Consequently, the specific experimental-control group differences were analyzed through the use of two univariate, repeated measures analyses of variance for groups of unequal sizes (Winer, 1962, pp. 374-378). One univariate analysis was used with the near transfer data, and the other univariate analysis was used with the far transfer data. Of the sources of variation analyzed by the two univariate analyses, the Treatment main effect for near transfer data and the Treatment main effect for far transfer data were of prime interest since they were the sources associated with the following two research hypotheses developed in Chapter II:

H_3: The experimental group's mean for the near transfer measures will be greater than the control group's mean for the near transfer measures.

H_4: The experimental group's overall far transfer mean will be greater than the control group's overall far transfer mean.

Whereas the multivariate analysis could be used with unaltered percentage scores,[3] in the use of univariate, repeated measures analysis it is suggested that percentage scores be transformed according to an arcsin function (Winer, 1962, p. 221). The score which results from the arcsin transformation is an angle measured in radians.

The univariate repeated measures analyses were performed with both the percentage and the transformed data. With respect to the tests of significance, the analyses performed with percentage data were the same as those analyses performed with the transformed data. Since percentage scores are more meaningful than scores given in terms of radians, the analysis of percentage data is reported here.

---

[3]Dr. William Schmidt, Educational Psychology, Michigan State University, personal communication, February, 1971.

For the univariate repeated measures analysis of near transfer data there was a Treatment factor (A) with levels of experimental and control and a repeated measures or Posttest (B) factor consisting of near transfer measures from the three posttests. Table 1 shows the near transfer means and standard deviations. Table 4 summarizes the results of the univariate repeated measures analysis of near transfer data.

Table 4

Repeated Measures Analysis: Near Transfer (Percentages)

Source of Variation	SS	df	MS	F
Between Subjects				
Treatment (A)	24771.56	1	24771.56	12.99**
Subjects w. groups	57171.88	30	1905.73	
Within Subjects				
Posttests (B)	408.39	2	204.19	0.40
A X B	2101.57	2	1050.78	2.09
B x Subjects w. groups	30339.27	60	505.65	

**$p < .01$

The univariate, repeated measures AVOVA of the near transfer data revealed a significant Treatment main effect (A) ($F = 12.99$, $df = 1/30$, $p < 0.01$) but neither a significant Posttest effect (B) ($F < 1.00$) nor a significant Treatment x Posttest (A x B) interaction ($F = 2.09$, $df = 2/60$, $p < 0.25$). The Treatment main effect was interpreted to mean that the experimental group's overall near transfer mean (80.2%) was significantly greater than the control group's overall near transfer mean (48.0%). Consequently, support was found for research hypothesis H_3, which stated that the experimental group's overall near transfer mean would be greater than the control group's overall near transfer mean.

Since the experimental subjects received training on tasks and materials very similar to those used in the near transfer tests and since the time between the training and the testing was great enough to allow loss of the induced capabilities, the significant near transfer Treatment effect can be interpreted as evidence for the experimental group's retention of the specific capabilities acquired during the training. Furthermore, the lack of a Treatment x Posttest interaction would suggest that the differences between the performances of the experimental and control groups, and hence, the experimental group's superiority, remained relatively unchanged across posttests.

The univariate, repeated measures analysis of the far transfer data was performed the same way as the univariate analysis of the near transfer data. Table 2 shows the far transfer means and standard deviations. Table 5 summarizes the results of the analysis of far transfer data.

Table 5

Repeated Measures Analysis: Far Transfer (Percentages)

Source of Variation	SS	df	MS	F
Between Subjects				
Treatment (A)	4354.82	1	4354.82	2.3680
Subjects w. groups	55169.94	30	1838.99	
Within Subjects				
Tests (B)	3963.70	2	1981.85	3.9764**
A X B	2736.96	2	1368.48	2.7457
B x Subject w. groups	29903.59	60	498.39	

**p <.05

The results of the far transfer analysis revealed a Posttest main effect ($F = 3.9764$, $df = 2/60$, $p < 0.05$) but neither a Treatment main effect ($F = 2.3680$, $df = 1/30$,

$p < 0.25$) nor a Treatment x Posttest interaction ($F = 2.7457$, $df = 2/60$, $p < 0.10$). Since there was no Treatment main effect, no support was found for the research hypothesis H_4, which predicted that the experimental group's overall far transfer mean would be greater than the control group's overall far transfer mean. Failure to find a significant treatment group difference on the far transfer tests leads to the conclusion that the training procedures lacked the necessary aspects to insure transfer of training to unfamiliar materials.

Post hoc comparisons (Winer, 1962, pp. 377-378) associated with the observed far transfer Posttest main effect revealed that the far transfer mean (54.0%) for posttest 2 (across both treatment groups) was significantly greater than the far transfer mean (38.5%) for posttest 1 ($F = 7.7223$, $df = 1/60$, $p < .01$) but not significantly greater than the far transfer mean (43.9%) for posttest 3 ($F = 3.2555$, $df = 1/60$, $p < .10$). No difference was found between the means for posttests 1 and 3 ($F < 1.0$).

Experimental versus Control versus Special Control Group Analysis

For this part of the overall analysis, multivariate techniques were used to compare the performances of the experimental, control, and special control groups on the near, far, and far-far transfer measures of posttest 3.

It should be noted that the experimental and control groups' data for the near and far transfer tests of posttest 3 have been used in the previous analyses. The experimental and control group subjects were in seriation stage II (could order but not insert sticks) at the beginning of the experiment. The special control group subjects, on the other hand, were in seriation stage III (could both order and insert sticks) at the beginning of the experiment. Since the purpose of the training was to induce stage III capabilities in the experimental group and since the training was hypothesized to be successful, the performances of the experimental and special control groups should not differ significantly on each of the three measures of posttest 3 research (hypotheses H_5 (near, H_6 (far), and H_7 (far-far) developed in Chapter II). Furthermore, since the special control group already possessed stage III capabilities at the beginning of the experiment and since the control group was given no training to induce stage III capabilities, the special control group should outperform the control group on each of the three measures of posttest 3 (research hypotheses H_8 (near), H_9 (far, and H_{10} (far-far) developed in Chapter II).

The analysis was necessarily performed in two parts. One part compared the performances of the experimental and special control groups while the other part compared the performances of the special control and control groups. See Appendix N for a thorough display of information (sample correlation matrices, variances, hypothesis mean products matrices) relevant to the multivariate analysis reported here.

The results of the experimental versus special control group analysis revealed that the two groups did not differ in their performances of posttest 3 (Multivariate $F = 0.2795$, $df = 3/40$, $p < 0.8399$). With the step-down F's (Table 6) showing no group differences for each posttest measure, research hypotheses H_5 (near), H_6 (far), and H_7 (far-far), which predicted the absence of group differences on the three measures of posttest 3, were supported.

Table 6

Multivariate Analysis: Experimental(E) versus Special Control(SC) on the Three Measures of Posttest 3

Posttest 3 Measures	Means E(n=15)	SC(n=13)	df	Step-down F	p less than
Near Transfer	76.7%	88.5%	1,42	0.4670	0.4987
Far Transfer	46.7	46.2	1,42	0.0115	0.9151
Far-Far Transfer	30.0	51.9	1,42	0.3780	0.5422

A multivariate analysis revealed that the special control and control groups differed in their performances of posttest 3 (Multivariate $F = 5.6897$, $df = 3/40$, $p < 0.0025$). The step-down F ratios (Table 7) showed that significant differences existed between the two groups' performances on the near and far-far transfer measures but no difference existed between their performances on the far transfer measures. The means shown in Table 7 indicate that where differences did exist, the special control group outperformed the control group. Thus, although no support was found for research hypothesis H_9, which predicted that the special control group would outperform the control group on the far transfer measure of posttest 3, support was found for research hypotheses H_8 and H_{10}, which predicted, respectively, that the special control group would outperform the control group on the near (H_8) and far-far (H_{10}) transfer measures of posttest 3.

Table 7

Multivariate Analysis: Special Control(SC) versus Control(C) on the Three Measures of Posttest 3

Posttest 3 Measures	Means		df	Step-down F	p less than
	SC(n=13)	C(n=17)			
Near Transfer	88.5%	52.9%	1,42	6.2504	0.0165
Far Transfer	46.2	41.2	1,42	1.1109	0.2981
Far-Far Transfer	51.9	20.6	1,42	8.3658	0.0062

Correlational Data

A regression analysis (previously described in Chapter III, p.123) revealed that the four covariates (age, field independence (EFT), reflectivity (MFF), and impulse control (IC) were not significantly associated with the dependent variables (chi square for the test of the hypothesis of no association between dependent variables and covariables = 30.4855, df = 28, $p < 0.3404$). Nevertheless, some of the correlations between the covariate scores and the dependent variable scores lead to some interesting suggestions.

For example, the correlations between age and the various posttest scores for the experimental group (Table 8) tended to be negative and tended to increase in magnitude over time (from posttest 1 to posttest 3). This pattern of correlations did not occur for the control group (Table 9). Also, age and the performance of the other covariate measures (EFT, IC, and MFF) tended to be negatively correlated for the experimental group but not for the control group (note, a higher impulse control score (IC) indicated less impulse control). Although subjects were randomly assigned to the experimental and control groups, the correlations mentioned above give rise to the notion that the experimental group, unlike the control group, contained older subjects which were less bright (with respect to the measures taken) than the younger subjects.

Table 8

Correlations Between Covariables and Between Covariables and Dependent Variables for the Experimental Group

	Age	EFT	IC	MFF
NT1	-0.067	0.594	-0.180	0.424
FT1	-0.390	0.069	0.093	-0.161
NT2	-0.505	0.470	-0.007	0.605
FT2	-0.362	0.411	0.055	0.009
NT3	-0.588	0.410	0.177	0.206
FT3	-0.568	0.665	0.154	0.374
FFT3	-0.612	0.260	0.516	0.526
Age	1.000	-0.455	0.247	-0.371
EFT	-0.455	1.000	-0.160	0.510
IC	0.247	-0.160	1.000	0.260
MFF	-0.371	0.510	0.260	1.000

Legend: EFT = Embedded Figures Test (field independence), IC = Impulse Control, MFF = Matching Familiar Figures (reflectivity), NT = near transfer, FT = far transfer, FFT = far-far transfer. The number after the designation for transfer type refer to the posttests. For example, NT1 refers to the near transfer measure of posttest 1.

Table 9

Correlations Between Covariables and Between Covariables and Dependent Variables for the Control Group

	Age	EFT	IC	MFF
NT1	0.068	0.048	0.064	0.069
FT1	0.076	0.102	0.174	0.365
NT2	0.212	0.186	0.111	0.117
FT2	-0.219	-0.133	-0.142	0.052
NT3	0.336	0.163	0.050	-0.071
FT3	-0.564	0.010	-0.010	0.224
FFT3	0.057	0.006	-0.218	0.040
Age	1.000	0.142	0.195	0.417
EFT	0.142	1.000	-0.076	0.463
IC	0.195	-0.076	1.000	0.205
MFF	0.417	0.463	0.205	1.000

See Legend in Table 8.

Another suggestion arises from the correlations between the field independence scores (EFT) and the post-test scores. For the experimental group, field independence (EFT), compared to theother covariates, seemed to be the covariate most positively related to the performance of seriation tasks. Since the control group did not reflect a similar pattern of correlations, the experimental group's comparatively high positive correlations between field independence and seriation performances suggests that those experimental subjects with high field independence scores might have benefited most from the training.

The "suggestions" gleaned from the correlational data should in no way be taken as solidly based observations. They are of some value, however, in considering relevant experimental manipulations for future experiments. From these suggestions, future studies might reasonably include blocking variables of age, mental alertness, and field independence.

Summary

For the Experimental versus Control Group Analysis a repeated measures, multivariate analysis of variance was used to analyze the experimental and control subjects' near and far transfer scores from each of the three posttests. From that analysis the following observations were made:

1) There existed a significant Treatment x Posttest x Test Type interaction. The locus of that interaction was thought to be the experimental group's elevated second far transfer mean.

2) There was a Treatment main effect and the direction of the effect favored the experimental group, as expected. Hence, support was found for research hypothesis H_1.

3) A Test Type main effect was observed and the direction of this effect, as predicted, favored the near transfer test type. Hence, support was found for research hypothesis H_2.

4) A significant Posttest main effect test revealed that on the average more correct responses occurred on posttest 2 than on either posttest 1 or posttest 3.

5) No significant Posttest x Test Type interaction was observed. This result was interpreted to mean that the difference between near and far transfer means (calculated across treatment groups) did not change significantly from posttest to posttest.

6) There existed no Treatment x Posttest interaction. Hence, the differences between experimental and control groups' means

(calculated across test types) did not change from posttest to posttest.

7) A significant Treatment x Test Type interaction revealed that the difference between the experimental and control groups' near transfer means (calculated across posttests) was greater than the difference between the experimental and control groups' far transfer means (calculated across posttests). This result was taken to suggest that treatment group differences on the near transfer tasks contributed more to the Treatment main effect than did the treatment group differences on the far transfer tasks.

Two univariate repeated measures analyses were used to examine the specific experimental-control group differences which comprised the overall difference indicated by the Treatment main effect. One repeated measures analysis was used to examine the experimental and control subjects' near transfer scores from the three posttests. The other repeated measures analysis was used to examine the experimental and control subjects' far transfer scores from the three posttests. The results from the two analyses are summarized as follows:

1) Repeated measures -- near transfer data.

There was a simple Treatment main effect and the direction of that effect was in the predicted direction (Experimental > Control). Therefore, since the experimental group's overall near transfer mean was greater than the control group's overall near transfer mean, support was found for the research hypothesis H_3. There was neither a Posttest main effect nor a Treatment x Posttest interaction. The failure to index a Treatment x Posttest interaction was interpreted to mean that the experimental group's performance superiority on near transfer tasks remained relatively unchanged across posttests.

2) Repeated measures -- far transfer data.

No Treatment main effect was observed for the far transfer data. Consequently, no support was found for the research hypothesis H_4. The absence of a Treatment main effect was interpreted to mean that the procedures used to train the experimental subjects lacked the necessary aspects to insure transfer of training to unfamiliar materials. The repeated measures analysis of far transfer data further revealed a Posttest main effect but no Treatment x Posttest interaction.

For the Experimental versus Control versus Special Control Group Analysis, a multivariate analysis of variance was used to compare the performances of the experimental, control, and special control groups on the near, far, and far-far transfer measures of posttest 3. From that analysis the following observations were made:

1) Since for each measure of posttest 3 there was no difference between the performances of the experimental and special control groups, support was found for research hypotheses H_5, H_6 and H_7. Thus, approximately 132 days after training, the experimental group subjects, who began the study with seriation stage II capabilities, performed seriation tasks just as well as the special control group subjects, who began the study with seriation stage III capabilities.

2) A comparison of the special control and control groups' performances of posttest 3 measures showed that the special control group outperformed the control group on the near and far-far transfer measures of posttest 3 but not on

the far transfer measures of posttest 3.

Thus, support was found for research hypotheses H_8 and H_{10} but not for research hypothesis H_9.

A cursory survey of some correlational data revealed that, for the experimental group, age was negatively correlated with both seriation performance and with the other individual difference factors. These negative correlations gave rise to the idea that the experimental group contained older subjects which were less bright (with respect to the measures taken) than the younger subjects.

To summarize, the experimental subjects, for the most part, acquired and retained the specific target capabilities of the training. They, however, did not acquire the ability to transfer those acquired capabilities to seriation tasks requiring the use of unfamiliar materials.

CHAPTER V

DISCUSSION AND SUMMARY

This chapter includes: 1) a brief summary of the study, 2) a discussion of the transfer data and the two hypotheses generated to explain that data, 3) a discussion of the relationship between the results of this seriation study and the results of other Piagetian training studies (conservation and seriation), and 4) a concluding statement. Implications for future research are described as they arise out of the discussions.

Fifteen kindergarten children, who began the study with stage II seriation capabilities (i.e., could serial order sticks but could not insert a number of sticks into an already ordered set), were individually given 30 minutes of seriation training on three consecutive days. Cue fading and reinforcement were used in the training to help the children meet the successive performance criteria leading to the acquisition of stage III capabilities (i.e., both order objects and insert objects into an already ordered set). Posttests, each consisting of a retention and transfer measure, were given approximately one, eight, and 132 days after training. In general, the results revealed that the subjects acquired and retained the specific target capabilities of the training, but failed to substantially transfer those acquired capabilities to the performance of seriation tasks involving unfamiliar materials.

Although the analysis of far transfer data did not reveal the expected substantial transfer effects, there are some interesting transfer data which should be discussed. Since the ability to perform a task often tends to gradually decline in time from the end of instruction, the slight gradual decrease in the experimental group's performance of near transfer tests was not unexpected. In a similar manner, the experimental group's performance of the far transfer tests might be expected to follow a similar decline, thus showing a decrease from maximum transfer on the first posttest to minimum transfer on the third posttest. However, contrary to expectation the experimental group showed a relatively low performance (mean = 41.7%) on the first far transfer test and a relatively high performance (mean = 68.3%) on the second far transfer test. The low third far transfer mean (46.7%) for the experimental group, although not desired, was probably not an unreasonable outcome.

Why did the experimental group's unexpected first and second far transfer performances occur? Before considering some experimentally based reasons for the unexpected transfer performances, possible extra-experimental influences will be discussed. As previously mentioned, in carrying out the experiment there were two groups, groups I and II, each with an experimental and control subgroup. The experimental subjects of group I acquired far transfer mean percentages of 35.7% and 60.7% for the first and second posttests, respectively. The experimental subjects of group II acquired the corresponding far transfer mean percentages of 46.9% and 75.0%. Hence, both groups of experimental subjects showed low performance on the first far transfer test and high performance on the second far transfer test. Since these tests were administered to the two experimental groups at different times, the unexpected far transfer means were not likely to have been associated with any specific facilitative or inhibitive extra-experimental events which might have occurred on the days of testing. Furthermore, if extra-experimental events were the basis for the unexpected far transfer means, those events would likely be reflected in the corresponding near transfer means. However, since, for each group, the near and far transfer means followed different patterns, extra-experimental events were probably not the cause for the unexpected nature of the experimental groups' far transfer effects.

Eleven of the 15 experimental subjects performed better on the second far transfer test than they did on the first. The performances of the remaining four subjects were the same on both tests. Thus, the experimental group's unexpected performance on the first two far transfer tests does not appear to be attributable to the spurious performances of a few individuals.

Since the data does not seem to support extra-experimental events and spurious performances as causitive factors underlying the experimental group's unexpected far transfer means, there is more reason to believe that the nature of those means might be explained in terms of task related, experimental phenomena. For example, it could be speculated that the experimental subjects' low performances on the first far transfer test reflected, in part, a spread of attention to the "novel" irrelevant stimuli rather than a focusing of attention on the relevant stimuli. With the "novelty" of the far transfer materials having worn off from the first posttest, the experimental subjects could have then attended more to the task relevant stimuli and used the capabilities acquired in training to foster the production of the elevated second far transfer mean.

This "novelty" hypothesis is not totally unfounded since Gelman (1969) has provided evidence indicating that children sometimes attend to task irrelevant stimuli and as a consequence perform the tasks poorly even though they are capable of performing the tasks. The "novelty" hypothesis, however, should still be considered with caution since inherent within it is the precarious assumption that the experimental group actually possessed the ability to perform the far transfer tasks of posttest 1.

Why would the "novelty" effect foster the production of such a sizable difference between the experimental group's near (83.9%) and far (41.7%) transfer means of posttest 1 and fail to foster such a difference for the control group (near=38.7%, far=35.3%)? It could be argued that "novelty" played no part in producing the control group's difference since the control subjects, for the most part, had been exposed to neither set of materials and consequently had no opportunity for developing differential perceptions of novelty with respect to those materials. The experimental group, on the other hand, had been exposed to the near transfer materials during training but had not been exposed to the far transfer materials. Therefore, by comparison to the near transfer materials the far transfer materials most likely seemed "novel" to the experimental subjects, thus distracting them from the tasks they were capable of performing.

Since the same tasks and materials were used on the first and second far transfer tests, test-retest learning, in addition to the attentional factors, could have contributed to the experimental group's relatively high second far transfer mean. From their experience with the first set of far transfer tasks, the experimental subjects, aided by the capabilities acquired in training, could have learned: 1) to more clearly distinguish the relevant from the irrelevant stimuli, 2) to compare the objects so that meaningful decisions about the order of objects could be made, 3) to identify and correct errors, and, in general, 4) to apply the newly acquired seriation capabilities to the unfamiliar set of far transfer materials. If indeed this learning did occur, it could very well be the basis for the improved performance of the second far transfer test.

It should be stressed that the test-retest learning effect was probably not as pronounced for the control subjects since they lacked the underlying seriation capabilities acquired by the experimental subjects in training and therefore could not learn as well from the testing experiences. Unlike the "novelty" hypothesis, the test-retest hypothesis does not rest upon the assumption that the experimental subjects possessed transfer capabilities at the time of the first posttest.

The "novelty" and the test-retest hypotheses have both implied that the transfer effects were genuine. To the contrary, it might be suggested that even the observed treatment group difference on the second far transfer test was artificial and, hence, indicated no genuine underlying transfer capability. At this time, however, no substantial reasons can be contrived for explaining why the effect occurred without at least some transfer capability being present.

Both the "novelty" and the test-retest hypotheses suggest interesting ideas concerning transfer. The experimental group did outperform the control group on the second far transfer test. The "novelty" hypothesis suggests that the experimental group had the transfer capability on the first far transfer test but failed to reveal the capability. Hence, the suggestion of the "novelty" hypothesis and the observation of the second far transfer performances together imply that the experimental subjects did possess the ability to transfer at least up to the time of the second posttest. If indeed the experimental subjects did possess the transfer capabilities up to the time of the second posttest but not up to the time of the third posttest, then the results of the seriation study would suggest that the training successfully induced the transfer capabilities but did not insure the long-term retention of those capabilities. From the suggestions initiated by the "novelty" hypothesis, it would seem that future research might very well be directed toward improving the seriation training techniques in such a way that children learn to cope with "novel" stimuli and at the same time retain the transfer capabilities for longer periods of time.

According to the test-retest hypothesis, the experimental subjects, although unable to perform the far transfer tasks of posttest 1, learned (as a consequence of the training) from their experience with posttest 1 tasks and thus produced an elevated far transfer performance on posttest 2. Since transfer, in part, refers to the ability to use acquired capabilities to learn from new experiences, the test-retest hypothesis suggests that transfer was taking place during the performance of the first posttest. It is important to note that the test-retest hypothesis not only suggests that the acquisition of far transfer capabilities took place but also that the experimental subjects retained the acquired transfer capabilities over the seven day interval between posttests 1 and 2. If, indeed, the test-retest hypothesis has a reasonable foundation, then the results of this seriation study might suggest that once a specific seriation capability has been induced through somewhat intense training, the difficulty of learning to perform seriation tasks with different materials will likely be reduced considerably. Should the test-retest hypothesis find support in the future, this suggestion arising from the seriation study might lead to a worthwhile training strategy. That training strategy would probably consist of two phases. In the first (or induction) phase, specific seriation capabilities would be induced through the use of a carefully designed, somewhat intense, short-term training sequence, much like the sequence used in this study. In the second (or generalization) phase, opportunity for using the induced capabilities with many different kinds of materials would be periodically given.

It should be made clear that the "novelty" and the test-retest hypotheses and the implications drawn from them are speculations. To test the hypotheses and bring the implications out of speculation this seriation experiment might reasonably be replicated with the inclusion of two additional groups of trained subjects. All three groups would receive the same seriation training and all three groups would be given the same far transfer test eight days after training. The groups would differ in the following ways (see Figure 9): 1) Group 1 would receive a far transfer test one day after training. This test would be identical to the one given eight days after training, 2) One day after training, group 2 would perform tasks with the materials used in the far transfer test.

The tasks, however, would not be seriation tasks, 3) Group 3 would be given only the far transfer test eight days after training.

TRAINED GROUPS	INTERIM TRAINING	FAR TRANSFER TEST(0)
(All groups receive the same training)	(Given one day after training)	(Given 8 days after training)
Trained Group 1	0 (far transfer test)	0
Trained Group 2	Nonseriation tasks performed with far transfer materials	0
Trained Group 3	No treatment	0

Fig. 9. An illustration of the basic experimental design for testing the "novelty" and test-retest learning hypotheses.

The groups' performances of the far transfer test given eight days after training should yield information concerning the "novelty" and the test-retest learning hypotheses. For example, if group 1 does better than group 2, there is evidence for the test-retest learning hypothesis since a differential novelty effect is eliminated by having both groups interact with the materials. If groups 1 and 2 do better than group 3 and groups 1 and 2 perform no differently, there is evidence for the "novelty" hypothesis since involvement with materials (both groups 1 and 2) facilitated subsequent performance but prior performance of the actual tasks (group 1) produced no advantage indicative of learning. Of course, if groups 1 and 2 perform no differently than group 3, neither of the two hypotheses is supported.

Both the "novelty" hypothesis and the test-retest hypothesis suggest that transfer did occur. Neither hypothesis suggests, however, that transfer effects were substantial. According to the "novelty" hypothesis, the transfer capability was present at the time of the first posttest but apparently was not strong enough to overcome the distraction of the "novel" task irrelevant stimuli. According to the test-retest hypothesis, at the time of the first posttest there was only the ability to learn from the transfer tasks, no ability to perform them. Therefore, regardless of which hypothesis finds support (possibly neither will), the results of this seriation study imply that future research needs to be directed toward finding ways of improving the acquisition and retention of seriation stage III transfer capabilities. One possible approach to improvement is the previously described two phase strategy in which an intense training session (induction phase) with only a few kinds of materials is followed by a number of occasional, less intense, training sessions (generalization phase) in which many different materials are used. Another approach, of course, would be to include many different kinds of materials and tasks in the intensive training session and possibly minimize or eliminate the occasional, less intense follow-up sessions. Research might very well be directed toward investigating the relative effectiveness of these two approaches for improving transfer.

In reviewing the conservation training studies the criterion for success was statistical significance in which the trained group showed performance superiority over the untrained or control group. If that same criterion is used with the seriation training study reported here, then the seriation training would be judged a success. The question then becomes: How do the results of the thesis compare with the results of the successful conservation studies (length and number)?

All of the successful conservation studies reviewed in Chapter II included retention and transfer measures. Retention of the specific induced capabilities was found in all of those studies. The maximum retention interval used in the conservation studies ranged from approximately three weeks (Gelman, 1969; Wallach and Sprott, 1964; and Wallach, Wall, and Anderson, 1967) to 16 weeks (Kingsley and Hall, 1967). The maximum retention interval used in this seriation study was approximately 19 weeks. Since retention of induced capabilities was found in the seriation study, the conservation studies and the seriation study seemingly correspond with respect to retention. All but one of the successful conservation studies (Wallach, Wall, and Anderson, 1967) revealed transfer of training. In the seriation study, however, there was only a hint of transfer on the second posttest and absolutely no transfer on the first and third posttests. Thus, with respect to transfer, conservation training was superior to the seriation training.

The difference between the success of the seriation training and the success of the conservation training should be put in proper perspective. The success of the conservation studies undoubtedly rests, in part, upon the groundwork laid down by the many unsuccessful attempts to induce sustained conservation capabilities. The seriation study, on the other hand, without the aid of previous seriation training research, laid down some groundwork for future seriation training and at the same time was, in part, successful.

Coxford's (1964) study, which was the only seriation training study reported in the literature review, revealed that it was possible to induce conservation of serial and ordinal correspondence in children who, before training, could construct serial correspondence. Similarly, the results of this thesis revealed that it was possible to induce specific seriation stage III capabilities (order and insert) in children who, before training, exhibited seriation stage II capabilities (order but not insert). The two seriation training studies taken together agree in support of the hypothesis that the ability to perform seriation tasks can be changed, at least to some extent, by relatively short periods of specific training. Although the evidence from the seriation training studies hardly refutes Piaget's notion that cognitive capabilities cannot be substantially induced by specific training, the evidence does cast some doubt on that notion and hence, suggests that it might be worthwhile to look further for methods that produce substantial changes in cognitive capabilities associated with seriation.

This thesis went beyond Coxford's study by showing that the specific target capabilities, in addition to being induceable, were durable but not, in general, transferable. Whereas the results of the thesis point to the specific need for improving instruction relative to transfer, the results of Coxford's study give no specific directions for improving instruction.

The results of this study revealed that the training sequence, which involved cue fading and reinforcement, produced the induction of specific stage III seriation capabilities in kindergarten children who, prior to training, possessed stage II capabilities. The results further revealed that the acquired capabilities were rather durable, but not substantially transferable to new materials. It should be pointed out, however, that although the transfer effects were not substantial, the trained subjects (experimental) performed the far and far-far transfer tests of posttest 3 no differently than those subjects (special control) who had acquired stage III capabilities "naturally" or without specific instruction. Since a reasonable degree of success has been found, there may be the tendency to draw implications from the study which concern seriation training in the classroom. Without alterations, however, the training sequence used in this study would require such large amounts of time, space, and material that its use in the classroom would be prohibitive. In addition, any attempt to merely shorten the training sequence for classroom use would likely lead to less success than that found in this study. Although this study does not seem to prescribe any specific, reliable, seriation training techniques for immediate use in the classroom, it does offer a method of cueing and cue fading which might be subsequently used in developing the more practical, classroom-like, training techniques. In fact cueing and cue fading methods might be reasonably used in attempts to improve the seriation training offered in some of the elementary science and mathematics programs. Since in this study, less than outstanding training effects were produced with a somewhat extensive and carefully designed training sequence, those elementary science and mathematics lessons which give children only brief opportunities for practicing seriation are not likely producing significant changes in seriation capabilities and hence should probably be revised.

Hunt (1961) has contended that it is now reasonable to believe that early childhood experiences can be governed to produce increased rates of intellectual development as well as expanded adult levels of intellectual capacity. Bloom (1964) has suggested that adult intellectual characteristics might be most easily influenced during childhood when those characteristics are developing most rapidly. Seriation capabilities likely play important roles in school learning and intellectual functioning. For example, concepts of seriation are necessarily involved in the concept of number. In addition, serial ordering is used in science to organize information and hence facilitate discovery. Undoubtedly, seriation is inherent in logical processes (if...then) and is an important element in the information processing system of the mind. From the propositions of Hunt and Bloom and from the contention that seriation plays vital roles in intellectual functioning and school learning, it seems reasonable to undertake a thorough, systematic search for those seriation training methods which produce significant changes in intellectual development and school learning. The seriation training study reported in this thesis was an attempt to initiate that search.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adams, J. A. Multiple versus single problem training in human problem solving. Journal of Experimental Psychology, 1954, 48, 15-19.

Anderson, R. C. Educational psychology. Annual Review of Psychology, 1967, 18, 129-164.

Banta, T. J. Tests for the evaluation of early childhood education: The Cincinnati autonomy test battery. Cognitive Studies, Vol. 1, 1968.

Bijou, S. W. Systematic instruction in the attainment of right-left form concepts in young retarded children. In J. G. Holland and B. F. Skinner (Eds.), An analysis of the behavioral processes involved in self-instruction with teaching machines. (Final Report, USOE, NDEA Title VII, Project No. 191) Washington, D.C.: United States Government Printing Office, 1965.

Bloom, B. S. Stability and change in human characteristics. New York: Wiley, 1964.

Bock, R. D. Multivariate analysis of variance of repeated measurements. In C. W. Harris (Ed.), Problems in measuring change. Madison, Wisconsin: The University of Wisconsin Press, 1963.

Bourne, L. E., Jr., & Pendleton, R. B. Concept identification as a function of completeness and probability of information feedback. Journal of Experimental Psychology, 1958, 56, 413-420.

Brainerd, C. J. & Allen, T. W. Experimental inductions of the conservation of "first-order" quantitative invariants. Psychological Bulletin, 1971, 75, 128-144.

Bugelski, B. R. & Cadwallader, T.C. A reappraisal of the transfer and retroaction surface. Journal of Experimental Psychology, 1956, 52, 360-366.

Callantine, M. R. & Warren, J. M. Learning sets in human concept formation. Psychological Reports, 1955, 1, 363-367.

Campbell, D. T. & Stanley, J. C. Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally, 1963.

Coxford, A. F. The effects of instruction on the stage placement of children in Piaget's seriation experiments. The Arithmetic Teacher, 1964, 11, 4-9.

Dallett, K. M. The role of response similarity in proactive inhibition. Journal of Experimental Psychology, 1962, 64, 364-372.

Desse, J. & Hulse, S. H. The psychology of learning. (3rd ed.) New York: McGraw-Hill, 1967.

Elkind, D. Discrimination, seriation, and numeration of size and dimensional differences in young children: Piaget replication study VI. The Journal of Genetic Psychology, 1964, 104, 275-296.

Finn, J. D. Univariate and multivariate analysis of variance and co-variance: A Fortran IV program. Computing Center of the State University of New York at Buffalo, 1968.

Flavel, J. H. The developmental psychology of Jean Piaget. Princeton, New Jersey: Van Nostrand, 1963.

Gagné, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, 1965.

Gagné, R. M. & Rohwer, W. D., Jr. Instructional psychology. Annual Review of Psychology, 1969, 20, 381-418.

Gelman, R. Conservation acquisition: A problem of learning to attend to relevant attributes. Journal of Experimental Child Psychology, 1969, 7, 167-187.

Glaser, R. Some implications of previous work on learning and individual differences. In R. M. Gagné (Ed.), Learning and individual differences. Columbus, Ohio: Charles E. Merrill Books, 1967.

Goldschmidt, M. L. The role of experience in the acquisition of conservation, Proceedings, 76th Annual Convention of the American Psychological Association, 1968, 361-362.

Grant, D. A. & Berg, E. A. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a weigh-type card sorting problem. Journal of Experimental Psychology, 1948, 38, 404-411.

Grant, D. A. & Cost, J. R. Continuities and discontinuities in conceptual behavior in a card sorting problem. Journal of General Psychology, 1954, 50, 237-244.

Harris, F. R., Johnston, M. K., Kelley, C. S., Wolf, M. M. Effects of positive social reinforcements on regressed crawling of a nursery school child. Journal of Educational Psychology, 1964, 55, 35-41.

Harris, F. R., Wolf, M. M., and Baer, D. M. Effects of adult social reinforcement on child behavior. Young Children, 1964, 20, 8-17.

Hively, W. Programming stimuli in matching to sample. Journal of the Experimental Analysis of Behavior, 1962, 5, 279-298.

Hively, W. Parametric experiments with a matching-to-sample program. In J. G. Holland & B. F. Skinner (Eds.), An analysis of the behavioral processes involved in self-instruction with teaching machines. (Final Report, USOE, NDEA Title VII, Project No. 191) Washington, D.C.: United States Government Printing Office, 1965.

Jenkins, W. O., McFann, H., & Clayton, F. L. A methodological study of extinction following a periodic and continuous reinforcement. Journal of Comparative and Physiological Psychology, 1950, 43, 155-167.

Kagan, J. Reflectivity-impulsivity and reading ability in primary grade children. Child Development, 1965, 36, 609-628.

Kagan, J. & Moss, H. A. Psychological significance of styles of conceptualization. Monographs of the Society for Research in Child Development, John C. Wright, ed., 86, 1963.

Kendler, T. S. & Kendler, H. H. Reversal and nonreversal shifts in kindergarten children. Journal of the Experimental Analysis of Behavior, 1959, 58, 56-80.

Kendler, T. S., Kendler, H. H., & Wells, D. Reversal and nonreversal shifts in nursery school children. Journal of Comparative and Physiological Psychology, 1960, 53, 80-87.

Kingsley, R. C. & Hall, V. C. Training conservation through the use of learning sets. Child Development, 1967, 38, 1111-1126.

Kohlberg, L. Early education: A cognitive-developmental view. Child Development, 1968, 39, 1013-1062.

Lintz, L. M., & Brackhill, Y. Effects of reinforcement delay during learning on the retention of verbal material in adults. Journal of Experimental Psychology, 1966, 71, 194-199.

Mermelstein, E., and Meyer, E. Conservation training techniques and their effects on different populations. Child Development, 1969, 40, 471-490.

Moore, R. & Goldiamend, I. Errorless establishment of visual discrimination using fading procedures. Journal of the Experimental Analysis of Behavior, 1964, 7, 269-272.

Morrisett, L. & Hovland, C. I. A comparison of three varieties of training in human problem solving. Journal of Experimental Psychology, 1959, 58, 52-55.

Olton, R. M. The effect of a mnemonic upon the retention of paired-associate verbal material. Journal of Verbal Learning and Verbal Behavior, 1969, 8, 43-48.

Osgood, C. E. The similarity paradox in human learning. Psychological Review, 1949, 56, 132-143.

Parker, R. K. & Nunnally, J. C. Association of neutral objects with rewards: Effects of reward schedules on reward expectancy, verbal evaluation, and selective attention. Journal of Experimental Child Psychology, 1966, 3, 324-332.

Piaget, J. The psychology of intelligence. London: Routledge, 1950.

Piaget, J. The child's conception of number. New York: Humanities Press, 1952.

Piaget, J. Development and learning. Journal of Research in Science Teaching, 1964, 2, 176-185.

Piaget, J. Three lectures. In R. E. Ripple and V. M. Rockcastle (Eds.), Piaget rediscovered. Ithaca, New York: Cornell University Press, 1964.

Piaget, J. The child's conception of number. New York: W. W. Norton, 1965. (Original English Publication: Humanities Press, 1952.)

Prentice, M. C. The effects of variation in task difficulty, materials, and instructions on children's capacity to seriate. Unpublished doctoral dissertation. Cornell University, 1963.

Rothkopf, E. Z. & Coke, E. U. Variations in phrasing, repetition intervals, and the recall of sentence material. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 86-91.

Sassenrath, J. M. & Yonge, G. D. Delayed information feedback, feedback cues, retention set, and delayed retention. Journal of Educational Psychology, 1968, 59, 60-73.

Schafer, L. E. The effects of various materials on the young child's serial ordering capabilities. Unpublished study. Michigan State University, 1969.

Shantz, C. U. The effects of redundant and irrelevant information on children's seriation ability. Journal of Experimental Child Psychology, 1967, 5 208-222.

Sidman, M. & Stoddard, D. The effectiveness of fading in programming a simultaneous form discrimination for retarded children. Journal of the Experimental Analysis of Behavior, 1967, 10, 1-19.

Sigel, I. E. & Hooper, F. H. (Eds.). Logical thinking in children: Research based on Piaget's theory. New York: Holt, Rinehart and Winston, 1968.

Skinner, B. F. Learning theory and future research. In J. P. Lysaught (Ed.), Programmed learning: Evolving principles and industrial applications. Ann Arbor, Michigan: Foundation for Research on Human Behavior, 1961.

Smedslund, J. Patterns of experience and the acquisition of conservation of length. Scandanavian Journal of Psychology, 1963, 4, 257-264.

Staats, A. W., Staats, C. K., Schutz, R. E., & Wolf, M. The conditioning of textual responses using "extrinsic" reinforcers. Journal of the Experimental Analysis of Behavior, 1962, 5, 33-40.

Staats, A. W., Minke, K. A., Finley, J. R., Wolf, M., & Brooks, L. O. A reinforcer system and experimental procedure for the laboratory study of reading acquisition. Child Development, 1964a, 35, 209-231.

Staats, A. W., Finley, J. R., Minke, K. A., & Wolf, M. Reinforcement variables in the control of unit reading responses. Journal of the Experimental Analysis of Behavior, 1964b, 7, 139-149.

Terrace, H. S. Discrimination learning with and without errors. Journal of the Experimental Analysis of Behavior, 1963a, 6, 1-27.

Terrace, H. S. Errorless transfer to a discrimination across two criteria. Journal of the Experimental Analysis of Behavior, 1963b, 6, 223-232.

Travers, R. M. W., Van Wagenen, R. K., Haygood, D. H., & McCormick, M. Learning as a consequence of the learners task involvement under different conditions of feedback. Journal of Educational Psychology, 1964, 55, 167-173.

Underwood, B. J. Speed of learning and amount retained: A consideration of methodology. Psychological Bulletin, 1954, 51, 276-282.

Underwood, B. J. & Richardson, J. Studies of distributed practice XVII: The influence of meaningfulness and intra-list similarity of serial nonsense lists. Journal of Experimental Psychology, 1958, 56, 213-218.

Underwood, B. J. & Richardson, J. The influence of meaningfulness, intra-list similarity and serial position on retention. Journal of Experimental Psychology, 1956, 52, 119-126.

Underwood, B. J. & Keppel, G. The effect of degree of learning and letter-sequence interference on retention. Psychological Monographs, 1963, 77, 16-20.

Underwood, B. J. Degree of learning and the measurement of forgetting. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 112-129.

Wallach, L. & Sprott, R. Inducing number conservation in children. Child Development, 1964, 35, 1057-1071.

Wallach, L., Wall, J., & Anderson, L. Number conservation: The roles of reversibility, addition-subtraction, and misleading perceptual cues. Child Development, 1967, 38, 425-442.

Whitlock, C. & Bushell, D. Some effects of "back-up" reinforcers on reading behavior. Journal of Experimental Child Psychology, 1967, 5, 50-57.

Wimer, R. Osgood's transfer surface: Extension and test. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 274-279.

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

Witkin, H. A. Individual differences in ease of perception of embedded figures. Journal of Personality, 1950, 19, 1-15.

Wohlwill, J. F., & Lowe, R. C. An experimental analysis of the development of the conservation of number. Child Development, 1962, 33, 153-167.

APPENDICES

APPENDIX A

Description of Pretest Materials

Only sticks were used in the pretest. Each stick had a 0.75 x 0.75 inch, square cross-section. The sticks used in each task are identified in Table A by the lengths in inches.

TABLE A

Pretest Materials

Kind of Task	Objects to be Inserted	Objects to be Ordered or Ordered set of Insertion Task
Ordering 4 Objects	-	2, 3, 4, 5
Inserting 3 into 4	3.5, 4.5, 5.5	2, 3, 4, 5
Ordering 6	-	2, 3, 4, 5, 6, 7
Inserting 3 into 6	1.5, 3.5, 6.5	2, 3, 4, 5, 6, 7
Ordering 8	-	2, 3, 4, 5, 6, 7, 8, 9
Inserting 3 into 8	2.5, 5.5, 8.5	2, 3, 4, 5, 6, 7, 8, 9

APPENDIX B

Verbal Instructions for Seriation Pretest

<u>Illustrations of Ordering and Inserting for Pretest</u>

"Here are some sticks that have different lengths. We want to put these sticks side by side on this line so they go from the shortest all the way to the tallest to make stairs. Watch - (Experimenter orders sticks by successively choosing the shortest stick from those sticks to be considered.) See how the sticks go from the shortest one <u>here</u> all the way to the tallest <u>here</u>. Here is a stick that was forgotten. We want to put this stick with the others so all of the sticks make stairs. The forgotten stick would go between these two sticks. (Experimenter places the forgotten stick into the ordered set.) See how all of the sticks go from the shortest to the tallest to make stairs." (Experimenter removes sticks used in example.)

<u>Instructions Given During Pretest</u>

The instructions were the same for each pair of ordering and inserting tasks. "Here are some other sticks. (Sticks are disarranged before the subject.) Place these sticks on the line from the shortest to the tallest to make stairs. (The subject makes an attempt and the instructions are continued.) Look at your sticks to see if you have made any mistakes. You may make changes if you wish. (After the child indicates he is satisfied with the sticks as they are, or after he makes changes, the performance is judged correct or incorrect. If the set of sticks are not properly ordered, the experimenter arranges them in the proper order for the insertion task which follows. The sticks to be inserted are brought out, and the instructions are continued.) Here are some sticks that have been forgotten. You put these sticks along with the others so all of the sticks go from the shortest to the tallest to make stairs. (The subject makes an attempt and when he is finished the instructions continue.) Look at the sticks to see if any mistakes have been made. You can make changes if you find mistakes." (After any corrections have been made, the result is judged correct or incorrect, the sticks are removed, and a new set of sticks is presented for the next ordering task.)

APPENDIX C

Description of Materials Used During the First Training Session

Sticks of various lengths were used during the first training session. Each stick had a 0.75 x 0.75 inch, square cross-section. The sticks used at each training station are identified in Table C by the lengths in terms of 0.75 inch units. The discrimination levels are indicated by the difference between the lengths of the two ordered sticks (1 unit = 0.75 inch).

TABLE C

Materials Used During the First Training Session

Discrimination Level	Training Station	Sticks in the Ordered Set	Sticks in the Disarranged Set	
			Sticks to be Inserted	Other Sticks
	1	1.5, 4.5	3.0	6.0
3	2	6.0, 9.0	7.5	4.5
	3	10.5, 13.5	12.0	9.0
	4	4.00, 6.50	5.25	7.75
2.5	5	7.75, 10.25	9.00	6.50
	6	11.50, 14.00	12.75	15.25
	7	1.5, 3.5	2.5	4.5
2.0	8	6.5, 8.5	7.5	5.5
	9	9.5, 11.5	10.5	8.5

Discrimination Level	Training Station	Sticks in the Ordered Set	Sticks in the Disarranged Set	
			Sticks to be Inserted	Other Sticks
	10	3.00, 4.50	3.75	2.25
1.5	11	5.25, 6.75	6.00	7.50
	12	9.00, 10.5	9.75	11.25
	13	1.5, 4.5	3.0	6.0, 9.0
3	14	4.5, 7.5	6.0	3.0, 9.0
	15	7.5, 10.5	9.0	12.0, 15.0
	16	2.75, 5.25	4.00	1.50, 6.50
2.5	17	7.75, 10.25	9.00	4.00, 6.50
	18	9.00, 11.50	10.25	12.75, 15.25
	19	4.5, 6.5	5.5	1.5, 3.5
2.0	20	6.5, 8.5	7.5	5.5, 9.5
	21	10.5, 12.5	11.5	9.5, 13.5
	22	3.75, 5.25	4.50	1.5, 3.00
1.5	23	6.00, 7.50	6.75	8.25, 9.75
	24	12.00, 13.50	12.75	9.75, 11.25
	25	1.5, 2.5	2.0	3.0, 4.0
1.0	26	5.0, 6.0	5.5	3.5, 4.5
	27	11.0, 12.0	11.5	9.5, 10.5
	28	3.0, 6.0	4.5	1.5, 7.5, 10.5
3	29	4.5, 7.5	6.0	9.0, 12.0, 15.0
	30	10.5, 13.5	12.0	9.0, 6.0, 15.5
	31	7.75, 10.25	9.00	6.50, 4.0, 1.5
2.5	32	1.50, 4.00	2.75	5.25, 7.75, 10.25
	33	9.00, 11.50	10.25	5.25, 8.00, 12.75
	34	2.5, 4.5	3.5	1.5, 5.5, 7.5
2.0	35	7.5, 9.5	8.5	2.5, 4.5, 6.5
	36	8.5, 10.5	9.5	7.5, 11.5, 13.5
	37	6.00, 7.50	6.75	5.25, 3.75, 2.25
1.5	38	6.75, 8.25	7.50	6.0, 4.5, 9.0
	39	1.5, 3.0	2.25	3.75, 5.25, 6.75
	40	4.5, 5.5	5.0	4.0, 3.0, 2.0
1.0	41	8.0, 9.0	8.5	9.5, 10.5, 11.5
	42	12.0, 13.0	12.5	9.5, 10.5, 11.5

APPENDIX D

Description of Materials Used During
the Second Training Session

Sticks of various lengths were used, and each stick had a 0.75 x 0.75 inch, square cross-section. The sticks used at each training station are identified in Table D by the lengths in terms of 0.75 inch units.

In many of the tasks locations between adjacent sticks were cued. Where no cues appeared in an ordered set the sticks increased in length by regular 0.75 inch increments; however, wherever a location between two sticks was cued, the difference in length ("cued" increment) between those two sticks was greater than the regular increment of 0.75 inch. Cue levels are indicated in Table D by the lengths of the "cued increments (1 unit = 0.75 inch).

TABLE D

Materials Used During the Second Training Session

Cue Level	Training Station	Sticks to be Inserted	Sticks in the Ordered Set
	1	3, 7	1.5, 4.5, 5.5, 8.5
3	2	3, 6	1.5, 4.5, 7.5, 8.5
	3	4, 7	1.5, 2.5, 5.5, 8.5
	4	2.75, 6.25	1.5, 4, 5, 7.5
2.5	5	2.75, 5.25	1.5, 4, 6.5, 7.5
	6	3.75, 6.25	1.5, 2.5, 5, 7.5
	7	2.5, 5.5	1.5, 3.5, 4.5, 6.5
2.0	8	2.5, 4.5	1.5, 3.5, 5.5, 6.5
	9	3.5, 5.5	1.5, 2.5, 4.5, 6.5
	10	2.25, 4.75	1.5, 3, 4, 5.5
1.5	11	2.25, 3.75	1.5, 3, 4.5, 5.5
	12	3.5, 4.75	1.5, 2.5, 4.0, 5.5
	13	2, 4	1.5, 2.5, 3.5, 4.5
1.0	14	3, 4	1.5, 2.5, 3.5, 4.5
	15	2, 3	1.5, 2.5, 3.5, 4.5
	16	3, 7, 11	1.5, 4.5, 5.5, 8.5, 9.5,12.5
3	17	4, 7, 11	1.5, 2.5, 5.5, 8.5, 9.5,12.5
	18	3, 6, 9	1.5, 4.5, 7.5,10.5,11.5,12.5
	19	2.75,5.25,9.75	1.5, 4.0, 6.5, 7.5, 8.5, 11
2.5	20	2.75,6.25,8.75	1.5, 4.0, 5.0, 7.5,10.0,11.0
	21	4.75,7.25,9.75	1.5, 2.5, 3.5, 6, 8.5, 11
	22	2.5, 6.5, 8.5	1.5, 3.5, 4.5, 5.5, 7.5, 9.5
2.0	23	3.5, 6.5, 8.5	1.5, 2.5, 4.5, 5.5, 7.5, 9.5
	24	2.5, 4.5, 6.5	1.5, 3.5, 5.5, 7.5, 8.5, 9.5
	25	2.25,3.75,7.25	1.5, 3, 4.5, 5.5, 6.5, 8.0
1.5	26	2.25,3.75,6.25	1.5, 3, 4.5, 5.5, 7.0, 8.0
	27	3.25,4.75,6.25	1.5, 2.5, 4.0, 5.5, 7.0, 7.0
	28	2, 4, 6	1.5, 2.5, 3.5, 4.5, 5.5, 6.5
1.0	29	2, 3, 5	1.5, 2.5, 3.5, 4.5, 5.5, 6.5
	30	3, 4, 5	1.5, 2.5, 3.5, 4.5, 5.5, 6.5

APPENDIX E

Description of Materials Used During

the Third Training Session

Lined cards with evenly spaced, parallel black lines were used during the third training session. Each of the lined cards was constructed by cementing a photograph of parallel, evenly spaced black lines to a rectangular piece of posterboard (4.5 x 11.5 cm.). The black lines were parallel to the short side of the card as shown in Figure E.

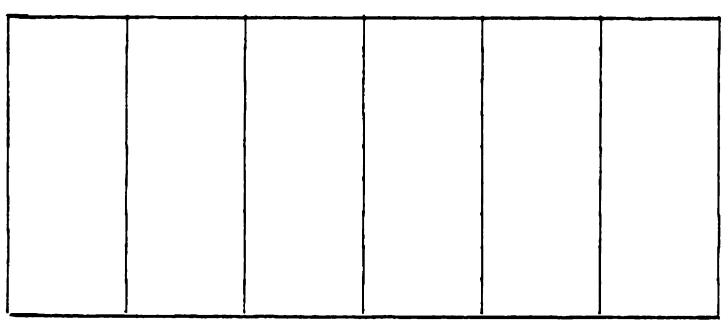


Fig. E. An example of a lined card (drawn to actual size).

Thirty-seven different line widths were required in the construction of the lined cards. Since it was almost impossible to manually draw and accurately reproduce 37 different line widths, large scale productions, in which line width could be easily controlled, were constructed and then photographically reduced in size to make the lined cards. The dimensions of the large scale productions were 47 x 122 cm., and the dimensions of the lined cards were approximately one-tenth (.094) the dimensions of the large scale productions. More specifically, this means that the lines on the cards were approximately one-tenth as wide as the lines of the large scale productions.

The cues for the lined cards took the form of redundant relevant information. The redundant relevant dimension was line width. Line width did not vary on the individual cards but did vary from card to card. In cued sets of cards, line width increased as the number of lines increased. The subjects could use both line width and the number of lines (or line density) in determining serial position. Since line width was more readily discriminable when the variability of line width was greater, it was assumed that greater line width variability would produce greater degrees of cueing.

To obtain the four different cue levels four sets of 12 cards each were constructed. The sets were identical with respect to the number of lines on the cards; that is, the cards in each set had the following number of lines: 3, 4, 5, 7, 9, 11, 14, 17, 20, 24, 28, 32. The four sets differed with respect to the variability of line width. In the fully cued set, line widths varied from 0.2mm. to 2.4 mm. The variation of line width in the second set of cards was two-thirds the variation of the fully cued set. In this second set line width varied from 0.41 mm. to 2.19 mm. Variation of line width in the third set was one-third the variation of the fully cued set. The line widths in this third set varied from 0.71 mm. to 1.89 mm. The fourth set of cards, the uncued set, showed no variation in line width. All lines on each card of the uncued set were 1.3 mm. in width. The cards used in training and testing were chosen from these four sets of cards.

Table E shows the cards used at each training station. Two numbers are required to describe a card. The first number, the one to the left of the hyphen, refers to the number of lines of the card, and the second number, the one to the right of the hyphen, refers to the millimeter width of the lines.

TABLE E

Materials Used During the Third Training Session

Cue Level	Training Station	Cards to be Inserted	Cards in the Ordered Set
	1	7-.8	4-.4, 14-1.4
Full	2	5-.6	9-1.0, 24-2.0
	3	20-1.8	11-1.2, 28-2.2
	4	28-2.03	9-10.6, 20-1.71
2/3	5	20-1.71	5-.73, 14-1.38
	6	11-1.22	17-1.54, 32-2.19
	7	5-.92	4-.82, 9-1.14
1/3	8	3-.71	5-.92, 11-1.25
	9	32-1.89	14-1.35, 24-1.67
	10	24-1.3	20-1.3, 28-1.3
None	11	5-1.3	4-1.3, 7-1.3
	12	20-1.3	11-1.3, 17-1.3
	13	5-.6	3-.2, 11-1.2, 17-1.6, 24-2.0
Full	14	14-1.4	3-.2, 5-.6, 9-1.0, 24-2.0
	15	5-.6, 20-1.8	3-.2, 9-1.0, 14-1.4, 28-2.2
	16	11-1.22	4-.57, 7-.9, 17-1.54, 24-1.87
2/3	17	4-.57, 32-2.19	7-.9, 11-1.22, 17-1.54, 24-1.87
	18	4-.57, 17-1.54	7-.90, 11-1.22, 24-1.87, 32-2.19
	19	5-.92, 20-1.57	4-.82, 9-1.14, 14-1.35, 24-1.67
1/3	20	11-1.25,32-1.89	5-.92, 9-1.14, 17-1.46, 24-1.67
	21	3-.71, 20-1.57	5-.92, 9-1.14, 14-1.35, 24-1.67
	22	4-1.3, 17-1.3	3-1.3, 5-1.3, 9-1.3, 14-1.3
None	23	5-1.3, 14-1.3	7-1.3, 11-1.3, 17-1.3, 24-1.3
	24	20-1.3, 28-1.3	11-1.3, 17-1.3, 24-1.3, 32-1.3

APPENDIX F

Description of Materials Used in Posttests 1 and 2

Sticks, lined cards, cars, and colored blocks were the four different kinds of materials used in posttests 1 and 2. The general characteristics of the sticks and lined cards have been described in Appendices D and E. Cars of various lengths were essentially sticks with 0.75 inch dowels glued to one side to give the appearance of wheels. Colored blocks were blocks of wood painted various shades of blue. Tasks performed with sticks and lined cards were considered near transfer tasks. Tasks performed with cars and colored blocks were considered far transfer tasks. The tasks were performed in the order described below.

Sticks

The numbers in Table F1 refer to the lengths of the sticks in 0.75 inch units (1 unit = 0.75 inch).

Table F1

Sticks Used in Posttests 1 and 2

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Inserting	2, 4	1.5, 2.5, 3.5, 4.5
Inserting	3, 4, 6	1.5, 2.5, 3.5, 4.5, 5.5, 6.5
Inserting	2, 4, 5, 7	1.5, 2.5, 3.5, 4.5,5.5,6.5,7.5,8.5

Lined Cards

The numbers in Table F2 refer to the number of lines on a card. The line widths were the same on all cards (1.3 mm.).

TABLE F2

Lined Cards Used in Posttests 1 and 2

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Ordering	-	4, 7, 11, 17
Inserting	9, 28	4, 7, 11, 17
Ordering	-	3, 5, 9, 14, 20, 28
Inserting	4, 11, 24	3, 5, 9, 14, 20, 28

Cars

The numbers in Table F3 refer to the lengths of the cars in terms of 0.75 inch units (1 unit = 0.75 inch).

TABLE F3

Cars Used in Posttests 1 and 2

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Ordering	-	2, 3, 4, 5, 6, 7
Inserting	2.5, 5.5, 7.5	2, 3, 4, 5, 6, 7

Colored Blocks

The blocks used in the testing tasks were chosen from a set of 12 blocks. Each block (representing a particular shade of blue) was given the number corresponding to its ordinal position in the set of 12 blocks ordered from the lightest (1) to the darkest (12) blue. The numbers in Table F4 refer to the various blocks used in the testing tasks.

TABLE F4

Colored Blocks Used in Posttests 1 and 2

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Ordering	-	2, 4, 6, 7, 8, 10, 12
Inserting	1, 5, 11	2, 4, 6, 7, 8, 9, 10, 12

APPENDIX G

Description of Materials Used in Posttest 3

Sticks, cars, "happies," and story cards were the four different kinds of materials used in posttest 3. Sticks and cars have been described in Appendices D and F, respectively. "Happies" were rectangular pieces of posterboard which varied in height and width. Smiling faces were drawn on the posterboard pieces to suggest the proper orientation. Width of the card was the relevant ordering dimension used with the "happies." Story cards were rectangular pieces of posterboard, each showing a picture of a stick man, the ground, a diving board, and water. The cards showed the man at different stages in the process of climbing up the diving board and diving into the water. The cards were to be ordered according to the time sequence. Tasks performed with sticks were considered near transfer tasks, tasks performed with cars were considered far transfer tasks, and tasks performed with "happies" and story cards were considered far-far transfer tasks. Tasks were performed in the order described below.

Sticks

The numbers in Table G1 refer to stick lengths in terms of 0.75 inch units (1 unit = 0.75 inch).

TABLE G1

Sticks Used in Posttest 3

Kind of Task	Objects to be Inserted	Objects in Ordered Set
Inserting	1.5, 4.5, 6.5, 7.5, 10.5	2, 3, 4, 5, 6, 7, 8, 9, 10, 11
Inserting	5.5, 8.5, 10.5, 11.5, 13.5, 15.5	4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15

Cars

The numbers in Table G2 refer to car lengths in 0.75 inch units (1 unit = 0.75 inch).

TABLE G2

Cars Used in Posttest 3

Kind of Task	Objects to be Inserted	Objects in Ordered Set
Inserting	2.5, 4.5, 8.5	2, 3, 4, 5, 6, 7
Inserting	3.5, 5.5, 6.5, 8.5	2, 3, 4, 5, 6, 7, 8, 9

"Happies"

A pair of numbers is required to describe each "happie." In Table G3 the number to the left of the hyphen refers to the width (the relevant dimension) and the number to the right of the hyphen refers to the height (irrelevant dimension) (1 unit = 0.75 inch).

TABLE G3

"Happies" Used in Posttest 3

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Ordering	-	1.5-7.0, 2.7-5.5, 4.3-2.5, 7.3-5.0, 8.3-3.0, 12.5-4.5
Inserting	2.1-3.5, 9.7-1.5	1.5-7.0, 2.7-5.5, 4.3-2.5, 7.3-5.0, 8.3-3.0, 12.5-4.5

Story Cards

Figure G shows the stick man in the 12 positions shown on the individual cards. The numbers in Figure G refer to one particular card in which the stick man is in the particular position shown by the number in Figure G. The numbers are used in Table G4 to indicate the particular cards used in the various testing tasks.

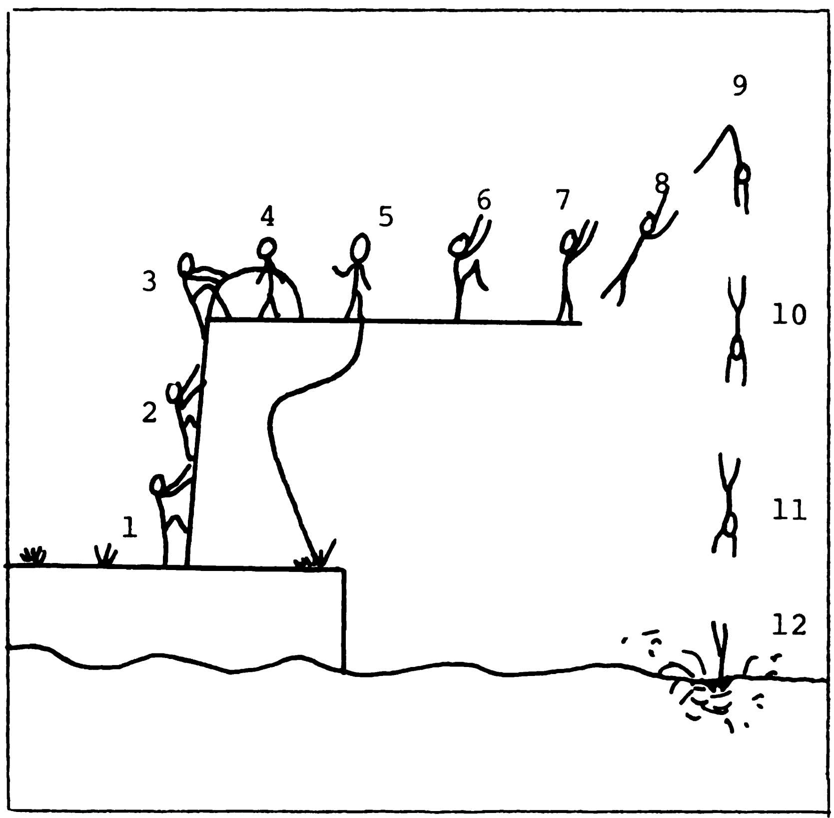


Fig. G. An illustration of the stick man positions shown on the story cards used in posttest 3. (Drawn to actual size.)

TABLE G4

Story Cards Used in Posttest 3

Kind of Task	Objects to be Inserted	Objects to be Ordered (Ordering Task) or Objects in Ordered Set (Insertion Task)
Ordering	-	1, 3, 5, 6, 7, 8, 9, 11, 12
Inserting	2, 4, 10	1, 3, 5, 6, 7, 8, 9, 11, 12

APPENDIX H

Verbalization Used During Posttests 1 and 2

Prior to testing with a particular kind of material an orientation session was given to acquaint the subjects with the materials and with the tasks to be performed with the materials.

General Introduction to Posttests 1 and 2

"Today we are going to play many different games. It is important that we do our best. We should check our work to make sure we do not make mistakes."

Orientation with Sticks

"These sticks go from the shortest stick here (E points) all the way to the tallest stick here (E points). The sticks make stairsteps. These sticks (E points) have been forgotten. We want to put these sticks along with these sticks to make more stairsteps. (E inserts the sticks and says...) See how all the sticks make stairsteps. See how we have the shortest one here, then the next tallest, then the next tallest (etc.)... all the way to the tallest. (E takes out the inserted sticks, scrambles them, and says...) Now you try to put these sticks with these to make stairsteps. (If S is correct, E says...)

Very good! The sticks go from the shortest to the tallest to make stairsteps. (E takes S to testing with sticks) (If S performed the orientation task incorrectly, E says...) These sticks do not make stairs. This one is not right. This one is not right (etc.). Let's see what we can do to make it right. (As E makes corrections he says...) See how this one fits here to make stairsteps (etc.). (When E has made all the corrections, he says...) Now all the sticks go from the shortest to the tallest to make stairs. Try putting these sticks with these to make stairs. (E helps S make the correct responses and then says...) Very good! The sticks go from the shortest to the tallest to make stairs. (E takes S to testing with sticks.)

Testing with Sticks

Here are some sticks that go from the shortest one here to the next tallest, to the next tallest, all the way to the tallest. These sticks have been forgotten. You put these sticks along with these sticks to make stairs. (When S stops, E says...) Let's check our work to see if we have made any mistakes. You may make any changes you want. You may not have made any mistakes. (After S checks and makes changes, E says...) Now let's go to the next game. (The same verbalization was used with all the testing tasks in which sticks were used.)

<u>Orientation with Lined Cards</u>

These cards have many black lines on them. This card has the most lines and this one has the fewest. (While ordering the cards, E says...) We want to put these cards side-by-side so we have the one with the fewest lines here, the one with the next most lines here (etc.), and finally the one with the most lines here. See how we have the fewest black lines here, the next most lines here, all the way to the most lines here. (E scrambles the cards and says...) Now you try it. Put the one with the fewest lines here, then the one with the next most lines here and so on. (If S is correct in performing the orientation task, E says...) Very good! You have the one with the fewest lines here, then the one with the next most lines (etc.), and finally the one with the most lines. (E takes S to testing with lined cards. (If S performed the orientation task incorrectly, E says...) This is not right, but let's see what we can do to make it right. This one does not belong here because .... It should go here (etc.). (When E has finished with the corrections, he says...) See how we have the one with the fewest black lines here, then the one with the next most lines (etc.) and finally the one with the most black lines. (E scrambles the cards and says...) Try putting the cards side-by-side so we have the one with the fewest lines here, then the one with the next most lines (etc.). (E helps S whenever necessary and then says...) Very good! You have the one with the fewest lines here, then the one with the next most lines here, all the way to the one with the most lines. (E takes S to testing with lined cards.)

Testing with Lined Cards

Verbalization for Ordering Tasks. Here are some cards with black lines. You put the card with the fewest lines here, then the one with the next most lines, and so on all the way to the one with the most lines. (When E stops, E says...) Let's check our work to see if any mistakes have been made. If you find any mistakes, please make changes to correct them. You may not have made any mistakes. (If S makes any mistakes, E makes the corrections and then says...)

Verbalization for Insertion Tasks. Here are some lined cards that were forgotten. You put them along with these cards so that we have the one with the fewest black lines here, then the one with the next most lines, and so on all the way to the one with the most black lines. (After S stops, E says...) Let's check our work to see if any mistakes have been made. You make any changes you wish. You may not have made any mistakes.

Orientation with Cars

We are going to pretend that these are cars of a train and that this line is a railroad track. (While ordering the cars on the track, E says...) We want to put the cars on the track so we have the shortest car first, then the next longest, then the next longest (etc.), until we get to the longest. (When the cars are ordered, E says...) See how we have the shortest car, then the next longest, all the way to the longest. (E brings out a car and says...) Here is a car that was forgotten. Watch how I put it with the others so all the cars go from the shortest to the longest. (When E has finished putting the car into the ordered set, he says...) See how all the cars go from the shortest to the longest. (E scrambles the cars and then says...) You put the cars on the track so the shortest car is first, then the next longest car is second, and so on all the way to the longest car. (If S performs the orientation task correctly, E says...) Very good! You have the shortest car here, then the next longest, then the next longest (etc.), all the way to the longest. (E then takes S to testing with lined cards.) (If S performs the orientation task incorrectly, E says...) Some mistakes have been made. Let's make the corrections. This one is not right. It should go here.(etc.). (When all corrections have been made, E says...) See how our cars go from the

shortest to the longest. You try it again. (E scrambles the cars) Put the shortest car first, then the next longest, then the next longest, all the way to the longest. (E gives help where necessary and then says...) Very good! That's right. You have the shortest car first, then the next longest, and so on all the way to the longest.

Testing with Cars

Verbalization for Ordering Tasks. Put these cars on the track from the shortest to the longest. Start with the shortest one here, then the next longest, and so on. (When S finishes, E says...) Let's check our work to see if any mistakes have been made. If you find any mistakes please correct them. Maybe no mistakes have been made. (E correctly orders the cars if S has made any mistakes.)

Verbalization for Inserting Tasks. These cars have been forgotten. Put them along with the others so we have the shortest car in front, then the next longest, all the way to the longest car. (When S is finished, E says...) Let's check our work to see if any mistakes have been made. You may make any changes you wish. Maybe no mistakes have been made.

Orientation with Colored Blocks

These cards are all painted blue. This card is the lightest blue, this card is the darkest blue. We want to arrange these cards so we have the lightest blue first, then the next darkest blue (etc.), until we have the darkest blue. See how we have the lightest blue here, then the next darkest, then the next darkest, all the way to the darkest. (E scrambles the blocks and says...) You put the lightest blue here, then the next darkest, then the next darkest, all the way to the darkest. (If S is correct, E says...) Very good! See how all the blocks go from lightest to the darkest. (E takes S to testing with colored blocks) (If S performed the orientation task incorrectly, E says...) We've made some mistakes. This one should go here (etc.). Now, see how the blocks go from the lightest to the darkest blue. (E scrambles the blocks and says...) You try it again. Put the lightest blue here, then the next darkest blue, and then the next darkest, all the way to the one with the darkest blue. (E gives help where necessary) Very good! See how the blocks go from the lightest blue to the darkest blue. (E takes S to testing with colored blocks.)

Testing with Colored Blocks

Verbalization for Ordering Tasks. Put these blocks side-by-side so we have the lightest blue first, then the next darkest, then the next darkest, and so on. (When S is finished, E says...) Let's check our work to see if any mistakes have been made. You may make any changes you wish. Maybe no mistakes have been made. (If S leaves mistakes, E makes the corrections and then goes on to the insertion task.)

Verbalization for Inserting Tasks. Here are some colored blocks we forgot. You put them along with these so all of the blocks go from the lightest to the darkest. (When S is finished, E says...) Let's check our work to see if any mistakes have been made. Make any changes you wish. Maybe no mistakes have been made.

APPENDIX I

Verbalization Used During Posttest 3

The verbalizations used with sticks and cars were the same as those used in posttests 1 and 2. Only the verbalizations used with Happies and story cards are presented here.

Orientation with Happies

We are going to call these Happies. See how the Happies are all smiling. Happies become sad when they are put on their sides. (E moves a Happie on its side.) Happies become sad when they are put on their heads too. (E puts a Happie on its head.) We want to make sure the Happies stay happy, so we always keep them up, like this. (E moves the Happie to the upright position.) Some Happies are wide. Some Happies are thin. Some are short, and some are tall. We only care about how wide they are. We do not care how tall they are.

(The orientation session with Happies included some discrimination training designed to get the S's to distinguish the width of Happies from the height of Happies. Four cards, each showing a picture of two Happies, were used in the discrimination training. The cards were presented one at a time, and the S's task was to point to the widest Happie. The S was cycled through the cards until he correctly performed four discrimination problems in a row. The verbalization used in the discrimination training was as follows.)

Point to the widest Happie in the picture. (If S is correct, E says...) Very good! Yes that is the widest Happie. (If S is incorrect, E says...) You pointed to the thinnest Happie. This one is the widest Happie. See how it is wider than this one. Now point to the widest Happie. (E presents the next picture and goes through the same dialogue until criterion performance is reached. Once criterion is reached E goes on by saying...) We are going to play a game with Happies. (E orders the Happies while saying...) We want to put the Happies side-by-side so that we have the thinnest Happie first, then the next widest Happie, then the next widest, all the way to the widest. See how all the Happies go from the thinnest to the widest. (E scrambles the Happies and says...) Now you try it. Put the thinnest first, then the next widest, then the next widest, all the way to the widest Happie. (If S performs correctly, E says...) Very good! The Happies go from the thinnest to the widest. (E then takes the subject to testing with Happies.) (If S performs the orientation ordering task incorrectly, E says...) Some mistakes have been made. This one should go here (etc.).

(After E makes the necessary corrections he says...) See how the Happies go from the thinnest to the widest. Now you try it again. (E helps S where necessary and then says...) Good! See how all the Happies go from the thinnest to the next widest, to the next widest, all the way to the widest. (E takes S to testing with Happies.)

Testing with Happies

Ordering with Happies. Here are some Happies. You put the thinnest on here, then the next widest, then the next widest, and so on. (When S stops E says...) Let's check our work to see if any mistakes have been made. If you find mistakes, you may make any changes. Maybe no mistakes have been made. (E makes sure the Happies are in order for the insertion task and then says...)

Inserting with Happies. Here are some Happies that were forgotten. Put these Happies along with the other Happies, so that all of the Happies go from the thinnest to the next widest, to the next widest, all the way to the widest. (When S stops E says...) Let's check our work to see if any mistakes have been made. If you find mistakes, you may make any changes. Maybe no mistakes have been made.

Testing with the Story Cards

Ordering with Story Cards. Here are some pictures showing a man climbing up a diving board. He dives off the diving board into the water. This is the first picture. Put the other pictures side-by-side so that they show the story of the man climbing up to the diving board and diving into the water. (When S stops E says...) Let's check our work to see if any mistakes have been made. You may make any changes. You may not have to make any changes.

Inserting with Story Cards. Here are some pictures that were forgotten. Put these pictures in with the others so that all of the pictures show the story of the man climbing up to the diving board and diving into the water. (When S is finished E says...) Check your work and see if you want to make any changes.

APPENDIX J

Individual Difference Measures--Three Parts of the Cincinnati Autonomy Test Battery

All of the subjects, except the special control subjects, were individually given the following three parts of Banta's Cincinnati Autonomy Test Battery (CATB) (1968); The Early Childhood Embedded Figures Test (EC-EFT), The Draw-a-Line Slowly Test (EC-MFF), and The Early Childhood Matching Familiar Rigures Test (EC-MFF). The CATB was designed for the use with children from three to six years of age. Reliability measures were given for each part. The three tests used in this study are discussed below in the order in which they were administered.

The Early Childhood Embedded Figures Test (EC-EFT)

The EC-EFT was designed to test field independence. Field independence is defined by the test designer as the tendency to separate an item from the field or context of which it is a part.

In the EC-EFT, fourteen test pictures embed a figure which looks like an ice cream cone. The subject is given a paper cut-out figure of an ice cream cone which is the same size and shape as the embedded cone. The subject is instructed to put the paper cone on top of the picture of the cone in the various figures. Prior to the testing, the subject is subjected to a training session in which three pictures are used to assess the subject's comprehension and readiness to perform the task.

Banta (1968) has reported an average reliability coefficient (internal consistency) of .54 for the EC-EFT. The subjects used in obtaining the mean reliability coefficient were lower class Negro children who were between the ages of three and six years.

The Draw-a-Line Slowly Test

This test was used to measure the ability of the child to restrain impulsive motor activity. It is assumed that the child who can draw a line slowly has the ability to restrain impulsive motor activity.

The tester begins the Draw-a-Line Slowly Test by drawing a fast and a slow line to give meaning to the words "fast" and "slow." The child then draws three lines slowly. He draws the first two lines as slow as he can, and then he is asked to draw the third line even slower than the previous two.

Since the rate of drawing the lines is the dependent variable in this test, both the lengths of the lines and times taken in drawing the lines are recorded. To calculate the average rate of line drawing, the total length of the three lines in inches is divided by the total time in hundredths of a minute.

Banta (1968) has reported correlations between the two members of each possible pair of line rates. The correlations between lines 1 and 2, lines 1 and 3, and lines 2 and 3 were respectively .55, .47, and .80 for a sample of 72 lower class Negro children. An inter-rater reliability coefficient of .96 was found using a sample of 33 children. In addition, a test-retest reliability coefficient of .43 was reported for 33 children who experienced a two month time interval between tests.

The Early Childhood Matching Familiar Figures Test (EC-MFF)

This matching familiar figures test is designed to measure the child's tendency to behave reflectively when confronted with conditions of uncertainty. As the title of the test implies, the subject's task is to match a sample figure with the same figure which is among similar figures. The child is first presented with a sample figure. He is then presented with an array of similar figures, one of which is the same as the sample. He is instructed to find the sample figure in this array. It is assumed that the reflective child tends to make less mistakes.

Banta styled the test after the original test which was devised by Kagan (1965) to study reflectivity in first grade children. Kagan (1965) used time latencies and error scores and found that the probability of error increased

as the time latencies shortened. In other words, children who take time to reflect over their decisions tend to make fewer errors. Banta omitted the measurement of time latencies and used only the number of correct responses as a measure of reflectivity. He maintains that nursery and kindergarten school children, for whom this test is designed, very often take long periods of time to respond because they are easily distracted, and tend to fantasize and talk with the tester. Therefore, these behaviors which take the young child away from the task render the time variable meaningless. The only score the child receives on the EC-MFF is the number of correct responses. It is assumed that the reflective child will make more correct responses.

Banta (1968) reports that data on the EC-MFF is sparse. From a sample of 62 subjects, a reliability coefficient of internal consistency (odd vs. even items) was found to be .37.

APPENDIX K

Descriptions of the Two Lenient Methods of Scoring (L1 and L2)

The need for two different methods of lenient scoring becomes apparent when a decision has to be made as to which of the two incorrect arrangements shown in Figure K is more correct.

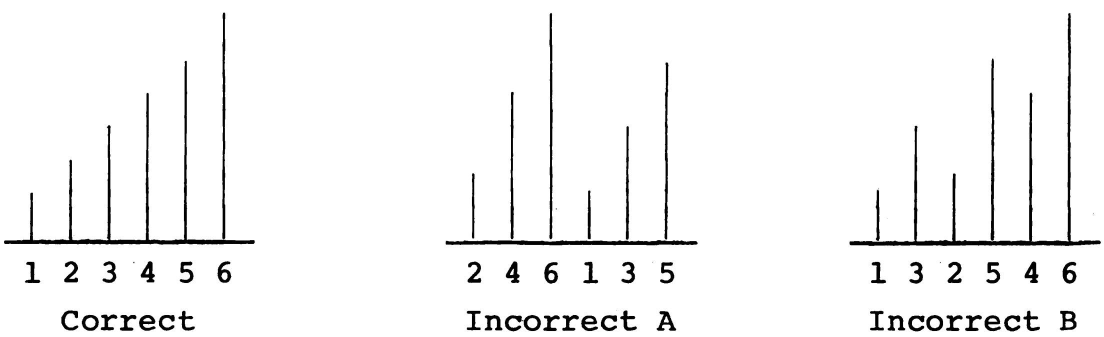


Fig. K. Different arrangements used to illustrate the need for different scoring systems. The number associated with each object refers to the ordinal position of that object in the properly ordered set.

The incorrect arrangement labeled "A" shows two separate three member sequence (2-4-6 and 1-3-5) which seemed to be placed side-by-side. It should be noted that in the "A" arrangement, small objects do not necessarily appear to the left and neither do large objects necessarily appear to the right. The incorrect arrangement labeled "B," on the other hand, has three, two member sequences (1-3, 2-5, and 4-6) and in general tends to show small objects toward the left and larger objects toward the right. If having small objects toward the left and larger objects towards the right ("B" in Figure K) is judged to be more correct than having two or more correctly ordered sequences incorrectly placed side-by-side ("A" in Figure K), then performance resulting in arrangement "B" would demand a higher score. If, on the other hand, sequencing ("A") is judged to be more correct than having smaller objects to the left and larger objects to the right ("B"), then performance resulting in arrangement "A" would demand a higher score. Rather than arbitrarily choose one partial credit scoring method which would favor one kind of arrangement over the other, two scoring methods were used, each designed to primarily reflect one of the two kinds of arrangements.

One partial credit or lenient scoring method, labeled L1, yielded a sequence score. The sequence score was tailored to give arrangements like "A" (Figure K) a higher score than arrangements like "B." To calculate

the sequence score for any arrangement the following formula was used: sequence score = $\frac{M-A}{M-1}$ where:

M = maximum number of possible sequences, which was equal to the total number of objects used in the seriation task since the maximum number of sequences occurs when each object appears as a "sequence of one"; and

A = actual number of sequences appearing in the particular arrangement being scored.

A sequence was defined in terms of the numbers used to record the positions of objects in an arrangement. Each object was given the number which corresponded to its ordinal position in the correctly ordered set of objects. These numbers were used to record the actual positions of the objects in the arrangements constructed by the subjects. For example, 1-3-2 referred to an incorrectly ordered arrangement which should have been ordered 1-2-3. In terms of the numbers used to describe an arrangement, a sequence was defined as: 1) a series of numbers which increased in value from left to right but not necessarily at regular intervals (...-<u>2-4-6</u>...; <u>1-2-3</u>...;), or 2) a number which was immediately preceded by a larger number and immediately followed by a smaller number (...-3-<u>2</u>-1...; ..-6-<u>3</u>-1...), or 3) a number which came at the right end of the arrangement and which was preceded by a larger number (..-5-<u>6</u>; 2-3-4-<u>1</u>), or 4) a number which was at the left end of an arrangement and which was larger than the number immediately following it (<u>4</u>-3-1..; <u>6</u>-5-2..).

For an example of calculating the L1 sequence scores, consider the arrangements labeled "Incorrect A" and "Incorrect B" in Figure K. The sequences in "A" are 2-4-6 and 1-3-5. Therefore, A equals two, and since there are six objects, M equals six. The sequence score for "A" then becomes $\frac{6-2}{6-1}$ or 0.8. The sequences in "B" are 1-3, 2-5 and 4-6 making A equal to 3. With M equal to 6, the sequence score for "B" becomes $\frac{6-3}{6-1}$ or 0.6. Note that the fewer the sequences, the larger the sequences and hence the larger the sequence score. Note also, that when all objects are ordered correctly, as in the "Correct" arrangement of Figure K, there is only one sequence and the sequence score becomes $\frac{6-1}{6-1}$ or 1.0. With the sequence formula and the definition of a sequence, the maximum and minimum scores were respectively 1.0 and 0.0.

The second method of lenient scoring, identified as L2, yielded scores which reflected the general tendency to have small objects toward the left and larger objects toward the right. This tendency was indexed by calculating a rank order correlation (Kendall Tau) between the order which appeared and the correct order. So the distribution of L2 scores could be assumed to be normal for analysis purposes, the correlation coefficients were transformed into Z values through the use of Fisher's r to Z transformation. The Z value calculated for each particular arrangement was used to derive the near, far, and far-far transfer scores for the data analysis. The Z values used in the study ranged from 0.00 which corresponded to a correlation of 0.00, to 5.00, which corresponded to a correlation of 1.00.

The L2 scores for the incorrect arrangements "A" and "B" are respectively, 0.203 and 0.933. Note that the L2 scoring method reverses that order of "A" and "B" correctness established by scoring method L1. Where the L1 or the sequence scoring method yielded a higher score for "A" than for "B," the L2 method of scoring yielded a higher score for "B" than for "A."

Thus far two different methods for scoring any one arrangement have been described. Since a number of tasks were used to assess a particular kind of transfer ability, a number of arrangements had to be considered in deriving near, far, and far-far transfer scores for an individual. A near, far, or far-far transfer score for either lenient scoring method (L1 or L2) was found by first summing the attained lenient scores for the set of tasks and then calculating the per cent attained of the maximum attainable sum. These percentages were the scores used in the data analysis.

APPENDIX L

Compilation of Data Used in the Analysis of Results

Initially, there were two large groups (Groups I and II) of subjects, each with an experimental (E) and a control (C) subgroup. Each subject, therefore is designated by two numbers and one or two letters. The number to the left of the letter(s) designates the large group number (1 or 2), the letter designates the treatment group (E = experimental, C = control, SC = special control), and the number to the right of the letter designates the specific individual. For example, 1E7 would indicate subject number 7 from the experimental subgroup of large group 1. The special control subjects were not members of either large group; hence, no number precedes their letters SC.

Seven dependent variable scores are given for each subject. NT refers to a near transfer measure, FT refers to a far transfer measure, and FFT refers to a far-far transfer. The number following the test type designation (NT, FT, or FFT) indicates the posttest number. For example, FT3 refers to the far transfer measure of posttest 3. Each dependent variable score is a per cent of tasks performed without error. For example, a score of 50.0 under NT1 would mean that 50% of the near transfer tasks of posttest 1 were performed without error.

Individual difference measures of age, field independence (EFT), motor control (IC), and reflectivity (MFF) are given. Age is in terms of months determined at the time of the first posttest. The other three ID measures are from Banta's CATB (Appendix J). EFT refers to an embedded figures test, IC refers to an impulse control test, and MFF refers to a matching familiar figures test.

TABLE L

Individual Subject Scores on the Dependent and ID Measures

Subject	Dependent Measures							ID Measures			
	NT1	FT1	NT2	FT2	NT3	FT3	FFT3	Age	EFT	IC	MFF
1E1	85.7	75.0	100.0	100.0	100.0	50.0	50.0	62.37	10	.129	6
1E2	100.0	50.0	100.0	100.0	100.0	50.0	0.0	68.06	12	.185	9
1E3	100.0	0.0	100.0	0.0	0.0	0.0	25.0	66.60	10	.098	10
1E4	87.5	25.0	71.4	100.0	100.0	100.0	0.0	66.53	13	.121	7
1E5	100.0	50.0	100.0	75.0	100.0	100.0	50.0	62.66	12	.093	10
1E6	42.9	25.0	57.1	25.0	100.0	50.0	50.0	61.73	11	.167	8
1E7	100.0	25.0	42.9	25.0	0.0	0.0	0.0	69.99	11	.177	6
1C1	14.3	50.0	28.6	25.0	0.0	0.0	25.0	65.99	13	.131	8
1C2	57.1	25.0	85.7	25.0	50.0	50.0	25.0	71.16	13	.257	9
1C3	57.1	25.0	62.5	25.0	100.0	50.0	0.0	69.37	12	.136	7
1C4	71.4	25.0	71.4	75.0	100.0	50.0	75.0	62.60	12	.096	6
1C5	14.3	25.0	42.9	50.0	100.0	50.0	25.0	69.33	11	.274	8
1C6	42.9	25.0	57.1	25.0	100.0	50.0	0.0	68.86	12	.201	7
1C7	42.9	25.0	57.1	50.0	50.0	100.0	75.0	69.16	8	.180	6
1C8	28.6	25.0	42.9	25.0	0.0	0.0	0.0	64.16	7	.090	6
2E1	57.1	50.0	42.9	75.0	50.0	0.0	0.0	65.99	8	.081	6
2E2	100.0	50.0	100.0	100.0	100.0	100.0	50.0	60.76	13	.189	11
2E3	71.4	50.0	71.4	75.0	100.0	0.0	0.0	68.76	8	.310	6

TABLE L (cont'd)

Subject	Dependent Measures							ID Measures			
	NT1	FT1	NT2	FT2	NT3	FT3	FFT3	Age	EFT	IC	MFF
2E4	71.4	50.0	85.7	75.0	100.0	100.0	100.0	66.73	10	.704	10
2E5	57.1	25.0	42.9	25.0	0.0	0.0	0.0	69.73	9	.261	8
2E6	100.0	50.0	85.7	100.0	100.0	50.0	50.0	64.12	13	.300	10
2E7	85.7	75.0	100.0	100.0	100.0	100.0	50.0	63.32	12	.123	8
2E8	100.0	25.0	100.0	50.0	100.0	0.0	25.0	65.09	11	.218	10
2C1	14.3	25.0	42.9	25.0	0.0	0.0	0.0	61.32	12	.170	6
2C2	28.6	0.0	28.6	0.0	0.0	0.0	0.0	62.93	10	.414	6
2C3	28.6	25.0	42.9	25.0	0.0	50.0	50.0	65.06	13	.116	10
2C4	14.3	50.0	57.1	25.0	50.0	50.0	0.0	66.32	12	.186	7
2C5	14.3	25.0	28.6	25.0	100.0	50.0	0.0	63.83	11	.105	6
2C6	71.4	100.0	71.4	50.0	100.0	100.0	0.0	69.47	11	.346	9
2C7	28.6	25.0	14.3	75.0	0.0	50.0	0.0	64.40	10	.154	8
2C8	28.6	50.0	57.1	50.0	50.0	0.0	25.0	63.60	10	.325	8
2C9	100.0	75.0	100.0	100.0	100.0	50.0	50.0	61.50	11	.184	7
SC1					50.0	0.0	50.0				
SC2					100.0	50.0	100.0				
SC3					100.0	0.0	0.0				
SC4					100.0	100.0	50.0				
SC5					100.0	50.0	50.0				
SC6					100.0	100.0	50.0				

TABLE L (cont'd)

Subjects	Dependent Measures							ID Measures			
	NT1	FT1	NT2	FT2	NT3	FT3	FFT3	Age	EFT	IC	MFF
SC7					50.0	0.0	50.0				
SC8					100.0	0.0	0.0				
SC9					100.0	100.0	100.0				
SC10					100.0	50.0	100.0				
SC11					50.0	50.0	50.0				
SC12					100.0	100.0	50.0				
SC13					100.0	0.0	25.0				

APPENDIX M

Repeated Measures, Multivariate Analysis Information for the Experimental versus Control Group Analysis (E vs. C)

The input dependent variables were: near (NT1) and far (FT1) transfer measures from posttest 1, near (NT2) and far (FT2) transfer measures from posttest 2, and near (NT3) and far (FT3) transfer measures from posttest 3. These input dependent measures were determined from the stringent (S) method of scoring (proportion of tasks performed without error).

The input dependent variables were transformed to new variables by the transformation matrix shown in Table M1. The variable labeled "const" refers to a constant. The variable A1A2 is associated with the difference between posttests 1 and 2. The variable A2A3 is associated with the difference between posttests 2 and 3. The variable B1B2 is associated with the difference between near and far transfer test types. The variables, inter1 and inter2 refer to interaction contrasts.

TABLE M1

Transformation Matrix - New x Old Variables (E vs. C)

		NT1	FT1	NT2	FT2	NT3	FT3
1	CONST	0.408248	0.408248	0.408248	0.408248	0.408248	0.408248
2	A1A2	0.577350	0.577350	-0.288675	-0.288675	-0.288675	-0.288675
3	A2A3	-0.000000	-0.000000	0.500000	0.500000	-0.500000	-0.500000
4	B1B2	0.408248	-0.408248	0.408248	-0.408248	0.408248	-0.408248
5	INTER1	0.577350	-0.577350	-0.288675	0.288675	-0.288675	0.288675
6	INTER2	-0.000000	-0.000000	0.500000	-0.500000	-0.500000	0.500000

TABLE M2

Experimental and Control Group Means Calculated from the New Variables

	CONST	A1A2	A2A3	B1B2	INTER1	INTER2
Experimental	1.621780	-0.059159	0.125000	0.342602	0.123668	-0.091667
Control	1.062262	-0.110630	-0.009971	0.113685	-0.051197	0.004735

TABLE M3

Sample Correlation Matrix (Within Cell) for the New Variables (E vs. C)

	CONST	A1A2	A2A3	B1B2	INTER1	INTER2
CONST	1.000000					
A1A2	-0.604552	1.000000				
A2A3	-0.478229	0.745036	1.000000			
B1B2	-0.008031	0.074739	0.030051	1.000000		
INTER1	-0.019792	0.017287	0.147123	-0.166142	1.000000	
INTER2	-0.290155	0.265639	0.072580	-0.079142	0.401092	1.000000

TABLE M4

Variances for the New Variables (E vs. C)

VARIABLE		VARIANCE
1	CONST	0.305637
2	A1A2	0.053269
3	A2A3	0.069156
4	B1B2	0.059987
5	INTER1	0.033647
6	INTER2	0.053586

TABLE M5

Symbolic Basis Vectors (E vs. C)

CO, Basis Vector =	1.0000	1.0000
C1, Basis Vector =	0.5000	-0.5000

Four major multivariate hypotheses were tested. The multivariate F's, the step-down F's, the mean squares between, and the hypothesis mean product matrices are presented below for the four hypotheses. The uses made of the various components associated with the hypotheses are described.

<u>Hypothesis 1 (CO)</u>

F-ratio for the multivariate test of equality of mean vectors = 0.0531, df = 5 and 26, p less than 0.0001.

The step-down F's for variables A1A2 and A2A3 (Hypothesis 1) were used as evidence for a Posttest main effect. The step-down F for variable B1B2 was used as evidence for a Test Type main effect. No Posttest x Test Type interaction was observed, as shown be the step-down F's for variables Inter1 and Inter2.

TABLE M6

Mean Squares Between, Step-Down F's, and Probability Statements Associated With the Variables Under Hypothesis 1 (CO) (E vs. C)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
A1A2	0.2395	4.4951	0.0424
A2A3	0.0909	14.0421	0.0008
B1B2	1.5628	16.1776	0.0004
INTER1	0.0303	0.3834	0.5410
INTER2	0.0524	0.0964	0.7588

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 30

TABLE M7

Hypothesis Mean Products Matrix Associated with Hypothesis 1 (CO) (E vs. C)

	A1A2	A2A3	B1B2	INTER1	INTER2
A1A2	0.23945				
A2A3	-0.147531	0.090898			
B1B2	-0.611723	0.376898	1.562769		
INTER1	-0.085177	0.052480	0.217602	0.030299	
INTER2	0.111979	-0.068995	-0.286071	-0.039833	0.052367

Hypothesis 2 (C1)

F-ratio for the multivariate test of equality of mean vectors = 4.1688, df = 5 and 26, p less than 0.0065.

TABLE M8

Mean Squares Between, Step-Down F's, and Probability Statements Associated with the Variables Under Hypothesis 2 (C1) (E vs. C)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
A1A2	0.0211	0.3963	0.5338
A2A3	0.1452	2.0586	0.1621
B1B2	0.4176	6.0730	0.0202
INTER1	0.2437	5.7668	0.0235
INTER2	0.0741	3.2314	0.0839

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 30

TABLE M9

Hypothesis Mean Products Matrix Associated with Hypothesis 2 (C1) (E vs. C)

	A1A2	A2A3	B1B2	INTER1	INTER2
A1A2	0.021112				
A2A3	0.055360	0.145167			
B1B2	0.093893	0.246211	0.417586		
INTER1	0.071723	0.188076	0.318987	0.243669	
INTER2	-0.039540	-0.103685	-0.175855	-0.134332	0.074056

Step-down F's for variables A1A2 and A2A3 (Hypothesis 2) indicated no Treatment x Posttest interaction. The step-down F for variable B1B2 indicated a significant Treatment x Test Type interaction. A significant Treatment x Test Type x Posttest was shown by the step-down F's for variables Inter1 and Inter2.

Hypothesis 3 (CO)

F-ratio for the multivariate test of equality of mean vectors = 42.1651, df = 6 and 25, p less than 0.0001.

TABLE M10

Mean Squares Between, Step-Down F's, and Probability Statements Associated with the Variables Under Hypothesis 3 (CO) (E vs. C)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
CONST	56.1407	183.6840	0.0001
A1A2	0.2395	7.8893	0.0089
A2A3	0.0909	2.4734	0.1271
B1B2	1.5628	2.0525	0.1635
INTER1	0.0303	0.0777	0.7827
INTER2	0.0524	1.1215	0.2998

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 30

TABLE M11

Hypothesis Mean Products Matrix Associated with Hypothesis 3 (CO) (E vs. C)

	CONST	A1A2	A2A3	B1B2	INTER1	INTER2
CONST	56.14066					
A1A2	-3.66646	0.23945				
A2A3	2.25900	-0.14753	0.09090			
B1B2	9.36669	-0.61172	0.37690	1.56277		
INTER1	1.30423	-0.08518	0.05248	0.21760	0.03030	
INTER2	-1.71461	0.11198	-0.06899	-0.28607	-0.03983	0.05237

None of the tests associated with Hypothesis 3 were used in interpreting the results of this study.

Hypothesis 4 (C1)

F-ratio for the multivariate test of equality of mean vectors = 5.1300, df = 6 and 25, p less than 0.0015.

TABLE M12

Mean Squares Between, Step-Down F's, and Probability Statements Associated with the Variables Under Hypothesis 4 (C1) (E vs. C)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
CONST	2.4947	8.1623	0.0077
A1A2	0.0211	6.6518	0.0153
A2A3	0.1452	1.6925	0.2039
B1B2	0.4176	3.3209	0.0795
INTER1	0.2437	3.8701	0.0600
INTER2	0.0741	1.0706	0.3108

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 30

The significant step-down for the const variable (Hypothesis 4) was interpreted as evidence for a Treatment main effect.

TABLE M13

Hypothesis Mean Products Matrix Associated with Hypothesis 4 (C1) (E vs. C)

	CONST	A1A2	A2A3	B1B2	INTER1	INTER2
CONST	2.494699					
A1A2	0.229493	0.021112				
A2A3	0.601788	0.055360	0.145167			
B1B2	1.020662	0.093893	0.246211	0.417586		
INTER1	0.779667	0.071723	0.188076	0.318987	0.243669	
INTER2	-0.429823	-0.039540	-0.103685	-0.175855	-0.134332	0.074056

APPENDIX N

Multivariate Analysis Information for the Experimental versus Control versus Special Control Group Analysis (E x C x SC)

There were two major parts to this analysis. One part (experimental versus special control) focused on the comparison of experimental and special control groups' performances of the near (NT3), far (FT3), and far-far (FFT3) transfer tests of posttest 3. The other part (special control versus control) focused on the comparison of the control and special control groups' performances of the same three tests. The dependent measures were determined from the stringent (S) method of scoring (proportion of tasks performed without error).

TABLE N1

Experimental, Control, and Special Control Group Means

	NT3	FT3	FFT3
Experimental	0.766667	0.466667	0.300000
Control	0.529412	0.411765	0.205882
Special Control	0.884615	0.461538	0.519231

TABLE N2

Sample Correlation Matrix (Within Cells) (E x C x SC)

	NT3	FT3	FFT3
NT3	1.000000		
FT3	0.525585	1.000000	
FFT3	0.227875	0.497719	1.000000

TABLE N3

Variances (E x C x SC)

VARIABLE	VARIANCE
NT3	0.148704
FT3	0.156708
FFT3	0.088979

TABLE N4

Symbolic Basis Vectors (E x C x SC)

C0. GRAND MEAN

Basis Vector = 1.0000   1.0000   1.0000

C1. EXPERIMENTAL VS. SPECIAL

Basis Vector = 0.6667   -0.3333   -0.3333

C2. CONTROL VS. SPECIAL

Basis Vector =-0.3333   0.6667   -0.3333

Part 1 - Experimental versus Special Control

F-ratio for the multivariate test of equality of mean vectors = 0.2795, df = 3 and 40, p less than 0.8399.

TABLE N5

Mean Squares Between, Step-Down F's, and Probability Statements Associated with the Variables Under Part 1 (E vs. SC)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
NT3	0.0694	0.4670	0.4982
FT3	0.0111	0.0115	0.9151
FFT3	0.0174	0.3780	0.5422

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 42

TABLE N6

Hypothesis Mean Products Matrix Associated with Part 1 (E vs. SC)

	NT3	FT3	FFT3
NT3	69.444444-003		
FT3	27.777778-003	11.111111-003	
FFT3	-34.722222-003	-13.888889-003	17.361111-003

Part 2 - Special Control versus Control

F-ratio for the multivariate test of equality of mean vectors = 5.6897, df = 3 and 40, p less than 0.0025.

TABLE N7

Mean Squares Between, Step-Down F's, and Probability Statements Associated with the Variables Under Part 2 (SC vs. C)

VARIABLE	BETWEEN MEAN SQ	STEP-DOWN F	P LESS THAN
NT3	0.9294	6.2504	0.0165
FT3	0.0183	1.1109	0.2981
FFT3	0.7233	8.3658	0.0062

DEGREES OF FREEDOM FOR HYPOTHESIS = 1
DEGREES OF FREEDOM FOR ERROR = 42

TABLE N8

Hypothesis Mean Products Matrix Associated with Part 2 (SC vs. C)

	NT3	FT3	FFT3
NT3	0.929449		
FT3	0.130241	0.018250	
FFT3	0.819928	0.114894	0.723313