69-7612

SHIRK, Ethel J., 1920-

POST-CRITERIAL PERFORMANCE IN ODDITY LEARNING AS A FUNCTION OF STIMULUS SIMILARITY, INTERPOLATED REST INTERVAL, GRADE LEVEL AND SEX.

The City University of New York, Ph.D., 1969
Psychology, experimental

University Microfilms, Inc., Ann Arbor, Michigan

POST-CRITERIAL PERFORMANCE IN ODDITY LEARNING AS A FUNCTION OF STIMULUS SIMILARITY, INTERPOLATED REST INTERVAL, GRADE LEVEL AND SEX

by

ETHEL J. SHIRK

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1968

This manuscript has been read and accepted for the University Committee in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 29, 1968
Date

Alfred Castaneda
Chairman of Examining Committee

Sept 13 1968
Date

Executive Officer

Alfred Castaneda

Louis D. Costa

Sam F. Korn

Supervisory Committee

ACKNOWLEDGMENTS

I wish to express my gratitude to Dr. Alfred Castaneda, Chairman, and Drs. Sam J. Korn and Louis D. Costa, members of my committee, for their generous guidance, encouragement and criticism in the preparation of this dissertation. I wish also to express my thanks to my former mentor, Dr. Eugene S. Gollin, for his direction and guidance into the areas of experimental and developmental psychology.

The permission granted by the New York City Board of Education to carry out this research with students of the New York City public school system, and the cooperation of the principals and faculty of the two schools involved, is greatly appreciated.

At this time I wish to thank the National Institutes of Health, the City University of New York, and the National Science Foundation for their financial assistance which made my graduate training possible. And last, but not least in my affection, I wish to express my deep appreciation to the faculty and staff of Queens College for their encouragement and assistance throughout my undergraduate and graduate years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

I

INTRODUCTION

It is generally observed (McConnell,1965), though often unreported, that when training is extended beyond the designated criterion of mastery in a learning task, marked individual differences in the ability to maintain the level of mastery appear. On the assumption that younger Ss are more susceptible to fatigue, boredom, and distraction, it is possible then that their post-criterial performance might be expected to exhibit a lower level of performance than in the case of older Ss. Similarly, if the task requires greater attentional or orienting responses, etc., to the discriminanda, such as would be the case where the relevant stimuli are less discriminable, post-criterial performance might be expected to be deleteriously affected, especially in the case of child Ss.

For example, in a previous unpublished study (Shirk, 1964, see Appendix) which employed an oddity learning problem, and in which the differences between the three discriminanda were highly discriminable, the post-criterial performance of first and second grade Ss was found to be highly variable. Thus, out of a group of 12 Ss, five (42%) were observed to make errors on the first post-criterial trial. Even more pertinent to the present proposal was the fact that only one S of the five succeeded in re-achieving the criterion.

Observations of the type described above are particularly relevant to the transfer of training type of design in which the effects of overlearning, i. e., the amount of post-criterial training, are varied (Spiker & Holton, 1958). In other words, these observations suggest that the nature of transfer may also be a function of the characteristics of post-criterial performance on the initial task. In general, most studies concerned with this type of training transfer do not report analyses based on post-criterial performance.

Studies in discrimination reversal also enploy overlearning as a variable and the same variable, level of performance in post-criterial trials, might be of importance here. These studies (Furth & Youniss, 1964; Gollin, 1965; Stevenson & Weir, 1959; Youniss & Furth, 1964), like the transfer studies, do not report analyses of post-criterial performance. The observations made in the Shirk (1964) study suggest that the level of post-criterial performance might be an important variable where child populations are employed, particularly where the task requires diligent attention to the relevant stimulus properties.

Oddity Problem Learning

Oddity learning has been selected to study this phenomenon since it is a form of multiple-sign learning requiring the subject to respond to relations among the simultaneously presented stimuli. This is more complex than a simple discrimination problem where response to a specific stimulus, such as red, in a red-blue simultaneous comparison procedure, is consistently reinforced.

Two types of oddity problem are found in the literature. There is the standard three-position oddity task for which Levinson (1958) denotes three defining conditions: (1) the odd or singly represented stimulus may be located in any of three positions and is always reinforced, (2) all the specific stimuli (e.g., all hues or intensities in this study) are correct an equal number of trials; and (3) the subject is permitted only one choice on each trial.

In another type of oddity task, sometimes referred to as two-position oddity, the odd or singly represented stimulus is never located in the middle position, so that the two like stimuli are always adjacent to each other, never separated by the odd stimulus, as they are in the standard three-position problem.

The two-position task was employed by House (1964) in her work with retardates, and has been used in other studies (Gollin et al, 1967). Harlow also uses this two-position technique in his oddity studies and states (1951) that the Wisconsin oddity and matching studies have shown that arrangements of stimuli which separate the identical objects by the non-identical object produce the most frequent and most persistent errors. However Riopelle (1959) found that the anticipated greater difficulty of the ABA array did not occur with his linear arrangement of stimuli in a study using rhesus monkeys as subjects.

Lipsitt and Serunian (1963) consider oddity-problem learning "... a type of transverse patterning (Spence, 1952) in which consistently correct performance can be achieved by responding to

a specific cue only in the presence of certain other cues, i.e., to the stimulus that is unlike the others present (p.201)." House (1964) explains differences in effectiveness of training methods in the oddity problem in terms of an attention theory (Zeaman & House, 1963) which postulates that the visual discrimination learning of subjects requires a chain of two responses, (1) attending to or observing the relevant dimension, and (2) choosing the correct cue of that dimension. Since the oddity dimension is a relation among cues, it requires two observing responses. "To see oddity, S must see the dimension(s) carrying the oddity-- the vehicle dimension-- and must then see the oddity relation among the cues of the vehicle dimension. Thus oddity requires the formation of a chain of two observing responses and one terminal instrumental response (approach the odd cue) (House, 1964a, p.642)."

The oddity problem might also be explained in terms of the Acquired Distinctiveness and Acquired Equivalence of Cues (Miller & Dollard, 1941). Oddity thus becomes a discrimination problem in terms of the concepts "same" and "different." In each trial the subject sees two matching stimuli and one different stimulus. While the actual stimuli change from trial to trial, the configuration in each trial is equivalent with that in every other trial. Therefore the subject chooses the odd stimulus because it is equivalent with all other odd stimuli, regardless of their other object characteristics. The odd stimulus becomes distinctive because it has been encountered in previous trials.

Harlow (1951, p.281) states, "On any given trial there is reward of the position, the object, and single representation, but over the series of trials only single representation is rewarded 100 per cent of the time." In the present study the child must concentrate upon the various aspects to ascertain the relevant dimension, then see the oddity relation in order to respond correctly to the odd cue. Thus if a breakdown in performance in overlearning might be anticipated under conditions involving greater attention to select the relevant stimulus, it could be expected that this would result with higher frequency in oddity rather than in simple discrimination learning, because of the complexity oddity embodies.

Infra-human Performance

The oddity problem was first utilized with infra-human organisms. Oddity problem solution by monkeys was first demonstrated by Robinson (1933) and further investigated by workers in the University of Wisconsin laboratory in the early 1940s. These studies found that monkeys take from 400 to 1500 trials to reach a criterion of 90% correct responses. While perseverative positional and perseverative object errors decrease rapidly in acquisition, Harlow (1951) found that both were present to some extent even after the learning criterion had been attained. Nissen & McCulloch (1937) reported efficient performance by chimpanzees on the oddity problem. However they used what is essentially a matching technique, presenting a single odd stimulus with eleven identical stimuli. (The matching problem was first

studied by Kohts in 1928, when she trained a chimpanzee to select from many objects on a tray, the object identical with the one she held in her hand.)

Subprimates seem to have difficulty with the oddity problem. Lashley (1938) attempted to train rats using the jumping stand apparatus. The rats were trained initially to choose a cross presented with two circles, then a circle presented with two crosses. After the third or fifth reversal, the animals either refused to jump or jumped persistently to one figure in spite of repeated falls. Pastore (1954) found that canaries failed to learn to select the odd stimulus consistently and he attempted to facilitate learning by substituting nine stimulus objects.

Warren (1960) trained five young cats on 20 three-choice discrimination and reversal problems before training on the oddity problem. Only the cat most successful on the discrimination and reversal series solved the oddity problem. This cat formed a learning set for efficient solution over a series of 30 transfer problems. Its final performance approximated that of rhesus monkeys and exceeded that of any infra-primate species on which data have been reported.

Harlow (1958, p.281) stated that the oddity problem is "beyond the intellectual capacity of the young child, although data defining the minimum chronological age level for oddity problem solution are lacking." The majority of the studies of human oddity problem solution have utilized retardate subjects and the two-position rather than the standard three-position

oddity task. Because it is felt that the two-position problem is simpler than standard three-position oddity, and also that differences between retardate and normal subjects are qualitative as well as quantitative, the standard three-position problem and normal Ss were selected for this study in an effort to clarify many of the contradictions in oddity literature.

The four variables of stimulus similarity, interpolated interval, grade level and sex, which have received considerable attention in the literature, but still remain unclear, will be considered in this study.

Stimulus Similarity

An important variable that should further amplify a post-criterial breakdown is that of stimulus similarity. In using two levels of discriminability, red-green-blue, and red-pink-orange, we should expect more difficulty with the latter since it is less discriminable and would require greater attention on the part of the children, as has been found in other studies utilizing these same dimensions (Lipsitt, 1961; Lipsitt & LoLordo, 1963).

Interpolated Interval

A performance decrement in overlearning might be explained in terms of Piaget's functional invariants of assimilation and accommodation. If the child is not ready to master a particular concept, even though we may train him in a particular learning situation and produce criterial performance, we should not expect this level of performance to be maintained, since the newly-accommodated-to feature cannot fit into any of the existing schemas

the child possesses at the time, thus there is no structure to which it can be assimilated. It is possible that the child might attain criterial performance through a conditioning or rote learning type of technique, but this does not necessarily mean he has an understanding of the concept underlying his performance. Katona (1940) found that memorization of a solution tended to narrow rather than increase the range of problem-solving. He concluded that learning with understanding not only improved retention of what was learned but also better qualified the learner to move forward to new learning. This stresses the importance of understanding for transfer of learning.

An interpolated interval between original learning and the post-criterial trials should produce interference if this conditioning or rote learning type of technique does obtain in this study. It would then be expected that children who have not mastered the concept underlying oddity performance would show a decrement in performance in overlearning if they are in the delay group, whereas those who have mastered the concept should maintain performance in the post-criterial trials. Another theoretical possibility, in line with Hull's concept of reactive inhibition (1943), would be that the effects of the interpolated interval tend to dissipate the reactive inhibition built up in the performance of the learning trials from repetition, boredom, etc., thus we would expect that the performance of the interpolated interval group would be superior to that of the group not having the interval, because of the change of pace and the novelty

embodied in the interpolated task.

Sex of Subjects

The present study is also concerned with oddity problem learning *per se*, as well as using it as a vehicle to investigate post-criterial performance. The variables of grade level and sex have been introduced to perhaps clarify these aspects of oddity performance, since the literature with regard to oddity performance in normal *Ss* is relatively sparse.

Martin and Blum (1960,1961) have found a significant sex difference in normals and familial types of mentally subnormal children in favor of boys, but with Mongoloids, the difference was in the opposite direction. Sex differences in maturation and verbal ability have been established with girls showing superior performance, thus we should anticipate other differences in the normal child population. Since the variable of sex differences is somewhat unclear, due to these apparently contradictory findings, it was included in this study in an attempt to establish whether these differences do apply to oddity problem learning in the normal child population.

Grade Level

The grade levels have been introduced to plot the oddity function. The general conclusions reached so far are that oddity learning is a function of MA (Ellis & Sloan, 1959; Martin & Blum, 1960, 1961; House, 1964, a.b.). Lipsitt and Serunian (1964) and Gollin and Shirk (1966) have found oddity solution to improve with CA. However, Weir (1964) found performance in probability learning over ages 3 - 18 to have an inverted U-shaped function, and

Friedman (1963) found a depression at the fourth grade level in the expected direct positive relationship between intellectual skill and school grade level in a series of sequential pattern problems. A comparison of the data from Lipsitt and Serunian (1963), Lipsitt and LoLordo (1964) and Gollin and Shirk (1966), as well as unpublished data on second graders (Shirk, 1964), raises the question whether the performance curve in oddity learning is a straight line or some other shape.

In terms of Piaget's developmental theory of cognitive functioning, we would expect performance to be other than linear, since development is felt to consist of fairly stable states with abrupt transitions between them. Gestalt theory also would predict other than linear performance over the wide grade range encompassed by this study. Dodge (1931) speaks of trial-and-error in insight experiments in terms of approximation and correction. Thus the "trial" is something more than just any "random" action in the behavior repetory and might be considered a plausible hypothesis which has to be tested or tried. Many such hypotheses may be tested before the appropriate solution is found. Thus the more intelligent subject may actually take longer to solve a given problem because of the greater number of hypotheses he can generate. This type of varied behavior implies an organization missing from mere trial-and-error behavior.

In order to relate the laboratory findings on learning covered in this experiment with scholastic achievement, correlations of oddity performance scores, both pre-and post-criterial, with scores on the New York City reading test were computed.

II

METHOD

Subjects

Ss were 546 public school children in grades 1 through 9. The number of children per grade was variable (see Table 1) because some children did not attain criterion. Only those children within the proper CA for their grade were utilized. A correlation of +.98 between age and grade level was obtained for these subjects, therefore age and grade level will be treated as functional equivalents in this study. All Ss were students in adjacent elementary and junior high schools in Flushing, Queens County, New York. The number of Ss in each sub-group in terms of the variables of Grade, Sex and Difficulty are presented in Table 1.

Apparatus

The apparatus is a minor variation of that employed by Gollin and Shirk (1966), a dark grey plywood box 12" high x 12½" deep x 25½" wide. The front panel has three stimulus apertures, each faced with a milk glass window on the upper part and three response buttons on the lower. The three different colored stimulus lights are set in metal reflecting units behind each window. Correct responses are reinforced by the sounding of a door chime activated by the two-phase response buttons.

Red, blue and green colored bulbs (GE C-6 series) set in metal reflecting units are behind each of the windows for the

TABLE 1

NUMBER OF SUBJECTS BY SUB - GROUPS

GRADE	MALE			FEMALE			TOTAL
	Easy (R-B-G)	Difficult (R-P-O)	Total	Easy (R-B-G)	Difficult (R-P-O)	Total	
1	18	21	39	23	21	44	83
2	20	26	46	18	26	44	90
3	17	23	40	19	24	43	83
4	17	19	36	16	22	38	74
5	16	17	33	19	20	39	72
7	16	18	34	21	18	39	73
9	16	18	34	17	20	37	71
TOTAL	120	142	262	133	151	284	546

Easy condition. Stimulus lights for the Difficult task are red, pink and orange. Selector switches, which enable E to preselect the colors to appear in each window, and to indicate the position of the correct response; and a master toggle switch for presentation of the stimulus array, are on the control panel attached to the right side of the box by a long cable. Any response turns off the lights, while the correct response turns off the lights and also activates the door chime.

Procedure

There are 18 possible position-counterbalanced combinations of the three colors employed in each condition, e.g., red-red-green, red-green-red, green-red-red, etc. From these combinations a series of 18 randomly-constructed trials, corrected as far as possible to eliminate the patterning of responses found by Weir (1964) on apparatus with three response positions, was assembled (see Appendix). This series of 18 trials was repeated four times, for a total of 72 trials, and the same stimulus series was employed for all Ss, regardless of condition. On all trials the odd color was the reinforced stimulus. All Ss were run individually.

When S was seated at the apparatus, E gave the following instructions:

> Now I am going to tell you about my puzzle. I am going to show you some colored lights in these windows (pointing). Each time the correct light will make the bell ring. If you think this light (pointing left) is the correct light, you push its button (indicating left button); if you think this light is the correct light (pointing to middle light), push its button (indicating middle button); and if you think this light is the correct light (pointing right) you push its button (indicating right button). Remember, the correct light makes the bell ring. Are there any questions?

For those Ss who failed to reach the criterion of six successive correct responses, trials were terminated after 72 trials. Children who achieved criterion were immediately continued for a block of 18 trials, post-criterion, if they were in the no-interval group. If they had been randomly assigned to the interpolated interval group however, then the interpolated task was administered before the 18 overtraining trials. The inter-trial intervals were approximately six seconds. This was the time needed to reset the selector switches for the next trial.

Procedure for the interval group was as follows: Upon S's reaching criterion, E said, "Before we go on with this puzzle, I have some pictures I want to show you." E then crossed to the other side of the classroom, indicating the chair E wished the child to occupy. When S was seated, E said, "Tell me the names of the objects you see." The interpolated activity consisted of naming the 16 picture vocabulary cards of the 1960 revision of the Stanford-Binet, Form L-M. This series was repeated three times with an interval of three seconds between each picture. If S could not name the object, E named it. Upon completion of this activity, E returned to the apparatus saying, "Now, let's go back to the puzzle." The elapsed time for this entire procedure was five minutes.

Subjects who did not attain criterion in 72 trials were returned immediately to their classroom. All Ss were shown the final trial and asked to verbalize the correct solution, upon completion of the 18 post-criterial trials, or the 72 trials, for those who did not reach criterion. E did not comment as to

whether or not the verbalized response was correct. All children were cautioned about discussing the puzzle or its solution, with their friends, before they were returned to their classroom.

III

RESULTS

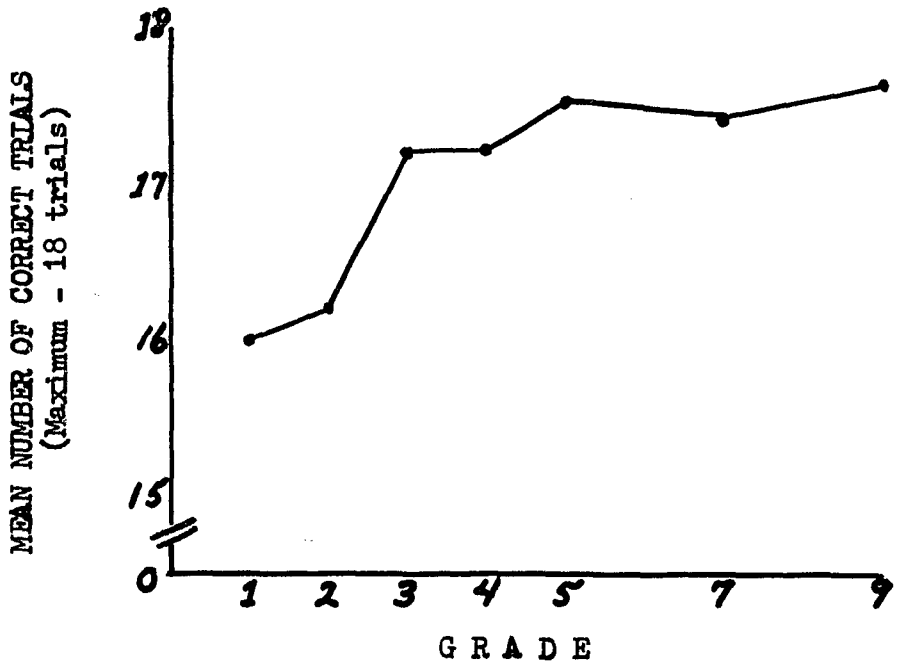
The original plan of this study was to test performance on the oddity task on children from grades 1 through 5. However, as the original oddity data indicated that performance increased steadily over these grades, it was decided to add two further grades in an effort to ascertain whether an asymptote in performance had been reached. Since the task imposed during the interpolated interval was not an age-appropriate one for these additional students from junior high school grades 7 and 9, no interpolated task was used. In addition, because the interpolated interval was not a significant main effect in original learning for either trials to criterion (Table 7, Appendix), or percent correct responses (Table 8, Appendix), the data for grades 1 through 5 for groups with and without the interpolated task were combined.

Post-Criterial Trials

Figure 1 indicates the mean number of post-criterial trials correct for the criterion groups over grades 1 - 9. These data were analyzed by analyses of variance in terms of grades 1 - 5 (Table 9, Appendix) and grades 1 - 9 (Table 10, Appendix). These two analyses are given to indicate that the results of the study were not significantly changed by the addition of the junior high school grades 7 and 9. For both these analyses, the main effects of Grade, Difficulty and Sex were significant at the .01 level.

FIGURE 1

MEAN NUMBER OF POST-CRITERIAL TRIALS CORRECT
FOR THE GROUPS OF CHILDREN



These analyses support the prediction regarding the lower level of functioning of the younger subjects on post-criterial performance. Figure 1 shows that the mean number of post-criterial trials correct increases from a low of 16.0 for Grade 1 to a high of 17.6 for Grade 9.

The prediction regarding stimulus similarity is also supported by the data. Subjects in the Difficult (red-pink-orange) group had fewer post-criterial trials correct (16.2) than did those having the more discriminable stimuli (red-blue-green), who had a mean of 17.6 trials correct. An analysis of errors for grades 1 - 9 indicates that almost three times as many errors were made by the Difficult groups as compared with the Easy groups, and that a significantly greater number of subjects in the Difficult groups made errors. This is shown in Table 2.

Sex has also been shown to be a significant main effect in the post-criterial trials series. A correlation between Sex and post-criterial trials correct was found to be significant at the .01 level, but it was small (+.11). An analysis of post-criterial errors indicates that more girls made more errors than boys in these trials (Table 2). These sex differences are in line with previous findings in the literature (Martin & Blum, 1960, 1961).

An analysis of covariance was computed to partial out the effects of pre-criterial on post-criterial performance. This analysis of number of trials to criterion and post-criterial trials correct, yielded the data shown in Table 3. It can be

TABLE 2

NUMBER OF POST-CRITERIAL ERRORS (GRADES 1 - 9)
AS A FUNCTION OF GRADE, DIFFICULTY AND SEX
(Maximum Errors - 288/Cell)

EASY (R-B-G)					
	FEMALE			MALE	
Grade	Errors	N	Grade	Errors	N
1	35	4	1	9	3
2	25	4	2	1	1
3	11	5	3	2	2
4	0	0	4	0	0
5	1	1	5	2	2
7	2	2	7	1	1
9	14	6	9	22	8
Total	88	22		37	17
TOTAL, EASY: Errors= 125; N= 39					
DIFFICULT (R-P-O)					
1	51	12	1	31	12
2	55	11	2	32	7
3	25	16	3	16	9
4	29	11	4	25	12
5	15	10	5	12	9
7	0	0	7	0	0
9	11	8	9	13	6
Total	186	68		130	55
TOTAL, DIFFICULT: Errors=316; N= 123					
TOTAL FEMALE	274	90	TOTAL MALE	167	72

TABLE 3

ANALYSIS OF COVARIANCE

MEAN NUMBER OF TRIALS TO CRITERION AND

POST - CRITERIAL TRIALS CORRECT

GRADES 1 - 9 (N = 448)

Source	d. f.	Mean Square	F
Total	447	-	-
Grade	6	1555.97	6.21 **
Difficulty	1	9807.15	39.12 **
Sex	1	1659.22	6.62 *
Trials to Criterion	1	4253.30	16.96 **
Grade x Difficulty	6	225.03	0.90 +
Grade x Sex	6	0.00	0.00 +
Difficulty x Sex	1	6.67	0.03 +
Grade x Difficulty x Sex	6	59.35	0.24 +
Error	419	250.71	-

* $P < .05$

** $P < .01$

+ 1/F

seen that the main effects of Grade and Difficulty are significant at the .01 level, while that of Sex is significant at the .05 level. The covariate, trials to criterion, is also significant at the .01 level. These indicate that the significant main effects found in the analysis of post-criterial trials correct are independent of the effects of pre-criterial performance. Since no significant interactions were found in the covariance analysis, we may conclude that the interactions obtained in the post-criterial analysis for grades 1 - 5 (Table 9, Appendix) were a function of interactions persisting from pre-criterial performance.

Original Oddity Learning

Criterion groups: The data on original learning for the criterion groups were analyzed in terms of two measures: number of trials to criterion, and percent correct responses.

Figure 2 indicates the mean number of trials to criterion for the criterion groups in grades 1 - 9. It seems to indicate a shift in performance level between Grades 1 and 2, followed by little or no change over Grades 2, 3 and 4. A similar process is indicated by the shift from Grades 4 to 5, where the performance level of Grade 5 does not depart significantly from that of Grades 7 and 9 in terms of number of trials to criterion. Figure 3 shows this same data in terms of cumulative percent of students attaining criterion over blocks of six trials. More students in Grades 5, 7 and 9 attain criterion performance sooner than those in Grades 1 through 4.

The trials to criterion data for Grades 1 - 5 and 1 - 9

FIGURE 2

MEAN NUMBER OF TRIALS TO CRITERION
FOR THE GROUPS OF CHILDREN

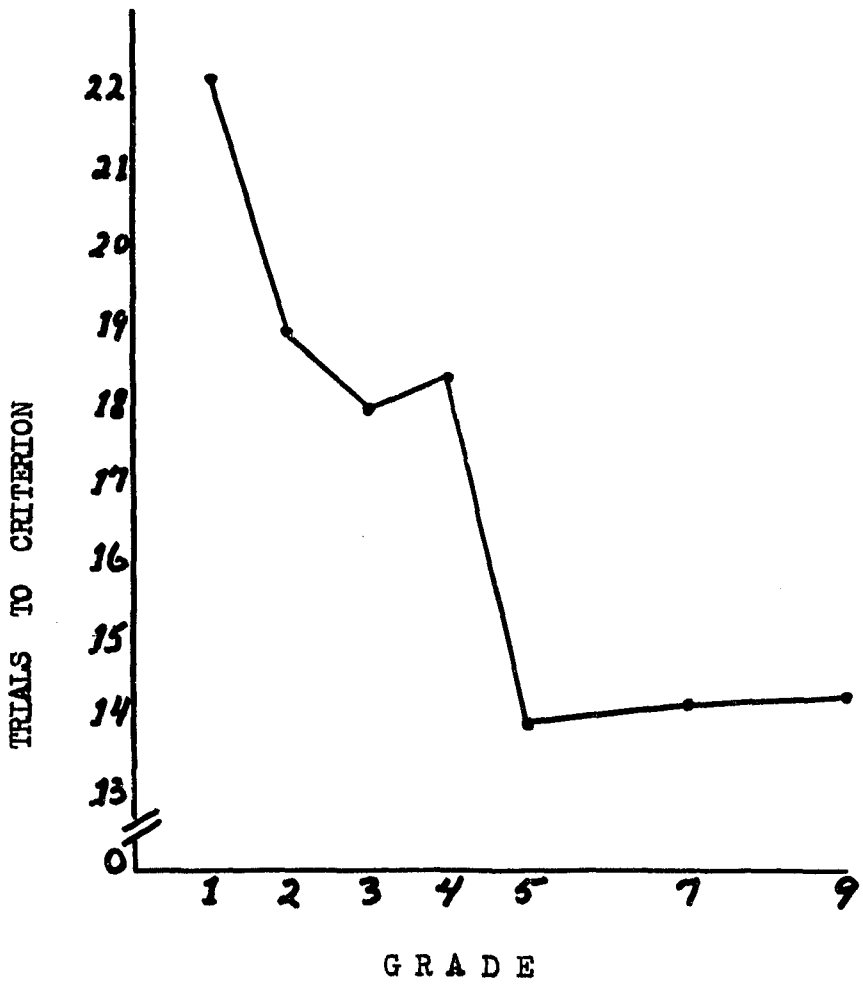
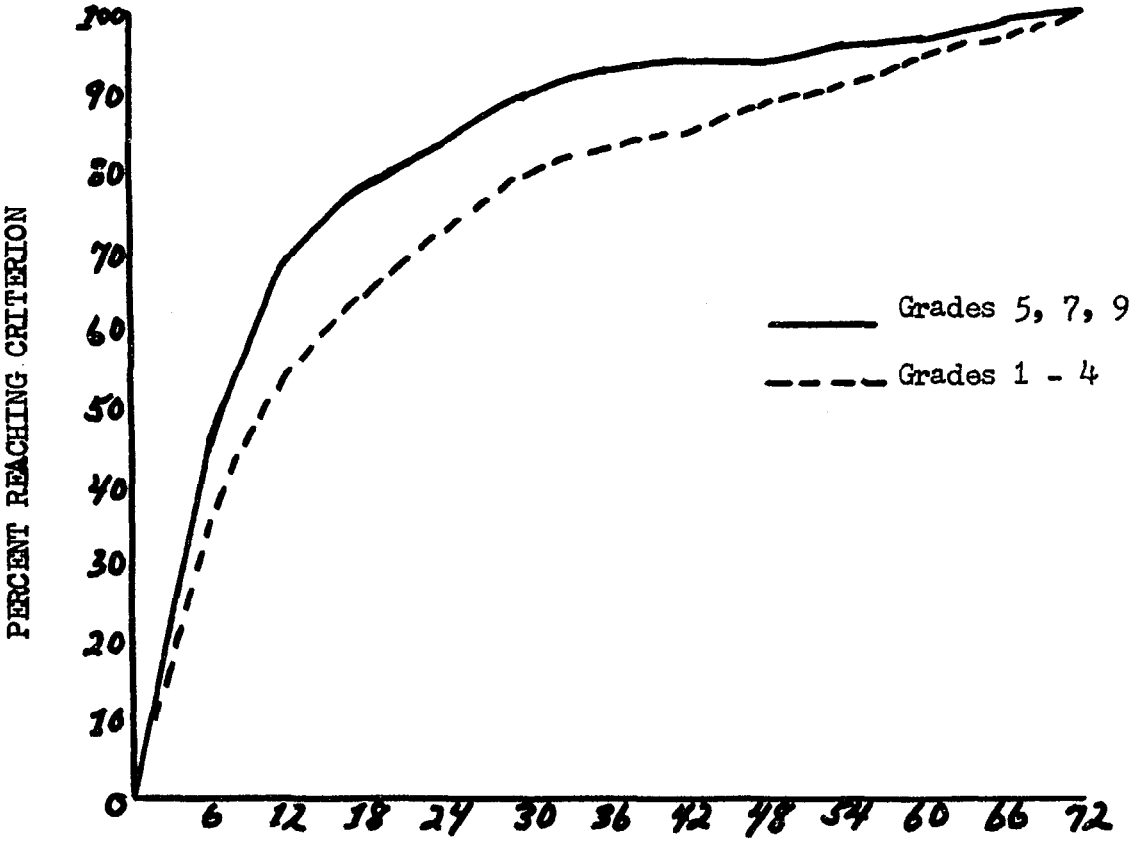


FIGURE 3

CUMULATIVE PERCENTAGE CURVE
MEAN NUMBER OF TRIALS TO CRITERION



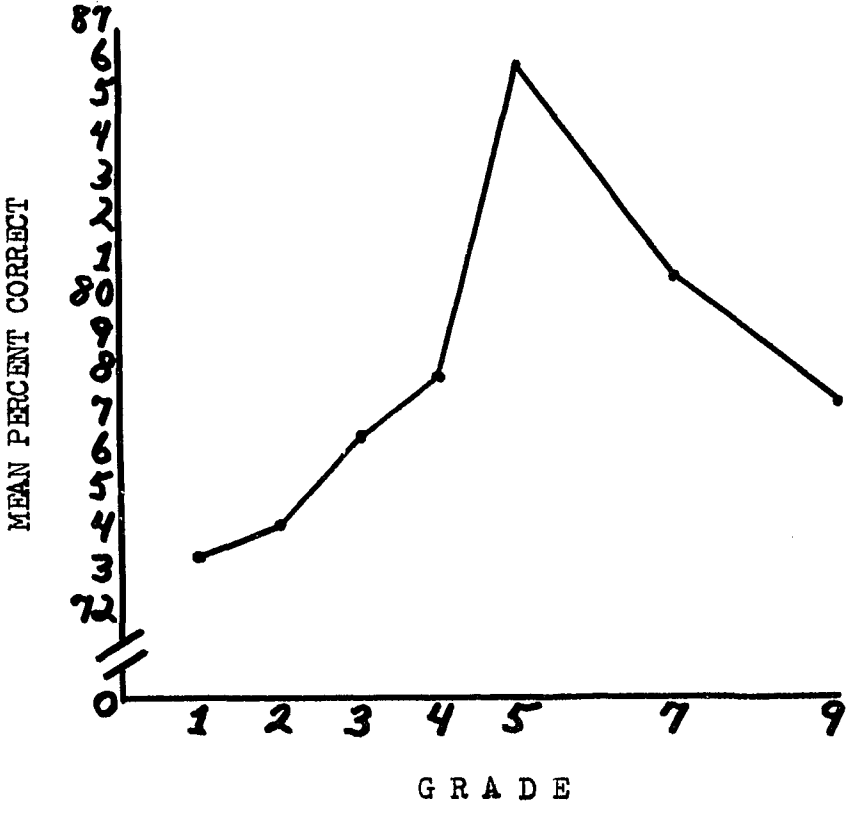
were analyzed by the analysis of variance technique (Tables 7 and 11, Appendix). While Grade was not found to be a significant main effect for Grades 1 - 5, it was significant for Grades 1 - 9 at the .05 level. Difficulty was a significant main effect at the .01 level in both analyses. For Grades 1 - 5 the means are as follows: Easy (red-blue-green) 15.5 trials; Difficult (red-pink-orange) 21.0 trials. For grades 1 - 9 the Easy condition had a mean of 14.2 trials whereas the Difficult group had a mean of 19.9 trials.

Figure 4 depicts the mean percent of trials correct for the criterion group over grades 1 - 9. It can be seen that these data depart from the linear function predicted in the oddity learning literature (Ellis & Sloan, 1959; Martin & Blum, 1960,1961; House, 1964, a.,b.; Lipsitt & Serunian, 1964; and Gollin & Shirk, 1966). While performance shows an increase over grades 1 - 5, reaching a mean of 86.1% trials correct at Grade 5, Grade 7 attained a mean of 80.7% and Grade 9 a mean of 77.8% correct. Thus not only was an asymptote not reached at Grade 5 in terms of this measure, but poorer performance was found in Grades 7 and 9.

This measure, percent of trials correct, has also been analyzed by the analysis of variance technique in terms of grades 1 - 5 and 1 - 9 (Tables 8 and 12, Appendix). Both analyses find Grade to be a significant main effect at the .05 level, and Difficulty a significant main effect at the .01 level. While no significant interactions were found for grades 1 - 9, a rather confusing interaction was found for grades 1 - 5 as a function of Difficulty and Interval ($p < .01$). While there was no signifi-

FIGURE 4

MEAN PERCENT CORRECT TRIALS, PRE-CRITERION,
FOR THE CRITERION GROUP
ACCORDING TO GRADE.



cant difference between means for the No Interval group (77.3% Easy; 76.1% Difficult), wide variability in performance existed for those groups having the interpolated interval (86.2% Easy; 71.2% Difficult).

With regard to this pre-criterial oddity learning, the effect of Sex on performance is somewhat different from that in post-criterial trials. Using the measure of number of trials to criterion, Sex is not a significant main effect. No significant sex differences were found in terms of percent of trials correct either. Here, in dealing with original oddity learning, we find that the data differ from the findings in the literature that males generally perform better at this task (Martin & Hlum, 1960, 1961), and also the differences found in the post-criterial trials of this study.

Both measures used, trials to criterion, and percent correct trials, yielded a significant main effect for Difficulty ($p < .01$). Thus the prediction with regard to stimulus similarity has also been supported by the data in original learning, with students having the less discriminable stimuli (red-pink-orange) performing in a significantly poorer manner than those with the easily discriminable stimuli (red-blue-green). This effect was also apparent in the post-criterial trials.

Grade was found to be a significant main effect for the percent correct measure only. Performance, as depicted in Figure 4, increases steadily for grades 1 - 5, but performance for Grades 7 and 9 falls to significantly lower levels than Grade 5 performance. Performance in oddity, in terms of number of trials to

criterion, reaches an asymptote at Grade 5 (Figure 2). This finding, taken in conjunction with that for percent correct, seems to indicate that this task is optimal for Grade 5, but beyond that grade, some extraneous factor confounds performance.

Interpolated Interval was found to interact significantly with Difficulty in terms of percent correct trials. There is wide variability in performance between the Easy and Difficult groups having the interval (86.2% and 71.2% respectively) as compared with groups not having the interval (77.3% and 76.1%). This wide variability among the groups in original learning was undoubtedly the major factor in the Interval x Difficulty x Sex ($p < .01$) interaction in post-criterial trials, since no significant interactions were obtained in the analysis of covariance. The interpolated interval was not introduced until after the criterion was met, thus it was not anticipated that it would be a significant main effect, nor that it would interact significantly with any other variable in pre-criterial performance. Since subjects were randomly assigned to Interval and No Interval groups, it is difficult to account for this interaction in original learning. The Interval x Difficulty interaction is sufficient to cancel Interval as a variable in this study since some obvious, inadvertent confounding is present.

All subjects: A group of 546 Ss was utilized to fill the cells in order to meet the statistical convention for the analysis of variance technique. It was found that the number of subjects who had to be tested to attain an N of eight subjects meeting criterion per cell varied over grade (see Table 1). There was

a decrease in the number of subjects tested at each grade after Grade 2. It is also evident that fewer male subjects were needed at each grade level, except for Grade 2. More Second Grade subjects were needed to obtain the criterion group than any other grade used. This confirms some of the other findings of the present study with regard to sex and grade differences.

Table 4 shows a comparison of mean percent correct responses as a function of Grade for the criterion and non-criterion groups, and for these groups combined. For the non-criterion group, all of the grades attained mean scores just above chance performance (33-1/3%). Those in the criterion group have attained percents of 72 and over. It can also be seen that all Grade 2 students had the lowest mean percent correct, although Grade 2 criterion group performance was superior to that of Grade 1.

Table 5 indicates the number and percent of subjects at each grade level who were able to verbalize knowledge of solution of the oddity problem for both the criterion and non-criterion groups. The children were asked, following their final trial (72 for the non-criterion group, and the last post-criterial trial for those achieving criterion), why they had chosen a particular stimulus light as correct. A greater percent of the children in the criterion group were able to verbalize the oddity concept than in the non-criterion group, and there was no overlap in terms of these values.

Table 6 is a fourfold table showing whether verbalization of solution was related to the occurence of post-criterial errors. A chi square analysis shows the differences to be significant

TABLE 4

COMPARISON OF MEAN PERCENT CORRECT RESPONSES BY GRADE FOR CRITERION, NON-CRITERION AND ALL SUBJECTS

Grade	Criterion (N = 448)	Non-Criterion (N = 98)	All (N = 546)
1	71.64%	42.03%	66.41%
2	74.26	41.61	64.83
3	76.58	38.38	67.83
4	78.12	40.69	73.06
5	86.08	39.58	80.91
7	80.70	35.34	75.11
9	77.80	44.44	74.51

TABLE 5

NUMBER AND PERCENT OF SUBJECTS VERBALIZING KNOWLEDGE OF SOLUTION FOR CRITERION AND NON-CRITERION GROUPS

Grade	Criterion		Non-Criterion	
	N	%	N	%
1	56	88	1	5
2	54	84	2	8
3	60	94	2	11
4	63	98	0	0
5	64	100	2	25
7	62	97	0	0
9	64	100	2	29

TABLE 6

VERBALIZATION OF SOLUTION AS RELATED TO POST-CRITERIAL ERRORS
GRADES 1 - 9

		Verbalization of solution		
		Yes	No	Total
Post-Criterial Trials	Errors	151	24	175
	No Errors	272	1	273
	Total	423	25	448

Chi Square = 34. ($p < .01$)

at the .01 level. One-third of the subjects who did verbalize the concept made post-criterial errors, while 24 of the 25 Ss who did not verbalize the concept made post-criterial errors. With regard to the criterion group for grades 1 - 9, a significant correlation ($p < .01$) of +.64 was obtained between the number of post-criterial trials correct and verbal report of the oddity concept. This supports the data shown in Table 6.

Possible factors determining performance: In order to amplify oddity problem learning in normal children, the data were subjected to correlational analyses in terms of all the variables. High significant correlations for all subjects were obtained for the following variables:

Grade and age	+.98
Grade and reading score on Metropolitan Achievement Test	+.76
Age and reading score on Metropolitan Achievement Test	+.75.

Statistically significant correlations were found between the Metropolitan Achievement Test and trials to criterion (-.22), post-criterial trials correct (+.23), and knowledge of the oddity concept (+.25). Since these correlations are of relatively small magnitude, their meaningfulness is very limited, although they are consistent with expectations that led to the use of the Metropolitan Achievement Test.

IV

DISCUSSION

The results of this study suggest that criterion measures require further examination, especially when young children are used as subjects. In studies concerned with original learning only (Lipsitt & Serunian, 1963; Lipsitt & LoLordo, 1963; Gollin & Shirk, 1966), when a subject attained criterion it was assumed that all subsequent trials to the maximum would be correct, and the measure of performance used, number of correct trials, included all those trials between the final criterial trial and the maximum. The findings of the present study, which show that errors are often made in post-criterial performance, suggest that percent correct trials of the trials administered might be a more sensitive measure. This is based on the actual performance of the subjects and requires no assumptions as to the S's subsequent performance.

The data also seem to indicate that the criteria imposed in learning studies should vary with the group being tested. Possibly one should persist as long as the subject will tolerate the test situation, for early termination of testing, as in a criterion of six successive correct trials, does not seem to give the fullest picture of performance, especially with young children. Continuing for an indefinite number of trials beyond the first successful trial can, however, also reflect the consequences of boredom, fatigue, lack of attention, etc..

In studies concerned with transfer of training, the findings

of the present study indicate that the adequacy of the transfer will be a function of the criterion used in original learning. A criterion such as the one used in this study, six successive correct trials, might not indicate whether the concept involved has been learned and understood. The high correlation obtained between number of post-criterial trials correct and knowledge of the oddity principle (+.64), indicates that perhaps some method of ascertaining whether or not the concept is understood should be included in transfer studies.

One method of assessing whether the subject is cognizant of the concept involved would be to have S verbalize the solution, as in this study, following completion of training performance. Many child studies involving transfer employ tasks which require little more than rote memorization of the solution. Katona (1940) found that this tended to narrow rather than increase the range of problem solving or transfer. His conclusions, that learning with understanding not only improved retention of what was learned, but also qualified the learner to move forward to new learning, stresses the importance of understanding for transfer. Asking S to verbalize the solution prior to transfer also gives the experimenter a means of ascertaining whether or not experimental artifacts play a part in the transfer performance.

The data (Table 6) show that if S can't verbalize the solution even after the performance criterion is achieved, S makes errors in post-criterial trials. If S can verbalize the solution, S is less likely to make errors. However, of the 423 subjects who can verbalize the oddity principle, about one-third

make incorrect responses in the post-criterial trials. Why does this happen? Obviously the ability to verbalize the solution is not the complete answer to the maintenance of 100% performance.

We are also faced with another question regarding our research objective. What is our real interest in studying the subject's performance-- his knowing what principle may be involved, or his 100% accuracy? This reflects again on the importance of criteria to evaluate learning, and seems to indicate that rather than a criterion of a certain number of successive correct responses, one based upon a percentage of the total trials correct might be more reliable or consistent. It is a question of what we assume has been learned, and its persistence and transfer.

Figure 5, depicting mean percent correct trials for the criterion group, indicated that performance of grades 1 - 5 increases steadily, but performance for Grades 7 and 9 falls to significantly lower levels than Grade 5 performance. While oddity performance, in terms of number of trials to criterion, reaches an asymptote at Grade 5 (Fig. 3), it is possible that the older subjects, with their greater level of intellectual sophistication, might be generating and testing more hypotheses such as object or position (Harlow, 1951), or position patterning of responses (LMR,RML; LMR,LMR; LRRL;LRLR; etc.) which Weir (1964) has indicated. This might also account for the increase in number of post-criterial errors and the larger number of subjects making them in Grade 9, following a steady decline in both number of errors and subjects making them for Grades 1 - 7 (Table 2). Further study of the oddity problem with 11th grade and college

sophomore groups might provide us with further useful information about this trend.

Another possibility that might account for the poorer performance is the attitude of the subject toward the experiment. In a problem-solving situation, the subject may be testing the limits the experimenter will permit. "If the bell rings when I press the button under the odd-colored light, will it also ring if I press the button under one of the matching lights?"

Boredom might be a factor in the increase in post-criterial errors and the number of subjects making them in Grade 9. Does the subject feel he has solved this problem to his own satisfaction, thus the remaining trials do not matter and he ceases to attend to the stimuli? We speak of neural satiation for stimuli. Perhaps we should also consider the possibility of intellectual satiation for a response.

Lewin's study (1935) of normal and feebleminded children drawing "moon faces" is concerned with the effect on performance of satiation for a particular task. Lewin calls this "psychical satiation." Both the normal and feeble-minded children (ages 8 - 11 years old), reaching satiation for the drawing of "moon faces," manifested secondary actions, which Lewin defines as "actions which are carried out on the side without interruption of the major activity (p. 200)." Mean satiation time for the normal group declined from 55 minutes drawing time for 8 - 10 year olds, to 45 minutes for the 10 - 11 year olds. However, the mean number of moon faces drawn declined steadily from 8 to 11 year olds. Perhaps in a more complex task such as oddity problem

solving, the satiation effect might be manifested at a later age, corresponding to the 9th grade in the present study. The satiation effect may indicate more about the subjects than their knowledge of the problem. For the organism with a low threshold for satiation, too many post-criterial trials, or over-extended criteria, may lead to the breakdown in performance that has been found in the post-criterial trials of this research, and in those studies mentioned by McConnell (1965). Thus there is some disadvantage in using overlearning as a variable. Because of the satiation effect, it may be inappropriate for some subjects. Perhaps in studies of overlearning, we might be able to counteract the effects of satiation through changes in the stimulus objects or the response, the use of increased or novel reinforcement, or through other manipulations to heighten motivation and attention.

The findings of this study with regard to sex have done nothing to clarify this variable in oddity performance. In original oddity learning, no significant sex differences were found, but in the post-criterial trials, performance of male subjects was superior to that of females. The latter finding is in keeping with the Martin and Blum findings (1960, 1961). On the basis of the differences between training and post-criterial performance, it is possible that the motor response (button pressing) might have proved fatiguing to the girls after a number of trials, thus making them more distractable because they were uncomfortable. Possibly other sex differences in the literature of oddity performance might be traced to the physical effort in-

volved in the task (although this does not seem to be the case in the Martin and Blum studies), and thus be due to differences in musculature and strength between males and females. It is also possible that the satiation effect may affect males and females differentially, due to their unequal rates of maturation. This, of course, would need some direct test to substantiate the tentative conclusions.

The most fruitful approach to studying the development of cognitive behavior in normal human subjects might be an analysis of the types of errors emitted by S in his responding. This research strategy might yield clues to the hypotheses S is generating as he proceeds to what the experimenter considers the correct solution to the problem. An attempt was made to analyze performance in terms of error factor analysis (Moon and Harlow, 1955) but, because of the three-alternative nature of the present task, the number of hypotheses which could be generated in this study far exceeded those considered in their studies. Not only is error analysis an exceedingly complex issue, but it might also give us an untrue picture of what is going on within the subject. Moon and Harlow are classifying errors in terms of the experimenter's orientation, rather than the subject's orientation, e.g., a perseverative error, implying position, does not necessarily mean the subject is working in terms of positional cues. Perseveration may be a function of anxiety about the task, or some other factor in the S's environment on the day of the test, rather than a positive response to any aspect of the test. It could also be interpreted cognitively

as an inability to shift the frame of reference and reorient to each new choice situation as a separate entity. Thus we should expect the younger child, functioning on a concrete level and not utilizing logical operations (Inhelder & Piaget, 1958) to be more bound by what he first perceives as the stimulus situation. In criterion performance, the correct response in young children may be a function of stimulus perseveration, in that the child has been fortunate enough to have selected the correct response, and because of his stimulus boundedness or perseverative tendencies, is unable to shift to what is deemed an incorrect response.

Sigel (1963) in speaking of possible approaches to IQ scoring, lists styles of categorizations, qualitative changes in type of response as a function of age and sex, and finally, the analysis of errors. He states that errors are manifestations of the quality of intellective operations, rather than merely indications of the child's lack of knowledge, especially in situations where choices are provided, as is the case in the present study. Sigel argues that the number of correct responses does not provide a complete measure of what the subject can and does do. Experience with objective tests such as multiple choice and true or false types, has shown that the student with the more comprehensive and advanced knowledge is sometimes penalized (Hoffman, 1962). So, too, in a complex learning situation, the more mature individual may be penalized by the number of possible solutions he can generate to the problem because of the greater knowledge he possesses. In this study, data showing that 5th, 7th and 9th graders take approx-

imately the same number of trials to reach criterion, but that 7th and 9th graders have a lower percentage of correct responses (80 and 78% respectively) than 5th graders (86%) may be indicative of this process at work. The findings seem to indicate that the oddity task is optimal for Grade 5, but that beyond that, some extraneous factor (satiation, intellectual sophistication, cognitive style, etc.) confounds performance.

Hebb (1958, p.454) points out that in a simple learning task "the lower animal gets there first." We might paraphrase Hebb's statement and say that perhaps the higher animal gets there last because he interprets the simple learning task as being more complex, or, in terms of satiation, prefers to make a simple, boring task more complex and interesting.

Sigel (1963) states that error analysis of IQ tests could contribute to our understanding of cognitive development. Developmental curves based on errors might reveal the quality of intellectual functioning and the kinds of difficulties the child is having. These curves could also reveal how children, with increasing age, alter their approaches to the cognitive problems involved. A conceptualization of the types of errors would be required, however, so that an incorrect response could be identified as having particular significance.

Another way to address this problem of mapping cognitive development might be in terms of a developmental-experimental approach such as that propounded by Gollin (1964). This Levels x Levels design would vary the organismic variable, stage or development (or age) with specially-programmed levels of stimulus

patterns which would delineate alternative hypotheses or "errors" and reveal the effects of ontogenesis upon cognitive behavior.

V

SUMMARY

An oddity task employing colored lights as stimuli was administered to 546 public school children in grades 1 - 9. Performance in 18 post-criterial trials was analyzed for the 448 Ss who attained criterion, to ascertain whether or not children continue to maintain the level of mastery in overlearning. Ss in grades 1 - 5 were tested to a criterion of six successive correct trials in terms of two conditions of the variables of stimulus similarity, interpolated interval and sex; interpolated interval was omitted for grades 7 and 9.

Oddity learning was selected to study the phenomenon of post-criterial performance since it is a form of multiple sign learning requiring the subject to respond to relations among the simultaneously-presented stimuli. This is more complex than a simple discrimination problem where response to a specific stimulus is consistently reinforced. Thus, if a breakdown in performance in overlearning might be anticipated with conditions involving greater complexity and attention, it should be expected that this would result with higher frequency in oddity learning.

It was further hypothesized that the less discriminable of the two stimulus conditions, red-pink-orange lights, as opposed to red-blue-green lights, would present greater difficulty for the child Ss. With regard to interpolated interval, a prediction in terms of Piaget's functional invariants of assimilation and accommodation would state that the interval should prove confounding

to those children who had not mastered the concept underlying oddity performance. The variables of grade level and sex were utilized to clarify oddity performance in normal children, in view of conflicting reports in the literature.

With regard to post-criterial performance, the prediction regarding the lower performance level of the younger Ss was borne out by the data, as was that for stimulus similarity. Sex was also found to be a significant main effect. The effect of interpolated interval was not clarified by the data since it was not a significant main effect. It is possible that the interval was not of sufficient length. However, a significant correlation was found between number of post-criterial trials correct and knowledge of the oddity principle as verbalized by the Ss (+.64) and lends support to the prediction made. Sex differences found in post-criterial trials indicated that the performance of males surpassed that of females.

In pre-criterial learning, the effect of sex upon performance differed, since there was no significant difference between male and female performance in terms of the two measures utilized-- number of trials to criterion and percent correct trials. Difficulty was found to be significant for both measures, indicating that Ss having the less discriminable stimuli performed in a poorer manner. While grade was a significant main effect for percent correct trials only, the curves for both measures indicate the shape of the oddity function for normal children is not a straight line.

The results suggest that criterion measures require further examination, especially when young children are Ss. They further

suggest percent correct trials as a more sensitive measure of performance, in comparison with the more common measure of number of correct trials, since the former includes only S's actual performance, and does not assume all trials between criterion and the final trial to be correct.

Results were discussed in terms of the concept of satiation. Analysis of types of errors emitted was advanced as a fruitful method of studying the development of cognitive behavior in normal human Ss.

VI

REFERENCES

Dodge, R. Conditions and consequences of human variability. New Haven: Yale University Press, 1931.

Ellis, N. R. & Sloan, W. Oddity learning as a function of M. A.. Journal of Comparative & Physiological Psychology, 1959, 52, 228-230.

Friedman, S. R. Developmental level and concept learning: confirmation of an inverse relationship. Psychonomic Science, 1965, 2, 3-4.

Furth, H. G. and Youniss, J. Effect of overtraining on three discrimination shifts in children. Journal of Comparative & Physiological Psychology, 1964, 57, 290-293.

Gollin, E. S. A developmental approach to learning and cognition. In Lipsitt, L. P. & Spiker, C. C. (Eds.), Child development and Behavior. New York: Academic Press, 1964, Pp. 159-186.

Gollin, E. S. Reversal learning and conditional discrimination in children. Journal of Comparative & Physiological Psychology, 1965, 58, 441-445.

Gollin, E. S. & Shirk, E. J. A developmental study of oddity-problem learning in young children. Child Development, 1966, 37, 213-217.

Gollin, E. S., Saravo, A., & Salten, C. Perceptual distinctiveness and oddity-problem solving in children. Journal of Experimental Child Psychology, 1967, 5, 586-596.

Harlow, H. Primate learning. In C. P. Stone (Ed.), Comparative psychology. New York: Prentice-Hall, 1951, Pp. 183-238.

Harlow, H. The evolution of learning. In A. Roe & G. G. Simpson (Eds.), Behavior and evolution. New Haven: Yale University Press, 1958, Pp. 269-290.

Hebb, D. O. Biological & biochemical bases of behavior. Madison: University of Wisconsin Press, 1958.

Hoffmann, Banesh The tyranny of testing. New York: Crowell-Collier Press, 1962.

House, B. J. Oddity performance in retardates: I. Acquisition and transfer. Child Development, 1964, a, 35, 635-643.

House, B. J. Oddity performance in retardates: II. Size discrimination functions from oddity and verbal methods. Child Development, 1964,b, 35, 645-651.

Hull, Clark L. Principles of behavior. New York: D. Appleton-Century Co., Inc., 1943

Inhelder, B. & Piaget, J. The growth of logical thinking. New York: Basic Books, 1958.

Katona, G. Organizing and memorizing. New York: Columbia University Press, 1940.

Lashley, K. S. Conditional reactions in the rat, Journal of Psychology, 1938, 6, 311.

Levinson, B. Oddity learning and its relation to discrimination learning set. Unpublished doctoral dissertation, University of Wisconsin. 1958.

Lewin, K. A dynamic theory of personality. New York: McGraw-Hill, 1935.

Lipsitt, L. P. Simultaneous and successive discrimination learning in children. Child Development, 1961, 32, 337-347.

Lipsitt, L. P. & LoLordo, V. M. Interactive effect of stress and stimulus generalization on children's oddity learning. Journal of Experimental Psychology, 1963, 66, 210-214.

Lipsitt, L. P. & Serunian, S. A. Oddity-problem learning in young children. Child Development, 1963, 34, 201-206.

Martin, W. E. & Blum, A. Effects of irrelevant cues on intertest generalization in discrimination learning of normal children. Psychological Reports, 1960, 7, 399-404.

Martin, W. E. & Blum, A. Intertest generalization and learning in mentally normal and subnormal children. Journal of Comparative & Physiological Psychology, 1961, 54, 28-32.

McConnell, J. V. Cannibals, chemicals and contiguity. Animal Behavior, Supplement I, 1965, 61-68.

Miller, N. E. & Dollard, J. Social learning and imitation. New Haven: Yale University Press, 1941.

Nissen, H. W. & McCulloch, T. L. Equated & non-equated stimulus situations in discrimination learning by chimpanzees. II. Comparison with limited responses. Journal of Comparative Psychology, 23, 365-376.

Pastore, N. Discrimination learning in the canary. Journal of Comparative & Physiological Psychology, 1954, 47, 389-390.

Piaget, J. The origins of intelligence in children. New York: International Universities Press, 1952.

Robinson, E. W. A preliminary experiment on abstraction in a monkey. Journal of Comparative Psychology, 1933, 16, 231-236.

Shirk, E. J. Oddity problem learning as a function of grade and previous school experience. Unpublished research, 1964.

Sigel, I. E. How intelligence tests limit understanding of intelligence. Merrill-Palmer Quarterly, 1963, 9, 39-56.

Spence, K. W. The nature of the response in discrimination learning. Psychological Review, 1952, 59, 89-93.

Spiker, J. S. & Holton, R. B. Associative transfer in motor paired-associate learning as a function of amount of first-task practice. Journal of Experimental Psychology, 1959, 56, 123-132.

Stanford-Binet Intelligence Scale, Form L-M. Boston: Houghton Mifflin, 1960.

Stevenson, H. W. & Weir, M. W. Response shift as a function of overtraining and delay. Journal of Comparative & Physiological Psychology, 1959, 52, 327-329.

Warren, J. M. Oddity learning set in a cat. Journal of Comparative & Physiological Psychology, 1960, 53, 433-434.

Weir, M. W. Developmental changes in problem-solving strategies. Psychological Review, 1964, 71, 473-490.

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

Youniss, J. & Furth, H. G. Reversal learning in children as a function of overtraining and delayed transfer. Journal of Comparative & Physiological Psychology, 1964, 57, 155-157.

Zeaman, D. & House, B. J. An attention theory of retardate discrimination learning. In N. R. Ellis (Ed.) Handbook of mental deficiency. New York: McGraw-Hill, 1963, Pp. 159-223.

VII

APPENDIX

ABSTRACT

Unpublished Study, E. J. Shirk, 1964

Oddity Problem Learning as a Function of Grade and Previous School Experience.

SUBJECTS: Ss were 24 first and second grade elementary school students in two newly-integrated schools of a suburban New York public school system.

APPARATUS: This was identical with that used in Gollin and Shirk, 1966, and the present study.

PROCEDURE: Procedure was originally identical with that used in Gollin and Shirk, 1966. However it was later decided to run Ss beyond the previous criterion of six successive correct trials because of performance irregularities observed. Therefore only 12 of the Ss were run to a criterion of more than six successive correct trials.

RESULTS: Of the 12 Ss run to a criterion of more than six successive correct trials, 5 Ss (42%) made errors on the first post-criterial trial. Only one S out of the five showing a breakdown in performance, succeeded in re-achieving the criterion during the maximum number of trials (54).

Due to the performance breakdown observed, it was decided to discontinue the use of the oddity problem as a cognitive measure in the larger study.

ORDER OF PRESENTATION OF STIMULI

Trial	EASY(red-blue-green)	DIFFICULT(red-pink-orange)
1.	R G R	R P R
2.	R R B	R R O
3.	R G G	R P P
4.	G B G	P O P
5.	B B G	O O P
6.	R B R	R O R
7.	B G B	O P O
8.	R B B	R O O
9.	R R G	R R P
10.	B G G	O P P
11.	G G R	P P R
12.	G R G	P R P
13.	B R R	O R R
14.	G R R	P R R
15.	B R B	O R O
16.	G B B	P O O
17.	G G B	P P O
18.	B B R	O O R

TABLE 7

ANALYSIS OF VARIANCE
MEAN NUMBER OF TRIALS TO CRITERION
GRADES 1 - 5 (N = 320)

Source	d.f.	Mean Square	F
Total	319	-	-
Grade	4	536.60	1.83
Difficulty	1	2425.50	8.26 *
Interval	1	502.50	1.71
Sex	1	52.00	5.64 +
Grade x Difficulty	4	126.46	2.32 +
" x Interval	4	429.29	1.46
" x Sex	4	22.43	13.09 *+
Difficulty x Interval	1	735.08	2.50
" x Sex	1	0.004	73382.25 *+
Interval x Sex	1	147.15	2.00 +
Grade x Diff. x Int.	4	192.16	1.53 +
" " x Sex	4	625.73	2.13
" x Int. x Sex	4	404.27	1.38
Diff. x Int. x Sex	1	367.65	1.25
Grade x Int. x Diff. x Sex	4	353.79	1.21
Error	280	293.53	-

* $P < .05$
** $P < .01$
\+ 1/F

TABLE 8

ANALYSIS OF VARIANCE
MEAN PERCENT CORRECT TRIALS - ORIGINAL LEARNING
GRADES 1 - 5 (N = 320)

Source	d.f.	Mean Square	F
Total	319		
Grade	4	15984.05	3.30 *
Difficulty	1	53080.97	10.96 **
Interval	1	3087.02	1.57
Sex	1	4893.08	1.01
Grade x Difficulty	4	3938.11	1.23 +
" x Interval	4	4235.12	1.14 +
" x Sex	4	1088.94	4.45 +
Difficulty x Interval	1	37970.79	7.84 **
" x Sex	1	7359.88	1.52
Interval x Sex	1	122.85	39.44 +
Grade x Diff. x Int.	4	6657.20	1.37
" " x Sex	4	5745.53	1.19
" Int. x Sex	4	8113.29	1.67
Diff. x Int. x Sex	1	6096.84	1.26
Grade x Diff. x Int. x Sex	4	8282.06	1.71
Error	280	4844.86	

* $P < .05$
** $P < .01$
\+ 1/F

TABLE 9

ANALYSIS OF VARIANCE —
MEAN NUMBER OF POST-CRITERIAL TRIALS CORRECT OUT OF 18
GRADES 1 - 5 (N = 320)

Source	d.f.	Mean Square	F
Total	319	-	-
Grade	4	26.80	5.46 **
Difficulty	1	132.62	27.01 **
Interval	1	0.32	15.34 +
Sex	1	43.52	8.86 **
Grade x Difficulty	4	3.92	1.25 +
" x Interval	4	13.29	2.71 *
" x Sex	4	7.38	1.50
Difficulty x Interval	1	13.60	2.77
" x Sex	1	0.00	0.00
Interval x Sex	1	0.10	49.10 +
Grade x Difficulty x Interval	4	5.08	1.03
" " x Sex	4	0.31	15.84 *+
" x Interval x Sex	4	6.36	1.30
Difficulty x Interval x Sex	1	32.53	6.63 **
Grade x Diff. x Int. x Sex	4	3.14	1.56 +
Error	280	4.91	-

* $P < .05$
** $P < .01$
\+ 1/F

TABLE 10

ANALYSIS OF VARIANCE
MEAN NUMBER OF POST-CRITERIAL TRIALS CORRECT
GRADES 1 - 9 (N = 448)

Source	d.f.	Mean Square	F
Total	447	-	-
Grade	6	25.20	6.21 **
Difficulty	1	155.57	38.37 **
Sex	1	26.04	6.42 **
Grade x Difficulty	6	19.49	1.25 +
" x Sex	6	7.98	1.97
Difficulty x Sex	1	0.22	18.18 +
Grade x Difficulty x Sex	6	0.40	10.14 **+
Error	420	4.06	-

* $P \leq .05$
** $P \leq .01$
\+ 1/F

TABLE 11

ANALYSIS OF VARIANCE
MEAN NUMBER OF TRIALS TO CRITERION
GRADES 1 - 9 (N = 448)

Source	d.f.	Mean Square	F
Total	447	-	-
Grade	6	609.06	2.28 *
Difficulty	1	3577.58	13.40 *
Sex	1	255.01	1.05 +
Grade x Difficulty	6	87.49	3.05 +
" x Sex	6	72.59	3.68 *+
Difficulty x Sex	1	61.51	4.34 +
Grade x Difficulty x Sex	6	444.54	1.66
Error	420	267.08	-

* $P < .05$
** $P < .01$
\+ 1/F

TABLE 12

ANALYSIS OF VARIANCE
MEAN PERCENT CORRECT TRIALS - ORIGINAL LEARNING
GRADES 1 - 9 (N = 448)

Source	d.f.	Mean Square	F
Total	447	-	-
Grade	6	11453.91	2.39 *
Difficulty	1	103669.76	21.61 **
Sex	1	4092.75	1.17 +
Grade x Difficulty	6	3644.26	1.32 +
" x Sex	6	1961.37	2.45 +
Difficulty x Sex	1	8057.16	1.68
Grade x Difficulty x Sex	6	3954.41	1.21 +
Error	420	4797.80	-

* $P \leq .05$
** $P \leq .01$
\+ 1/F