



CONCEPT ATTAINMENT IN CHILDREN WITH TYPE
OF PRACTICE AND MEASURES OF SUCCESS VARIED

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
Charlene A. O'Reilly
1963

LIBRARY
Michigan State
University

C 345

L _____

CONCEPT ATTAINMENT IN CHILDREN
WITH TYPE OF PRACTICE AND MEASURES OF SUCCESS VARIED

By

Charlene A. O'Reilly

A THESIS

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

MASTER OF ARTS

Department of Psychology

1963



ABSTRACT

CONCEPT ATTAINMENT IN CHILDREN WITH TYPE OF PRACTICE AND MEASURES OF SUCCESS VARIED

by Charlene A. O'Reilly

The present experiment tests two basic hypotheses: (1) that subjects given practice defining while learning a concept graphically will do significantly better on a terminal defining task than will those who have not had this practice, and (2) subjects learning the concept verbally will do better on the terminal defining task than those who learn graphically, but not as well as those who learn graphically and also have practice defining. The first hypothesis was supported by the data. The test of the second hypothesis showed no significant differences, although the mean differences were in the predicted direction.

The combined data was analyzed further by utilizing both of the two available measures of concept mastery (converging operations) and the agreement between the two measures was found to be significantly greater than chance. The differences noted above were no longer present when this method was used, but it is possible that the methodology used introduced experimental error that obliterated the differences.

by Charlene A. O'Reilly

The methodological conclusions to be drawn from this experiment are:

(1) Ratings can supply an effective dependent variable when collected under proper conditions.

(2) Type of practice as well as amount of practice must be controlled.

(3) Ability to define a concept after labeling it correctly may be a function of the type of conceptual materials used.

Donald M. Johnson
July 19, 1963

ACKNOWLEDGEMENTS

Sincere appreciation and thanks are extended to Dr. Donald M. Johnson for his guidance, encouragement, and consideration during the course of this investigation. Gratitude is expressed to Drs. John R. Hurley and Alfred G. Dietze for their criticisms of this manuscript. Special thanks are extended to the staff members of the schools involved in this study for their assistance in the acquisition of the subjects.

TABLE OF CONTENTS

	Page
INTRODUCTION	1
RECRUITMENT OF SUBJECTS	4
SUBJECTS	5
STIMULI	5
INSTRUCTIONS	8
PROCEDURE	12
RATINGS	13
RESULTS	16
DISCUSSION	22
SUMMARY	26
BIBLIOGRAPHY	28

TABLE OF TABLES

Table		Page
1	Matching data by school systems	7
2	Comparison of frequency of correct defining responses in three groups of 20 subjects each	18
3	Comparison of three groups of 20 subjects each on mean trials to criterion	18
4	Agreement between two measures of concept attainment	19
5	Number of subjects who learned concept as indicated by converging operations and by definition measure alone	21



INTRODUCTION

In previous experiments on concept mastery and concept formation researchers have shown that when subjects have proven their mastery of a concept by sorting or labeling the materials given them, there has been little assurance that they would perform successfully on a defining or other verbal test for mastery of the same concept. For example, Smoke (1932, p. 20) noted that "a considerable number of subjects wrote defective definitions after having gone through the test series without error," more generally, that "individuals who have learned concepts may be unable to give an accurate verbal formulation of them."

Logic would indicate that we cannot be sure that the subject has actually learned the concept involved if he cannot successfully meet two or more criteria of concept mastery. Only after he can describe, label, classify and pass other tests of concept attainment can we be sure that the concept in its entirety has been assimilated. The necessity for the use of "converging operations" is discussed by Garner, Hake, and Eriksen (1956) in regard to perceptual tasks, and Johnson (1963) discusses its implications for research on concept formation. "Beyond the general requirement of

explicitness the principal recommendation of Garner, Hake, and Eriksen, extended to the present topic would assert that, since a concept is a theoretical construct, its validation requires 'converging operations.'

When converging operations have been used, the lack of agreement between measures such as labeling and defining has been an important point of discussion. We cannot expect the difficulties of any two measures to be exactly equal, but in the past, few if any, attempts have been made to equalize the difficulty of the materials and measures used. In the present experiment an attempt is made to equalize the difficulty of the materials for all three groups. This study also investigates the effects of varying types of practice on performance on a terminal definition task. The attempt to equalize the difficulty of the materials consisted mainly of trying to present verbal and graphic materials of equal difficulty. The manipulation of performance on the terminal definition task by varying the type and amount of practice brings us to the primary concern of this study.

The typical format of previous studies in this area is as follows. The subjects learn a concept such as "DAX" pictorially or verbally and after reaching a criterion of eight or ten consecutive correct responses they are asked to define

the concept. Experimenters have found that the subjects have often failed to give an acceptable definition under these conditions (Smoke, 1932). Why? Our contention is that at least a part of these results are an artifact created by the method used by the experimenters. Our thinking runs as follows. While the subject was learning the concept he was also, in a sense, "learning to learn" pictorially. He had perhaps fifty or a hundred exposures to a pictorial or verbal learning situation and no exposures to a defining task until the end of the experiment. Thus the results may be a reflection of the fact that the subject had practice on one operation and no practice on the other. Our first hypothesis, then, is that subjects given practice defining while learning a concept pictorially will do significantly better on a terminal defining task than will those who have not had this practice. A second hypothesis is that subjects who learn the concept verbally will do better on the terminal defining task than those who learn pictorially, but not as well as those who learn pictorially and also have practice defining. The basis for this hypothesis is that the verbal materials are more closely related to the defining task than are the graphic materials.

A brief discussion of the use of converging operations is included. The idea of converging operations implies a relationship between measures. The relationship between the task and the transfer task is examined in terms of amount of agreement in determining whether and to what extent a subject has mastered a concept.

RECRUITMENT OF SUBJECTS

A letter requesting permission to run this experiment was sent to twenty-five school superintendents in the Lansing, Michigan area. The names of these superintendents were taken from a list of Michigan school systems located within thirty miles of Michigan State University published by the Michigan State University Placement Bureau. Of the twenty-five letters, seven brought a favorable response: Grand Ledge Public Schools, St. Johns Public Schools, Laingsburg Community Schools, Dansville Agricultural School, DeWitt Public Schools, Elsie Public Schools and Durand Public Schools. The Dansville and Durand school systems were not visited. The four school systems represented in this experiment lie in predominantly rural communities, although Grand Ledge and DeWitt might more accurately be termed suburban communities since they are butten miles from the city of Lansing, Michigan.

SUBJECTS

The subjects were children from four public elementary schools in the Lansing area. There were 60 subjects, 39 boys and 21 girls, who ranged in age from eleven years to twelve years and eleven months. The matching of subjects was done on the basis of sex and on a total percentile score or a grade equivalent score on a standardized achievement test. The test score used varied from one school system to another. A summary of the matching data is shown in Table 1. The criterion for matching children was that they be within .5 units of each other on a grade equivalent score or within five percentile points of each other on an average percentile rank on an achievement test. Within each block the subjects were randomly assigned to three matched groups of twenty each. This was accomplished by putting the names of the three children into a box and drawing them out one at a time. The name drawn first was assigned to Group 1. The name drawn second was assigned to Group 2, and the remaining child was assigned to Group 3.

STIMULI

The stimuli consisted of fifty-four 4 X 5 cards. Twenty-seven of the cards had simple drawings of birds on them and

the other twenty-seven cards had four phrases describing the birds printed on them. On each of the fifty-four cards three features were varied, each in three ways. The beaks were either short and pointed, long and pointed, or hooked. The wings were either red, yellow, or blue. The tails were either orange, green, or black.

The twenty-seven bird picture cards were made such that no two cards were alike in respect to the three variable properties, but all cards were identical in every other respect. Since there were three different possibilities for each of the three significant features, there were twenty-seven possible ways in which they could be combined without duplication. Thus each varying element was combined with each other varying element an equal number of times in the set of twenty-seven cards.

In an attempt to equalize the difficulty of the verbal and graphic modes of presentation a matching phrase card was constructed for each of the twenty-seven picture cards. For example, if the bird pictured on picture card #1 had a hooked beak, a red wing, and a black tail, the bird described verbally on phrase card #1 also had a hooked beak, a red wing, and a black tail. Because the picture cards contained many irrelevant cues that the phrase cards did not, such as size

TABLE 1

Matching data by
school systems

School System	Criteria	Group 1	Group 2	Group 3
Grand Ledge	5th grade Stanford	53	58	58
	Achievement average	90	85	85
	percentile rank,	65	62	60
	Form J	20	21	24
		$\bar{X} = 57$	$\bar{X} = 56.5$	$\bar{X} = 56.75$
St. Johns	6th grade Metro-	7.9	8.2	7.9
	politan Achieve-	6.9	6.9	7.2
	ment Test, Form C,	6.6	6.7	6.7
	grade equivalent	5.4	5.2	5.2
	score	$\bar{X} = 6.7$	$\bar{X} = 6.75$	$\bar{X} = 6.75$
Laings- burg	5th grade Stanford	6.8	6.6	7.2
	Achievement Test,	6.5	6.5	6.6
	Form J, grade equ-	6.3	6.1	6.2
	ivalent score	6.2	6.3	6.2
		6.5	6.5	6.6
		5.8	6.1	5.8
		4.5	4.4	4.3
		5.3	5.3	5.2
		4.8	4.7	4.7
		5.6	5.3	5.4
		4.6	4.6	4.5
	$\bar{X} = 5.71$	$\bar{X} = 5.67$	$\bar{X} = 5.70$	
DeWitt	5th grade Total	125	122	123
	Intelligence score	120	121	121
	on California Short-Form Test of Mental Maturity	$\bar{X} = 122.5$	$\bar{X} = 121.5$	$\bar{X} = 122.0$

and shape of the birds, it was decided to add a fourth phrase,

"two feet with toes," as a constant to each phrase card in a

second effort to equalize the difficulty of the two sets of cards.

INSTRUCTIONS

Instructions were different for each group and are reported below. It will be noted that the instructions to Group 1 and Group 3, whose subjects learned the concept pictorially, include the statement that "it is either something about the beak, the toes, the wing, or the tail." It was felt that the additional distracting cues that were unavoidable in the pictorial presentation would make learning by this method more difficult, i.e., it would take the subjects more trials to reach the criterion by this method. In an attempt to equalize the difficulty of the two modes of presentation the subjects in groups 1 and 3 were told that the correct answer was something about either "the beak, the toes, the wing, or the tail." Thus the pictorial learning situation became more nearly equivalent to the task in which the subjects were asked to read the four phrases descriptive of the beak, toes, wing, and tail.

Group 1 Instructions

"I am going to show you several pictures of birds. No two of these birds look exactly alike. Some of the birds



are gunkle birds and some are bunkle birds. I want you to look at each picture carefully and then make a guess as to whether it is a gunkle bird or a bunkle bird. I will tell you whether you are right or wrong. I want you to try to learn how to tell a gunkle bird from a bunkle bird. Only one thing about the birds is important. In other words, the correct answer is a single thing rather than a combination of two or more things. It is either something about the beak, the toes, the wing, or the tail. Do you have any questions as to what you are to do?"

After ten consecutive correct responses: Ask "how do you think you tell a gunkle bird from a bunkle bird?"

Then: "Now I am going to show you several cards with phrases written on them describing the bird pictures you have just been looking at. Some of these birds are gunkle birds and some of them are bunkle birds, just as was true of the bird pictures. The same thing that made a bird a gunkle bird on the picture cards makes him a gunkle bird on these cards. Are there any questions?"

Group 2 Instructions

"I am going to show you several cards with phrases describing birds printed on them. Some of the cards describe

gunkle birds and some describe bunkle birds. I want you to read each card carefully and then make a guess as to whether the card describes a gunkle bird or a bunkle bird. I will tell you whether you are right or wrong. I want you to try to learn how to tell a gunkle bird from a bunkle bird. Only one thing about the birds is important. In other words, the correct answer is a single thing rather than a combination of two or more things. Do you have any questions as to what you are to do?"

After ten consecutive correct responses: Ask "how do you think you tell a gunkle bird from a bunkle bird?"

Then: "Now I am going to show you several cards with pictures on them of the birds you have just been reading about. Some of these birds are gunkle birds and some of them are bunkle birds, just as was true of the cards with the phrases written on them. The same thing that made a bird a gunkle bird on the written cards makes him a gunkle bird on these picture cards. Are there any questions?"

Group 3 Instructions

"I am going to show you several pictures of birds. No two birds look exactly alike. Some of the birds are gunkle birds and some are bunkle birds. I want you to look at each



picture carefully and then make a guess as to whether it is a gunkle bird or a bunkle bird. I will tell you whether you are right or wrong. I want you to try to learn how to tell a gunkle bird from a bunkle bird. Only one thing about the birds is important. In other words the correct answer is a single thing rather than a combination of two or more things. It is either something about the beak, the toes, the wing, or the tail. After you have guessed five times I will ask you to tell me how you think you can tell a gunkle bird from a bunkle bird. I will ask you the same question after each set of five guesses. Do you have any questions as to what you are to do?"

After ten consecutive correct responses: Ask "how do you think you tell a gunkle bird from a bunkle bird?"

Then: "Now I am going to show you several cards with phrases written on them describing the bird pictures you have just been looking at. Some of these birds are gunkle birds and some of them are bunkle birds, just as was true of the bird pictures. The same thing that made a bird a gunkle bird on the picture cards makes him a gunkle bird on these cards. Are there any questions?"

In the few cases where there were questions following the instructions the instructions were paraphrased by the

experimenter but no information not specifically stated in the original instructions was added.

PROCEDURE

On entering the experimental room the subject was invited to take a seat on the left of the experimenter, who was seated at a table. After a brief conversation in which the experimenter attempted to put the child at ease, the instructions were read aloud to the subject. When the subject indicated that he understood the instructions, the trials began.

The cards were shuffled thoroughly before each subject started his trials and also each time the subject went completely through the set of twenty-seven cards. This was done to eliminate any serial effect that might have been present and to minimize the possibility of rote memorization of the correct responses. A card was held up by the experimenter at a comfortable distance from the subject. The subject was given an upper limit of ten seconds to look at the picture or to read the phrases, although this was not mentioned in the instructions. Only two or three subjects waited the full ten seconds before responding. The response given was the subject's guess as to whether the card in question represented a "gunkle bird" or a "bunkle bird." The experimenter told the subject whether the guess was correct or



incorrect by giving a simple "yes" or "no," or by saying, "yes (or no), that's a gunkle (or bunkle) bird." The card was then turned face down on the table, and the experimenter recorded a plus (correct) or a minus (incorrect) sign on the record sheet. This procedure constituted one trial and the subject was ready for the next trial. In Group 3 the procedure varied from that described above in that after every set of five trials the subject was asked how he thought he could tell the difference between a gunkle bird and a bunkle bird. He was not told whether his response to this question was correct or incorrect. After the subject reached the criterion of ten consecutive correct responses he was asked to respond to the question "how do you tell a gunkle bird from a bunkle bird?" Then he was given ten trials with the opposite set of cards from those he had been using. Thus those who originally learned from the picture cards were switched to the verbal cards and vice versa.

RATINGS

Ratings were used to determine whether the childrens' responses were correct or incorrect, more precisely, to what degree the individual responses were correct. Three raters were used. The first rater has a master's degree in college

personnel work and is currently a head resident adviser in a womens' residence hall at Michigan State University. The second rater is a female graduate student, six credit hours away from a master's degree in counseling and guidance at Michigan State University. The third rater is a male Ph.D candidate in the anatomy department at Michigan State University. Each rater was given the set of sixty responses in a random (shuffled) order and was then given the following set of instructions.

Raters' Instructions

"I am going to give you a list of sixty responses that were given in answer to the question, "how do you tell a gunkle bird from a bunkle bird?" The question was asked as a part of an experiment in which several children participated dealing with concept formation. It is now necessary that the responses of these children be scored; that is, we must decide whether or not they had learned the appropriate concept at the end of the experiment."

"On the next page you will find a scoring scale ranging from 0 for a completely incorrect response, to 4 for a totally correct response. Categories 1, 2, and 3 are included for those partially correct responses which do not fall into the

0 or 4 categories of 'totally incorrect' and 'totally correct' respectively. An 'irrelevancy' is defined as any response or portion of a response that does not include or directly refer to the correct response. The correct answer is that gunkle birds have black tails and bunkle birds have orange and green tails, although a limiting response such as 'only gunkle birds have black tails' or 'no bunkle birds have black tails' would be scored completely correct. Thus an irrelevancy would be any response or part of a response that did not pertain directly to the color black or the word 'tail,' i.e., red wing or hooked beak. The response 'black wing' would be scored a '1' since it is a partially correct response (black) with one or more irrelevancies (wing)."

Rating Scale

When the rater had read the instruction sheet he was given the rating scale given below and a sheet on which to record his ratings. No rater had access to the rating of any other rater before or after completing his ratings.

0. Totally incorrect response

Example: "The gunkle birds have hooked beaks and bunkle birds have straight ones."

(or)

"I don't know. I was just guessing."

1. Partially correct response with one or more irrelevancy

Example: "Bunkles have lighter tails on them than gunkles, and shorter beaks."

(or)

"By the color of the wings and tails."

2. Partially correct response with no irrelevancies

Example: "It's something about the colors, I think."

(or)

"It's something about the tails, but I can't figure out what."

3. Correct response plus one or more irrelevancy

Example: "Gunkle birds have black tails and red wings."

(or)

"Bunkles have curved beaks and orange tails and gunkles have straight beaks and black tails."

4. Completely correct response, no irrelevancies

Example: "The gunkle birds have black tails and the bunkle birds don't."

(or)

"Gunkles have black tails."

RESULTS

The reliability of the mean of the three raters' ratings was computed according to the procedure suggested by Guilford

(1956, p. 281). The results showed that the reliability of the average of the three ratings was very high in all groups (Group 1 = .946, Group 2 = .924, Group 3 = .933). The three ratings were summed for each subject, leaving scores ranging from zero to twelve. A score of nine was used as the cutting point in determining whether a response was correct or incorrect. With a rating of three representing a minimally correct response, a score of nine was needed to insure that the response in question had been given an average rating of three by the combined raters. Thus a response with a summed rating of eight was scored "incorrect," while a nine was scored "correct."

The data, as summarized in Table 2, support the hypothesis that subjects given practice defining while learning a concept pictorially do significantly better on a terminal defining task than those who do not have this practice. Nine subjects in Group 1, fourteen subjects in Group 2, and sixteen subjects in Group 3 gave adequate definitions of the concept. The only significant difference between the three groups on performance on the definition task was between groups 1 and 3. Those subjects who did receive practice defining did significantly better on the definition task than those who did not.

TABLE 2

Comparison of frequency of correct
defining responses in three groups
of 20 subjects each

Groups	No. Correct Definitions	Chi Square	
		Group 1	Group 2
1	9	-----	-----
2	14	2.55	-----
3	16	5.22*	.53

*Significant at .05 level

The three groups were also compared on the basis of trials
to criterion. The results are summarized in Table 3.

TABLE 3

Comparison of three groups
of 20 subjects each on mean
trials to criterion

Group	Mean	SD	t	
			Group 1	Group 2
1	181.35	80.49	-----	-----
2	87.35	57.12	4.13**	-----
3	140.25	70.18	1.95	2.27*

*Significant at .05 level

**Significant at .01 level

Had Group 3 taken longer than Group 1 to reach the crit-
erion of ten consecutive correct responses it might have been

hypothesized that the greater success of Group 3 subjects on the defining task was a function of the greater amount of labeling practice they received. Table 3 shows the t-ratios obtained by the matched groups technique (Edwards, p. 169) from which it can be seen that the actual mean difference is in the opposite direction. Although Group 1 subjects took longer to learn the concept, Group 3 subjects did significantly better on the defining task.

The advantage of the use of converging operations in determining whether a given subject has actually mastered a concept was discussed earlier in this paper. Table 4 summarizes the agreement between the two measures of concept attainment used--the number of transfer trials correct and the correctness of the definition given.

TABLE 4
Agreement between two measures
of concept attainment

Transfer Task	Definition Task	
	Correct	Incorrect
Correct	29	8
Incorrect	10	13

To see how much agreement there was between the defining task and the transfer task in terms of determining whether or

not a given subject had mastered the concept, the standard error of a proportion was computed (.065) according to the formula suggested by Guilford (1956, p. 175). Using an estimate of .5 as the hypothesis of chance agreement, the difference between the obtained proportion of agreement, .7, and .5 was tested. The t obtained was 3.07, which is significant at the .01 level. Thus we can see that the agreement between the two measures is significantly greater than chance.

To determine what, if any, differences the use of converging operations would produce in the results of this study, the experimenter combined the results of the two tests of concept attainment used--the number of transfer trials correct and the correctness of the definition given. The correctness of the definitions was previously determined by the ratings and was discussed on pages 10-12. Of the ten transfer trials, the subject was judged to have mastered the concept if he responded to at least nine correctly. Using converging measures, then, a subject was scored as having learned the concept if, and only if, he was successful on both tests of concept attainment. Thus if a subject was scored correct on the basis of the definition task, but responded correctly on only seven of the ten transfer trials, he was not scored

as having learned the concept. Combining these two measures produced the results summarized in Table 5. The results using only the definition measure are also included, for contrast.

The significant differences between groups 1 and 3 diminished to a nonsignificant level when converging operations were used to determine whether a given subject had mastered a concept. It must be remembered, however, that these groups had been previously equated in that all subjects had performed

Table 5

Number of subjects who learned concept as indicated by converging operations and by definition measure alone

	Group	Correct	Incorrect
Converging Operations	Group 1	7	13
	Group 2	12	8
	Group 3	10	10
Definition Measure Only	Group 1	9	11
	Group 2	14	6
	Group 3	16	4

successfully to a criterion of ten consecutive correct responses. Comparing the two groups after equating them in

respect to the labeling task introduced experimental error that may have obliterated real statistical differences.

DISCUSSION

The procedure as outlined on pages 9-10 was developed after many lengthy sessions with 11-12 year-olds in the East Lansing area. Originally the question "how do you tell a gunkle bird from a bunkle bird?" was phrased "how do you tell a good bird from a bad bird?" and it was asked after each trial for subjects in Group 3. This was found to be unwise as well as time consuming. Subjects soon became annoyed, highly anxious or bored, and tended to give the same response repeatedly. It was decided to try asking the question only after every three trials, but this too proved to be inefficient. Next questioning after every five trials was attempted, and this turned out to be quite satisfactory. The subjects now appeared to be stimulated rather than bored by the questioning, and it appeared to generate greater interest in the task. Only two subjects appeared annoyed because the experimenter would not indicate whether their responses to the "how" question were correct or incorrect.

Finding suitable materials to use was also a lengthy process. The materials varied in three ways and it was originally

decided to use the terms "good bird" and "bad bird" rather than "gunkle" and "bunkle" birds. The correct response was "red wing." The children in the pilot study consistently ran over two hundred trials using this combination. The experimenter observed that these children most frequently began by looking at the beaks, so it was decided to try "hooked beak" as the concept to be learned. This proved to be too easy a task, for most children associated "bad" bird with a hooked beak, presumably because birds of prey notably have hooked beaks. Next it was decided to try the combination "black tail" and retain the "good bird"--"bad bird" response. This combination also turned out to be a poor one. The children this time seemed to be associating "bad" with the color black and were learning the concept with little trouble. At this point the experimenter decided that the "good bird"--"bad bird" responses were the main problem, and changed the required response alternatives to "gunkle bird" and "bunkle bird," retaining the property "black tail" as the correct response. This appeared to set the problem at a suitable level of difficulty for these subjects.

The problems involved in using complex but familiar conceptual materials such as the birds used in this experiment

appear to be quite different from those involved in using abstract materials such as nonsense syllables or geometrical forms, where difficulty level and previous contact with the material can be relatively easily controlled. Many previous associations have been formed to familiar conceptual material such as "hooked beak" or the color black--associations that may aid or hinder the process of learning a specific label for the concept or learning to sort the material in a new way. Although the use of familiar concepts introduces new problems, the experimenter suggests that one of the largest problems--that of manipulating the difficulty of the materials--is not an insurmountable one. Modifying the instructions offers the easiest method of control, and the introduction or elimination of irrelevant cues is also an effective means of controlling the difficulty of the materials.

As mentioned on page 1, Smoke found that his subjects could not define after having gone through the test series without error. Our results show that 39 of the 60 subjects in this experiment defined satisfactorially to a relatively strict criterion. The experimenter suggests that the use of familiar conceptual materials makes the definition task more reasonable than it was in Smoke's experiment where the subjects were asked to define a concept dealing with complex geometrical

forms. Further experimentation is needed to determine whether the inability to define that Smoke found to be so prevalent is a real inability or whether it is a function of the type of materials used.

The attempt to equalize the difficulty of the graphic and verbal materials used in this experiment was not entirely successful, as is shown by the large mean differences (significant at the .01 level) between groups 1 and 2 on trials to criterion. It appears that introducing one irrelevant cue on the verbal cards (two feet with toes) was not sufficient to make up for the numerous irrelevant cues on the picture cards. Judging from the comments and responses of the subjects, the "two feet with toes" response was eliminated from consideration within the first fifteen trials whereas the subjects who received the picture cards examined the toes at some length, taking several trials to count them, noting whether the slant of the legs was different, etc. For future studies, it is suggested that one of two procedures be followed. Either more irrelevant cues could be added to the verbal cards or the instructions for the subjects who learned pictorially could be further modified to exclude the toes variable.

The high reliability of the average rating in each group indicates that such ratings supply an effective dependent variable when collected under proper conditions. Few experimenters have reported their criteria for scoring a given response, and in such cases one may assume that the decision was arbitrary and subject to the bias of the experimenter. By using a carefully prepared rating scale and giving adequate instructions the problem of scoring responses can be nearly eliminated. In this experiment the three raters were not previously trained. One rater is in anatomy and has had no rating experience or training in psychology. The other two are in a field related to psychology and have had no previous rating experience, yet the reliability of the three combined sets of ratings is very high. The prime considerations appear to be adequate instructions and a scale with clearly defined categories rather than previous rating experience or training in psychology.

SUMMARY

The present experiment tests two basic hypotheses: (1) that subjects given practice defining while learning a concept graphically will do significantly better on a terminal defining task than will those who have not had this practice,

and (2) subjects learning the concept verbally will do better on the terminal defining task than those who learn graphically, but not as well as those who learn graphically and also have practice defining. The first hypothesis was supported by the data. The test of the second hypothesis showed no significant differences, although the mean differences were in the predicted direction.

The combined data was analyzed further by utilizing both of the two available measures of concept mastery (converging operations) and the agreement between the two measures was found to be significantly greater than chance. The differences noted above were no longer present when this method was used, but it is possible that the methodology used introduced experimental error that obliterated the differences.

The methodological conclusions to be drawn from this experiment are:

(1) Ratings can supply an effective dependent variable when collected under proper conditions.

(2) Type of practice as well as amount of practice must be controlled.

(3) Ability to define a concept after labeling it correctly may be a function of the type of conceptual materials used.

BIBLIOGRAPHY

BIBLIOGRAPHY

- EDWARDS, ALLEN L. Experimental design in psychological research (Rev. Ed.). New York: Holt, Rinehart and Winston, 1962.
- GARNER, W. R., HAKE, H. W. & ERIKSEN, C. S. Operationism and the concept of perception. Psychol. Rev., 1956, 63. 149-159.
- GUILFORD, J. P. Fundamental statistics in psychology and education (3rd ed.), New York: McGraw-Hill, 1956.
- JOHNSON, D. M. Cognitive structures as hypothetical constructs. Unpublished paper, 1963.
- SMOKE, KENNETH L. An objective study of concept formation. Psychol. Monogr., 1932, 42. No. 4.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03146 3205