THE RELATIVE ROLE OF POSITIVE AND NEGATIVE INSTANCES IN CONCEPT FORMATION

Accepted

by

John Edward Wood

A Thesis submitted in Partial Fulfilment of
The Requirements for the Degree of
M A S T E R O F A R T S
in the Department
of
EDUCATION

THE UNIVERSITY OF BRITISH COLUMBIA

April, 1943

ACKNOWLEDGMENT

The writer wishes to express his thanks to Mr. H. Straight for permission to conduct this investigation in the Vancouver Schools, and to Mr. A. Rines, Principal of Lord Roberts School, Mr. L.H. Robb, Principal of General Wolfe School, Mr. S.D. Meadows, Principal of Simon Fraser School, Mr. T.B. Spencer-Baynes, Principal of Charles Dickens School, Mr. W.R. Johnson, Principal of Livingstone School, Mr. C.R. Messenger, Principal of Dawson School, Mr. T. Woodcock, Principal of MacKenzie School, and to Mr. F.A. Armstrong, Principal of McBride School, for their generous assistance in the preliminary testing, and to the Grade Six classroom teachers who co-operated so willingly throughout the experiment.

TABLE OF CONTENTS

CHAPTER V (Cont'd) Page

INDEX TO ILLUSTRATIONS

INDEX TO ILLUSTRATIONS (Cont'd)

INDEX TO TABLES

INDEX TO TABLES (Cont'd)

CHAPTER I.

THE EXPERIMENTAL FIELD OF CONCEPT FORMATION

1. The Value of Laboratory Studies

"No greater harm can be done to education than to believe that laboratory workers solve the problems of field workers". Continuing from these words, A.S. Barr (4) states that laboratory studies make important contributions toward the solution of these problems, but that the problems themselves must be solved through the concerted efforts of many persons including laboratory workers and field workers alike. Barr finds that teachers and field workers in education demand methods which will actually produce results in practical situations; that the professor, on the other hand, free as he is from the pressure of having to produce immediate results, is interested in the search for truth for its own sake; e.g., in problems which have persisted over long periods of time and over extensive areas. He finds also that the validity of the findings in the controlled laboratory cannot be taken for granted in the practical situation where there are many uncontrollable factors.

Field workers in education are greater consumers than producers of research. The studies of a field worker, by the very nature of his position, must be practical though not necessarily superficial (4). Research in the field is no primrose path. Lack of teachers trained for research,

inadequate tools, time filled with daily duties, and the unwillingness of parents to having their children "experimented" upon, all contribute to the difficulty of producing research.

Brownwell (9) points a searching finger at experimental studies in teaching methods to expose several of their weaknesses. He says that pedagogical research is fragmentary and sporadic; that learning studies have been confined too much to elementary levels; that teaching rather than learning has been studied; that fear of being called unscientific has paralysed the attack on many vital problems; that there exists too blind a faith in statistics and other techniques; and that problems are seen too narrowly.

Brownwell also shows, however, that pedagogy has benefited from experimental studies. He states that in Reading, Arithmetic and Social Studies, techniques and materials have been improved through experiments; that diagnostic techniques have pointed the way to remedial procedures; that ways of measuring intangibles are being devised; and that many new methods are available for studying the learning and teaching processes.

Laboratories must assist in the initial stages of educational problems. The writings of such men as Freeman, Buswell, Judd and Thorndike show that the laboratory method provides one of the chief bases of theorizing. The fact that psychological laboratory studies have proved of value to

classroom teaching is justification for the present investigation. This study may suggest problems which can be attacked in the classroom, or it may result in generalizations of direct value to classroom procedure. It clarifies, to a certain extent, the importance of negative instances in the formation of new concepts, and suggests certain definite limitations to verbalization as a means of communication.

2. The Importance of Concept Formation

To observe and to reason about observation is of major importance. Warren (73) argues that rational behaviour is possible only after concepts have been formed.

To Warren (73: 426) a concept is "an experience built up in the individual as the result of many past experiences in which the relatively universal in those experiences becomes emphasised and the particular eliminated". Warren (73: 284,285) further states that "a concept is a thought which includes only the characteristic elements of meaning or value". He defines a judgment as a thought which combines two concepts. The language equivalent of a concept he calls a term, and of a judgment, a proposition. Concepts and judgments are rational thought, and rational behaviour is often called reason (73). Reasoning is a series of thought experiments. Few aspects of human behaviour are more important than the ability to reason.

Thus, the factors influencing concept formation

become of major interest to educators and psychologists alike. The present study investigates certain aspects of the role played by positive and negative instances in concept formation.

3. What is Concept Formation?

"We see our problems too narrowly", said Brownwell, in criticising experimental education. Critics of other experimenter's definitions of concept formation may well pause to consider Brownwell's words. "There are as many definitions of concept as there are schools of psychology", states Smoke (66) in commenting on the great confusion in the psychological literature regarding the meaning of the term "concept formation".

Since the term "concept formation" is nothing more than an arbitrary verbal symbol, it has a technical meaning as narrow or as varied as the situations to which it is applied by those philosophers and psychologists who are recognized as authorities in their respective fields. Accordingly, a review of the use of the word "concept" is necessary in order to clarify its meaning.

"The process of concept formation may be treated from two points of view--the philosophical and the psychological", says Pratt (57) in his summary of the experimental work done between 1922 and 1926.

Wolff (83) stated in 1939 that much contemporary

philosophy views the concept as a pointer directed toward objects which are sensational; that the concept is said to be true if it leads to a percept; that all philosophies are based on indefinables; and that divergences are due to incongruity in the sets of indefinables assumed in the discourse and those recognized by the reader. Many are the philosophies on which the various schools of psychology are based. Accordingly, in the case of the meaning of concept formation, it is not surprising to find a host of interpretations entering through the numerous portals of indefinables. Added confusion is to be expected concomitant with the selection of experimental criteria by individual psychologists who crystalize their theories into the concrete form of psychological experiment.

The Oxford English Dictionary defines a concept as: (1) a thought--an idea; (2) the product of the faculty of conception, an idea of a class of objects, a general notion. For a psychologist, defining a concept as a thought makes the application of the word so broad that the term becomes nearly useless. If a concept is the "product of the faculty of conception" as stated above in (2), then Vedenov, and Berkenblit, can be supported in their view that concepts must be sensory. On the other hand, if a concept is "an idea of a class of objects", as in (2) above, the opposing views of Welch and Long may be held. The further definition above, that a concept is "a general notion", is so broad

that it becomes useless as a definition for a psychologist.

Further evidence of overlapping of definition, and definition in terms of itself, is given in the Oxford Dictionary in its meaning of "generalize"; viz., (1) to form into a general concept, to reduce to general laws; (2) to designate by a general name; (3) to infer inductively from particulars; (4) to form general notions by abstraction from particular instances. According to (1), concept formation becomes a process of generalizing. Smoke, Yerkes, Hull, Kuo, Welch and Long either implicitly or explicitly accept this point of view. Defining generalizing as "to reduce to general laws" supports the thought that the induction of rules or principles is generalization. This latter statement is rejected by Smoke in these words: "Concepts are indispensable constituents of rules, formulae, and principles, but there is no justification for identifying them".

Definition (2) allows workers like Welch (77), Long (45), and Berger (6) to speak of generalizing in terms of abstraction of "classes", "genus" and "hierarchical development of concepts". Definition (3), which speaks of generalization as induction, is in harmony with both Thurstone's and Warren's usage. Definition (4) speaks of generalization in terms of abstraction, and could be quoted to support Hull's and Kuo's criteria, the abstraction of common elements. According to the Oxford Dictionary, therefore, induction, generalization, formulation of rules or principles, discrimination between genus or classes, and abstraction are

properly considered as aspects of concept formation.

Warren's Dictionary of Psychology gives the following definition for concept: "A mental state or process which means or refers to more than one object or experience, or to one object in relation to others. (When it represents different individuals or items, it is called a class concept or general concept; when it represents a common aspect or attribute of the class, it is an abstract idea). The formulation of a concept into words is called a term. (Conception refers to the process, concept to the product)."

Warren's Dictionary defines generalizing as follows: (1) (introspective)--The process of perceiving or conceiving a general characteristic or fact or meaning in single or in complex situations or things; (2) (behaviouristic)--Responding to the common aspects (from any point of view) of the specific elements in a complex situation.

Warren's Dictionary defines induction as follows: (1) (logic)--The process of reasoning from the particular to the general; (2) The end result of such reasoning. Reasoning is explained as "The process of solving a problem by means of a concept or general principle".

Tyler (71) states that "an induction may be defined as a generalization based upon observed facts" (p. 1). Tyler speaks of the "ability of subjects to form a generalization from serially-presented different examples of a common rule" as one example of rational learning (p. 3).

He defined rational learning as "The ability to make generalizations or to do inductive reasoning" (p. 4). He says that "Yerkes multiple-choice method can be used to investigate generalizing ability". Thus Tyler points out the similarity between induction, generalization, rational learning, and multiple-choice techniques.

In 1923, Kuo pointed out the close relationship between his study of inductive inference, Hull's experiment on concept formation, and Hamilton's (1911) and Yerkes' (1921) studies using multiple-choice techniques. Kuo classes the multiple-choice problems as inductive inference.

In 1932, in reporting their study, Ewert and Lambert considered concept formation and generalization to be synonymous terms. In 1932, Smoke also (67) identified concept formation with generalization.

"Inductive problems consist of specimens, and the result is to obtain a definition, or at least a working knowledge of the 'class' represented by the given specimen", says Woodworth (84:800). "In the laboratory a problem in induction or concept formation calls for the development of an effective response to a class of objects and a different response to objects not belonging to this class". With these words Woodworth identifies induction and concept formation.

From the opinions quoted, it would seem that a complete review of the studies in concept formation would

include a history of the experiments involving rational learning, inductive inference, generalizing abstractions, multiple-choice techniques, generalizing and concept formation.

4. Criteria of Concept Formation

(a) Introspective:

In 1916, Fisher (23) adopted as a criterion of concept formation the ability to discover class characteristics from a series of ten drawings having certain common characteristics of shape, and to formulate a definition.

(b) Abstraction of a Part:

Hull (40) in 1920, and Kuo (44) in 1923, considered that the ability to abstract common elements (radicals in Chinese characters) from a complex pattern was a criterion of concept formation. In 1934, Drever (17) used the recognition of the presence of a right-angle in certain polygons as his criterion of the presence of a concept. In 1932, Smoke (68) criticised and avoided the use of common "elements" in the sense of parts such as Chinese radicals, and made the common characteristic consist in relation between the parts. Since Smoke's work has a close relationship to the present study, Smoke's definition and criterion of a concept will be considered at greater length later.

(c) Classification:

Woodworth (p. 804) lists classification problems

as studies in concept formation. Ach (1) in 1921, Huper (29) in 1928, and Hanfmann and Kasanin (35) in 1937, considered a concept to have been formed when a subject was able to discover the principles by which a collection of blocks could be sorted into four classes.

Woodworth (p. 806) classifies the induction of a rule of action as an aspect of concept formation. In this type of experiment the concept to be formed is not of a class of objects, but the correct formula for meeting a class of situations. The experiment is like a game played between the experimenter and the subject, with subject's task to discover and to verbalize the rules of the game. This type of mental activity is induction.

(d) Induction of a Rule:

Thurstone (29) accepts that induction is a type of generalization, for he speaks of inductive generalization. He speaks of the verbalization of some rule or principle from a series of observations as being an induction.

Peterson (55) in 1920 used as his criterion of concept formation the induction of discernable arithmetical principle in his arithmetic game. In 1934, Heidbreder's (39) subjects were asked to determine the rules of a game which she played with them. She considered the concept to be formed when the subjects had discovered the rules. In 1939, Tyler used a complicated lighting apparatus which rendered highly objective a situation requiring the subject to induce rules

by which lights were to be controlled. Although several of these workers did not call their experiments studies in concept formation, they may be considered to be in that field, according to Woodworth's classification.

(e) 1936 and Since:

In spite of Smoke's protestations in 1932 against the loose and indefinite usage of the term concept, and his discussion and definition of the term concept, later experimenters have continued to use the term broadly and to adopt many and varied criteria of its presence.

In 1937, Meyer and Piaget (29) studied "children's conceptions of speed and time". They accepted an "idea" (as did Oxford Dictionary) of speed and time as a "concept" and they adopted verbalized definitions (as did Gibson, Heidbreder, Smoke, Welch and others) as their criterion. In 1937, Berger (6) accepted "abstract ideas" as concepts and used as criteria verbalizations in terms of: (a) definitions of objects (Oxford's "general notion of"); (b) verbalized classifications such as genus (Oxford's idea of a class of objects); (c) verbalizations of principles such as cause (Oxford's definition of generalize, as "to reduce to general laws"); and (d) recognition of color (like both Berkenblit's and Vedenov's usage of concepts as sensory experience). In 1937, Hanfmann and Kasanin (35) assumed that a relationship between meaningless objects and meaningless words is a concept, and accepted verbalization as a criterion of the

presence of "superordinate" concepts between unrelated words. Graham (29) speaks of concept formation as generalizing, and generalizing as a form of learning. Vedenov (72) speaks of isolated concepts and general concepts. "Isolated concepts", he says, "are sensory, and if they are excluded from the structure of perception, they become primary abstractions". Thus Vedenov used the term, abstraction, when many psychologists like Yerkes, Kuo, Smoke and Tyler, and when, according to Wolff, many contemporary philosophers, would employ the terms concept or generalization.

Nineteen thirty-nine added its tribute to the broad usage of the term concept. Berkenblit (7) speaks of the "single concept" in the same type of situation where Mott (51) uses the words "abstract idea", and where the Oxford Dictionary defines generalization as a "general notion by abstraction from particular instances".

In 1940, Welch (78) speaks of abstract thought and concepts in the same breath. He also discusses "hierarchical development of concepts" as though classification generalizations were concepts. Long and Welch (80) apply the term concept to both objects and classes.

(f) Summary:

Thus it is seen that even yet, experimental psychologists use the term concept in a variety of ways. The criteria used for the detection of concepts are numerous. Verbalization remains the usual criterion of concept formation.

As will be shown later, the use of verbalization renders even an otherwise objective experiment highly subjective.

5. Review of Studies of Concept Formation

The following discussion is divided into three arbitrary sections; namely, (a) Introspective; (b) Studies Approaching Objectivity, and (c) 1936 and Since.

(a) Introspective Studies:

Introspective techniques were applied in the pioneer studies of concept formation. Early exploration of this type yielded comparatively barren results. Fisher (23) in 1916 confronted her trained psychological observers with a series of 10 drawings belonging to a class called, for example, "zalof". Subjects were to discover the characteristics of this class and formulate a definition. Full introspective reports were required. Fisher concluded that a concept is a readiness to respond appropriately to any member of a class of objects; and that the response may be an overt movement or a verbal or visual image, so long as it conforms to the essential characteristics of the class.

Other introspective studies by English (19), Stevanovic (69) and Ghant (12) were published in 1922, 1927, and 1933, respectively. These studies agreed in finding two main lines of attack in forming concepts of novel objects: (1) assimilation by the new object to some familiar object, and (2) analysis of the new object into parts which are familiar.

The first method may give quicker mastery of the single specimen, but since the extrinsic resemblance no longer holds good, it is likely to break down when the variations come into view. The analytical process is more dependable in reaching a new concept.

(b) Studies Approaching Objectivity:

Important objective studies of concept formation have been made by Hull (41), Ach (1), Kuo (44), Smoke (66-68) and Drever (17). In 1920, Hull used a paired association technique in which Chinese characters were paired with nonsense names. The subjects were required to abstract common radicals from the Chinese characters in order to name them. This abstraction of common elements was called concept formation by Hull. In his experiments on concept formation, Kuo adapted Hull's technique of testing for the identification of common elements.

In 1932 and 1933, Smoke (66) published his studies of concept formation. These are of great importance to the present paper. Smoke criticizes previous objective experiments in concept formation (such as those made by Kuo and Hull), which define the concept in terms of the common element to be found in a series of geometrical figures, Chinese characters, and the like, on the grounds that such studies deal with a purely analytical process; and that they impose upon the subject the task of discovering some hidden or camouflaged element. Smoke's viewpoint is that

"the sine qua non of concept formation is a response to relationships common to two or more stimulus patterns". Acting upon this theory, he constructed a series of geometrical designs which represented patterns rather than common elements, and used nonsense syllables to designate them. For example, a "Dax" was "a circle and two dots, one dot being outside the circle and the other being on the inside of it". This series, together with other series of designs, made up his exposure material, which was presented to the subjects by an electrically driven exposure apparatus.

Thus Smoke avoided common "elements" in the sense of parts such as Chinese radicals, and made the common characteristic consist in certain relations between the parts. When the subject, after examining a number of specimens, one at a time, believed himself able to define the class, he was asked for three things, namely: (1) to give his definition (Verbalization); (2) to distinguish in a test series those figures which did and those which did not belong to the class (Recognition); (3) to draw some specimens (Reproduction), and in agreement with Hull, Smoke found his subjects sometimes able to pass the other tests while still unable to give an adequate Verbalization. The faulty definitions were usually too inclusive. In one form of the experiment Smoke introduced "negative" instances. He concluded that the negative instances were: (1) of very little assistance; (2) of use to some subjects; and (3) distracting to others.

A frequent process of reaching the concept involved the formulation, testing and rejection of hypotheses, till one was found that stood up for a series of specimens. Definite recall of previous members of the series played a part, as was shown by the "thinking out loud" required in this experiment. By aid of memory, similar specimens were grouped as a step toward definition of the whole class.

Smoke gives a technical meaning to the term "elements" and "relationship". He criticizes the use of the term "common elements" in the studies of concept formation because, following Hull and Kuo, the experimental literature has used "common elements" to refer not to a condition but to a section of a pattern; e.g., Kuo's radicals. Accordingly, Smoke insists that the term "common elements" be used technically to designate only some section of a pattern. In studies of concept formation, Smoke uses the term "relationship" only in a technical sense, ignoring the "elements" which are required in order that there be a relationship. Tyler (71) in clarifying Smoke's terminology points out that a relationship without elements to be related would be impossible. Tyler states that "generalization or concept formation involves both elements and relations between these elements". Although on one hand Tyler speaks of concept formation as involving both elements and relationships, and on the other hand, Smoke defines concept formation as involving only relations, both Tyler and Smoke are actually in agreement in their implicit definition of concept formation. Smoke's

technical use of the term "relation" includes but is less clear than Tyler's two terms "relations" and "elements".

"By 'concept formation', 'generalization' and 'concept learning'", says Smoke (67), "we refer to the process whereby an organism develops a symbolic response (usually but not necessarily linguistic) which is made to the members of a class of stimuli patterns but not to other stimuli." The present study considers Smoke's definition as an adequate and a fruitful vehicle for an experimental approach to an investigation of concept formation.

In common with Smoke's work, the present study defines triangles, rectangles, circles, lines and dots as elements, while abstract ideas, such as inside, outside, in a direct line with, touching and near are defined as relations. Guesses as to what the concept is, in a given learning situation, are defined as hypotheses.

In the classification problem introduced by Ach (1) in 1921 and used by several experimenters including Huper (29) in 1928, and Hanfmann and Kasinin in 1937, a collection of blocks is to be sorted into 4 classes. The basis of classification must be discovered by the subject. In one form the blocks are 5 different colors and of 6 different shapes, color and shape being irrelevant. They are tall or short, large or small, and by cross-classification, fall into 4 classes. Grouping or "convergence" appears to be an

important part of the process of reaching the concepts.

In another type of experiment, the concept to be formed is not that of a class of objects, but that of the correct formula for meeting a class of situations. The experiment has somewhat the form of a game played between the experimenter and the subject. The task of the latter is to discover the rules of the game. He discovers them by playing and meeting with success or failure. The rules may be entirely arbitrary as in Yerkes' "multiple-choice" problems. The rules may depend upon some discernible principle, as in the Peterson's arithmetical game. Tyler (71, p. 9) states that Kinnamon (43) in 1902, and Hamilton (32) in 1911, were among the first to report studies employing the multiple-choice technique. (p. 9). In 1923, Brown and Whittell (8) used a multiple-choice technique with adults. In 1932, Roberts modified the technique to investigate the ability of pre-school children to see and apply a principle of relationship. In 1933, Arons (3) studied the "generalizing abstraction" ability of under-graduates, using the multiple-choice technique.

Extensive experiments performed by Heidbreder in 1924 required the subject to discover the rules of a game played between the subject and the experimenter. The materials consisted of geometrical figures and check marks. The subject endeavored to discover the rules by which composite geometrical figures were to be marked. After making a trial by

marking the figure in the way that he thought it might be correct to mark it, the subject was informed that he was "Right" or "Wrong". At the end of each trial and before being informed of its rightness or wrongness, the subject reported "everything that went on in his mind" during the trial. Other double figures of this class are treated similarly, till the subject discovered the rule as proved by correct statement (Verbalization), and response (marking the figures). When the first game was mastered, the second game followed. Heidbreder's outstanding result was the demonstration of "spectator" behaviour. The more usual "participant" behaviour consists in trying out hypotheses. In "spectator" behaviour, the subject has no hypothesis; all his guesses have been proven erroneous and he can only make some random response and remain on the watch for some new hypothesis to emerge. Says Woodworth (p. 807), "Spectator behaviour affords a clue toward answering the question of how hypotheses arise. The receptive attitude may at times be just what is necessary to get us out of a rut and allow some hitherto neglected aspect of the situation a chance to exert its effect. The receptive phase of the inductive process is less observable, either objectively or introspectively, than the more active phase of trying out the hypothesis, but it may be no less essential".

Peterson, in 1920, defined rational learning as purposive thinking in which ideas were used in the effective

solution of problems and devised a letter-number memory-reasoning test, which made a pioneer contribution to studies of generalization, but which was not entirely satisfactory due to the possibility of problem solutions by other than rational behaviour.

Haught (36) subjected Peterson's findings and other related data to statistical analysis. Whereas in certain tests modified from Peterson's problems the best measures of learning were time and unclassified errors, in Peterson's tests themselves, perseverative errors constituted the significant measures. In his discussion Haught pointed out that in studies where the experimenter determines to some extent the rate at which the subject works, time as a criterion becomes of little value. This conclusion has a bearing on the present study.

(c) 1936 and Since:

Roslow (62) in 1936 reported a study requiring subjects to induce a principle, but Tyler (71) suggests several ways in which his statistical analysis could have been improved. The use of Tyler's methods would have produced statistical results of greater inclusiveness and conclusiveness than those employed by Roslow.

Tyler (71) in 1939, investigated the ability of junior high school pupils to solve rational learning problems of the type requiring the subject to generalize or to formulate a rule from a succession of problems involving a common

principle. The apparatus used to study rule induction consisted essentially of a panel of switches for the subject to manipulate and a group pattern of light bulbs for him to observe. A pattern of lighting was set up by the experimenter and from these the subject was to formulate the rule which would indicate how the key which turns out all the lights could be selected without error. Tyler found: (1) that the majority of cases were not consistent in using exclusively either an exploratory or am analytic approach to the problems but that most cases employed both methods; (2) that many subjects appear to wait to verify a tentative implicit hypothesis before verbalizing it; (3) that many subjects gave evidence of ability to select the proper key but were unable to give the corresponding verbalization; and (4) that verification may be done by using either positive or negative instances.

In 1936 Heidbreder (38) conducted a series of experiments related to this study. In these, she gave evidence of the role of language in the acquisition and use of concepts. The procedure was a modification of that used by Hull in what was apparently a memory experiment. Each subject learned to associate given nonsense syllables with given situations under conditions which permitted him to discover that the many different situations to which a given syllable was applied possessed a common characteristic. The subjects, 220 college students, were studied individually, but were classified into groups on the basis of variations in the experimental

procedure and the materials presented. The process of concept formation was measured in terms of repetitions, prompts, and opportunities to apply the concept to new situations. At the close of the experiment each subject was required to write definitions (Verbalizations) or descriptions of the concepts he had formed, and also to take an objective examination of the single choice type. In view of the findings of the present study, the conclusions of Heidbreder are of interest.

She suggests that there are several different ways in which language is used as a tool during concept formation and that these methods vary with the situation; that a concept may be used with consistent correctness even though the subject cannot formulate (verbalize) it; that the ability to formulate (verbalize) a concept is more closely related to the nature of the referent than to the readiness with which it is acquired or to the accuracy with which it is applied; that the readiness with which a concept is formed is determined not by the ease or difficulty with which its name (nonsense syllable) can be memorized but by the relation between its referent and the perceptual situation in which it is presented.

In 1937 Hanfmann and Kasanin (35) published a paper describing a test of concept formation adapted from Sakharov and Vigotsky who employed it for the study of conceptual thinking in psychotic patients. The test had its origin in a

method employed by Ach for studying the development of concepts. It consisted in general in confronting the subject with a number of meaningless objects and a number of meaningless words between which the subject is made to discover relationships through prolonged manipulations and demonstrations. The study is of interest to this paper chiefly as an example of one of the many techniques employed to study concept formation.

In 1937, Berger (6) examined the role of age, sex and environmental differences in the ability of school children to form abstract judgments. Children were asked to state in writing the difference between fifteen pairs of objects or ideas, such as milk and water, a mistake and a lie. Berger found that older children, more often than younger ones, differentiate in terms of genus, cause, and generalized ideas, while younger ones refer to external features such as colour and form, and that city children more often make their distinctions in terms of values and consequences, while rural children more often resort to examples and description. This study contributes to the present paper not only some conclusions about concept formation, but also an example of the prevalent practice of using the subject's verbalizations as an experimental criterion. A later chapter indicates the high subjectivity inherent in this procedure.

Sibano's experiment (64) in 1938 was an introspective investigation of concept formation as productive thinking.

Two stimulus words were given from which the subjects made "superordinate" concepts common to these two words. Subjects were asked for introspective reports of their experiences during the process of concept formation. Results seem to be of little value as they depend upon the experimenter's selection and interpretation. Sibano concluded that the inevitable factors in concept formation are adequacy of stimulus words, ground and figure of thinking, images of things and words, emotion and will. (Whatever these things mean!). This review was included to show that introspective techniques have been used even lately, and to contrast the type of conclusion given in introspective studies with those of objective ones.

Vedenov's (72) experimental investigation of the structure of concepts takes the opposite view, concluding (1) that isolated concepts must be sensory, but that when excluded from the structure of perception they go over into a primary abstraction; and (2) that the structure of general concepts and the process of their evolution is determined by the object, so that all thought has its basis in sensation. "The problem of evolution from sensory to abstract is one of the main problems in psychology", says Vedenov.

Graham's study (29), in 1938, considered generalizing to be a form of learning and examines some special conditions imposed by certain types of learning. He developed a general pattern for the construction of exercises in developing skills

in generalizing, and illustrates this pattern with detailed examples and references to other tasks employed. In support of the observations, inferences, and conclusions given, the experimental evidence is utilized largely in an illustrative manner. In his study of generalizing, Graham points out the lack of ability to obtain part scores for analysing the skill components. His subjects reported difficulty in reporting their planful manipulations, sets, techniques or processes for the illuminating of their activities. Graham suggests that the crucial aspects of learning are essentially subjective mental discriminations and in his treatment of data he points out their lack of objectivity.

In 1938 Crudden (15) investigated form abstraction by children. In his experiment the material was the counterpart of the negative instance in the present investigation. Crudden used simple learned geometrical figures which were imbedded in relatively unknown geometrical figures of varying degrees of complexity. He used 65 children from 65 to 78 months old. He found that the degree of difficulty in abstracting the known material increases roughly in proportion to the degree in which it is imbedded in a more complex figure; that "that-which-is-to-be-avoided" in abstraction has almost as much influence in successful abstraction as "that-which-is-to-be-chosen"; and that knowledge of the "figure-to-be-abstracted" frequently results in successful response where previously no abstraction could be made.

In 1939 Berkenblit (7) reported on a study of the genesis of concepts. He investigated the formation of single as well as general concepts in children from 16 to 28 months old. Berkenblit concluded that single concepts connected with a certain object depend upon the child's sensory experience with, as well as on his emotional attitude toward, the object while playing with it. By generalizing single concepts of the acting object general concepts about action were formed. Sensory perception, and the child's motor manipulations with the object, says Berkenblit, play a decisive part in the formation of general concepts.

During the years 1938 to 1940, inclusive, Welch has published a greater number of studies on concept formation than any other experimenter. For one article during 1936 he collaborated with Davies; in 1940, jointly with Long, he published two experiments; and during the years 1938 to 1940 he published six articles of his own. Welch's subjects have been young children, practically all of them less than 7 years 6 months of age. Two of Welch's conclusions are related to this investigation. Welch found that the associational or memory development necessary for the first manifestation of genus-species phenomena at the linguistic level comes later than the requisite development as discrimination and generalization from the same behaviour; and that the genetic development of the structure of abstract thought may

be described as passing through rather definite stages. Except for these conclusions, his studies and findings have little explicit connection with the present experiment. It is advisable, however, that any worker in the field of concept formation be familiar with Welch's contributions.

6. The Present Experiment

(a) The Problem:

The present study is designed to investigate the part played by both positive and negative instances in the generalizing process whereby twelve-year-old boys form concepts.

(b) Definitions:

For purposes of this experiment, Smoke's definition is accepted. He said, "By concept formation, generalization, and concept learning, we refer to the process whereby an organism develops a symbolic response (usually, but not necessarily linguistic) which is made to the members of a class of stimuli pattern but not to other stimuli". More specifically, concept formation is the abstraction of the highest common factor found in a series of examples of the concept. This highest common factor is composed of both relations and the elements required for those relations. Guesses as to what the concept is are defined as hypotheses. Formulation of hypotheses similar to those which the experimenter had in mind when he constructed the instances is called generalization and concept formation.

Geometrical figures in which the relationship and the necessary and sufficient common elements are found, are defined as positive instances. Figures exhibiting any deficiency in the relationship or in the elements necessary and sufficient thereto are called negative instances. Triangles, rectangles, circles, lines, and dots are defined as elements, while abstract ideas, such as inside, outside, in a direct line with, touching, and near are referred to as relations.

(c) Criteria:

In this experiment, the three criteria employed to determine the nature of the subject's hypotheses are called Verbalization, Recognition, and Reproduction. Verbal definitions of the concept are called Verbalizations. Performance on a test in which the subject indicates the figures he considers to be, and those not to be, the concept is termed Recognition. Examples drawn by the subject of his idea of the concept are defined as Reproductions.

Recognition is the non-verbalized symbolic response accepted as a criterion of achievement toward concept formulation. Recognition is accepted as a more objective experimental criterion than Verbalization. In utilizing Recognition rather than Verbalization as the criterion to determine when a given amount of concept formulation had occurred, this study differs from many of the experiments in the same field. The justification for this selection of

criteria is given in Chapter III.

CHAPTER II.

SUBJECTS, APPARATUS AND PROCEDURE

1. Subjects

Sixty-five twelve-year-old Grade VI boys from the Vancouver school system were divided into 3 groups: 15 boys for pretesting and standardizing the experimental situation; and 50 boys divided into two comparable experimental groups--one receiving the positive presentation and the other the positive-negative presentation of material. Individual testing was employed throughout.

Comparable experimental groups were selected by matching boys for chronological age, intelligence quotient, school grade, sex, native tongue and socio-economic status. Subjects were chosen so that four groups of paired I.Q.'s could be made.

First, the whole group was split into two matched sets of I.Q.'s. The twenty-five pairs of I.Q.'s were arranged in descending magnitude. Then the pairs were split vertically producing two comparable I.Q. groups, each ranging from I.Q.'s of 88 to 133.

Second, the subjects were segregated horizontally into three paired groups, the dull normals with I.Q.'s of 95 plus and minus 7.5, the brighter normals with I.Q.'s of 110 plus and minus 7.5, and the superiors with I.Q.'s of 125 plus and minus 7.5. The number of subjects in each group was chosen to conform roughly to the normal distribution having

A.M. equal to 100, and S.D. equal to 16.

The I.Q. comparability of these groups is shown below:

TABLE I.- ALL SUBJECTS: I.Q.'s PAIRED VERTICALLY TO FORM TWO MATCHED EXPERIMENTAL GROUPS.

	Positive Presentation	Positive-negative Presentation
No. of cases	25	25
Arith. Mean (A.M.)	108.3	108.6
Standard Dev. (S.D.	10.7	10.6

TABLE II.-COMPARISON OF THE TWO MATCHED GROUPS FORMED BY VERTICAL SPLITTING OF THE LOWEST 9 PAIRS OF I.Q.'s (I.Q. RANGE 95 PLUS AND MINUS 7.5)

	Positive Presentation	Positive-negative Presentation
No. of cases	9	9
Arith. Mean (A.M.)	96.8	97.2
Standard Dev. (S.D.)	3.9	3.9

TABLE III.- COMPARISON OF THE TWO MATCHED GROUPS FORMED BY VERTICAL SPLITTING OF THE MIDDLE 10 PAIRS OF I.Q.'s. (I.Q. RANGE 110 PLUS AND MINUS 7.5)

	Positive Presentation	Positive-negative Presentation
No. of cases	10	10
Arith. Mean (A.M.)	109.9	110.7
Standard Dev. (S.D.)	4.3	3.0

TABLE IV.- COMPARISON OF THE TWO MATCHED GROUPS FORMED BY VERTICAL SPLITTING OF THE HIGHEST 6 PAIRS OF I.Q.'s. (I.Q. RANGE 125 PLUS AND MINUS 7.5)

	Positive Presentation	Positive-negative Presentation
No. of cases	6	6
Arith. Mean (A.M.)	123.0	123.0
Standard Dev. (S.D.)	2.8	4.6

On March 31, 1941, all subjects fell within the chronological age range of 12 years plus and minus 6 months. All subjects spoke English as their mother tongue. Socio-economic status was roughly controlled by matching pupils from the same school.

Experimental testing was carried out between May 27 and June 29, 1941.

2. Apparatus

The apparatus consisted of individual record booklets and 9 sets of concept cards, each set composed of 8 positive instances, 4 negative instances, and 16 test instances.

Two hundred eighty-eight white cards 3" x 5" shown in Figures 2 to 10 were employed. On each card an instance (either positive or negative) of one of the nine experimental concepts was drawn in India ink. The 32 cards for each concept were divided into 3 packs, two teaching packs of 8 cards each and one testing pack of 16 instances. All cards were numbered serially on the back; the teaching sets from 1 to 8, and the testing pack from 1 to 16.

The teaching pack of 8 positive instances was used with one experimental group, hereafter called the positive group. The other pack, consisting of 8 instances alternately positive and negative, was employed with the other experimental group, hereafter known as the positive-negative group. The same test pack composed of twelve negative instances and 4 positive instances, arranged originally in chance order,

was used for both groups.

In the two teaching packs of each concept, the odd-numbered cards (positive instances) were identical. The even-numbered cards in the positive teaching set were positive instances, while the even-numbered cards in the positive-negative teaching pack were negative instances. The negative examples (even numbered) in the positive-negative teaching series were designed to differ from the correspondingly numbered cards in the positive series only sufficient to allow them to become negative instances and to contribute clarifying characteristics. These negative cards were intended to eliminate the acceptance of a concept composed of common elements or relationships less than the concept (the highest common factor found in the positive instances).

In the test series, the 12 negative instances were designed to be diagnostic, so that if the subject failed to learn the concept, a glance at his response revealed whether it was elements, relations, or both, that he had not recognized.

The instances were composed and selected with the following principles of construction in mind. Integrated patterns in which the relationships between the elements gave a sense of unity to the whole pattern were desired, and the more interesting the design, the better. Simple figures with a minimum of elements and relations were employed with a view to diagnosis through systematic variation in both the

teaching and the testing packs. Each card made a unique contribution of at least one clarifying aspect.

The concepts were named after the nine nonsense syllables having the least associative value. This technique was borrowed directly from Smoke's experiments. The concepts used in this study were as follows:-

"Dax" - triangle with dot inside; (see Fig. 2)

"Mef" - a circle half black and half white; (see Fig. 3)

"Vec" - a line with a dot at one end and in direct line with the line; (see Fig. 4)

"Mib" - a circle touching a square; (see Fig. 5)

"Zum" - a circle and 2 dots, the one dot being inside the circle, and the other outside; (see Fig. 6)

"Tov" - a square having one cross near each of the four sides; (see Fig. 7)

"Pog" - two lines unequal in length; (see Fig. 8)

"Wez" - a circle and a triangle with the circle touching the triangle on its shortest side; (see Fig. 9)

"Zif" - a rectangle and a circle with the circle inside the rectangle and touching its two longest sides, but not touching either end. (see Fig. 10)

In the six-page record booklet, space was provided for the subject's name, age, birthday, grade, I.Q., school, and parents' nationality. Frames in the booklet were provided for the subject's responses on the three levels--Verbalization, Recognition, and Reproduction. A typical page is shown in Figure 1.

In the Verbalization section, under "Attempt Number", was recorded the serial number of the subject's attempt to

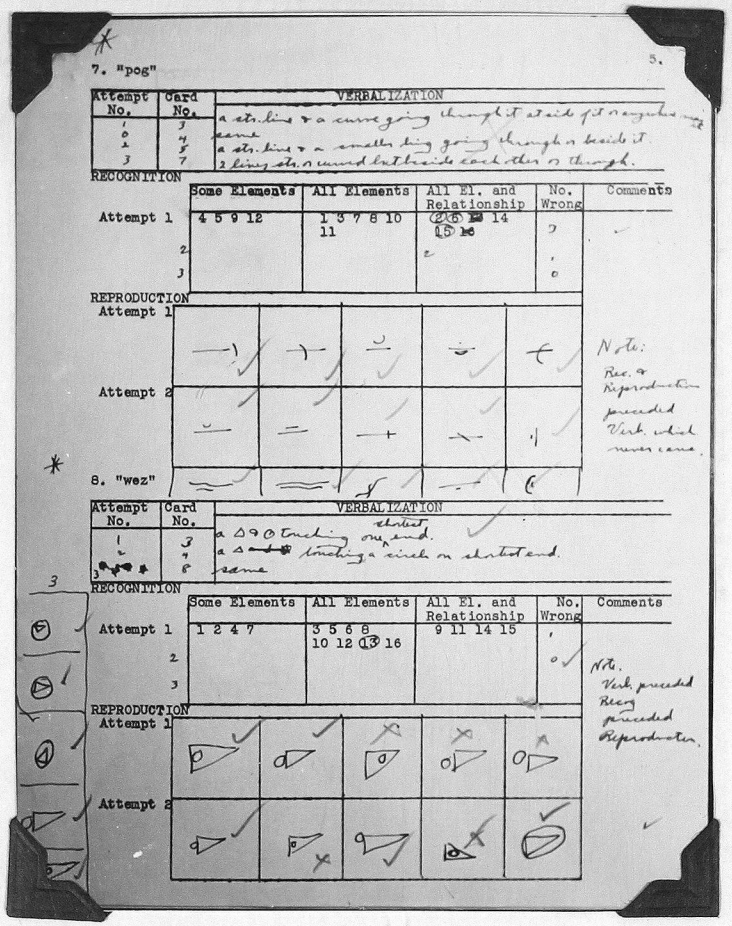

FIGURE 1.- A page from the record booklet showing a subject's performance on two concepts.

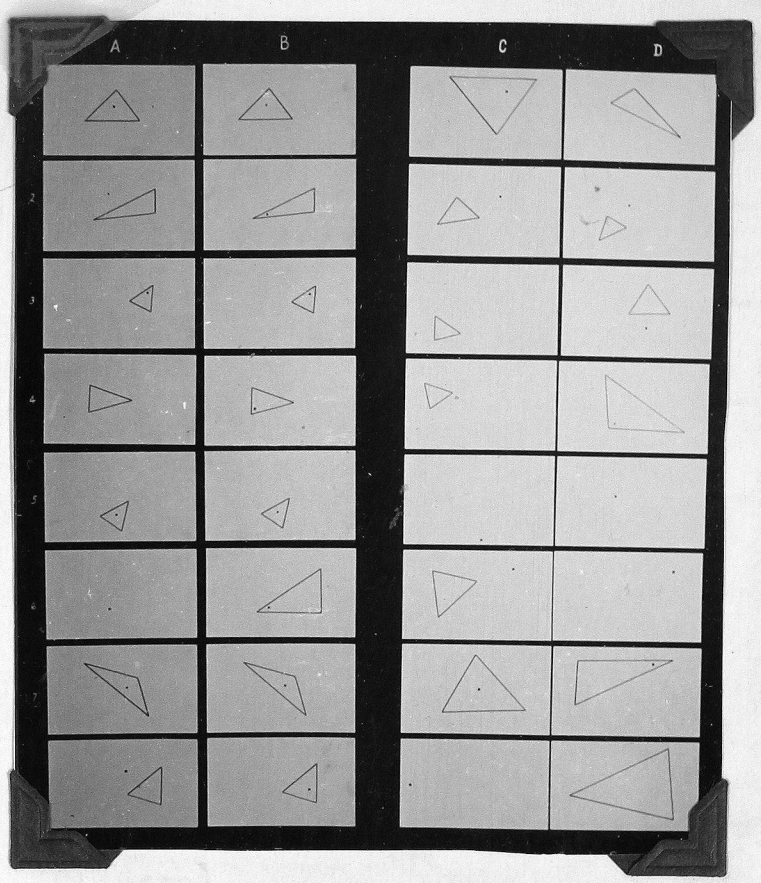

FIGURE 2.- DAX - 32 instances (both positive and negative).
Column A - Positive-negative teaching pack;
" B - Positive teaching pack;
" C & D - Recognition test series.

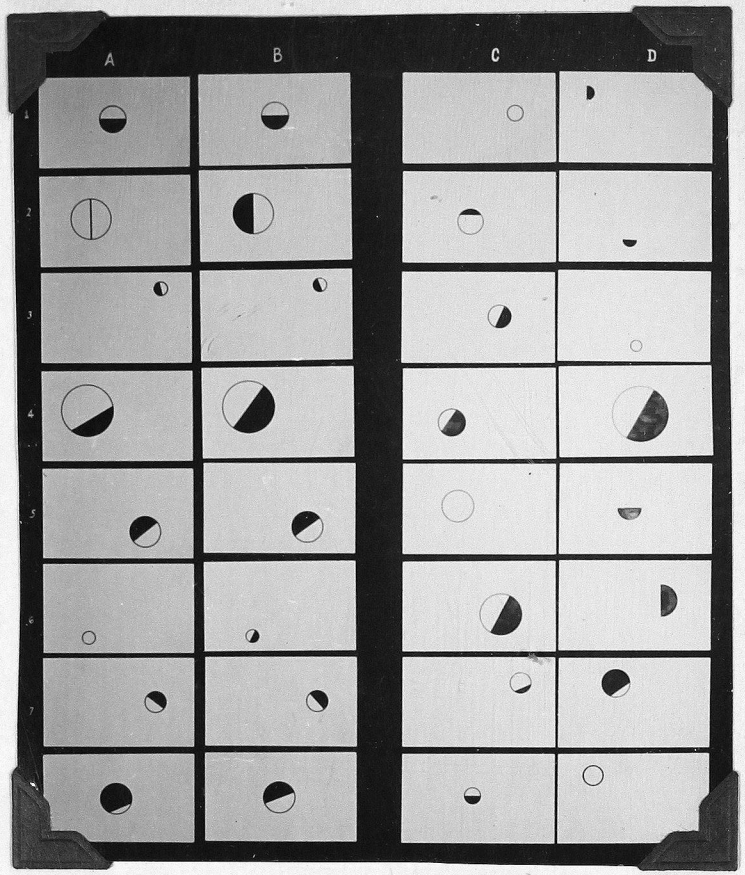

FIGURE 3.-MEF - 32 instances (both positive and negative).

FIGURE 4.- VEC - 32 instances (both positive and negative).

Column A - Positive-negative teaching pack.
" B - Positive teaching pack.
" C and D - Recognition test series.

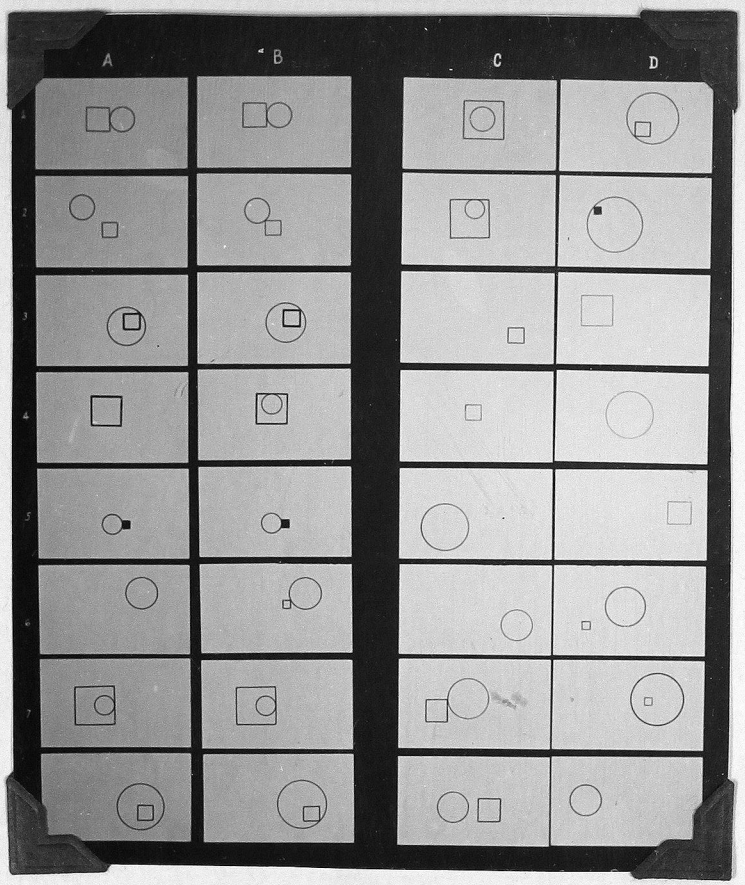

FIGURE 5.- MIB - 32 instances (both positive and negative).

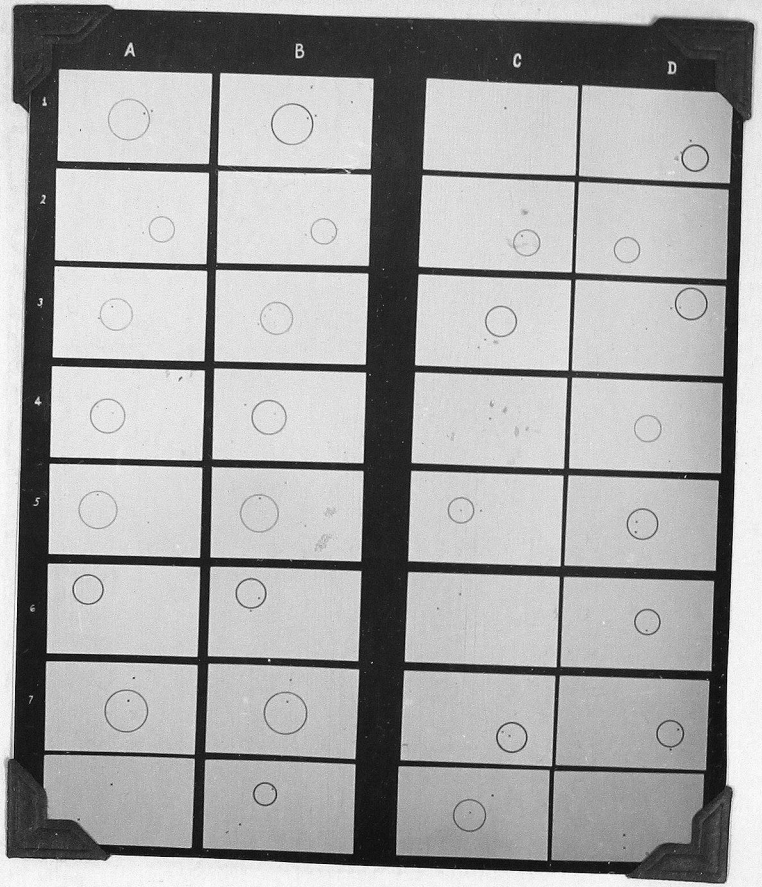

FIGURE 6.- ZUM - 32 instances (both positive and negative)

Column A - Positive-negative teaching pack.
" B - Positive teaching pack.
" C and D - Recognition test series.

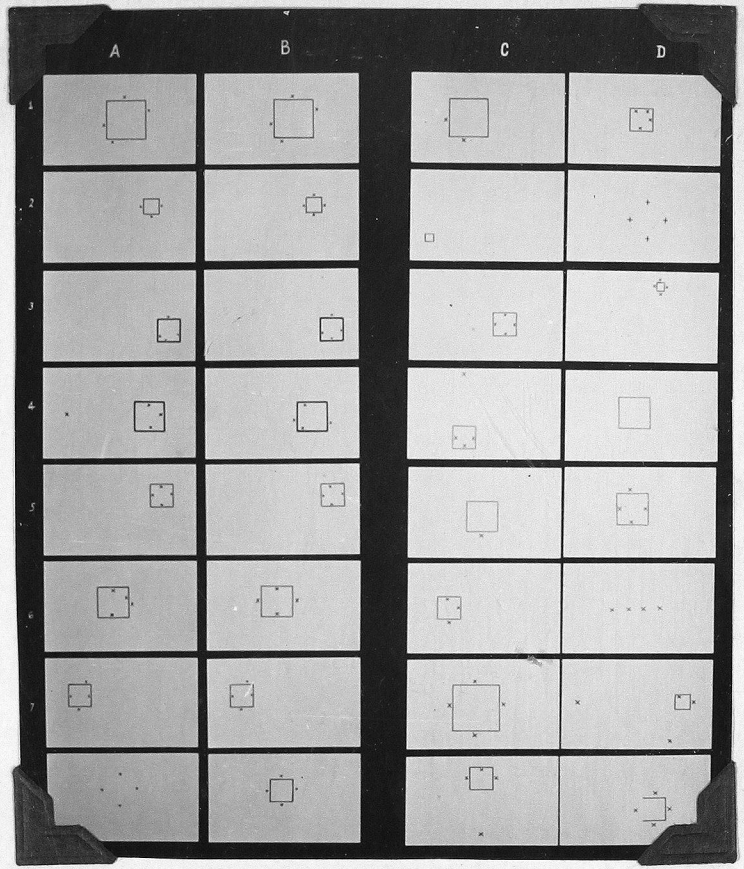

FIGURE 7.- TOV - 32 instances (both positive and negative).

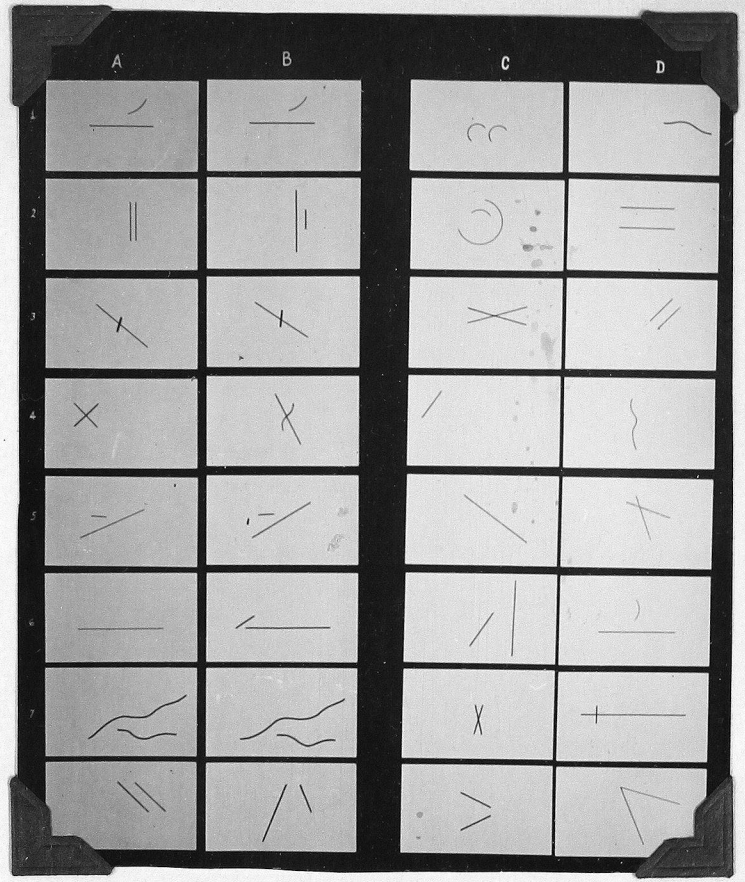

FIGURE 8.- POG - 32 instances (both positive and negative).

Column A - Positive-negative teaching pack.
" B - Positive teaching pack.
" C and D - Recognition test series.

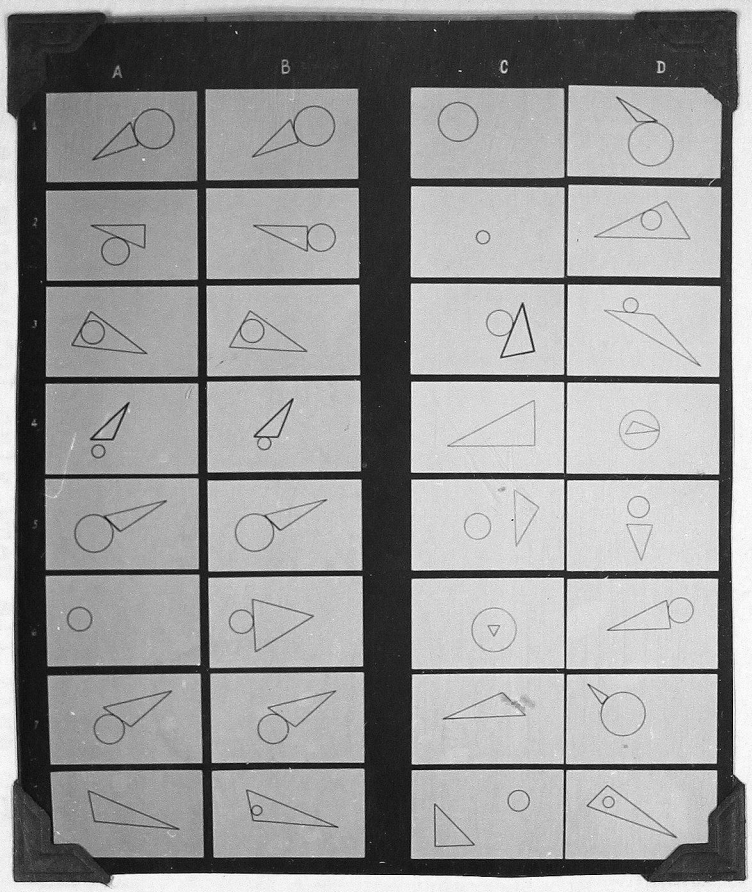

FIGURE 9.- WEZ - 32 instances (both positive and negative).

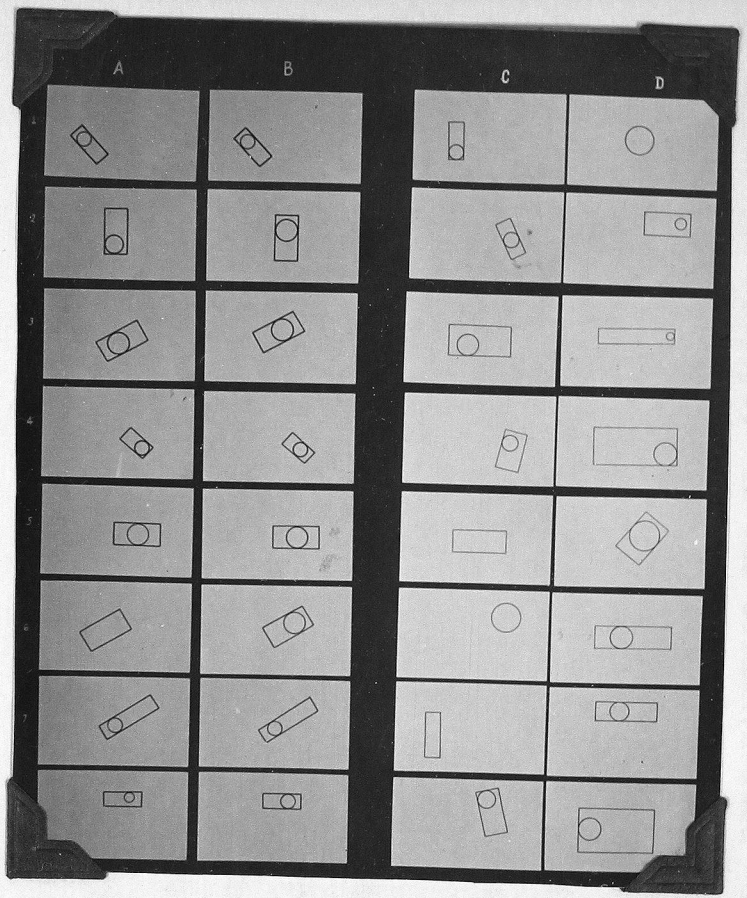

FIGURE 10.- ZIF - 32 instances (both positive and negative)

Column A - Positive-negative teaching pack.
" B - Positive teaching pack.
" C and D - Recognition test series.

verbalize the concept. In the column,headed "Card Number", was recorded the number of the last card shown before that attempt. Under "Verbalization", the pupil's word-for-word definition of the concept was recorded. Verbalization was scored as successful only when a subject made a generalization clearly applicable to the concept.

Recognition was recorded in the following manner: the serial numbers of the test cards in which only some of the necessary elements appeared are given in Column 1. Similarly, the serial numbers of the test cards which lacked some of the necessary relations but which included all the elements, are found in Column 2. In other words, the numbers shown in the first two columns are the serial numbers of cards which did NOT represent the concept. The third column contains the numbers of the cards representing the positive instances in the test pack. Under "Number Wrong" was recorded the number of errors made by the subject in the 16 trials. Recognition was scored as successful only when the recognition test cards had been identified with 100 per cent accuracy.

To test Reproduction, two sets of 5 squares were provided in which, for each Reproduction attempt, the subject drew five freehand pictures of his idea of the concept. In some cases the examiner found it necessary to provide for further attempts to reproduce the concept. Reproduction was marked as successful only when all five drawings satisfied

the conditions necessary for the concept.

3. Procedure

Each subject underwent one experimental period only. After rapport was established, and the subject was seated to the left of the experimenter at a table, a preliminary concept, "Dax", was employed to introduce and explain experimental procedure.

The preliminary teaching pack was placed face downward and in order from numbers 1 to 8, with 1 on top within the subject's reach. He was instructed to turn the cards over one at a time and place them face upward on the table where they remained in sight throughout both the learning and testing period for that concept. He was told that the puzzle was to figure out "what a Dax" was, and that the fewer cards he had to turn over to discover it the better his score would be.

Following the exposure of any card and as soon as he thought that he might know what the concept was, a subject was allowed to try the tests. On exposure of card numbers 4 and 8, regardless of previous test perfor-

mances, successful or otherwise, samples# of the developing## concept were taken by means of all three criteria--Verbalization, Recognition, Reproduction.

Procedure for presenting the instances was identical for both positive and positive-negative groups, except that the subjects were given to understand in the one case that all cards represented the concept and in the other that alternately the cards were and were not examples of it.

In sampling the concepts, three criteria representing the subject's ideas were recorded; i.e., Verbalization, Recognition and Reproduction. In the spaces provided in the booklet, word-for-word statements (Verbalizations), and responses of "Dax" or "not Dax" to the test cards (Recognition) were noted, and then the subject drew five examples of his idea of the concept. He was encouraged to draw examples of his own rather than to copy any positive cards in front of him. No rule that he must not copy was made. Such an instruction would raise the philosophical difficulty of defining "copy" - it being impossible for him to reproduce an identical replica, anyhow, no matter how hard he tried.

#The word "sample" is used to mean evidence available to the experimenter revealing the nature of the subject's concept. Since there may be other aspects of the concept not revealed by our testing situation, our criteria give a "sample" which may be partial. If through Verbalization, Recognition and Reproduction a complete picture of the subject's concept has been revealed to the experimenter, then the "sampling" was complete.

##The word "developing" is used to indicate the possibility of progressive growth of the concept from incomplete stages to more complete. At times the complete concept is formed very quickly, but at other times incomplete concepts are formed which are progressively altered to accommodate them to additional instances seen.

Smoke made an experimental error in his study with negative instances, and as a result did not find the chief contribution made by negative examples. In his instructions to subjects receiving the allegedly positive presentation, Smoke used the words, "That is part of the concept, but not all of it".[1] In saying "but not all of it", Smoke introduced a factor which invalidated those results which pertain to the effect of introducing negative instances. This statement will be substantiated in a later chapter.

At present it is sufficient to say that procedure must be carefully controlled to the end that "NO CLUE EXTERNAL TO THE TEACHING PACK MUST BE ALLOWED TO GIVE EVIDENCE TO THE SUBJECT AS TO THE ACCEPTABILITY (COMPLETE OR INCOMPLETE, CORRECT OR INCORRECT) OF HIS TEST RESPONSES. In other words, neither verbally nor by any change in procedure must the experimenter allow the subject to know whether his response to the test situation was successful or otherwise.

Once this point had been recognized it was easy to refrain from giving the subject a verbal clue. To refrain from giving this information to the subject by an insidious change of procedure following correct responses, however, requires constant vigilance. Following a test situation, strict adherence to identical procedure must be the rule, whether the response be successful or unsuccessful and whether the presentation be positive or positive-negative.

[1]Underlining not in the original.

The individual record booklet remained in the hands of the experimenter except while it was given to the subject for the purpose of recording his Reproduction responses. Since the subject could know what the experimenter had written in the record booklet, Verbalization and Reproduction had to remain unevaluated, until after the experimental period was over and the subject had left the room. The circles placed around the numbers in the Recognition frames were explained as being notes the experimenter wanted to make, but which notes had nothing to do with the subjects' responses. During the positive presentation, in the cases where subjects made errors, failure to observe these rules would have suggested to some subjects that the response was not correct. In this way, the "Not-a-Dax" idea would have been introduced into a situation which was supposed to be purely positive. In other words, to neglect to observe this precaution is to commit through the scoring method the same experimental error that Smoke committed verbally in his instruction already noted.

To allow repetition of the experiment and minute inspection of the present procedure, a word-for-word detailed account of the procedure has been included in the appendix.

CHAPTER III.

EXPERIMENTAL CRITERIA OF LEARNING

In experimental work on concept formation, it has been commonly assumed that the ability to formulate a concept into words comes later than the ability to recognize the presence of that concept in an observed example. Among many experimenters who have found that subjects could pass recognition tests when they could not verbalize correctly are Berkenblit (7), Long (45) and Smoke (67). The acceptance of the idea that Recognition precedes Verbalization has been a contributing factor to the use of Verbalization rather than Recognition as an experimental criterion of concept formation.

The present study utilizes recognition tests rather than verbal responses as the more useful experimental criterion. The chief arguments submitted in support of this action are presented under the following headings:-

(1) Is Verbalization a "level" which appears later than Recognition?

(2) Which criterion, Recognition or Verbalization, is more vulnerable to reversals of judgment?

(3) Which criterion, Verbalization or Recognition, is the more subjective?

1. Is Verbalization a "Level" which Appears Later than Recognition?

Assuming that the ability to formulate concepts into words comes later than the ability to recognize the presence of that concept in an observed example, and assuming further

that the criteria, Verbalization and Recognition, as employed in the present study, indicate the presence of those two abilities respectively; therefore -

(a) there must always be on the average at least as many Recognitions achieved as Verbalizations;

(b) the number of correct Recognitions should be significantly greater than the number of acceptable Verbalizations;

(c) cases of Verbalization preceding Recognition will be relatively rare;

(d) Recognition will on the average always precede Verbalization.

Should any or all of these four conclusions fail to be found in the results of the present experiment, it is questionable whether the assumptions are sound. The last two assumptions are strongly supported by definition. The weaker assumption is that Verbalization is a "level" which appears later than Recognition. Accordingly, should the conclusions be unsubstantiated in this experiment, doubt would be cast on this latter assumption.

(a) Are there always as many Recognitions achieved as Verbalizations?

Table V shows the statistics of the present study pertinent to this question.

TABLE V.- THE COMPARISON OF THE NUMBER OF RECOGNITIONS AND VERBALIZATIONS ACHIEVED

	Positive Presentation (25 subjects, 8 concepts each)		Positive-Negative Presentation (25 subjects, 8 concepts each)	
	Verbalizations Achieved	Recognitions Achieved	Verbalizations Achieved	Recognitions Achieved
Total Achieved	85	78	143	158

Table V shows that by the positive-negative presentation, 158 Recognitions and 143 Verbalizations were achieved. This group, therefore, had successful performances in 15 more Recognitions than Verbalizations. By the positive presentation, 78 Recognitions and 85 Verbalizations were successful. In this group, therefore, there were seven less Recognitions than Verbalizations. The results of the positive-negative group add supporting evidence to conclusion (a); i.e., that perhaps there must always be on the average at least as many Recognitions achieved as Verbalizations. However, in the positive presentation, more Verbalizations than Recognitions were successful. This suggests that it is not necessary that there always be on the average as many Recognitions achieved as Verbalizations.

Perhaps the explanation for this seeming contraduction is that sometimes one "level" appears first and sometimes the

other. Perhaps the order of the appearance of Verbalization and Recognition is traceable not to relative intrinsic difficulties in the criteria themselves so much as to the method by which the concept is presented, the nature of the concept itself, or the facility with which subjects can employ the criteria themselves. Perhaps one type of criterion lends itself more readily as a vehicle of expression in one type of conceptual situation, while the other may be the more appropriate in another.

That subjects can Recognize when they cannot Verbalize has been corroborated by several experimenters. The results of the positive-negative group of the present experiment (Table V) agree with this finding. Where then is the weakness in the conclusion that Verbalization is a higher "level"; i.e., a more difficult achievement? An answer might be that while it is true that Recognition does at times appear when Verbalization cannot be successfully accomplished, experimenters,in formulating this conclusion, have failed to take into account the number of cases where Verbalization can be achieved but Recognition cannot.

(b) Are the number of correct Recognitions significantly greater than the number of acceptable Verbalizations?

Table VI exhibits results dealing with this question.

TABLE VI.- COMPARISONS OF THE MEAN NUMBER OF VERBALIZATIONS AND RECOGNITIONS ACHIEVED PER SUBJECT, BY EACH METHOD OF PRESENTATION

	Positive Presentation (25 subjects, 8 concepts each)		Positive-Negative Presentation (25 subjects, 8 concepts each)	
	Verbalizations achieved	Recognitions achieved	Verbalizations achieved	Recognitions achieved
A.M. per Subject	3.40	3.12	5.72	6.32
σM	.94	1.31	1.15	1.09
$D_{M_r} - M_v$	--.28		.4	
σD	1.61		1.58	
$\frac{D}{\sigma D}$	-- .17 (57 chances in 100)		.25 (60 chances in 100)	

Table VI indicates that there is not a significant difference between the number of Recognitions achieved and the number of Verbalization successes. In the positive-negative presentation, the critical ratio is .25; i.e., there are only 60 chances out of 100 that on the average there will always be more Recognitions than Verbalizations correct. The critical ratio of the positive presentation is .17, showing no significant difference--in fact, showing that there are 57 chances in 100 that Verbalization will precede Recognition. These results

deny conclusion (b); i.e., that the number of correct Recognitions should be significantly greater than the number of acceptable Verbalizations.

(c) Are cases of Verbalization preceding Recognition relatively rare?

Table VII shows the number of times in this experiment that the two criteria preceded one another.

TABLE VII.- THE ORDER OF APPEARANCE OF THE CRITERIA

	Positive Presentation (25 subjects 8 concepts each)		Positive-negative Presentation (25 subjects 8 concepts each)	
	Verbalization preceded Recognition	Recognition preceded Verbalization	Verbalization preceded Recognition	Recognition preceded Verbalization
Total	24	8	26	26

If conclusion (c) were to be valid, there would be a relatively greater number of Recognition preceding Verbalization than vice versa. Table VII indicates that by the positive-negative method just as many Verbalizations preceded Recognitions (26) as vice versa. Furthermore, in the positive situation, the two criteria appeared with Verbalization preceding Recognition three times as often as in the opposite order. Clearly, therefore, the results of this experiment do not support the conclusion that Verbalization

rarely precedes Recognition. Furthermore, these results are in direct contradiction to the original assumption that Verbalization and Recognition are "levels" in the sense that one always precedes the other in their order of appearance and difficulty, and that Recognition is the lower "level".

(d) Does Recognition on the average always precede Verbalization?

Table VIII shows the results of the present study computed on a per subject basis.

TABLE VIII.- COMPARISON PER SUBJECT OF THE MEAN NUMBER OF TIMES EACH OF THE TWO CRITERIA PRECEDED THE OTHER.

	Positive Presentation (25 subjects, 8 concepts each)		Positive-negative Presentation (25 subjects, 8 concepts each)	
	Verbalizations preceded Recognition	Recognition preceded Verbalization	Verbalizations preceded Recognition	Recognitions preceded Verbalization
A.M. per subject	.96(M_1)	.32(M_2)	1.04(M_1)	1.04(M_2)
σ_M	1.00	.545		
$D_{M_1-M_2}$	.64		0	
σ_D	1.14			
$\frac{D}{\sigma_D}$	.56 (71 chances) (in 100)		0	

Since in the positive-negative situation both A.M.'s are 1.04, the conclusion to be drawn is that under the present experimental conditions, Verbalization and Recognition will, on the average, precede one another an equal number of times. Neither one seems to appear on the average more than the other. By the positive presentation the $\frac{D}{\sigma_D}$ of .56 indicates that in 71 cases out of 100, Verbalization will on the average always appear before Recognition. These results, therefore, contradict conclusion (d) and indicate that there is NOT a significant difference between the order of appearance of Verbalization and Recognition.

The evidence from this experiment bearing on the question of "levels" has been presented. It has been shown that if Verbalization is a "higher level" than Recognition, then four conclusions would follow. It has been indicated that the results of this study supported none of those four conclusions. In neither the positive nor the positive-negative situation can either Verbalization or Recognition be depended upon to precede the other. The order of their appearance, therefore, allows either criterion to be used equally advisedly. It may be concluded, therefore, that for this study it would be unwise to assume that the ability to verbalize a concept indicated a "higher level" of understanding of the concept, than did the ability to recognize its presence in an observed example.

There are several decided objections to Verbaliz-

ation as an experimental criterion. These disadvantages are not inherent in Recognition as a criterion of learning. Further discussion of the relative merits of, and objections to, the two criteria follow in the next section of this chapter.

2. Which Criterion, Recognition or Verbalization, is More Vulnerable to Reversals of Judgment?

Reversal of judgment for a given criterion (Verbalization, Recognition, Reproduction), is defined as retrogression to an incorrect response for a given concept after a correct response has been given. All other factors being equal, the fewer the reversals of judgment to which a criterion is subject, the better.

Table IX shows the number of times in which reversal of judgment occurred.

TABLE IX.- NUMBER OF REVERSALS OF JUDGMENT.

Concept	Verbalization		Recognition		Reproduction	
	Positive	Pos.-Neg.	Positive	Pos.-Neg.	Positive	Pos.-Neg.
Mef	2	1	-	1	2	-
Vec	-	-	-	-	-	-
Mib	-	-	-	-	2	2
Zum	-	-	-	1	-	-
Tov	-	-	1	-	1	-
Pog	-	1	-	-	4	4
Wez	-	-	-	-	3	-
Zif	-	-	-	-	1	-
Total:	2	2	1	2	13	6

Reversal of judgment in Verbalization was present twice in the positive and twice in the positive-negative situation. In Recognition, 2 positive-negative and 1 positive reversal of judgment occurred. In Reproduction, 13 positive and 6 positive-negative reversals took place. The vulnerability to reversals of judgment in Reproduction and Verbalization is small, while that of Recognition is comparatively great. With respect to reversals of judgment, therefore, either criterion, Recognition or Verbalization, serve equally well.

3. Which Criterion, Verbalization or Recognition, is the more Subjective?

Language

In experimental work on generalization, Verbalization has been used frequently as the chief criterion. The question arises, which of Verbalization or Recognition, as defined in the present experiment, is the better criterion to use in this study.

The author encountered considerable difficulty in marking, right or wrong, many of the Verbalizations given by the subjects. It is hard to know what a subject "means" by what he says. The experimenter evaluated the subject's responses found in the Record Book on two occasions; the first time was within a few days of giving the test, while the second time was about two months later. The experimenter's second evaluation of the Verbalizations did not always agree

with his first one. The difficulty was that the experimenter did not interpret the Verbalizations in the same way both times. If it was difficult for one experimenter to give consistent interpretations, how much greater would be the difficulty for different persons to agree on the interpretation of a given set of verbalizations!

The greater the amount of subjective interpretation required, the more objectionable is Verbalization as an experimental criterion. The present study yields evidence of at least four aspects of Verbalization which render it highly subjective. These four lines of evidence are submitted below.

(a) The "Meaning" of Verbalizations Given

In the following discussion, the "meaning" of the Verbalization was determined by the experimenter in terms of the Recognition cards which the subject chose as conforming to his idea of the concept. A consideration of the subject's selection of the Recognition cards and Reproduction responses is the only experimental justification of such a statement as, "the subject said 'round' when he meant 'oval'".

Table X shows the different ways in which the subjects attempted to verbalize the same thought. The column headed "Meaning" shows the thought that the subject's selection of the Recognition cards would lead the experimenter to conclude the subject was thinking. The words "positive" and

"positive-negative" indicate the type of learning situation which produced the Verbalizations shown in the table. The figure in the column headed "Frequency" shows the number of times a given Verbalization was used to express the thought indicated in the "Meaning" column.

TABLE X.- VERBALIZATION MEANING DISPARATE FROM RECOGNITION MEANING

"Meaning"	Frequency	Concept	Verbalization
"In direct line"			
Positive	7	Vec	"after"
	17	"	"at end"
	2	"	"and"
	1	"	"under"
	1	"	"in front of"
	1	"	"beside"
	1	"	"over the top"
Total:	30		
Pos.-Neg.	1	"	"above it"
	15	"	"at one end"
	1	"	"like an exclamation mark"
	1	"	"beside"
	1	"	"behind or in front of"
	7	"	"behind" and "after"
	1	"	"and"
Total:	27		
"Oblong"			
Positive	7	Zif	"square"
	1	"	"long square"
	1	"	"long shaped box"
Total:	9		
Pos.-Neg.	1	"	"long square"
	1	"	"triangle"
	1	"	"box"
	2	"	"square"
	1	"	"oblong square"
	1	Wez	"a long stick like a square with short ends"
Total:	7		
"Circle"			
Positive	5	Mib	"oval"
	12	"	"round circle"
	5	"	"ball"
	1	"	"globe"
Total:	23		
Pos.-Neg.	1	"	"round circle"
	3	"	"oval"
Total:	4		

TABLE X (Cont'd).-VERBALIZATION MEANING DISPARATE FROM RECOGNITION MEANING.

"Meaning"	Frequency	Concept	Verbalization
"Touching"			
Positive	2	Mib Wez	"joined together"
	1	Wez	"beside"
	1	Wez	"put together in any fashion.
	1	Mib	"connected"
	1	"	"hitched on to"
	1	"	"together"
	1	"	"fixed on the side"
Total:	8		
Pos.-Neg.	1	Wez	Doesn't say "touching" but Rec. and Ref. show this meaning.
	1	Zif	Doesn't say "not touching end" but both Rec. and Ref. showed this idea.
	1	Mib	"stuck together"
	2	Mib	"joined"
	1	Wez	"on"
Total:	6		
"Shortest end"			
Positive	1	Wez	"blunt end"
Total:	1		
Pos.-Neg.	1	Wez	"this end here", pointing.
	5	"	"the end part of the triangle".
	1	"	"the flat end of the triangle".
	1	"	"the top of the triangle".
Total:	8		
"Square"			
Positive	1	Mib	"box"
	1	"	"block"
Total:	2		
"Straight line"			
Positive	A great many		"line"
Pos.-Neg.	A great many		"line"
"Short sides of Oblong" (lying horizontally on its long sides)			
Pos.-Neg.		Zif	"bottom and top"

Table X, and the further evidence furnished below, suggest that no experimenter can know from the subject's words what thought the child has in mind. In other words, it is impossible to tell from the words a child uses, what he actually "means". The following examples support this conclusion.

"Vec" is defined as "a dot in a direct line with a straight line". At times when the Recognition performance indicated the existence of the idea "in a direct line", subjects verbalized as follows: "above it", "under", "like an exclamation mark", "beside", "in front of", "after", "and", "under", "over the top". Some subjects verbalized a "Vec" as "a straight line and a dot" and then properly identified all of the Recognition cards and produced drawings, all of which were "Vec". These subjects verbalized that "Vec" was "a line and a dot", but in the Recognition did not accept those cards on which a line and a dot appeared, if the dot was not in a direct line with the line. At another time, although a subject verbalized that a "Vec" was "a dot under a line", he identified as a "Vec" cards showing the dot horizontally to the left of a horizontal line, diagonally to the right and above a diagonal line, and vertically above a vertical line. The subject who verbalized a "Vec" as "a line with a dot over the top", identified as "Vec" those cards showing the dot directly under the vertical line. In the face of this

evidence the experimenter has reached the conclusion that it is very difficult to interpret what a subject may "mean" by a Verbalization. The evidence shown illustrates that Verbalization requires a great deal of subjective interpretation by the experimenter when it is used as a criterion.

Subjects expressed their idea of the "shortest side" of a triangle in a number of interesting ways. One subject attempted his Verbalization of "Wez" (a circle touching the shortest side of a triangle) with the statement, "A circle touching a triangle on this end here", pointing to the short side of several triangles which were in sight. Another subject expressed the idea of "shortest end" by saying, "the flat end of a triangle". Actually, it could be argued that both of the longer sides of a triangle are "flatter ends" of the triangle than is the short side. Another subject expressed the idea of the shortest end by the words "the top of the triangle", regardless of whether the triangle was lying on its side or standing on the so-called "top of the triangle". Another subject expressed the same idea with the words "blunt end". These methods of expressing the idea of "shortest end" are intelligible in the presence of other clues to their meaning in addition to Verbalization. Taken alone, without other clues, these Verbalizations would convey little definite meaning to an auditor.

The experimenter was interested to observe that "line" meant "straight line" to several subjects. In many cases

when the geometrical meaning of the word "line" was wanted, the boys indicated it by saying "line, which is either straight or crooked". In those cases where the subjects left the word "line" unqualified, it practically invariably meant a "straight line".

Several interesting labels for the idea of "touching" were given. Two subjects expressed the idea by the words "joined together"; other subjects by "beside"; "put together in any fashion"; "connected"; "together"; "fixed on the side"; "stuck together"; "joined"; and "on". "Wez" (a circle touching a triangle on shortest side) was defined as a circle <u>beside</u> a triangle. At the same time this subject identified as "not Wez" any figures in which a triangle and a circle lay in a horizontal position but not touching. He also correctly identified as "Wez" circles and triangles in a horizontal position tangent to one another. The Verbalization of Wez which said a "triangle and a circle fastened together in any fashion" could mean almost any relationship between a triangle and a circle. The subject, however, identified as "Wez" only those figures which were Wez and discarded as "not Wez" all those figures in which a triangle and a circle "were put together in any fashion" if they were not touching on the short side of the triangle.

Verbal labels applied by the subjects to triangles were varied and numerous. Simultaneously with Recognition

performance which to the experimenter would have indicated the word "rectangle", the subjects applied such verbal phrases as: "square", "long square", "long shaped box", "triangle", "box", "oblong square", "a long stick like a square with short ends". One subject indicated the short vertical sides of a rectangle lying horizontal as "bottom" and "top". It would be very difficult for any experimenter who relied on Verbalization alone and had no further clues of meaning to know what a subject was thinking who labelled his idea of "rectangle" with the verbal label, "triangle" or "square".

If Verbalization as defined in the present experiment is to be used as a criterion of learning, every misapplied word described in this section, plus the many more which these illustrate, must be evaluated. The Verbalizations must be marked as acceptable or not acceptable. Are they to be judged according to the standard set by the experimenter's vocabulary, or by the child's? It seems hardly reasonable to judge the definitions elicited from a twelve-year-old Grade 6 boy by the vocabulary standards of a university graduate. On the other hand, then, is it more satisfactory to evaluate the verbal response on the subject's own vocabulary level? If so, what is that level?

This description suggests that the great amount of interpretation required in evaluating Verbalizations,

renders Verbalization highly subjective as a criterion. In contrast, Recognition responses required no interpretation, being either correct or incorrect. Thus, Recognition is the more objective of the two criteria.

(b) Partial Verbalizations--Omissions "Understood"

One factor increasing the subjectivity of the evaluation of a Verbalization is the subject's frequent omission of part of the Verbalization as being "understood". The following examples illustrate this situation.

One subject who attained complete Recognition, verbalized the concept "Wez" (circle touching shortest side of triangle) as "circle touching the shortest line". This Verbalization was given both during the trial on which the subject achieved Recognition and also on subsequent trials. These subsequent trials produced correct Recognition and Reproduction. The subject omitted the idea of triangle from his Verbalization, probably because it was so obvious that he considered that the term "shortest line" would mean to the experimenter the "shortest line (side) of the triangle". In other words, the subject omitted to verbalize the idea, triangle--perhaps leaving it to be understood in the same way

that in common English the subjects of imperative sentences are left by the Verbalizer"to be understood". A further illustration was found in a Verbalization of "Zif". The Verbalization under consideration was acceptable on the second trial but subsequently, in a third trial, the Verbalization (according to the letter) was incomplete. Recognition and Reproduction were achieved on the second trial and consistently maintained in the third. It seemed to the experimenter that the subject intended parts of his previous Verbalization "to be understood". Another illustration is to be found in one subject's treatment of the concept "Mef" (a circle half black and half white). In this subject's Verbalization no mention is made of "circle", but in his Recognition he did not accept the blackened half-circle as being "Mef" even though this would have been the figure which would have corresponded to the actual words of his Verbalization. Although his Verbalization contained no mention of "circle", in Recognition the subject identified as "Mef" only those figures in which a complete circumference, half blackened, was shown. Furthermore, in drawing his Reproductions, all examples of "Mef" were complete circumferences, half blackened; i.e., "Mef" was correctly recognized and reproduced. Perhaps the subject omitted the idea of complete circumference from his Verbalization because he considered it obvious that it was to "be understood".

Since "to-be-understood" omissions seem to occur, how

should an incomplete Verbalization be interpreted? The presence of this problem suggests that (as a criterion of learning) there is a subjective quality in Verbalization. On the other hand, Recognition as a criterion of learning meets no such objection.

(c) Omissions Which May or May Not be Expected to be "Understood"

Another factor contributing to the subjectivity of Verbalization as an experimental criterion is that omissions which may or may not be expected to be "understood" appear in Verbalization simultaneously with achievement in Recognition and Reproduction. The concept "Zif" contributes many examples of this factor. "Zif" is defined as a "circle inside a rectangle, the circle touching both long sides of the rectangle, but not touching either end". Five subjects, in verbalizing this concept did not mention "touching both sides". However, they identified as "Zif" only cards containing this characteristic and rejected any cards which were an actual graphic representation of their Verbalization but in which the circle did not touch both sides. Their Reproduction as well as Recognition gave the idea of "touching both sides", even though the Verbalizations continued to omit this characteristic. In one case the subject in making a spontaneous comment about "Zif" stated that <u>it did not matter</u> whether the circle touched both sides or not, as long as it was in the square. Subsequent to making his comment he

rejected any Recognition cards in which the circle did not touch both sides. Further examples were found where both Recognition and Reproduction conveyed the idea of "not touching end", while Verbalization did not mention this limitation. In one case, although the Verbalization of "Tov" did not mention "near", the Recognitions and the Reproductions showed the existence of the idea of nearness.

In determining the acceptability of such responses, great dependence upon the experimenter's judgment is thus necessitated when using Verbalization as a criterion of concept formation. In contrast, the subject's responses to the Recognition test require no interpretation by the experimenter. Verbalization is shown, therefore, to require a greater amount of interpretation by the experimenter than does Recognition.

(d) Verbalizations--Relatively Incorrect or Incoherent

In view of vocabulary differences between experimenter and subject, where is the dividing line between acceptable and unacceptable Verbalizations? When does incoherent wording become coherent?

A fourth factor contributing to the subjectivity of Verbalization as a criterion of concept formation is the appearance of numerous incorrect and incoherent Verbalizations simultaneously with correct Recognition and Reproduction. It might be argued that such performance indicates that a concept

has not yet been formed completely but only in the Recognition and Reproduction levels and for this reason the achievement of Verbalization would be a good criterion of concept "formed". This argument carries little weight, however, in the light of the many "Verbalization-precedes-Recognitions reported in this experiment.

One subject verbalized "Mef" as a "circle with a black and white spot", but identified only those figures as "Mef" which were half black and half white. In the concept "Tov", many Verbalizations gave the idea of only one "X", while the corresponding Recognitions and Reproductions have the idea of four "X's". One Verbalization of "Vec" used the words "horizontal line" in the Verbalizations but recognized and reproduced his "Vec" lines in any position, horizontal, vertical or oblique. Many meaningless Verbalizations were given for the harder concepts, even though Recognition and Reproduction were correctly indicated. For example; for "Wez" (a circle touching triangle on shortest side) the incoherent Verbalization "a circle that touches the bottom" was offered with correct performances of Recognition and Reproduction. For the concept "Zif", the Verbalization "touching twice, both sides of square" was given to correspond to Recognition and Reproduction performances in which a "circle touched both sides of a rectangle once". The Verbalizations given by the subjects were often incoherent when divorced from any other clues to meaning. The following

is an example of this incoherence, "a square with a circle touching the two oblong sides of the square".

Verbalized responses have to be marked as definitely right or definitely wrong. There is no such definite line of demarcation in the Verbalizations themselves. Where the dividing line shall be, becomes a matter of judgment by the experimenter. This shows that Verbalization, when used as a criterion of learning, is inherently subjective.

4. Correct Verbalization Whose Significance is not Understood

Another factor contributing to the questionable nature of Verbalization as an experimental criterion of concept formation was the ability of subjects to express the correct Verbalization before they were aware of the significance of the words they were using. The correct Verbalization was often given while Recognition could not be achieved. The concept "Mef" was often verbalized correctly with the words "a mef is a circle half black and half white". Immediately after this Verbalization, several subjects accepted as "Mef" cards in the Recognition test which the experimenter would describe as "circles partly black". These instances accepted as "Mef" were circles some more than half black and some less than half black. In other words, in many cases the subjects said "half" when they 'meant' "part". Another subject in verbalizing "Zif" used the words "touching both edges", but the Recognition test

accepted six cards which did not touch both edges. One subject gave the acceptable definition that "Mib" was a "square with a circle touching it". In geometrical terminology this definition allows the circle to "touch" the square either internally or externally. On the Recognition test, the instances showing external tangency between circle and square were accepted as "Mib". The figures exhibiting internal tangency were rejected as "not a Mib". These are examples where the subject gave the acceptable experimental Verbalization. It seems, however, that he did not realize the significance of his words, and, consequently, he could not pass the Recognition Test. In these cases, the exigencies of poor habits of speech (saying "half" for "part" and omitting detail by saying "touching" instead of "touching on outside") reaped for the subject the fruits of the combination of generalizing ability and precise expression of ideas. If Verbalization is used as the criterion of concept formation, situations like this arise. This aspect of Verbalization detracts from its usefulness as a criterion of learning.

5. Static Verbalizations

Another factor contributing to the unreliability of Verbalization as an experimental criterion of concept formation is that the Verbalization sometimes remains static while performance on the Recognition cards progresses to achievement. To illustrate this, in one case while successive Verbalizations of the concept "Zif" remained

identical, Recognition progressed, expanding on the second trial to the idea of "touching both sides" and on the third trial to "not touching an end". Whereas on the third trial the Recognition Test and the Reproduction Test were passed successfully, Verbalization remained identical to earlier attempts. Thus Verbalization remained static, not indicating any growth or change in the development of the concept. Assuming that the instances seen by the subject after the first trial modified the concept to some extent, the conclusion follows that Recognition was the more sensitive indicator. This argument suggests that at times Recognition is a less static criterion of learning than is Verbalization.

6. An Interesting Case

The inability of subjects to recognize a change in their ideas was indicated in an interesting case where a subject learning the concept "Mef" volunteered information several successive times that he had not changed his mind at all. In this case, all levels--Verbalization, Recognition and Reproduction--had been changed from the initial incorrect responses to final correct responses for all criteria. The subject, however, after achieving the concept on all levels, still happily volunteered the information that in his last trials he thought exactly the same thing he had thought in his first ones.

7. Summary and Conclusions

1. Verbalization is not a "level" that can be depended upon to appear later than Recognition. In fact, in this present experiment, Verbalization preceded Recognition more often than vice versa.

2. Verbalization and Recognition are approximately equally subject to reversal of judgment. In neither case do reversals occur frequently.

3. Recognition is a more objective criterion of concept formation than is Verbalization. The latter requires a great deal of interpretation by the experimenter as to what the subject "means", while the former does not.

4. Verbalization and Recognition often give conflicting evidence regarding the nature of the subject's concept. Subjects misapply words, become incoherent, and make omissions which may, or may not, be intended to "be understood", while at the same time giving correct responses on the Recognition tests. On the other hand, many correct Verbalizations are made while Recognition responses are still incorrectly made.

5. Recognition, rather than Verbalization, is a more sensitive indicator of the degree to which a concept has been formulated. It can be used to measure the progress of concept formation at any stage, and does not require the concept to be completed before it yields numerical data convenient to statistical treatment.

6. For the accurate communication of ideas such as are used in the testing situation of this experiment, twelve-year-old Vancouver boys require other media of expression than Verbalization alone.

7. This study employs Recognition rather than Verbalization as its criterion of concept formation. Their advantages and disadvantages appear to be approximately even as regards (a) "levels" and (b) reversals of judgment. Recognition appears superior to Verbalization as an experimental criterion of concept formation because it is (a) more sensitive than Verbalization, and (b) far more objective than Verbalization.

CHAPTER IV.

SOME RESULTS AND CONCLUSIONS

1. Which Method, the Positive or the Positive-Negative was Antecedent to the Greater Tendency Toward Description Rather than Generalization?

Although in the following two paragraphs no hard and fast definitions are given, "generalization" is considered to be the recognition of the highest common factor found in the positive examples of a given concept, while "description" is a statement of many of the characteristics, common or otherwise, found in the positive instances seen by the subject. An example of a description might be, "A Dax is a long triangle and a short triangle with a dot near the bottom or a fat triangle with the dot in the middle or a dot in one end or a big triangle or a little triangle". Verbalizations did not have to be this long to be called a "description". In "description" frequent use is made of the conjunctions "and" and "or". In other words, description is defined as a running enumeration of the characteristics of the positive cards seen.

For each type of presentation, twenty-five subjects tried eight concepts each, and were allowed to verbalize each concept as many times as they wished. Any, all, or none of the Verbalizations could have been descriptions.

The positive presentation in this experiment was antecedent to twenty-eight descriptions, the positive-negative to twenty-seven descriptions. The difference between the A.M.'s of the number of descriptions per subject

was .04, while the standard error of the difference was .067. This indicates that in 73 chances out of 100, the positive presentation, on the average, will advance more description than will the positive-negative (24 p. 213). The difference, however, is not significant.

The results of this experiment do not indicate that either method of presentation is antecedent to any significant tendency toward either generalization or description.

2. Is the Recognition of a Concept Gradual or Sudden?

Matheson (47 p. 260) describes the existence of insight in these words "understanding of a situation exists in degrees and ..... the complete understanding which Kohler characterizes as insight represents one experiment of the distribution when all cases are considered". Alpert (2) reports three types of solution: (a) solution with the minimum insight; (b) solution with gradual insight which may be partial or complete; (c) solution with sudden insight which may mature during exposure or between exposures. Both Tyler (71) and Patton (54) in experimentation with problem situations whose solution depended upon the subject's discovery of a principle, concluded that the emergency of insight might be gradual as well as abrupt. Drever (17) similarly reported the discovery of a principle as gradual.

On the graphs shown in Figures 11 to 18, sudden insight would be indicated by a sharp upward acceleration in

the curve. A comparatively straight curve, whether its gradient be steep or shallow, indicates gradual recognition development. A straight steep curve, like the positive presentation of "Zum", indicates gradual development of an easy concept.

The sharp upward accelerations found in Figure 11 at Card 3 (both presentations), in Figure 12 at Card 3 (positive-negative method), in Figures 16, 17 and 18 at Card 3 (positive-negative presentation), cannot be interpreted as being due to sudden insight. Another factor, the compulsory discard of caution[1], is operative.

In the present study, the following figures exhibit comparatively straight curves:

(a) Figure 12, positive presentation, cards 1-4
(b) Figure 13, positive presentation, cards 1-8
(c) Figure 13, positive-negative presentation, cards 1-8
(d) Figure 14, positive presentation, cards 1-4
(e) Figure 15, positive-negative presentation, cards 1-8
(f) Figure 16, positive-negative presentation, cards 1-8
(g) Figure 18, positive presentation, cards 1-8

This suggests that during at least a part of their formulation, these concepts developed gradually.

This study, therefore, shows some evidence pointing to gradual development of concepts but gives no conclusive evidence on the question of the existence or absence of sudden insight.

---

[1]This factor is considered more fully in the next section of this chapter.

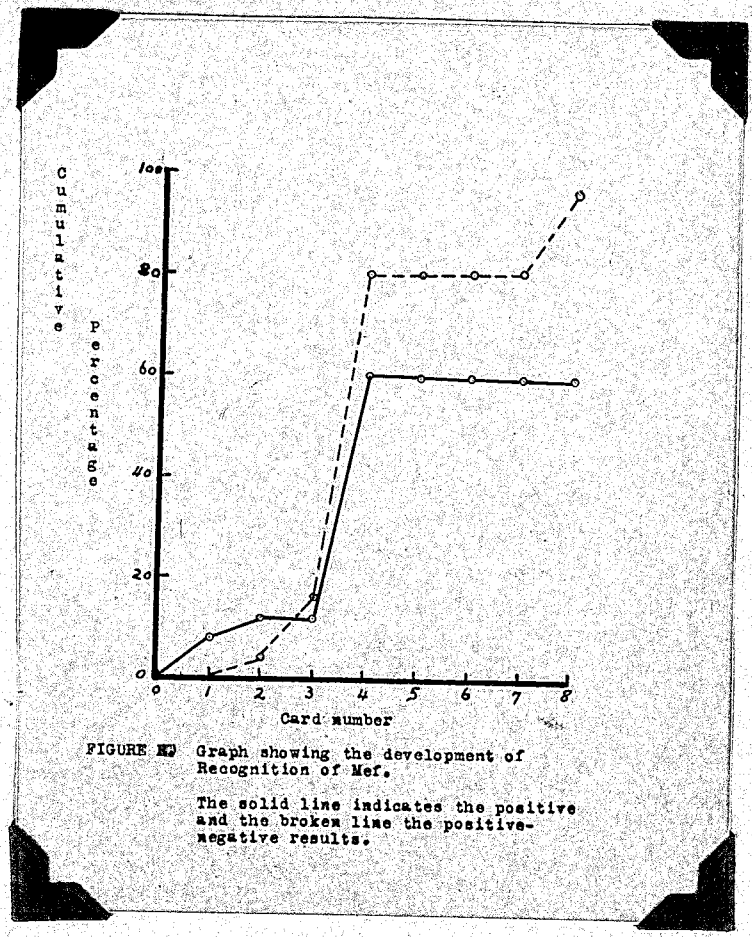

FIGURE [illegible] Graph showing the development of Recognition of Mef.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 11.- MEF - Cumulative Achievement is plotted against the number of instances shown.

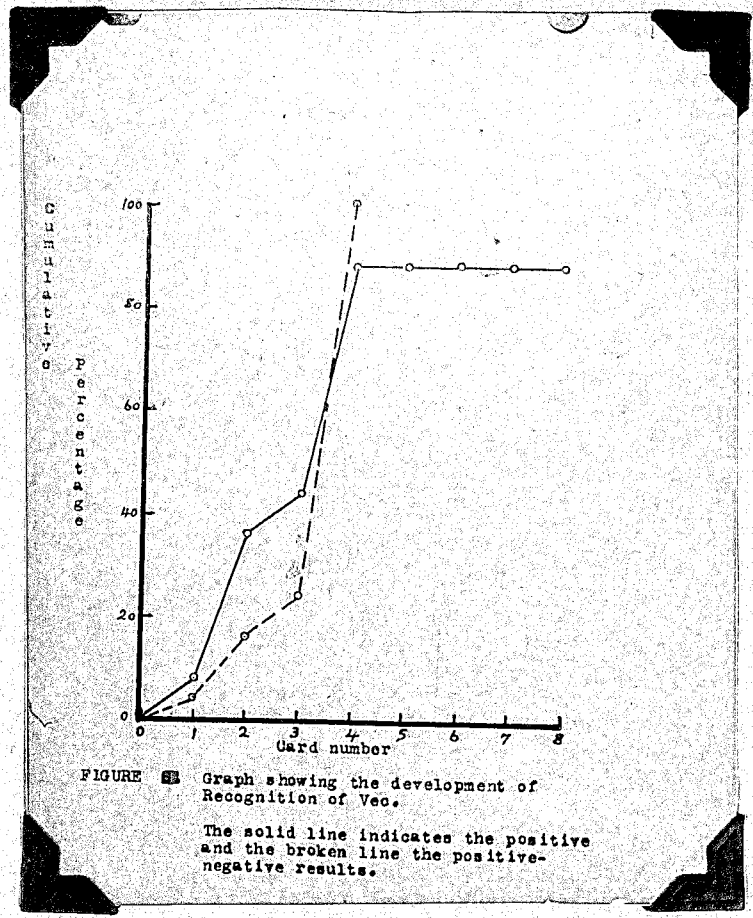

FIGURE [illegible] Graph showing the development of Recognition of Vec.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 12.- VEC - (see Figure 11 note)

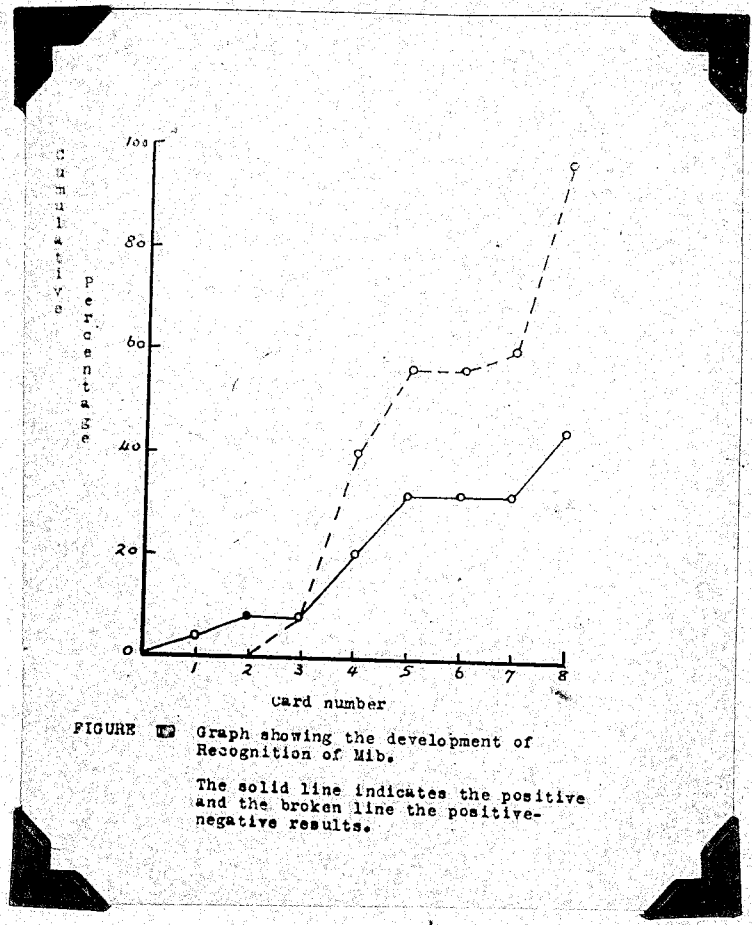

FIGURE 13 Graph showing the development of Recognition of Mib.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 13.- MIB - Cumulative Achievement plotted against the number of instances shown.

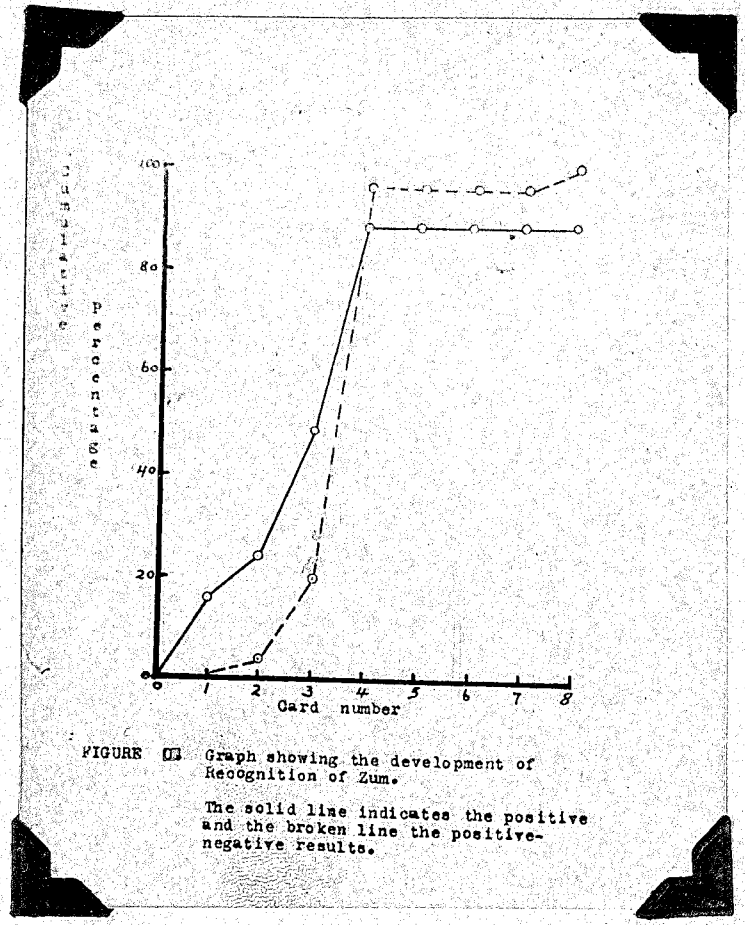

FIGURE 14 Graph showing the development of Recognition of Zum.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 14.- ZUM - (See note on Figure 13).

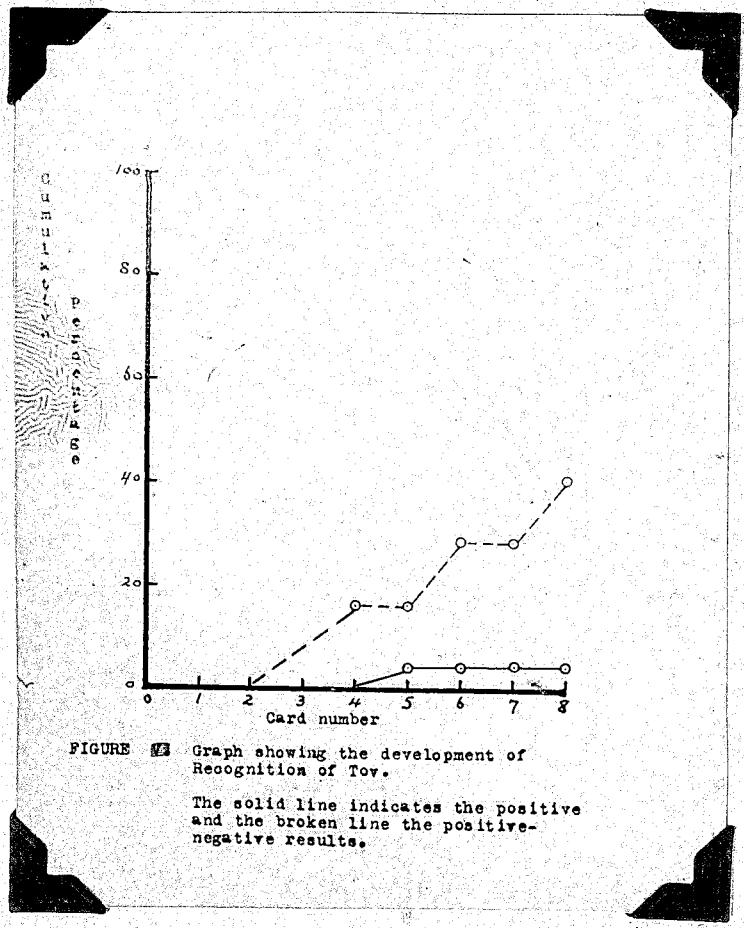

FIGURE [illegible] Graph showing the development of Recognition of Tov.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 15.- TOV - Cumulative achievement plotted against the number of instances shown. Only 4% of subjects achieved Recognition by the Positive Method.

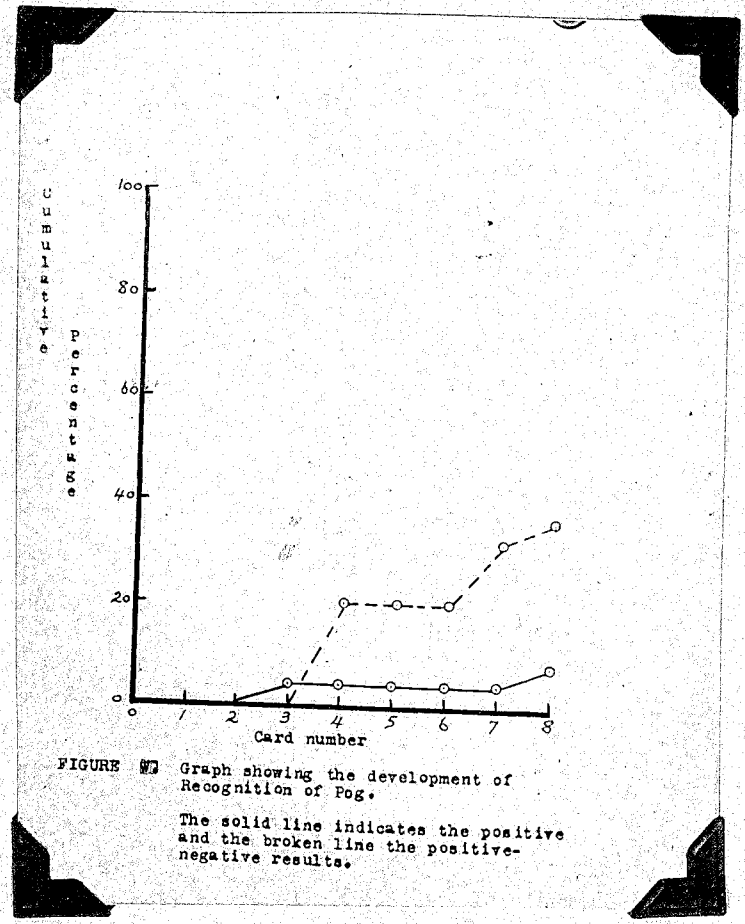

FIGURE [illegible] Graph showing the development of Recognition of Pog.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 16.- POG - (See note in Figure 13). By Positive Method only 8% achieved Recognition.

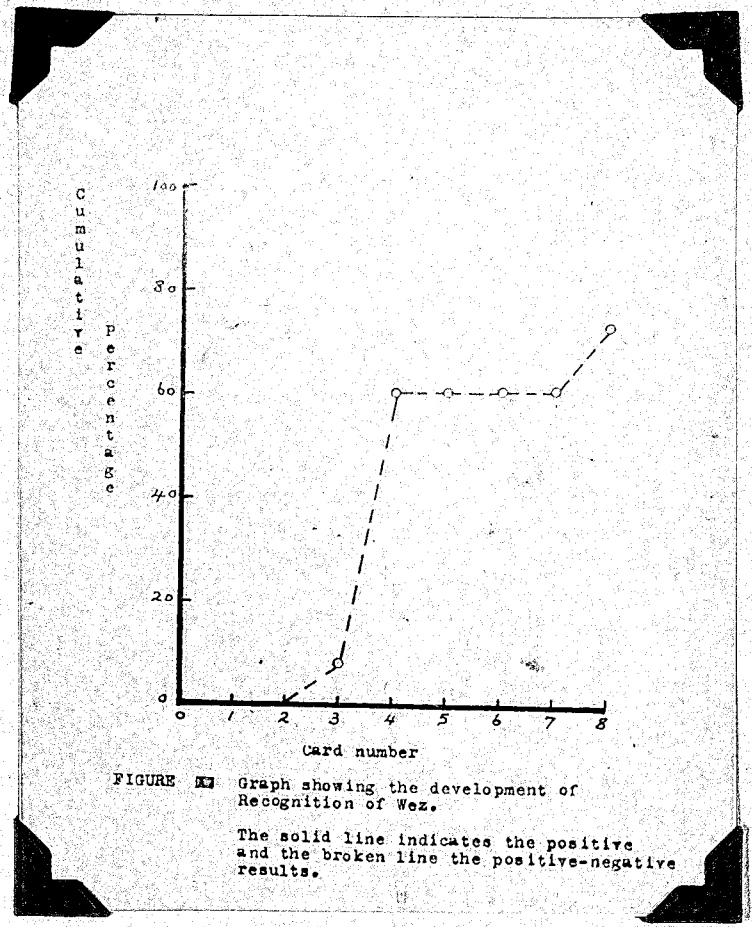

FIGURE [illegible] Graph showing the development of Recognition of Wez.

The solid line indicates the positive and the broken line the positive-negative results.

Figure 17.- WEZ - Cumulative achievement plotted against the number of instances shown. By Positive Method, no one achieved Recognition of WEZ.

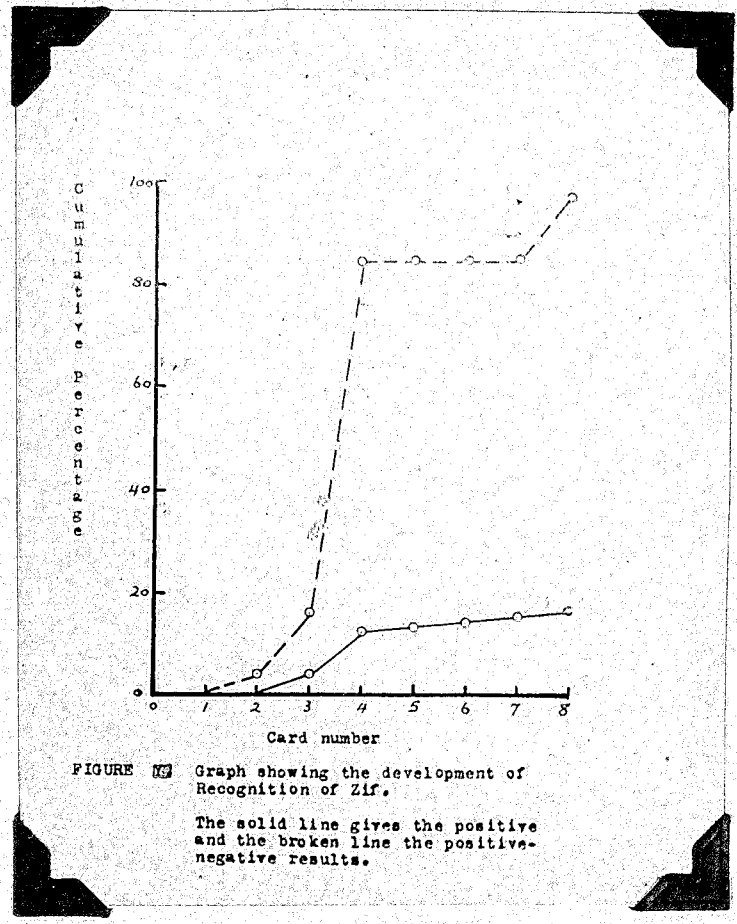

FIGURE [illegible] Graph showing the development of Recognition of Zif.

The solid line gives the positive and the broken line the positive-negative results.

Figure 18.- ZIF - (See note on Figure 17). Only 16% achieved Recognition in the positive situation.

3. Which of the Two Experimental Types of Presentation Gave the Greater Impetus to Caution?

The frequent appearance of a positive acceleration[1] in the curves at Card 3 show that many partial concepts existed which were not sufficiently well recognized to encourage the subject to volunteer them. On seeing Card 4, the subject was stripped of his caution, since he was required to give expression to his nebular impression, and frequently he was able to make the correct generalization.

If the curve between Card 3 and 4 continued at the same pitch as the curve up to Card 3, this would indicate that even when told to do so, the subjects could not dip below into their general hazy impressions. In other words, constant variation from card to card indicates there is no change in the amount of caution being exercised. The amount of the positive acceleration of the slope at Card 3 is a criterion of the extent to which the subject wanted to "make sure". The greater the amount of positive acceleration at Card 3, the greater was the subject's desire to test this hypothesis; i.e., the greater was his desire to wait and see whether following examples corroborated or refuted his generalization. Where the positive-negative show greater difference in pitch than the positive, indication is given of greater desire to test the hypotheses when presentation is

---

[1] Mathematically speaking, acceleration is defined as a change in velocity. The velocity is indicated by the slope of the curve. Accordingly, in speaking of the graph, "at" rather than "after" should introduce the phrase "at Card 3".

by the positive-negative method than when it is by the positive method.

In the concept "Mef" of the present experiment, the positive presentation and not the positive-negative gave the greater positive acceleration to the curve at Card 3 (see Figure 11). In every other concept except "Mib", however, the positive-negative presentation shows the greater corresponding acceleration (see Figures 12,13,14,15,16 and 18). The concept "Wez" cannot enter this discussion because no subject given the positive presentation achieved it. The ratio is, therefore, 5 to 2 in this experiment that the positive-negative presentation, rather than the positive method, will produce the greater positive acceleration on the graph at Card 3.

These results could be explained by the hypothesis that of the two types of presentation used in this experiment the positive-negative method gave the greater impetus to caution. Apparently, the inspection of negative examples made the subjects hesitate to hazard a guess before they had seen enough examples to "make sure".

Tyler's study (71) pointed out that the testing of hypotheses was a part of the generalizing process. The preceding paragraphs discuss the relative role of positive and negative instances in increasing the desire to test hypotheses, and implies that in this capacity the role of the negative is the greater. In Chapter V, further evidence that supports this conclusiom is presented.

4. Factors of Difficulty in Concepts

(a) "Mental Inertia" or "Rigidity" and the Influence of Negative Instances

Many factors have been suggested as contributing to the difficulty of concepts. Egger's (20) findings that "Mental inertia (fixation, direction, or perseveration) is the factor which most markedly interferes with successful reasoning behaviour", is in accord with the findings of many other experimenters. Welch (79,80) states that a relationship is harder to learn than an element. Tyler (71) indicates a lack of correlation between the difficulty of the concept and the combined number of elements and relationships. He suggests the following factors of difficulty:-

(a) The number of words required to express a rule;

(b) The number of hypotheses that can be formulated;

(c) The number of patterns that must be shown before all hypotheses are tested;

(d) The abstractness of the elements and relationships involved.

Crudden (15) found: (a) that the difficulty of the concept increased in direct proportion to the complexity of the figure in which it is imbedded; (b) that knowledge of the "figure-to-be-abstracted" increases the ease of abstraction.

Many psychologists in discussing generalizing and reasoning have pointed to rigidity as a factor of difficulty. Crudden (15) found "that-which-is-to-be-avoided" in

abstraction has almost as much influence in successful abstraction as "that-which-is-to-be-chosen". Maier (46) suggests that many errors in thinking are due to the rigidity of the reasoner's set. Duncker (18) thinks somewhat similarly to Maier. Rees and Israel (59) state that a rigid mental set often prevents the subject from seeing details which ought to make him reject a previous hypothesis. Chant (12) in suggesting that associations and meanings determine many incorrect responses, was suggesting that rigidity in the "centering points" was a contributing factor to error. Siipolo (65), Ewert and Lambert (21) and Sullivan (70) all emphasized the role of mental inertia in producing error.

In the foregoing review of experimenter's conclusions, "rigidity" and "mental inertia" have been named repeatedly as being antecedents to erroneous hypotheses. Crudden found that "that-which-is-to-be-avoided" was an important antecedent to successful abstraction. Why was it helpful? Its value lay in its ability to upset the tranquility of an erroneous hypothesis and so to overcome "mental inertia" and "rigidity". What is Crudden's "that-which-is-to-be-avoided" but a negative instance! According to Crudden's conclusions, therefore, a positive-negative presentation should be superior to a purely positive method, due to the influence of the negative instances in preventing "mental inertia" or "rigidity".

The conclusion that negative instances assist toward

correct abstraction is supported by still another line of reasoning. Gengerelli (28) found that it is more difficult to conceive an object a second way after having learned to conceive it a first way. The section on "Impetus to Caution" showed that the influence of the negative instances was toward a tendency of suspended judgment; i.e., a tendency not to have conceived an hypothesis hard and fast in a "first way". This tentative attitude would soften the effect of having had an incorrect hypothesis. Accordingly, it is to be expected that the introduction of negative instances will lessen the difficulty of concept.

Table XI shows the number of Recognitions achieved for each concept by each method.

TABLE XI.- ORDER OF DIFFICULTY OF THE CONCEPTS

Order of Difficulty Hardest to Easiest (out of 50)	Positive Presentation		Positive-Negative Presentation		Both Presentations	
	Name	Number Correct* (out of 25)	Name	Number Correct* (out of 25)	Name	Number Correct* (out of 50)
1.5	Wez	0	Tov	10	Tov	11
1.5	Tov	1	Pog	9	Pog	11
3	Zif	5	Wez	18	Wez	18
4	Pog	2	Zif	24	Zif	29
5	Mib	11	Mef	23	Mib	35
6	Mef	15	Vec	25	Mef	38
7	Zum	22	Mib	24	Zum	47
8	Vec	22	Zum	25	Vec	47
TOTALS:		78		158		

Some concepts, while being achieved by a high percentage of the subjects who saw both positive and negative

*Criteria of "correct" was the successful Recognition of the 16 test cards for each concept.

instances, were learned by only a small fraction of the subjects who saw only positive examples. These results, therefore, corroborate the conclusions of those experimenters who suggest that mental inertia and rigidity were factors of difficulty in concept formation. These results also strengthen the hypothesis that introducing negative instances tends to develop an attitude of suspended judgment, which, in its turn, increases concept achievement by preventing definite acceptance of an incorrect hypothesis.

(b) The Acceptance by the Subject of Incomplete Hypotheses and the Influence of Negative Instances:

In nearly all cases where the concept was not achieved, part of it, but not all of it, was accepted as the whole of the concept. Particularly was this glaringly evident in the group where generalization was done from positive examples only. Any factors which can be shown to encourage the acceptance of common factors less than the highest common factor, become major factors of difficulty.

The logical analysis submitted in Chapter V suggests that because the presence of the characteristics composing an incomplete hypothesis can be abstracted from every positive instance, positive examples tend toward being antecedent to incomplete hypotheses composed of part of the correct hypothesis, but not all of it. Further, the logical analysis to be found in Chapter V shows that a negative instance can exist to eliminate directly any and all incomplete hypotheses. Therefore, the inclusion or

exclusion of negative instances along with positive examples is a major factor related to difficulty in concept formation.

(c) Complexity of Concept:

The differences between the number of Recognitions achieved by the positive and by the positive-negative presentation are different for each concept. Since the inclusion or exclusion of negative instances constituted the experimental variable, these results would suggest that the efficacy of introducing negative examples varies with the concept. The logical analyses of negative instances submitted in Chapter V support this conclusion.

Table XI shows that when the results of both presentations are combined, the concepts "Tov" and "Pog" were recognized least frequently. Why were these generalizations harder than the others?

A "Tov" was defined as "a square with an X near each of the four sides". The writer suggests that in the "Tov" there were more probable combinations of common factors than in any other concept. Consequently, a greater number of bypotheses could be formulated regarding "Tov" than could be with the other concepts. In pointing to the number of possible hypotheses as a factor of difficulty, the present study is adding supporting evidence to one of Tyler's suggested factors.

On the other hand, however, there is a serious

objection to speaking of "the number of elements and relations". The present study does not give a psychological analysis of its figures into a number of elements and relations, for that would be extremely difficult if not impossible. Who is to say whether a subject sees a triangle as one area, or as three straight lines intersecting to form three angles? Is a segment of a circle one element or area, or is it two radii and one arc enclosing a space? Who is to say how many elements or relations there are in concepts such as were employed in the present experiment.

Crudden stated that "knowledge of the figure-to-be-abstracted" increases the ease of abstraction. The writer would suggest that the explanation of this is that "knowledge of the figure-to-be-abstracted" knits what would otherwise be several elements and relations into a unit and so reduces the "number of elements and relations" to be contended with.

(d) Abstractness of the Concept:

A "Pog" was defined as "two unequal lines". Table XI lists "Pog" as being one of the two hardest concepts. Why should "Pog" be hard? The writer suggests that the relationship of inequality in length is quite abstract when applied to an element which can be varied in as many ways as "line" can be. This evidence, therefore, supports another of Tyler's factors of difficulty, namely, the abstractness of the elements and relationships involved.

(e) "Minimum Number of Words Required":

Tyler suggests that "the minimum number of words required to express a rule" is a factor of difficulty. The writer would encounter two obstacles in attempting to use this suggested criterion of difficulty. First, how could "the minimum number of words required" be determined? Secondly, how can the number of words be counted?

First, how is "required" to be defined? Required by whom? The experimenter? Webster? Or the subject? If the answer is "required by the subject", how is the experimenter to know whether a given subject's vocabulary would have permitted him to define a "Dax" as "a triangle enclosing a dot", or whether that definition was beyond him and his vocabulary "required" him to define a "Dax" as "a triangle with a dot in it"?

Secondly, after the difficulty of defining "minimum number of words required" has been overcome, another obstacle in counting "the number of words" is encountered. Are "a's" and "the's" to be counted as words? If so, how would we deal with the definitions of two subjects one of whom said, "Dax is triangle with dot inside" and the other of whom said, "a Dax is a triangle with a dot inside"? In other words, the subject's optional inclusion or omission of "a's" and "the's" would present difficulty. If "a's" and "the's" should not be included, should one count words which could be replaced by "the" without obscuring the meaning of the definition? For

example, if "the" is not to be counted as a word, should "its" be counted in the following partial definition, "a circle touching a triangle on its shortest side"?

After considering the perplexities inherent in trying to use as a criterion of difficulty the number of words required to express a concept, the writer would not attempt to utilize it. However, the author is willing to concede that to the extent that the number of words indicate the complexity of the concept, it may be a criterion of difficulty.

5. Relationship of Presentation Method and Time

Both Haught (36) and Tyler (71) are agreed that in studies like the present one, "time" is not an important measure of learning. Smoke (67) found no significant difference between the time required to learn concepts from a positive-negative presentation and the time necessary when the presentation was from positive instances only.

Table XII compares the time factor in the two methods used in the present study.

TABLE XII.- RELATION OF TIME FACTOR TO METHOD OF PRESENTATION

	Positive (min.)	Positive-Negative (min.)	Critical Ratio
Total time for 25 S's	1275	1340	
Mean time	51	54	.174 (57 chances in 100) not significant
σ time	16	12	1.4 (92 chances in 100) not significant

The positive-negative took slightly longer than the positive presentation. Comparison of the mean time per subject, however, shows only three minutes difference. This gives a critical ratio of .174, i.e., 57 chances in 100, that there is a real difference in the time required. The critical ratio of 1.4 shows that there are 92 chances in 100 that the time taken in the positive-negative presentation will be less variable than the time required for the positive presentation.

The results of the present study, therefore, support the view that in experiments like this one, time is not an important measure of learning.

6. Summary and Conclusion

1. There is no significant difference between the two methods in the tendency toward description rather than generalization.

2. The evidence supplied by this experiment supports the hypothesis that concepts are formulated gradually rather than suddenly.

3. The introduction of negative, in addition to positive, examples increases the attitude of suspended judgment and thus gives an impetus to caution.

4. "Rigidity" or "mental inertia" tend to be antecedent to erroneous hypotheses. Negative instances tend to distrub the "rigidity" and "mental inertia" and thus contribute

to greater achievement.

5. Many errors are made by the acceptance of incomplete hypotheses. The introduction of negative instances lessens this difficulty.

6. The difficulty of a concept increases with -

   (a) Complexity,

   (b) Abstractness.

7. It is very difficult to count the number of elements and relations in a geometrical concept.

8. It is very hard, if not impossible, to determine the "minimum number of words required" in the definition of a concept. Thus this measure cannot be used as an experimental criterion of the difficulty of a concept.

9. In an experiment like the present one, time is not an important measure of learning.

CHAPTER V.

THE RELATIVE ROLE OF POSITIVE AND NEGATIVE INSTANCES

1. In the Present Experiment, Which Method of Presentation, the Positive or the Positive-Negative, Produced the Greater Achievement?

(a) Comparison of the two methods as to the achievement on each concept and also on the totals of all eight concepts

Table XIII exhibits experimental results which are crucial to the problem stated above. In the column headed $\frac{D}{\sigma D}$ is shown the critical ratio between the achievements for each concept by the positive and by the positive-negative methods of presentation. The critical ratio of totals, also, is shown. Garrett's short formula for the standard error of differences between percentages was used. A more accurate formula gives critical ratios greater than the results of Garrett's short formula.

It will be noted that the critical ratios of four concepts are significant, while the remaining critical ratios show that on the average, in at least 99.77, 99.45, 96.5 and 96.5 chances out of 100, respectively, the efficiency of learning is greater on each of the given concepts by the positive-negative than by the positive presentation.

The critical ratio of totals (3.3) shows that, under present experimental conditions, it is virtually certain that subjects will, on the average, always learn these experimental concepts with greater efficiency in the present positive-negative situation than in the experimental positive situation.

TABLE XIII.- CRITICAL RATIOS BETWEEN THE TWO METHODS FOR EACH CONCEPT.

	NO. OF CONCEPTS ACHIEVED[3]		% ACHIEVED		σ_p[1]		D_P	σ_{D_P}[2]	$\frac{D_P}{\sigma D_P}$	DIFFERENCE IS (chance in 100)
	Positive	Positive-Negative	Positive	Positive-Negative	Positive	Positive Negative				
VEC	22	25	88	100	6.50	0.00	12	6.50	1.846	96.5 chances
ZUM	22	25	88	100	6.50	0.00	12	6.50	1.846	96.5 chances
POG	2	9	8	36	5.43	9.60	28	11.04	2.535	99.45 chances
MEF	15	23	60	92	9.80	5.43	32	11.21	2.85	99.77 chances
TOV	1	10	4	40	3.92	9.80	36	10.57	3.405	Significant
MIB	11	24	44	96	9.93	3.92	52	10.67	4.87	Significant
WEZ	0	18	0	72	0.00	8.98	72	8.98	8.02	Significant
ZIF	5	24	20	96	8.00	3.92	76	8.91	8.53	Significant
TOTALS:	76	158	38	79	9.71	8.14	41	12.68	3.23	Significant

[1] $\sigma_p = \sqrt{\frac{pq}{N}}$ where σp is the standard error of a percentage; where p is the percentage of times a given event occurs; where q is 1 - p; where N is the number of cases. (Formula from Garret's Statistics (25)).

[2] σ_{D_P} where $\sigma_{D_P} = \sqrt{\sigma p_1 + \sigma p_2}$ is the standard error of the difference between two percentages.

[3] For purposes of this chapter, a concept was considered to be "achieved" or "learned" when all 16 test cards were correctly recognized.

(b) Comparison of the achievement of the whole positive group with the achievement of the whole positive-negative group:

Which method is the more efficient? Table XIV shows further statistics relevent to the question.

TABLE XIV.- CRITICAL RATIOS BETWEEN THE MEAN SCORES ACHIEVED BY EACH METHOD

	Positive	Positive-Negative
No. of Subjects	25	25
A.M. of concepts learned[1] by each subject	3.12	6.32
S.D. of concepts learned by each subject	1.31	1.09
S.D. of Mean (σ_M)	.27	.22
σ_D = .33 $D_{M_{P-N-P}}$ = 3.20	$\frac{D_{M_{P-N-P}}}{\sigma_D}$ = 9.7 (a significant difference)	

Those subjects who received the positive-negative presentation had a mean achievement (6.32) more than double that of the subject who saw only positive instances (3.12). The critical ratio of the difference between the means (3.2), and the standard error of their difference (.33), is significant (9.7). In other words, it is "virtually certain" that under present experimental conditions a pupil exposed to the positive-negative teaching cards will always score above a pupil learning from the positive instances alone.

Furthermore, the S.D.'s of each group (1.31 and 1.09) and the S.D.'s of the mean (.27 and .22)

[1] See footnote 3 of previous page.

suggest that in the group where negative instances were introduced there is less spread in achievement and consequently a greater stability of the mean.

Providing we can generalize from this experimental attack, these two conclusions have great significance to pedagogy. They indicate a two-fold superiority of the positive-negative over the positive method of presentation; i.e., (1) the positive-negative method produces an increased achievement; and (2) a greater uniformity of achievement in the group taught.

This first factor is of direct assistance to the learner, while the second factor, without retarding any bright learner, gives the teacher the advantage of producing a homogeneous achievement with all learners. This second advantage accrues from the relatively greater assistance rendered by the negative instance to the duller learner. This latter statement will be considered in following sections of the chapter.

(c) Comparison of the achievement of the group having both the highest I.Q.'s and the positive presentation, with the group having both the lowest I.Q.'s and the positive-negative presentation:

How do the number of concepts learned by the highest positive I.Q. group compare with the achievement of the lowest positive-negative I.Q. group? Table XV shows the statistics relevant to this question.

TABLE XV.- COMPARISON OF THE NUMBER OF CONCEPTS LEARNED BY THE POSITIVE HIGHEST I.Q. GROUP AND THE POSITIVE-NEGATIVE LOWEST I.Q. GROUP.

	Positive	Positive-Negative
No. of cases	6	9
A.M. of concepts learned	4.0	6.0
S.D. of concepts per subject	1.0	.82
S.D. of Mean	.45	.29
S.D. of Difference between means = .535	$\frac{\text{Difference}}{\text{S.D. of Diff.}}$ = 3.7 (significant)	

The table above shows that under the present experimental conditions, it is virtually certain that a group of dull normals (I.Q. 95 plus or minus 7.5) taught by the positive-negative cards, will on the average always have a higher Recognition achievement on this test than a superior group (I.Q. 125 plus or minus 7.5) exposed to positive instances only. The number of cases contributing these data was very small, only 6 in the one presentation and 9 in the other.

The statistics exhibited in Table XV, therefore, must be regarded only as barely suggestive. With this reservation, we may say that these results indicate that it may be that, within the I.Q. range considered, achievement in learning the kind of concept here presented depends more upon method of presentation than upon I.Q. The lowest I.Q. group in this experiment when given positive-negative presentation was on

the average more efficient at generalizing than was the highest I.Q. group when given the positive presentation only.

(d) Comparisons of the amount of advantage received from the introduction of negative instances by pupils of different I.Q. levels:

In the following sections, the evidence is analysed in two ways:-

A. In each of the three I.Q. levels, the pupils receiving the positive-negative presentation had higher achievements than those having the positive presentation. Section A compares, for each of the three I.Q. levels, the difference in achievement by each of the two presentation methods. This comparison is intended to determine whether it is the brighter pupils or whether it is the duller children who received the most advantage from the introduction of negative instances.

B. Section B deals with the same question as Section A, and analyses the same data, also, but it attacks the problem from a different angle. In both methods of presentation the group of bright pupils had a greater achievement than the group of dull pupils taking the same type of presentation. Section B compares the difference between the achievement of the bright and dull groups seeing the positive instances only, with the difference between the schievement of the bright and the dull groups exposed to both positive and negative instances. The purpose of this comparison is

to determine which method of presentation tends to give the more uniform amount of achievement.

If high I.Q. is associated with the ability to evaluate hypotheses critically, and if the introduction of negative instances is also an antecedent of the critical evaluation of hypotheses (as has been suggested both in this chapter and in the section on "Impetus to Caution"), then both high I.Q. and negative instances tend to be antecedent to the same thing. This "same thing" (critical evaluation of hypotheses) may be antecedent to achievement in this experiment. If these assumptions are correct, the negative instances compensate the dull learner for that which he otherwise lacks, and for that which the bright child enjoys. Accordingly, the introduction of negative instances into the learning situation would tend to obscure the influence of the difference in the I.Q.'s of the dull and the bright child.

A. Which groups, the dull or the bright, received the more advantage from the introduction of negative instances? and why?

The following table exhibits statistics showing in which of the three I.Q. groups (low, middle or superior) the introduction of negative instances made the greatest difference to achievement:-

TABLE XVI.- COMPARISON OF THE I.Q. GROUPS SHOWING THE AMOUNT OF ADVANTAGE IN EACH, GIVEN BY THE INTRODUCTION OF NEGATIVE INSTANCES

	I.Q. 95±7.5		I.Q. 110±7.5	
	Pos.	Pos.-Neg.	Pos.	Pos.-Neg.
A.M.	2.22	6.00	3.40	6.40
σ of group	.92	.82	1.28	1.11
σ_M [1]	.325	.29	.43	.37
σ_{D_M}	.436		.567	
Critical Ratio	8.7 (significant)		5.3 (significant)	

	I.Q. 125±7.5		I.Q. 110±22.5	
	Pos.	Pos.-Neg.	Pos.	Pos.-Neg.
A.M.	4.00	6.67	3.12	6.32
σ of group	1.00	1.23	1.31	1.09
σ_M [1]	.45	.55	.27	.22
σ_{D_M}	.71		.33	
Critical Ratio	3.8 (significant		9.7 (significant)	

Table XVI shows that subjects in the low I.Q. group learning from positive instances only, learned an arithmetic mean of 2.22 concepts. The group having corresponding I.Q.'s but seeing negative as well as positive instances, achieved

[1] Garret, H.E. "Statistics in Psychology and Education", p.201. Formula to apply if N<30.

an arithmetic mean of 6.00 concepts learned. The critical ratio of 8.7 shows the existence under experimental conditions of a significant difference between the means achieved by the two methods of presentation.

In the lowest I.Q. group, the difference between the means of the positive and positive-negative methods is 3.78 concepts. The corresponding statistic for the middle I.Q. group is 3.0 concepts, and for the superior group is 2.67 concepts. These figures suggest that the advantage enjoyed by introducing negative examples diminishes as the intelligence increases.

Whereas in the lowest I.Q. group the critical ratio between the achievement produced by the two methods is 8.7, this statistic dwindles to 5.3 for the middle I.Q. groups, and to 3.8 with the highest I.Q. group. This suggests the conclusion that the higher the I.Q. group the less significant is the difference between the number of concepts learned by the two methods. In other words, the higher the intelligence, the less useful are the negative instances, or vice versa, the lower the intelligence (within the limits of this experiment) the more useful are the negative cases.

These results strengthen the argument advanced elsewhere in this thesis that the chief role of the negative instance is its direct attack on incomplete hypotheses. Pupils with lower I.Q.'s have less ability for self-criticism than those with higher I.Q.'s. The imposition of an external check on the hypothesis of a pupil with a low

I.Q. supplies one of the missing factors necessary to a modified hypothesis. The dull child, given a positive presentation only, lacks this check. Accordingly, while the positive-negative presentation elevates the performance of a subject with a low I.Q. to proximity with the subject having a higher I.Q., the positive presentation leaves the low group with little achievement. The high group supplies its own self-criticism and so in the higher levels the negative instances are less needed. The conclusion is, therefore, that in this experiment the lower the I.Q. group, the more advantageous was the introduction of negative instances.

B. Which Method of Presentation, the Positive or the Positive-negative, is Antecedent to the more Uniform Amount of Achievement? And why?

A second avenue of attack on this question is supplied by comparing the critical ratio of the lowest and the highest I.Q. groups given the positive presentation with the corresponding critical ratio for the positive-negative method.

TABLE XVII.- COMPARISON OF THE ACHIEVEMENT MADE BY THE LOWEST WITH THE HIGHEST I.Q. GROUPS BY EACH TYPE OF PRESENTATION

	Positive		Positive-Negative	
	87 - 102	118 - 132	87 - 102	118 - 132
A.M.	2.22	4.00	6.0	6.67
$\sigma_{A.M.}$	.325	.45	.29	.55
* S.D. of $M_P - M_{P-N}$	.555		.621	
** $\frac{M_P - M_{P-N}}{\text{S.D. of Diff.}}$	3.2 (significant)		1.078 (86 chances out of 100)	

In Table XVII the critical ratio of 3.2 shows that when only positive instances are present it is virtually certain that the mean achievement of the highest I.Q. group of the experiment would be superior to that of the lowest I.Q. group. When negative instances were introduced, however, the chances dropped to 86 out of 100 that the highest I.Q. group would show superiority over the lowest.

In other words, when negative instances are absent, differences in I.Q. level are more likely to produce differences in achievement than when negative instances are shown. This result substantiates the conclusion that negative instances introduce an element which tends to obscure the influence of differences in I.Q. level.

*M_P is A.M. of positive presentation.

**M_{P-N} is A.M. of positive-negative presentation.

2. The Relative Roles of Positive and Negative Instances.

(a) Positive Instances:

In a series of positive instances, the concept to be formulated is the greatest aggregate of characteristics found in all positive instances, i.e., the highest common factor of the instances. The concept to be formulated is composed of more than one element and of the relation or relations between these elements.

Every positive instance is composed, therefore, as follows:-

$$I_p = F_{hc} + C_a \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$$

where

I_p represents any given positive instance;

$=$ means "is composed of" and has no reference to equality of magnitude;

F_{hc} is the abstracted concept (the highest factor or greatest aggregate of characteristics that the given positive instance has in common with all other positive instances);

C_a is any characteristic of the given instance, which characteristic is acceptable in, but not necessary to, the concept.

The C_a may vary from relative simplicity through many successive advances to great complexity. It is found at its simplest in the positive example requiring the smallest aggregate of characteristics comprising the medium from which an intellect can abstract the concept. It is found in its highest complexity in the medium containing F_{hc} and having the greatest aggregate of characteristics.

The writer suggests that, in addition to the concept (F_{hc}), some C_a is necessary in order to have a positive instance. If no C_a were present, the concept could not be expressed. No instance could be formed. No communication of the concept from one person to another would be possible. To illustrate this point, let us consider the concept "Dax" in the present experiment. Many methods of producing positive instances of this concept may be employed. The method used in the experiment was to draw a picture of a triangle with a dot inside it. The idea "triangle-enclosing-dot" was the F_{hc}. The actual lines and dot composed part of the C_a. These lines and dots were of different lengths, widths, and intensities in each individual instance. These lines and dots were characteristics which were acceptable in positive instances, but at the same time no given set of them was necessary to all positive instances - various sets of them could be used.

Thus the lines and dots of a given instance are recognized as being part of the C_a for that instance. Although no <u>given</u> set of lines and dots was necessary to all positive instances, nevertheless it is necessary to have some C_a in order that an instance of the concept could be present. Thus it is established that in order to have a positive instance, a medium must be presented from which an intellect might abstract the F_{hc}. At the same time, the very nature of medium requires <u>some</u> C_a as an integral part.

In no positive example is less than F_{hc} present..... (2)

Each different positive instance varies from every other positive instance by--and only by--differences in the C_a term. The most simple complete series of positive instances which contained every possible positive instance would be formed as follows:-

$$S_p = I_{p'} + I_{p''} + \ldots + I_{pr} \ldots\ldots\ldots\ldots (3)$$

where

S_p represents the most simple complete series of positive instances

= means "is composed of" and has no reference to equality of magnitude;

and the $I_{p'} + I_{p''} + \ldots + I_{pr}$ series represents every possible positive instance from the most simple ($I_{p'}$) to the most complex (I_{pr}).

The analysis of a positive instance given in formula (1) and the discussion thereafter, established that F_{hc} may be found in all positive instances of a given concept. It is the C_a that varies in each different positive example. Accordingly, in the S_p series, every C_a which can be an integral part of a positive instance of a given concept will be exhibited in association with that concept. In other words, a complete series of positive instances displays every situation in which the concept is present..... (4)

If statement (4) is correct, then one of the important roles of the positive instance is to enrich. Positive examples assist the student to orient and to recognize his

concept in the whole general field of which the concept may be a part.

(b) Negative Instances:

Any given negative instance could be represented as follows: $I_n = F_c \div C_a \div C_{na}$ .....(5)

where: I_n represents any given negative instance;

= means "is composed of" and has no reference to equality of magnitude;

F_c represents any factor common to the given negative instance and the concept to be formulated but

$0 \leq F_c < F_{hc}$;[#] F_c is necessary to, but not sufficient for F_{hc};

C_a is any characteristic of the given instance; which characteristic is acceptable in, but not necessary to, the concept.

C_{na} is any characteristic neither necessary to the concept nor acceptable in a positive instance, and varies between the limits $0 \leq C_{na} < Z$, where Z is the most complex aggregate of characteristics in which the complete concept (F_{hc}) is not found.

Values can be assigned to the variables in formula (5) such that I_n becomes F_c either by itself or with any other characteristic. Thus an instance corresponding to any partial concept can be displayed as not being the concept.(6)

Formula (5) demonstrates that in a negative example both C_a and C_{na} factors may be present. Thus from negative instances themselves, a subject cannot know whether a given characteristic may, or may not, be associated with the concept. Consequently, experience with negative instances can not

[#] $0 \leq F_c < F_{hc}$ is a common mathematical symbol meaning that F_c is equal to, or is larger than, 0, but is less than F_{hc}; and that F_c may vary within these limits.

enrich knowledge of characteristics associated with the concept. On the other hand, a complete positive series displays every situation in which the concept may be found (4). We conclude, therefore, that the positive rather than the negative instances assist in orientation and enriching.

Any instance may be a negative instance if it does not contain the abstracted concept. A complete negative series would, therefore, contain all combinations of terms except F_{hc}, and could be represented thus:

$$S_n = I_{n'} \div I_{n''} \div \ldots\ldots\ldots\ldots \div I_n{}^z \ldots\ldots\ldots\ldots \quad (7)$$

where

S_n represents the most simple complete series of negative instances;

= means "is composed of" and has no reference to equality of magnitude;

and the $I_{n'} \div I_{n''} \div \ldots\ldots \div I_n z$ series represents every possible negative instance from the most simple ($I_{n'}$) to the most complex ($I_n z$).

Formula (7) indicates that a complete negative series contains all characteristics, acceptable or not acceptable, necessary or unnecessary, except that no instance may contain all of the necessary factors; i.e., F_{hc} ..... (8)

F_c was limited thus: $0 \leq F_c < F_{hc}$#. Many subjects shown the positive cards only, recognized F_c as being common to all positive examples. F_c is not only common to, but also is necessary to all positive instances. Many subjects accepted an F_c as being the F_{hc}. If these did not recognize the remainder of the concent (F_{hc}-F_c), nothing in the positive

#See footnote on previous page.

series could directly point out the omission. As proof of this argument, we examine formula (3):

$S_p = I_{p'} + I_{p''} + \ldots\ldots\ldots + I_{pr}$.....(3)

Factor F_{hc} is found in every instance;

Factor F_c is present in F_{hc} by definition.

Therefore, the subject can find F_c in every positive instance and continue to think that F_c is the total concept.

Positive examples used to verify generalization hypotheses allow partial concepts to be accepted for the whole. In (2) the observation on the nature of positive instances foreshadowed this conclusion.

Compare the potentialities of the negative instance with the impotency of the positive in the role of rejector of partial conceptual hypotheses.

Formula (5) shows that for any incorrect hypothesis (including every F_c alone or in combination with any other characteristic) a corresponding negative instance can be created. Such a negative instance can make a direct attack on any incomplete or incorrect hypothesis by showing directly that the instance corresponding to the hypothesis is <u>not</u> a positive example.

In other words, <u>a complete negative series would make possible the direct elimination of all incorrect hypotheses</u>.

On the other hand, as was shown in statement (2), <u>a complete positive series can not directly eliminate any</u>

incomplete (F_c) hypothesis.

Furthermore, a subject who has accepted a F_c hypothesis can point to that common factor in every positive instance. The positive series is helpless to assist such an erring subject except by damning his hypothesis with faint praise and hoping that the subject will soon notice the so-far-unrecognized (F_{hc}-F_c) part of the concept. Thus we conclude that as an hypothesis-testing agent, negative instances can be superior to positive examples.

This conclusion supports two corrollaries, (1) that negative instances rather than positive are useful in concept analysis, and (2) that the finer the discrimination necessary, (the more complex the concept), the more necessary are the negative instances.

(c) The Process of Concept Formation:

Tyler defined induction as "generalization based upon observed facts", as opposed to the deduction used in demonstrative geometry. "The thinking process uses the two methods alternately, or even in effect simultaneously, so that it is difficult to say at what point the thinking is deductive and at what point primarily inductive." In the present experiment, inductive reasoning produced working hypotheses; deductive reasoning tested them. Inductive reasoning produced further modified working hypotheses, while deductive reasoning tested these in their turn. Since the role of the positive examples was to suggest and enrich, positive instances make

their greatest contribution to the inductive side of generalizing. Since the role of the negative instances was to test and reject, the negative examples make their greatest contribution to the deductive aspect of generalizing.

Since a completed generalization derived from observations is the result of a whole series of inductions and deductions, the presentation of both positive and negative examples could contribute more efficiently to the formulation of a generalization than could the presentation of instances all of which were positive.

The superiority of a learning situation embracing both positive and negative instances is demonstrated by the results of the present experiment.

In Chapter I it was shown that generalization is commonly identified with concept formation. We conclude, therefore, that a positive-negative, rather than a purely positive, presentation is of the greater assistance in concept formation.

3. Summary and Conclusions

1. The results with every concept employed in the present experiment demonstrated the superiority of the positive-negative learning situation over the purely positive presentation.

2. The harder the concept, the greater was the demonstrated value of the negative examples.

3. Under experimental conditions, the positive-negative presentation was virtually certain to be antecedent to a higher Recognition achievement than was the purely positive presentation.

4. Within experimental limits, the lower the intelligence, the more useful were the negative instances.

5. Within the limits of this experiment, the presence of negative instances proved to be a greater determiner of Recognition achievement than did difference in intelligence.

6. The positive-negative learning situation, rather than the purely positive, is antecedent to the greater group homogeneity of achievement.

7. In the generalizing process, the role of the positive instances is two-fold (a) to suggest hypotheses, and (b) to orient the concept in the whole field where it is to be found.

8. In the generalizing process, the role of the negative instance is to assist in the rejection of incomplete and incorrect hypotheses by being direct evidence that any such generalization is <u>not</u> the required concept.

9. Since the generalizing process is both inductive and deductive in nature, the roles of positive and negative instances are complementary, not overlapping.

10. Any experiment, such as the present one, is based on the assumption that a concept is the highest common factor of characteristics to be found in the positive examples exhibited and that such an abstraction is intangible.

CHAPTER VI.

SUMMARY AND CONCLUSIONS

1. The Present Experiment

In this experiment, the general problem attacked was concept formation. The particular aspect studied by this investigation was the part played by positive and negative examples in the process of generalization.

A survey of the experimental field reveals a great variety in opinion as to what constitutes a concept. The author's conclusion was that it is quite proper to use the term concept to designate a wide variety of psychological phenomena, but that he would limit this investigation within the confines of Smoke's experimental definition, that "By 'concept formation', 'generalization', and 'concept learning', we refer to the process whereby an organism develops a symbolic response (usually, but not necessarily linguistic) which is made to the members of a class of stimuli patterns but not to other stimuli".

Accordingly, the type of materials and method employed in this experiment was influenced by the tacit assumptions that concepts are intangibles, and that Verbalization, Recognition and Reproduction, as defined in this study, are methods of communicating the nature of the concept.

For the present experiment, fifty twelve-year-old boys, selected so that intelligence, grade level, native

language, and socio-economic status were controlled, were tested individually.

The apparatus was composed of eight sets of concept cards; each set made up of eight positive examples of the concept, four negative instances and sixteen test examples, together with individual record books in which were recorded each subject's responses.

In the positive presentation, the eight teaching examples were all positive instances, while in the positive-negative presentation alternate exposures were negative examples constructed to be systematic variations from their positive counterparts, but also made as nearly similar as possible to their corresponding positive cards.

All instructions and explanations were given to the subject before the experimental concepts were exhibited. A preliminary concept "Dax" was presented and administered to acquaint the subject with the experimental situation. The experimental concepts were not introduced until the subject understood the procedure thoroughly.

The experiment was conducted "without memory" as all the teaching instances for a given concept were left in the subject's sight during all test responses for that learning situation. The teaching cards for each concept were given serially, the subject being asked to make a "try" as soon as he thought he might know what the concept was, as well as every time he changed his mind as to what the concept might

be. Three types of response, Verbalization, Recognition and Reproduction, were recorded for each "try".

After Cards 4 and 8, the subject was required to make an attempt regardless of whether he had achieved the concept or not. At no time was any clue given to the subject as to whether his responses were acceptable or not. Word-for-word Verbalizations were recorded by the experimenter. Recognitions were recorded. The subject drew his Reproductions in appropriate rectangles in the record booklet.

2. Results and Conclusions

This experimental method produced both quantitative and qualitative results, among the most important of which were (1) the quantitative comparison of the achievement by the two methods, and (2) both the qualitative and the quantitative comparisons of the excellence of Verbalization with Reproduction as the criterion of concept formation.

It was found that Verbalization was not a higher "level" of achievement than was Recognition. The results of this experiment corroborated Heidbreder's conclusions, that while certain concepts can be expressed more easily through one criterion, other concepts may be expressed more easily through other criteria. Some graphical concepts, like Fisher's "Zalof", are composed of "elements" and "relations" for which there is no concise expression in a twelve-year-old boy's vocabulary. For such concepts, a rigidly accurate

Verbalization by a young boy would be nearly impossible even when he recognized the concept correctly. Usually, for such a concept, Recognition would precede Verbalization. In such cases, scores for Recognition achievement would be higher than scores for Verbalization achievement. On the other hand, in the present experiment, some graphical concepts like "Mef" (positive presentation) and "Pog" (positive-negative situation) seem to lend themselves more easily to Verbalization than to Recognition. Accordingly, the subject's achievement of the various criteria is more closely related to the nature of the referent and to the breadth of the subject's vocabulary than to any alleged difference in the inherent difficulties of the "levels" themselves.

Verbalization and Recognition often give conflicting evidence as to the nature of the subject's concept. Verbalization is highly subjective, while Recognition is objective. Recognition is the more sensitive of the two measures. This study corroborates Graham's findings that Verbalization lacks the ability to produce part-scores for analysing concept components. Recognition, in contrast, can be used to measure concept formation at any stage and yields numerical data convenient for statistical treatment. Accordingly, Recognition is the criterion used for computing the quantitative results of this experiment.

As a criterion of concept formation, neither the

minimum number of words required to define a concept, nor time, is of experimental value.

The difficulty of concepts varies with complexity and abstractness.

Many of Smoke's conclusions in his article on negative instances are particularly clear sighted and penetrating. He did, however, miss the great role played by the negative instances. This oversight was due to an experimental error of introducing, into an assumed purely positive learning situation, the important negative instance, "You have some of it, but not all of it".

The introduction of negative instances appears to combat both rigidity; i.e., mental inertia, and the acceptance of incomplete hypotheses. With every experimental concept the positive-negative presentation was antecedent to greater Recognition achievement than was the positive presentation. The experimental variable was the introduction of negative instances. The group which saw both positive and negative examples scored significantly higher than the group which saw the positive only.

While the role of the positive instance is to suggest hypotheses and orient the concept in the whole field where the concept is to be found, the role of the negative is to reject directly all incorrect hypotheses. Since the generalizing process requires both induction and deduction, the roles of the positive and negative are complementary,

both are necessary, neither is sufficient.

3. Educational Implications

Much stress in education has been placed upon the advisability of teaching from positive instances. Teachers should realize that direct evidence necessary for generalization is supplied by negative instances also. To illustrate this, the following problem may be helpful:-

Is the combination of letters "RSTT" a member of the same "class" as the following combination of letters, "LMNL", "FGHF", "BCDB"?

The answer is that there is insufficient evidence from which to give an answer. Evidently the three examples given are not the complete series of positive instances, because the question is asked as to whether or not "RSTT" is an additional positive instance belonging to the "class". In solving the problem, the following line of reasoning might be used. We note that the given examples have the following characteristics in common:-

(1) They are all capital letters;

(2) There are four letters in each group;

(3) In each group one letter occurs twice, but this repeated letter varies from group to group;

(4) The repeated letter appears at each end of the group;

(5) The first two letters are in alphabetical order;

(6) The middle two letters are in alphabetical order;

(7) The first three letters are in alphabetical order.

The example "RSTT" contains characteristics (1), (2), (3), (5), (6) and (7). Characteristic (4) is the only one listed which "RSTT" does not have in common with the examples given. If this characteristic is not only acceptable in, but also necessary to, any member of the "class", then "RSTT" is not a member. If, on the other hand, characteristic (4) is acceptable but not necessary, then "RSTT" is a member of the "class". If it had been shown that (a) "LMNL", "FGHF", and "BCDB" are members of the "class" and (b) "EFGG" is not, the problem could have been solved. Since "EFGG" and "RSTT" are similarly constructed, having all listed characteristics but (4), "RSTT" is not a member of the "class". This problem demonstrates that both negative and positive instances give direct evidence necessary in the process of generalizing.

One application of the use of both positive and negative examples in the classroom may be found in teaching the solution of any linear equation in one unknown. By the use of positive examples demonstrating all the necessary thought processes in solving such a problem as -

if $7Y = 21$

then $Y = 3$,

teachers attempt to lead pupils to the following generalization: to solve for the unknown, get rid of all other numbers or letters associated with it by undoing the operations which associated them with the unknown; that is, by applying

the processes which are the inverse (opposite) of those which bind these letters or numbers to the unknown.

Even after a thorough presentation of all the positive examples necessary to demonstrate all of the processes involved in such a problem, pupils make mistakes. A pupil might make the error of thinking that since the "7" on the left side of the problem above did not appear in the last line, it must have been "taken away". Accordingly, it should be "taken away" from the right side also. In working out another problem, such a pupil might make the following type of mistake:-

If $8x = 32$,

then $x = 24$

("taking away" 8 from both sides).

In such a situation as this, it would be advisable to inform the pupil that his "solution" was incorrect, i.e.,

If $8x = 32$

then $x \neq 4$.

The example just shown is a negative instance. When teachers mark a pupil's solution incorrect, and then ask the pupil to correct his work, the teacher is pointing out a negative instance to the learner and then asking him to produce a positive example.

Re-teaching after testing usually combines the presentation of negative instances with the presentation of positive instances. Though many teachers have given lip

service to the elimination of negative teaching and to the adoption of purely positive teaching, few of them have ceased entirely from pointing out to the younger generation the error of its ways. Accordingly, our teachers have not discarded all negative instances, even when pledging loyalty to the cause of the purely positive.

One great aim of education is the establishment of the attitude of suspended judgment. The positive-negative presentation has a great contribution to make to learning by supplying an impetus to caution. The use of the negative instance should be exploited to break up the rigidity of mental set. If we can generalize from experimental procedure to classroom teaching, the value of the negative instance to the pupil of low I.Q. demands that the teacher give attention to this factor.

The greater efficiency of the positive-negative presentation, over the purely positive, in assisting toward geometrical generalizations, warrants a very definite, persistent, and searching scrutiny of the educational potentialities of this method.

A negative example should be introduced at this point in the discussion. As defined by this experiment, rote learning, such as spelling achievement and mastery of foreign language vocabulary, does <u>not</u> come into the classification of concept formation. Accordingly, this thesis and its conclusions have not any necessary bearing on the pedagogy of such subjects.

4. Suggestions for Further Research

Before any great body of research into the process of concept formation can be undertaken, certain preliminary techniques must be perfected. The two most important among these are the development (a) of group testing techniques; and (b) of concepts for use with a wide range of both chronological age and intellectual ability.

The attempt to develop a group technique could be directed along two lines; namely (1) the use of film slides, and (2) the use of large sheets of cardboard on each of which are drawn individual instances.

In designing procedure and materials to hold the attention of the subject, film slides are likely to prove useful for older pupils, while the large cards may be more suitable for the less mature children. For the young subjects, a lighted, rather than a darkened, room minimizes errors in recording answers and makes supervision more easy. Each change of card at the front of the room can be employed to recapture wandering attention. Older subjects, requiring less alert supervision and having a longer attention span, could have the materials administered by film slide.

The same type of record book could be used with both types of presentation. To make it interesting, each page might be devoted to the Recognition responses for one concept only; each page should be of a distinctive colour; and each should contain generously spaced cells designed for

as many response attempts as desired.

(a) "With Memory"

The present experiment excluded the memory factor and found that the introduction of negative instances increased Recognition achievement. If memory had been introduced by removing the teaching instances before the administration of the test series, would the negative instances have proved as useful? "With memory", perhaps the negative instances would have caused confusion. Both the lecture method (auditory) and moving pictures (visual), as used in teaching, are a "with memory" type of presentation. The usefulness of negative examples in a "with memory" learning situation should be determined before making generalizations from the laboratory results to classroom procedures.

(b) Sex Differences

Is there any sex difference in the type of learning studied in this experiment and as suggested in (a) above?

(c) Readiness to Generalize

In the present experiment, it was found that negative instances gave an impetus to caution. Test responses were required, however, at Card No. 4 and Card No. 8. If subjects had been allowed to wait as long as they pleased before they volunteered a response, the Recognition achievement curves might have been of a much different shape.

What is the influence of positive and negative instances on readiness to generalize? An experiment to answer this question would include no compulsory trials like those of this experiment and might incorporate several arrangements of instances other than the alternate positive and negative arrangement.

(d) Incorrect Hypotheses and Achievement

How does the number of previous incorrect trials affect the speed of learning the concept? The answer to this question has a direct bearing upon the amount of teaching that pupils should receive before they are required to try to solve problems.

(e) Order and Frequency of Positive and Negative Instances

What are the principles linking optimal achievement with order and frequency of positive and negative instances? To determine the best order in which to present the positive and the negative examples, a number of learning situations should be prepared in each of which the orders of the instances can be arranged in a number of ways. Concepts such as those used in the present experiment could be employed.

(f) Maturity and Negative Instances

This experiment has shown that in formulating concepts under experimental conditions, twelve-year-old boys receive assistance from the introduction of negative instances, the amount of assistance varying inversely with intelligence.

The question arises as to the usefulness of negative instances at varying levels of maturity. Perhaps an immature mind needs negative instances more than a mature mind does. On the other hand, perhaps the immature are confused by negative cases.

To attack this problem, two groups of 180 subjects each might be employed. One group might be made up of Grade One seven-year-olds, while the other might be made up of Grade Eleven seventeen-year-olds. Each group should be formed so that it can be paired for intelligence and then split vertically to form two comparable experimental groups with mean I.Q. of 100 and range from 70 to 130. The subjects should be such that they can be divided horizontally into three groups of sixty subjects, paired for intelligence, the three groups to have I.Q. ranges of 70 to 90, 90 to 110, and 110 to 130, respectively. Comparisons of the results of these groups should provide reliable and enlightening information.

BIBLIOGRAPHY

1. Ach, N., Uber die Begriffsbildung, Bamberg: Buchner, 1921, viii and 343.

2. Alpert, A., "The Solving of Problem Situations by Preschool Children: An Analysis". T.C. Contr. to Educ., 1928, No. 323, 69.

3. Arons, L., "Serial Learning and Generalizing Abstractions", Am. J. of Psychol., XLV (July 1933), 417-432.

4. Barr, A.S., "Educational Research and the Research Worker". Journal of Educational Research, Sept. 1937.

5. Bedell, R.C. "The Relationship Between the Ability to Recall and the Ability to Infer in Specific Learning Situations", Ph. D. Thesis, University of Missouri, 1934.

6. Berger, D., Unterschiede der Abstraktion sfahigheit nach. Alter Geschlecht und Milieu der Schulkinder. (Age, Sex, and Environmental Differences Among School Children in the Ability of Form Abstract Judgments). Z Kinderforsch, 1937: 46, 250-282.

7. Berkenblit, Z.M., "The Genesis of Concepts". Trud Inst. Mozga Bekht. 1939, 10, 21-25.

8. Brown, W. and Whittell, F., "Yerkes' Multiple Choice Method with Human Adults", J. Psychol., 1932, 44, 409-453.

9. Brownwell, W.A., "Quantitative Research on Teaching and Learning". School and Soc., 1939; 50, 847-856.

10. Bulbrook, M., "An Experimental Inquiry into the Existence and Nature of 'Insight'". Amer. J. Psychol., 1932; 44, 409-453.

11. Bunch, M., "The Amoutn of Transfer in Rational Learning as a Function of Time". J. Comp. Psychol., 1936; 22, 325-337.

12. Chant, S.N.F., "An Objective Experiment on Reasoning". Amer. J. Psychol., 1933; 45, 282-291.

13. Claparede, E., "La Genese de l'Hypotheses". Arch. de Psychol., 1933; 24, Mos. 93-94, 155.

14. Claparede, E., "La Decouverte de l'Hypothese", J. de Psychol., 1932; 29, 648-656.

15. Crudden, C.H., "Form Abstraction by Children", Psychol. Bull. 1938; 45, 257-266.

16. Dewey, John, Studies in Logical Theory, Chicago: University of Chicago Press, 1903, 171.

17. Drever, J.E., "The Pre-Insight Period in Learning". Brit. J. Psychol., 1934, 25, 197-203.

18. Duncker, K., "Zur Psychologic des Produktiven" Denkins, Berlin Julius Springer, 1935, vii & 135.

19. English, H.B., "An Experimental Study of Certain Initial Phases of Abstraction". American Journal of Psychol., 1922,33,305-350.

20. Eggers, W.J., "An Experimental Study of Children's Reasoning by The Multiple Choice Method": M.A. Thesis, University of California, May 1933.

21. Ewert, P.H., and Lambert, J.F., Part II "The Effect of Verbal Instructions upon the Formation of a Concept". J. Genet. Psychol., 1932, 6, 400-413.

22. Field, Paul E., "Studies in Concept Formation" 1. The Development of the Concept of Triangularity by the White Rat. Comp. Psychol. Monog. 1932, 9 No. 2.

23. Fisher, Sara C., "The Process of Generalizing Abstraction; and its Product, the General Concept". Psychol. Monog. 1920, 28, No. 2.

24. Fowler, H.L., "The Development of Concepts": "An Investigation into the Methods of Teaching". Brit. J. Educ. Psychol., 1931, 1, 13-40.

25. Garrett, Henry E., Statistics in Psychology and Education Longman, Green & Co., Toronto.

26. Gellerman, L.W., "Form Discrimination in Chipanzus and Two-year-old Children". J. Genet. Psychol., 1933, 42, 3-27.

27. Gengerelli, J.A. "Studies in Abstraftion of White Rat". Ped. Seminar and J. Genet. Psychol., 1930, 38,171-202.

28. Gengerelli, J.A., "Mutual Interference in the Evolution of Concepts". Amer. J. Psychol., 1927, 38, 639-646.

29. Gibson, E.J., and McGarvey, H.R. "Experimental Studies of Thought and Reasoning". Psychol. Bull. 1937, 34, 327-530.

30. Graham, J.L., "Learning to Generalize". Psychol. Monog. 1938, 50, No. 5, 84-115.

31. Gray, J.S. "A Behavioristics Interpretation of Concept Formation". Psychol. Rev., 1931, 38, 72.

32. Hamilton, G.U. "A Study of Trial and Error Reactions in Mamals". Journal of Animal Behavior, I (Jan.-Feb., 1911), 33-66.

33. Hanfmann, E., "Concept Formation Test in Schuzophrenia". Psychol. Bull., 1936, 33, 296, Abstract.

34. Hanfmann, E., "Personal Patterns in the Process of Concept Formation". Psychol. Bull., 1940, 37, 515 Abstract.

35. Hanfmann, E., and Kasanin, J. "A Method for the Study of Concept Formation", J. Psychol., 1937, 3, 521-540.

36. Haught, B.F., "The Interrelation of Some Higher Learning Processes", Psychol. Monog. Whole No. 139, 1921, 71.

37. Hazlitt, V., "Children's Thinking", Brit. J. Psychol., 1929-30, 20, 354-361.

38. Heidbreder, E.F., "Language and Concepts". Psychol. Bull. 1936, 33, 724.

39. Heidbreder, E.F., "A Study of the Evolution of Concepts". Psychol. Bull., 1934, 31, 673.

40. Hull, C.L., "Quantitative Aspects of the Evolution of Concepts". Psychol. Bull., 1934, 31, 673.

41. Hull, C.L., Psychol. Monog., 1920, 28, #123.

42. James W., Principles of Psychology, Vol. I, H. Holt, N.Y., 1905.

43. Kinnamon, A.J., "Mental Life of Two Macacus Rhesus Monkeys in Captivity". Am. Journal of Psychol. XIII (Jan. & April, 1902), 98-148, 173-218.

44. Kuo, Z.Y., "A Behavioristic Experiment on Inductive Inference". J. Exper. Psychol., 1923, 6, 247-293.

45. Long, L., "Concept Formation in Children". Psychol. Bull., 1939, 36, 550-551 Abstract.

46. Maier, N.R.F., "An Aspect of Human Reasoning". Brit. J. Psychol. Bull., 1936, 33, 724.

47. Matheson, E., "A Study of Problem Solving in Pre-School Children". Child Develop. 1931, 2, 242-262.

48. McGeoch, J.A. and Oberachelp, V.J., "The Influence of Length of Problem and Transfer upon Rational Learning and its Retention". General Psychol., 1930, 4, 154-170.

49. Miller, I.E., The Psychology of Thinking, New York, MacMillan 1917, 192.

50. Moore, T.V., "The Process of Abstraftion". Univer. of Calif., Publ. in Psychol. 1910, 1, No. 2.

51. Mott, S.M., "The Growth of an Abstract Concept". Child Develpm., 1939, 10, 21-25.

52. Munn, N.L., and Stiening, B.R., "The Relative Efficiency of Form and Background in a Child's Discrimination of Visual Patterns". J. Garret Psychol. 1931, 39, 73-90.

53. Ovsepjan, G.T., "The Evolution of Observation in Children". Sci. Mim. Pedog. Inst. Herzen, 1939, 18, 21-59.

54. Patton, E.K., "The Problem of Insightful Behavior", Psychol. Monog., 1933, 44, No. 197, 198-124.

55. Peterson, J.C., "The Higher Mental Processes in Learning" Psychol. Monog., 1920, 28 No. 7.

56. Peterson, J.C., "Experiments in Rational Learning", Psychological Review XXV (Nov. 1918), 443-467.

57. Pratt, C.C., "Experimental Studies of Thought and Reasoning". Psychol. Bull., 1928, 25, 550-561.

58. Radlinslsa - Ostaszewska, J., (On the Nature of so-called Clear but Indistinct Concepts). Kwart. Psychol. 1935, 7, 393-470.

59. Rees, H., and Israel, H., "An Investigation of the Establishment and Operation of Mental Sets". Psychol. Monog., 1935, 46, No. 210, 1-26.

60. Roberts, K.E., "Ability of Preschool Children to Solve Problems in which a Simple Principle of Relationship is Kept Constant". Pedagogical Seminary, XL (March, 1932, pp. 118-135).

61. Roberts, K.E., "Learning in Pre-School and Orphanage Children". Univ. Iowa Stud., Stud. in Child Welfare, 1933, 7, No. 3, 94.

62. Roslow, S., "A Statistical Analysis of Rational Learning Problems". J. Genet. Psychol. 1936, 48, 441-467.

63. Sells, S.B., "The Atmosphere Effect: An Experimental Study of Reasoning". Arch of Psychol. 1936, 29, No. 200, 72.

64. Sibano, M., "Concept Formation as Productive Thinking". J. Psychol., 1938, 13, 389-400.

65. Siipola, E.M., "A Group Study of Some Effects of Preparatory Set". Psychol. Monog., 1935, 46, No. 210, 27-38.

66. Smoke, K.L., "Negative Instances in Concept Learning". J. Exper. Psychol., 1933, 16, 583-588.

67. Smoke, K.L., "The Experimental Approach to Concept Learning". J. Exper. Psychol., 1935, 42, 274-279.

68. Smoke, K.L., "An Objective Study of Concept Formation". Psychological Monographs XLII Whole No. 191, 1932, 8.

69. Stevanovic, B.P., "An Experimental Study of the Mental Processes Involved in Judgment". Brit. Journal of Psychol. (Monog. Supp.), 1927, XII, 138.

70. Sullivan, E.B., "Attutude in Relation to Learning". Psychol. Monog. 1927, 36, No. 169, 149.

71. Tyler, F.T., "Generalizing Ability of Junior High School Pupils: An Experimental Study of Rule Inductions": Unpublished Ph. D. Thesis, Univ. of Calif.

72. Vedenov, A.V., "On the Question of Experimental Investigation of the Structure of Concepts", Bekht. Inst. Brain Res. 20th Anniv. 1938, 65-66.

73. Warren, H.C. and Carmichael, L., Elements of Human Psychology, Houghton Mifflin Company, New York, 1930. (The Riverside Press, Cambridge).

74. Welch, L., "A Preliminary Study of the Interaction of Confliction Concepts of Children Between the Ages of 3 and 5 Years". Psychol. Rec. 1938, 2, 439-459.

75. Welch, L., "The Span of Generalization Below the 2-year Level", J. Genet. Psychol., 1939, 55, 269-297.

76. Welch, L., "The Development of Size Discrimination Between the Ages of 12 and 49 Months", J. Genet. Psychol. 1939, 55, 243-268.

77. Welch, L., "A Preliminary Investigation of Some Aspects of the Hierarchical Development of Concepts", J. Genet. Psychol. 1940, 22, 359-378.

78. Welch, L., "The Genetic Development of the Associational Structures of Abstract Thinking", J. Genet. Psychol. 1940, 56, 175-206.

79. Welch, L. and Long, L., "The Higher Structural Phases of Concept Formation of Children", J. Psychol., 1940, 9, 59-95.

80. Welch, L., and Long, L., "A Further Investigation of the Higher Structural Phases of Concept Formation", J. Psychol. 1940, 10, 211-220.

81. Wilhelm, W., "Bertrage Zur Psychologie des Schliessens", Arch f.d. ges. Psychol., 1933, 89, 371-428.

82. Wilkins, M.C., "The Effect of Changed Material on Ability to do Formal Syllogistic Reasoning", Arch of Psychol., 1928, 16, No. 102, 83.

83. Wolff, F.F., Concept, Precept and Reality. Phil. Rev., New York. 1939, 48, 398-414.

84. Woodworth, Experimental Psychology, H. Holt & Co., New York, 1938.

85. Yerkes, R.M., "A New Method of Studying the Ideal Behavior of Mentally Defective and Deranged as Compared with Normal Individuals", Journal Comp. Psychol., 1921, 1, 369-394.

APPENDIX.

PROCEDURE (in detail)

On coming for his one experimental period, the subject was given the following verbal instruction:

To establish rapport:

"Would you like to come over here? I have asked your teacher to let you come to help me. You would like to do that, wouldn't you? Here are some interesting puzzles for you to do. They are different from any you have ever done before. I think you will like them. You like working puzzles, don't you? I will show you how to do the first one."

Subject's introduction to the test situation:

"I have here the picture of a Dax. Now you do not know what a Dax is, nor have you ever heard of one. When I tell you to, you will turn these cards over", (pointing to the teaching pack which has been placed face downward in readiness on the table). "Some are Daxes and some are not. The puzzle is to figure out what a Dax is. Turn over as few cards as possible. The fewer cards you have to turn up before you figure out what a Dax is, the higher your score will be. However, the important thing is to get it right. If you do not get it right the first time you can change your mind and try again."

(The subject is seated to the left of the experimenter.

The cards are placed face downward and in order from #1 to 8, with #1 on top, and within the subject's reach between him and the experimenter).

POSITIVE-NEGATIVE PRESENTATION

"The first card you turn over is a Dax. Now turn it over and place it here."

The cards were then turned over one at a time so that all 8 examples were exposed to the subject's view. As the examples were turned over they were placed in front of the subject in two columns of four cards each with odd-numbered (positive) cards in the left-hand column, and even-numbered (negative) cards in the right, cards #1 and #2 farthest from the subject, and #7 and #8 nearest to him.

"Now turn over the next card (#2) and place it here. This is <u>not</u> a Dax. As you turn over the rest of the cards you will place the next one (#3, a positive instance) here. It will be a picture of a Dax. The next card you turn up you will place here, because it is <u>not</u> a picture of a Dax..." etc.

(Using this exact wording, the investigator leads the subject to understand that all examples in the left-hand column WERE Daxes, and that EVERY instance to the right as <u>not</u> a Dax.)

(Pointing to where the positive instances go:) "Will the cards you place here be Daxes, or not? (The experimenter

got the answer that they WOULD be Daxes.) "The cards that you place here (indicate the place where the other column would be) will they be Daxes, or not?" (The answer was that they would NOT be Daxes).

(1 and 2 were then replaced on the pack and the subject instructed as follows:)

"Try to find out what a Dax is. Do not turn up any more cards than you have to. Tell me as soon as you think you know what a Dax is. Now go ahead and turn up the first card. This is a Dax."

(As soon as the subject thought he knew what a Dax was, the experimenter instructed him as follows:)

"The answer that you are going to give me now may be right or it may be wrong. After you have given me your answer you will be able to look through the rest of the cards. If you are right, when you see these cards you will not want to change your mind, but if you are wrong, probably these cards will make you want to. Don't be afraid to change your mind. The important thing is for you to have found out what a Dax is."

(The responses on the Verbalization level were recorded word-for-word in the booklet, as already described).

(Taking the test pack, arranged in serial order, with #1 on top and #16 on the bottom, the experimenter exposed the cards in succession, instructing the subject to say either, "Dax" or "Not Dax". The author noted errors mentally

or if there were too many he remembered the right responses instead of the wrong, or tipped the cards so that the cards for which incorrect responses were given fell out of line with those correctly named. After the subject had responded to every card in the Recognition test, circles were placed in the booklet around the numbers of the cards incorrectly named. Entry in the "Number Wrong" column was deferred until the subject had finished the test and had left the room, because such recording in his presence would have given him an external clue to the acceptability of his responses.)

(Giving the record booklet to the subject, the experimenter said:)

"Draw one Dax in each of these squares", (pointing to the appropriate frames in the booklet). "If you can, make each picture different from any other pictures you draw, and make your pictures different from any of these examples" (indicating the positive instances exposed on the table).

(After taking a sampling of the subject's ideas by each of the criteria, Verbalization, Recognition, and Reproduction, and regardless of whether the responses were successful or not, the subject was instructed to inspect the remainder of the teaching pack.)

Your answer may be right or it may be wrong, I am not allowed to tell you which. But the rest of these cards will tell you. If the rest of these cards makes you change your mind as to what a Dax is, tell me. Each time you turn up a

card, say either, 'I think the same' or 'I change my mind'. Now turn up these cards."

(If after turning up any card, the subject did not spontaneously say whether he had changed his mind or not, the experimenter asked him the question, "Do you think the same, or do you change your mind?" This forced the subject to make a decision and to reveal a modification of his concept on the card at which he recognized the change.)

(Test attempts were made on any card at any time the subject wished. In addition to these voluntary testings, compulsory samplings were taken on Card 4 and Card 8 even though (1) the subject had already correctly formulated the concept at the three levels, and even though (2) he thought he did not know the concept.)

(If after exposing all 8 cards of the teaching pack, the subject was still unable to make the successful generalization of the "Dax", the writer told the subject what a Dax was and showed him that each positive instance was a Dax, and that each negative instance was not a Dax. When the experimenter was satisfied that the subject understood the experimental situation, he presented to the subject the 8 experimental concepts.)

"Now that you have learned a Dax, you will be shown some more puzzles. This first one has been used just to show you how to go about doing these others. If you do not understand any part of what you are to do, please ask me about it

now."

(After giving any necessary explanations, the 8 experimental concepts were administered similarly.)

POSITIVE PRESENTATION

In the positive presentation exactly the same instructions were given as in the positive-negative, except that all the instances in the teaching pack were positive, and a corresponding change of wording was necessitated.

As has been stated previously, careful precaution was taken that the experimenter gave the subject no indication as to whether his responses were correct or incorrect. The subject was made to rely upon the eight cards in the teaching pack to inform him whether or not his concept had been acceptably formed. To give uncontrolled external evidence on this question is to obscure one of the chief contributions of the negative examples.