B. JOAN MACKENZIE FACULTY OF EDUCATION M.A. DEGREE

SHORT TITLE

EFFECTS OF AGE AND TEST FORM ON CATEGORIZING BEHAVIOUR

A STUDY OF THE EFFECTS OF AGE AND TEST FORM ON CERTAIN ASPECTS OF CATEGORIZING BEHAVIOUR

BY

BETTY JOAN MACKENZIE

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of Master of Arts

Faculty of Education,
Macdonald College of McGill University,
Montreal.

August 1967

ABSTRACT

M.A. Education

Betty Joan MacKenzie

A STUDY OF THE EFFECTS OF AGE AND TEST FORM ON CERTAIN ASPECTS OF CATEGORIZING BEHAVIOUR

Two groups of girls, aged 12 and 18, were administered object, verbal and picture forms of Clayton and Jackson's sorting test according to a pre-arranged sequence.

Scoring by defined categories, unsorted objects, total categories and level of abstraction indicated that total categories was an imprecise measure of categorizing behaviour.

Differences in test forms significantly affected the number of categories sorted. Fewest categories were sorted on the object form of the test at both age levels.

Older subjects sorted into fewer defined categories for all test forms, and these categories were of a more abstract nature. Individual consistency as measured by between-test correlations was lower for the older group.

Findings tended to support an interpretation of age differences in categorizing style based on developmental theory and suggested increased attention should be paid to test form differences when assessing categorizing behaviour.

ACKNOWLEDGEMENTS

I wish to express my appreciation to the many people who have made this study possible.

Special thanks are due Professor Reginald Edwards, my advisor, whose stimulating instruction and wise counsel have proved invaluable.

To Mr. W. Grant Taylor, former Principal of Macdonald School; Mr. Murray Ellison, Principal, Dorset School; and Mr. Graham Smith, Principal, Oakridge School, go particular thanks for their enthusiastic cooperation in providing a supply of subjects and in scheduling the testing periods. I wish also to thank Dr. Gerald McKay, Chairman, Department of Educational Psychology, Macdonald College, for his assistance in obtaining the volunteer sample at the college level. And, of course, I want to express my gratitude to those girls who so willingly participated in this study.

Miss Frances Wren, Department of Household Science, Macdonald College, very skillfully fashioned several of the items for the object form of the test and her help, too, is gratefully acknowledged.

It is to my husband, "Gus," without whose sustaining optimism this investigation could never have been completed, that I owe most of all. I wish to specifically thank him for proof-reading parts of the manuscript.

Finally, I would like to thank our children, Douglas, Susan and Kenneth, who have waited so long and so patiently.

Betty Joan MacKenzie,
July 1967.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF TABLES	vi
Chapter	
I. INTRODUCTION	1
II. A REVIEW OF THEORY AND RESEARCH RELATING TO CATEGORIZING STYLE	4
A. Historical Antecedents of Categorizing Style	6
B. Studies of Stylistic Tendencies in Categorizing Behaviour	13
1. Category width	15
a) Measurement of category width	15
b) Interpretation of category width	18
i) Category width as ability to discriminate	19
ii) Category width as impulse expression	20
iii) Category width as a risk-taking strategy	21
c) Ambiguities and inconsistencies in studies of category width	23
2. Level of Response	28
a) Measurement of level of response	28
b) Level of abstraction as a dimension of style	29
C. Summary	31
III. A REVIEW OF THEORY AND RESEARCH RELATING TO EFFECTS OF AGE AND TEST FORM ON CATEGORIZING BEHAVIOUR	33
A. Effects of Age	33
1. Research on cognitive styles in children	34

Table of Contents (cont'd)

Chapter		Page
	2. Developmental theory and age differences in cognitive functioning	36
	a) Developmental theory and category width	36
	b) Developmental theory and level of response	38
	B. Effects of Test Form	39
	1. Studies regarding the effects of test form	40
	a) Test form and category width	41
	b) Test form and response level	41
	2. Developmental theory and test form effects	43
IV.	THE EXPERIMENT	44
	A. Statement of the Problem	44
	B. Rationale Behind the Experimental Design	47
	1. The subjects	47
	a) Sex	47
	b) Age	47
	2. The test and testing techniques	48
	a) The technique of free sorting	48
	b) Choice of test and test forms	50
	c) Test order effects	51
	3. Aspects of categorizing behaviour	52
	a) Number of categories defined	53
	b) Abstraction index	54
	c) Number of objects unsorted	54
	d) Total categories	55
	C. The Procedure	55
	1. Testing procedure	56
	2. Scoring procedure	59
	a) Scoring procedure for abstraction index	59
V.	HYPOTHESES	62

Table of Contents (cont'd)

Chapter		Page
VI.	RESULTS	65
	A. Age and Test Form Effects	65
	1. Number of categories defined	66
	2. Number of unsorted objects	68
	3. Total categories	69
	4. Percentage of abstract responses	71
	B. Effects of Age and Test Form on Individual Consistency	72
	C. Order Effects	73
VII.	DISCUSSION OF RESULTS AND THEIR IMPLICATIONS	75
	A. Relationship Among Total Categories, Defined Categories and Unsorted Objects	75
	B. Age Effects	76
	C. Test Form Effects	77
	D. Individual Consistency	81
	E. Age and Test Form Interaction Effects	82
	F. Order Effects	83
	G. Summary of Results in Terms of Hypotheses	83
VIII.	SUMMARY AND CONCLUSIONS	86
BIBLIOGRAPHY		89
APPENDICES		95
	A. The Test: Directions for Administering and Scoring	95
	B. Raw Data and Tables I - IX	104

LIST OF TABLES

Table		Page
1.	Mean Values of Scores Obtained by Two Age Groups on Three Forms of a Free-Sort Test	66
2.	Analysis of Variance for Number of Categories Defined on Three Forms of a Free-Sort Test by Two Age Groups	67
3.	Significant Mean Differences for Number of Categories Defined on Three Test Forms by Two Age Groups. Duncan Multiple Range Test	67
4.	Analysis of Variance for Number of Unsorted Objects on Three forms of a Free-Sort Test at Two Age Levels	69
5.	Analysis of Variance for Total Number of Categories Formed on Three Forms of a Free-Sort Test by Two Age Groups . . .	69
6.	Significant Mean Differences for Total Categories Formed on Three Forms of a Free-Sort Test. Duncan Multiple Range Test. Groups 2 and 1 Combined	70
7.	Analysis of Variance for Per Cent Abstract Responses on Three Forms of a Free-Sort Test at Two Age Levels (arc sine transformations)	71
8.	Correlations Between Scores on Test Forms for Number of Defined Categories	72
9.	Correlations Between Scores on Test Forms for Percentage of Abstract Responses	73
10.	Summary of Results on Order Effects. Mean Values of Scores Obtained by Two Age Groups on Three Consecutive Free-Sort Tests	74

CHAPTER I

INTRODUCTION

The opinion is widely held that virtually all cognitive activity involves and is dependent upon the process of categorizing. It has been shown to underly concept formation, perception, judgment, memory and problem solving (Bruner, Goodnow & Austin, 1956). In fact, the questions of how categories are formed and used seems basic to all problems concerned with the study of cognitive functioning.

A large body of research suggests that there are consistent and relatively permanent individual differences in ways of experiencing and organizing stimulus events. The inference has been that the individual has a preferential mode of categorizing, that he has, in other words, a categorizing style that is discernible over a wide range of cognitive tasks. These aspects of categorizing style are considered to be reflective of a more general cognitive style and therefore characteristic individual differences in categorizing behaviour may denote broad but significant differences in ways of acquiring and organizing knowledge. This could have great implications for the educative process.

As evidence accumulates to uncover preferential modes of response in an individual and to reveal individual differences, the obvious questions arise: How do these differences develop? How did the individual get "that way"? Little is known of how a characteristic categorizing style develops in a particular individual. Most investigations

have been designed to study the nature of categorizing styles in adults, with limited success. Even less has been confirmed about our knowledge of these styles in children. Yet, educational research is necessarily concerned, eventually, with the cognitive styles of children of all ages.

From this point of view, it becomes important to study categorizing styles at different age levels for in this way the evolution of such styles may be determined. However, before developmental studies can be carried out extensively, the relative effects of tests and procedures for determining and interpreting categorizing behaviour must be evaluated for use at different age levels. It is important, then, to explore the effects of test forms as well as age. An initial investigation has been designed, therefore, to make a modest beginning by studying the effects of three forms of a sorting test on certain defined aspects of categorizing behaviour at two age levels.

Two groups of girls, aged 12 and 18 respectively, are to be administered three adaptations of the Clayton and Jackson sorting test. These tests are the original verbal form, slightly modified to include the name of each item on a separate card, an object form composed of models or replicas of the items, and a picture form composed of photographs of these items. Subjects are to be required to sort the items into categories that seem "most natural, most comfortable and most logical" to them; the categories will be scored in terms of their size (width) and their content. The effects of age and test form on these specified aspects of the categorizing behaviour will be determined by analyses of variance.

It will be observed that of all the processes usually referred to as cognitive, emphasis is to be placed on categorization. In the review of the literature that follows, particular attention will be devoted to one facet of this process, namely categorizing style as defined by individual preferences in category widths and, to a lesser extent, level of response. One further point should, perhaps be acknowledged. Certain psychometric problems arise in the most suitable application of any test. In the process of categorization these factors have often been overlooked. Accordingly, some attention will also be devoted to the use of varying forms of object sorting tests by which these designated aspects of categorizing style may be measured.

CHAPTER II

A REVIEW OF THEORY AND RESEARCH RELATING TO CATEGORIZING STYLE

During the past decade there has been a resurgence of interest in the cognitive approach to the study of thinking and learning. Many factors have contributed to the growth of the movement, among which may be mentioned the reaction to the purely mechanistic emphasis of early behaviourists, the contributions of neo-Freudian psychology which centered on the role of the ego, the influence of the theories of intelligence stressing the complex nature of the interaction between personality and intellectual factors in the cognitive processes and the insights into the ontogenesis of cognition provided by Piaget and Werner. The event which may be said to have signalled the revival was the publication in 1956 of a book, A Study of Thinking, by Bruner, Goodnow & Austin which focussed on the categorization process as the fundamental phenomenon basic to all cognitive functioning.

The characteristics which identify the cognitive school of psychology today are somewhat elusive to define. It has been suggested that the cognitive approach to problems of psychology is not distinguished so much by its content as by its "flavor" (Van de Geer & Jaspars, 1967). Ausubel (Anderson & Ausubel, 1964) would probably agree as much of his book (co-edited with R. Anderson) is devoted to the convergence of the neo-behaviouristic and cognitive positions on the nature of thought processes and phenomenon.

Nonetheless, certain differences between the two theoretical orientations can be detected. Instead of focussing on the study of

habit acquisition through stimulus response connections and the role of mediation as does the neo-behaviourist, the cognitive psychologist directs himself to the discovery of psychological principles which govern the selection, organization and processing of states of conscious experience.

The expressions of this renewed interest in cognitive psychology are many and varied. One such manifestation is the current concern with the study of cognitive styles. Style is seen as an adaptational mechanism that enables the individual to control and organize the stimulus world (Klein, 1958; Witkin, 1964). By definition styles must be relatively permanent, enduring and self-consistent within the individual and must be discernible in a wide range of cognitive behaviour (Rapaport, 1957). While cognitive style is not the only path along which individual difference in cognitive functioning may be explored, it is one which, although still in the speculative stage, promises much for the future.

Most of the investigations in this field have been concerned with allowing individuals to reveal their preferential cognitive behaviour and then testing the consistency of this behaviour in other related areas of cognitive functioning. It has been assumed that there are certain aspects of behaviour which when picked up through tests can be used as "tracer elements" to detect and interpret an individual's characteristic mode of reacting to stimulus events (Witkin, 1964). The ubiquitous and basic categorizing process, usually tapped by a sorting task, has frequently served as the window through which such characteristic individual modes of response could be viewed.

The influences which converged to produce the concept of cognitive style, and more particularly categorizing style, as a means of determining and assessing individual differences in cognitive functioning were, originally at least, from widely differing theoretical orientations. Early attempts to study intellectual functioning centered on studies of intellectual ability viewed as capacity to abstract. It was from these studies that the tests and test procedures used to tap categorizing style in this investigation have evolved. Much of the theory on which the concept of style is based and by which it often is interpreted, however, grew out of the psychoanalytic approach to personality development and the eventual attempts to relate these theories to the intellectual sphere. A brief sketch of these merging trends may help to place this study and the theory from which it takes its direction in perspective.

A. HISTORICAL ANTECEDENTS OF CATEGORIZING STYLE

Sorting procedures have long been used to study higher mental processes and although the specific tests and free sort procedures used in this investigation were developed for the most part over the past decade, their roots can be traced through a long and varied history of research into intellectual functioning.

Many of the early workers engaged in the clinical study of thought processes came under the influence of the emerging Gestalt school and reflected the introspectionist mode of their predecessors as well. This was especially true in Europe where early work by Ach was developed by Vigotsky and extended by Hanfmann and Kasanin (Hanfmann & Kasanin, 1937). The latter used an adaptation of the Vigotsky Block

Test which required subjects to sort 22 blocks with four differing dimensions into four groups. The "best" answer involved ignoring both shape and colour and synthesizing two dimensions, top area and height, to form a single concept on which a sorting could be based. Although the categorizing (conceptualizing) process could be, and sometimes was, studied by this means, the main purpose was to find individual differences by means of qualitative assessments, rather than to formulate laws of cognitive behaviour.

Meanwhile Goldstein, first with Gelb (Goldstein & Gelb, 1918), and then with Scheerer (Goldstein & Scheerer, 1941), developed an object sorting procedure in which the subject when presented with familiar objects and was required to sort them in certain ways. This method of testing ability to classify objects according to specified criteria was useful in exploring the relationship between brain function and quality of thought because it revealed differences in cognitive functioning that were interpreted along the abstract-concrete dimension developed by Goldstein (Goldstein & Scheerer, 1941). By examining the groups formed, it was possible to determine whether the subject had sorted concretely (defined as classification on the basis of immediate sensory impressions), or abstractly (on the basis of common attributes in the face of disparate other properties). This test eventually found use as a diagnostic tool since certain pathological conditions of thought, for example, schizophrenia, were believed to be characterized by an inability to give an abstract sorting performance when required to do so (Rapaport, Gill & Schafer, 1946).

Weigl, a pupil of Goldstein's, also conducted testing sessions with schizophrenic, brain-damaged, and normal subjects (Weigl, 1927).

He used various sorting tests - the Holmgren wools, coloured shapes and everyday objects. In the object test all the items were placed within certain squares on a checkered table cloth. The subject was handed one single object and instructed, "Hand me all the objects which go with this." At a later stage, all the items were placed before him and he was asked to sort them into groups. Although Weigl did not develop this latter technique further, it is of particular interest because of the subsequent use made of it by others interested in the study of individual differences in modes of categorizing (for example, Gardner, 1953; Gardner et al., 1959; Gardner et al., 1960; Gardner & Schoen, 1962).

It can be noted that all the early investigations were characterized by their reliance upon clinical signs, verbal statements, perhaps even introspections. The clinical assessment of individual differences in mental functioning was a qualitative one. Partly, one suspects because the means of analysis defied absolute measurement and partly because these workers believed that intelligence was not truly subject to precise psychometric quantification. This early work gave some information about concept formation, object sorting and categorizing. It contributed an instrument and a technique from which grew the tests and procedures used in the experiment to be described here; it certainly popularized the terms abstract and concrete with reference to thought processes. Unfortunately, however, clear-cut definitions were not provided for those who were subsequently to use these terms, and much confusion was generated in this way.

The demonstrated use of the sorting test to detect pathology of thought led quite naturally to an interest in its use to determine

qualitative differences in normal brain functioning (Freeman, 1951). However, these tests did not appear to be sensitive enough to differentiate performances within the normal range. Besides, the psychometric approach to individual differences had flourished and so there were other more objective, less time consuming means of determining the level of intellectual functioning and intellectual capacity. Moreover, evidence had been accumulating which cast doubt on the usefulness of capacity measures as the prime criterion for understanding cognitive differences in individuals (Thurstone, 1948; Weschler, 1950). New and compelling views on the complexity of the intellect (Weschler, 1950; Guilford, 1959) had emphasized the need to expand the restricted view of intellectual functioning, defined as the capacity to abstract, and had heightened interest in the emphasis from what the individual can do to what he does do. Thus, the sorting test as an instrument for the study of thought may have fallen into disuse had it not been for the impact of Freudian, and later neo-Freudian, psychology which added a new dimension to the investigation and interpretation of object sorting behaviour.

Throughout the whole of this period, the Freudians had also been concerned with the development of theories about cognitive functioning. Many of these bore the direct imprint of Freud's own genius; some were accepted, and others rejected. Among those accepted were such viewpoints as the tripartite division of id, ego and superego and the concept of the defense mechanisms of the ego (Freud, 1925).

One of the most successful of the early attempts to reconcile the Freudian position on intellectual development with other psychometrically based accounts of intellectual performance was that of Rapaport,

Gill and Schafer (1946). Rapaport's work was an attempt to consider the Freudian approach to disorders of personality and thought and relate them to what was known of conventional tests of thought processes. This orientation showed clearly in his work with Gill and Schafer and led at a later date to a more ordered presentation, The Organization and Pathology of Thought (Rapaport, 1951).

From the Freudian primary processes, the concept of delay and its relation to repression, and to other defense mechanisms at the disposal of the ego, he attempted to make the transition from the drive-organization to the conceptual-organization of what came to be known as cognitive structures. Hence he proceeded to consider the role of thought organization in reality testing and drive mastery. Within this framework, concept formation appeared as a basic, fundamental aspect of thought and thus was the means by which thought processes in general could be studied. It is interesting to note that Bruner, taking a little finer-grained view, perhaps, further narrowed this basic common attribute to a process of categorization (Bruner, Goodnow & Austin, 1956).

Much of the early work relating these psychoanalytic theories to the intellectual sphere was quite speculative and was chiefly concerned with demonstrating how the rational processes of man were controlled by his environment, his drives and his defenses (Fenichel, 1945). However, with the publication of Anna Freud's book on the ego and the mechanisms of its defense (A. Freud, 1937), a shift in emphasis occurred. Increasingly, the view was accepted that man is not merely a victim of his needs and his environment, but is able to shape and control both or either. This centered attention on the individual as

a motivated and highly idiosyncratic being and heightened interest in his influence on the thought processes. This theory was further developed by Klein (1954, 1958) who defined certain aspects of the cognitive process as "adaptational control mechanisms" or cognitive styles of the ego that mediate between internal need states of the individual and the external world. These controls or styles were viewed as enduring and self-consistent and represented increasing control over action.

It should be pointed out that while the foregoing was the theoretical interpretation of cognitive styles espoused by the many researchers whose work is cited in this study, there is another point of view concerning the role of conceptual structures seen as style. Some investigators (Rogers, 1959, 1963) emphasize the self concept and see cognitive structures as major motivating influences.

The work of Gardner, first under Scheerer at the University of Kansas (Gardner, 1953) where the stress was on organizational modes, and later at the Menninger Clinic where it became tinged with the relationship to defense mechanisms and drive conditions embodied many of the theories and much of the rationale of the first group. Gardner noted that the early work of Rapaport (Rapaport, Gill & Schafer, 1946) had indicated that there were consistent individual differences in the number of objects that subjects were willing to group together in an object sorting test. Rapaport had referred to this phenomenon as "realm span." These observations were extended by Gardner into ideas of "equivalence range" which he studied by means of a test using free sort procedures adapted from Goldstein and Scheerer's object sorting test. The procedure was based on one originally developed by Weigl (Weigl, 1927). This instrument, known as a free sort test will be

discussed more fully later but it should be noted at this point that it was designed to allow the subject to sort the objects under free, relaxed conditions so that his preferential, characteristic way of grouping would have full opportunity to operate. The subject was simply asked to put into groups the items which seemed to him to belong together. Subsequently, it was shown that certain aspects of these preferential modes of categorizing were enduring and self-consistent (for example, Gardner et al., 1959). It thus appeared that the individual could be said to have a "style" of categorizing.

A number of other efforts have been concentrated in this area and one can see several paths along which research has been taken. Not all investigations are concerned with individualistic tendencies in categorizing. Witkin and his associates (Witkin et al., 1954) started out with a set of observations on a dimension they have named field dependence, and developed two tests for use in its determination, the Rod and Frame test and the Embedded Figures Test. More recently they have tried to show a constancy of performance in a variety of situations (Witkin et al., 1962). Hence they too have a concept of "style," by virtue of which the subject constructs his environment, reacts to it and so controls it.

In summary, it has been shown that theories from ego-psychology gave impetus to the search for examples and explanations of autism in cognitive functioning. Emerging views on the nature of the intellect directed that search towards manifestations of the personality as a whole rather than to indices of intellectual ability alone. And, when the Freudian position that a function of a person demands a structure was accepted, it became possible to examine the cognitive

structures (of which style is one) which were presumed to exist. It was consistent with the cognitive approach that attention be centered on the study of styles in organizing and acquiring knowledge. Thus, much research focussed on the defining and interpreting of individual styles in the categorizing process. Attention may now be directed to certain of these studies and their theoretical interpretations.

B. STUDIES OF STYLISTIC TENDENCIES IN CATEGORIZING BEHAVIOUR

Since categorizing is believed to be a basic attribute of thought, it is not surprising that much research has been undertaken to explore and interpret various aspects of categorizing behaviour. One avenue of approach has been to investigate individual differences in categorization in terms of consistent and enduring personal styles. Perhaps the most comprehensive studies have been carried out by Gardner and associates (Gardner, 1953; Gardner et al., 1959; Gardner et al., 1960; Gardner and Long, 1961; Gardner and Schoen, 1962). They have attempted to identify the adaptational cognitive controls (Klein, 1958) operative in various forms of cognitive functioning and to determine the complex interrelationships between them. The portion of their work which centered on the categorizing process has employed sorting procedures as the means whereby individual differences could best be revealed. The interpretations of their findings have been based on ego-psychology as formulated by Klein and Rapaport (Klein, 1949; 1954; 1958; Rapaport, 1951; 1957).

Another group, of whom Bruner appears to be the spiritual leader, has explored characteristic individual tendencies in categorizing from yet another vantage point. Within the broad framework of information

theory, they have viewed categorizing as a decision process and they have interpreted the various aspects of categorizing behaviour in terms of strategies or risk-taking tendencies. They seldom use the term "style" as such, but the concept of consistent and enduring individualistic modes of functioning is implicit in their analyses and their interpretations.

A third group represented by Messick and co-workers appears, on the whole, to have taken a neutral stance in the controversy concerning the theoretical interpretations of the categorizing process. In general, their approach may be said to be rather closely allied to the neo-behaviouristic school, and they have conducted studies chiefly designed to refine techniques for measuring the various dimensions of categorizing behaviour (Messick & Kogan, 1965; Messick & Fritsky, 1963; Messick & Kogan, 1963). They have also sought further understanding of these dimensions by investigating behavioural and personality correlates. However, their attempts to relate breadth of categorizing to intellectual variables such as capacity (Wallach & Kogan, 1965; Messick & Kogan, 1965), quantitative ability (Messick & Kogan, 1965), and memory (Messick & Damarin, 1964) have generally yielded inconclusive results. Their proposed relationships between category width and such personality variables as tolerance for ambiguity (Clayton & Jackson, 1961), anxiety (Kogan & Wallach, 1964), authoritarianism (Clayton & Jackson, 1961) and creativity (Wallach & Kogan, 1965) have not been unambiguous, nor have their results always found support elsewhere in the literature. In general, evidence has been clouded by the use of diverse measurement techniques, a weakness this last group has recognized and has been attempting to rectify.

Although each of these groups may be identified by its unique approach to the study of individual differences in cognition, they have at least one feature in common. Heretofore, all have placed major emphasis on a single aspect of the product of the categorization process - the size or width of the category formed. More recently, some investigators have begun to explore individual differences in the level of response as indicated by the labels used by an individual to define the contents of his categories. Thus, it is pertinent to review some of the work that has been done on these two stylistic dimensions of categorizing behaviour. Because of its importance as denoted by the volume of research, the larger share of attention will be directed to the category width measure.

1. Category Width

a. Measurement of category width

It has frequently been observed that individuals differ in the range of values that they are willing to call "the same" (Bruner, Postman & Rodrigues, 1951; Bruner & Rodrigues, 1953; Gardner, 1953; Fillenbaum, 1959; Rapaport, Gill & Schafer, 1946). For example, one person may happily call a diverse range of colors "blue". Another would prefer to differentiate these same colors as "navy", "turquoise" and "blue".

Attempts to measure this observed phenomenon have centered on sorting tests administered under free conditions and on specially designed tests such as the Pettigrew Scale of Category Width (Pettigrew, 1958). The choice of measuring instrument appears to be associated with the theoretical view point of the tester. Generally speaking, those who espouse a psychoanalytic interpretation of style have employed the

sorting test and those whose orientation comes from information theory have used a scale-type test such as Pettigrew's. Thus, as one might predict, the observed phenomenon has been variously interpreted and variously named: equivalence range (Gardner, 1953), category width (Bruner & Tajfel, 1961), conceptual differentiation (Gardner & Schoen, 1962). The term category width is used in this study because it appears to adequately describe the phenomenon regardless of the theoretical interpretation placed upon it.

Category width on free sorting tests is determined by the number of categories used when objects are sorted into "comfortable groupings". This instrument requires the individual to generate his own conceptual realms and to place and order over a complex of such realms. As the number of groupings increases, the number of objects placed within each obviously declines with broad categorizers producing few categories and narrow categorizers relatively more. Thus category width, an index of preferential categorizing breadth, is inversely proportional to the number of categories formed.

In contrast to the sorting test, each question in the Pettigrew Scale assesses an individual's preferential inclusion span in one conceptual area. A typical item from the Scale is as follows:

> The average muzzle to tail length of a sample of 1,000 German Shepherd dogs is 40.3 inches. What do you think:
>
> a) is the length of the longest Shepherd dog in the sample?
> 1. 60.4 inches 3. 44.7 inches
> 2. 47.8 inches 4. 54.2 inches
>
> b) is the length of the shortest Shepherd dog in the sample?
> 1. 34.6 inches 3. 19.7 inches
> 2. 28.4 inches 4. 36.9 inches

A thoughtful analysis of the two instruments reveals that the cognitive processes required of the individual are not identical. The

sorting test has been described as a multidimensional measure of category width; the Pettigrew Scale has been called a unidimensional measure (Bruner & Tajfel, 1965a; Gardner & Schoen, 1962). Although the two instruments have been used almost interchangeably, research into the relationship between the two has produced ambiguous results. Gardner & Schoen (1962) in a factor analysis obtained a strong factor on which both unidimensional and multidimensional tests had significant loadings. A subsequent study (Sloane, Gorlow & Jackson, 1963) indicated that the category width values obtained on the two test types were independent. That is, persons who formed few groups on the sorting tasks did not necessarily have broad category widths on the scale-type tests. Wallach & Kogan (1965) similarly found insignificant correlations between the two indices.

Tajfel, Everstine & Richardson (1964a) have reported positive but moderately low correlations between category width measures obtained from the two test types, and have stressed the need for specifying the variables involved in a task in which category width is to be assessed. In a following study (Tajfel, Everstine & Richardson, 1964b) they emphasized the importance of knowing whether the test involved cognitive risk and whether or not the risk was real or hypothetical. Bruner & Tajfel (1965) citing these works of Tajfel, Everstine and Richardson, have suggested that individual consistency in category width does not hold across multidimensional and unidimensional tests of categorizing breadth. They point out that the cognitive risk is not the same for both procedures. Recently, Gardner & Schoen (1962) have expressed the opinion that multidimensional tests of category width are superior to unidimensional tests because they more nearly approximate the complex

extra-laboratory conditions of conceptualizing. Clearly, the issue awaits further clarification.

In spite of the confusion and dissension, several positive results of the studies on categorizing breadth can be noted. Regardless of the test form used, it has been found that individuals characteristically differ from each other in what they are willing to include in a category (Bruner & Tajfel, 1961; Gardner, 1953). They have been found to be self consistent in their category width preferences when tests were of similar structure (Bruner & Tajfel, 1961; Fillenbaum, 1959; Gardner, 1953; Gardner et al., 1959; Gardner, Jackson & Messick, 1960; Gardner & Schoen, 1962; Glixman, 1965; Pettigrew, 1958; Wallach & Caron, 1959). Similarly, it has been determined that individual consistencies can be detected in a wide variety of tasks (Fillenbaum, 1959; Sloane, 1959), over lengthly periods of time (Gardner & Long, 1961; Gardner & Schoen, 1962) and in widely differing subject meaning domains (Glixman, 1965; Sherif & Hovland, 1953). Thus, an individual's preferential breadth of categorizing appears to be an integral part of his basic cognitive style.

b. Interpretations of category width

Much research has been done with a view to explaining the meaning and implications of the observed individual differences and personal preferences in category width. The usual procedure has been to determine the relationship between category width and other cognitive and personality variables through correlational studies and to make inferences from these findings about the nature of the category width phenomenon. Several interpretations have been offered, most notably those based on psychoanalytic and information theories, as has

been noted previously.

i. Category width as ability to discriminate.- The explanation of category width differences that first comes to mind is that the differences between broad and narrow categorizers are attributable to individual differences in ability to discriminate. It seems quite logical to argue that narrow categorizers tend to notice differences and to classify accordingly while broad categorizers, hampered by an inability to discriminate, lump everything similar together under the same rubric. Berkowitz (1960) has based his notions of the significance of category width on this argument, using as evidence a study by Gardner (1953) which noted that people who used relatively few categories were subject to size and shape constancy effects. From this observation it followed that broad categorizers were more subject to assimilation effects and were actually unable to form many categories chiefly because they did not discriminate differences.

Compelling as the simplicity of this interpretation may be, additional studies can be cited to cast doubt on it. Concerning the relationship between category width and ability to discriminate, Bruner & Tajfel (1961) found no relationship between category width and accuracy of perceptual judgments. Others have also reported negative findings on this hypothesized relationship (Fillenbaum, 1959; Gardner et al., 1959). Indeed, the original assumption on which Berkowitz' interpretation was based (Gardner's finding that broad categorizers are more susceptible to assimilation effects) has not been supported in subsequent studies (Gardner et al., 1960; Gardner & Schoen, 1962; Sloane, 1959). Thus, this view has been largely discarded.

ii. Category width as impulse expression.- Gardner's investigations, rooted in ego-psychology and the theorizings of Klein (1949; 1954; 1958) and Rapaport (1951; 1957), yield an "impulse expression" interpretation of category width. Evidence has been accumulating to indicate that breadth of categorizing is a very complex dimension, much more complex than was originally supposed. A sequential review of Gardner's position reflects this.

In 1953 Gardner suggested that broad and narrow categorizers differ in their preferred ways of knowing the world about them, and that differences in their breadths of categorizing reflect differences in their modes of reality testing and impulse expression. Broad categorizers, he contended, do not feel compelled to rely on exact information concerning the external world in order to respond to it. Instead, they rely on their own feelings and emotions. On the other hand, narrow categorizers respond only in terms of reducible, classifiable elements and maintain a vigilant, discriminative approach so as not to express inappropriate impulses. Their attention to detail, somewhat of a defensive mechanism, he suggested, is evidence of their inability to express impulses freely. Thus, he concluded that broad categorizers do not differ from narrow categorizers in ability to discriminate differences but that the narrow categorizers, being impelled to group in terms of differences are of necessity more difference oriented.

Some support for this view was found when it was noted that broad categorizers projected more human movement and color on the Rorschach than narrow categorizers, demonstrating that they have greater freedom of emotional expression (Gardner, Jackson & Messick, 1960). In addition, they gave more variable responses to the stimuli in a free

word association test which was interpreted as meaning that those with broad inclusion ranges are able to relax reality oriented rules for category inclusion when required to do so (Gardner, Jackson & Messick, 1960).

A subsequent study (Gardner & Schoen, 1962) forced a revision of Gardner's original position, however, when it was found that the replies of narrow categorizers were more comprehensive than those of broad categorizers when they were asked to define relationships among the 10 items on a Rokeach narrowmindedness test. This seemed to indicate that narrow categorizers were aware of similarities as well as differences. Additionally, the previously noted relationships between broad category widths and high scores on a free-association test were not confirmed by this study (Gardner & Schoen, 1962). Gardner, henceforth, reconceptualized the number of groups formed on a free sort test as being indicative of conceptual differentiation and hypothesized that those forming many groups were not simply difference oriented but were "in a much more general sense unusually responsive to questions concerning similarity-difference." (Gardner & Schoen, p.3)

iii. Category-width as a risk-taking strategy.- The categorization process has frequently been interpreted in the light of information theory. According to this theory, the organism receives input from the stimulus world. This input is assimilated and referred to existing categories with efficient and parsimonious strategies (Kogan & Wallach, 1964). Chief proponents of this view have been Bruner and associates (Bruner, Goodnow & Austin, 1956; Bruner & Tajfel, 1961; Kogan & Wallach, 1964; Wallach & Caron, 1959).

In this context, information theory relative to the categorization process states that broad categorization is reflective of a

willingness to accept a wide range of values as acceptable for inclusion in a category. Since this range is large and therefore the defining criterion not so critical, the broad categorizer is able to refer objects to a category on the basis of relatively scant information. Such subjects are called risk-takers. The narrow categorizers, on the other hand have small inclusion ranges for each category and require more information, more knowledge, more cues before they can refer input to the appropriate class. Thus, the narrow categorizer is cautious or conservative and has been branded a non-risk taker (Bruner & Tajfel, 1961; Wallach & Caron, 1959).

This interpretation of category width has not been wholly supported in the more recent literature (Tajfel, Richardson & Everstine, 1964a, 1964b). Results have indicated that the relationship is not so simple, nor so general, although some evidence has been presented to show that under certain conditions broad categorizers are risk takers. A large-scale investigation by Kogan & Wallach (1964) has pointed out the importance of moderator variables such as personality traits in revealing meaningful links between category width preferences and risk-taking tendencies.

Comparison of a risk-taking versus impulse expression interpretation of breadth of categorizing underlines the many similarities in the two views since the risk-taker and the impulsive categorizer appear to be quite analogous. Wallach & Kogan (1965) have suggested that a strategy interpretation may be merely a more exacting view of the style concept and that one interpretation does not completely rule out the other. However, there appears to be at least one important difference deserving of mention.

By virtue of the psychoanalytic theory on which it rests, Gardner's interpretation of category width suggests that he conceives of the process as being largely an unconscious one. On the other hand, a strategy or decision making interpretation allows for a more conscious and purposive process. This could have important implications for hypothesizing possible effects of test form differences on categorizing style. A theory which sees cognitive risk-taking as being conscious and purposive suggests a flexible cognitive style geared to immediate needs and external conditions. Hence, whether or not a person categorizes conservatively may depend on the nature of the stimuli being categorized. On the other hand, an impulse expression interpretation based on neo-Freudian psychology admits of less influence by immediate situational factors and favors a more consistent generalized response that is relatively unaffected by stimuli changes and motivational factors. Hence, according to this viewpoint, test form differences would not be expected to obliterate habitual categorizing styles.

c. Ambiguities and inconsistencies in studies of category width

In view of the divergent notions as to the meaning and definition of category width, it appears that the only generalization that can safely be made is that category width is an indicator of the way an individual comes to terms with the similarities and differences in his environment. Much of the confusion may have resulted from the fact that there are wide test and procedural differences in the studies investigating stylistic aspects of categorizing. Thus, results between experiments using differing tests may not be comparable and studies which appear to support or contradict each other may not actually do so

at all.

The differences between the unidimensional and multidimensional tests used to tap category width and the resulting discordance have already been indicated. Other testing problems may have contributed to the inconsistency in the field as well. For example, interpretation of the category width phenomenon has been hindered by differences in methods of scoring the sorting tests. When subjects are required to group an array of items into comfortable sortings they are told that any item which does not seem to belong with any of the others may be put in a category by itself. Gardner's procedure has been to score each unsorted item as a category (Gardner & Schoen, 1962). The index of category width, therefore, includes the total number of categories containing two or more items plus the number of categories containing a single item. This is known as a total category index. Messick and Kogan (1963) questioned this scoring procedure and indicated that the number of defined categories (those with two or more items) is independent of the unsorted items (those categories with just a single item). Thus, they proposed that the number of defined categories be adopted as the appropriate measure of category width.

Little is known about the significance of unsorted items. Messick and Kogan (1963) have indicated that for college males it appears to be an index of compartmentalization of thought. Wallach and Kogan (1965) supported the independence of unsorted items from defined categories in male children, but found the two to be correlated in the sortings of female children. Several recent studies have adopted the Messick and Kogan index of category width (defined categories) but, because findings were restricted to males, theirs was not a definitive

study and more work is needed.

What further complicates the comparisons and analysis of individual differences and consistencies in categorizing behaviour is the fact that within each of the two main test types, test form differences have arisen. The Pettigrew Scale has been modified for use with children (Wallach & Caron, 1959) and many adaptations and variations of the sorting tests have mushroomed.

The original free sort instrument, Gardner's Object Sort Test, is a heterogeneous array of 73 familiar, concrete objects, and has been a popular test for studying categorizing styles in the past (Gardner, 1953; Gardner et al., 1959; Gardner, Jackson & Messick, 1960; Sloane, 1959). Other sorting tests designed to tap breadth of categorizing are self descriptive adjectives (Fillenbaum, 1959), verbal statements pertaining to self and to war (Glixman, 1965), phrases describing human behaviour (Gardner & Schoen, 1962) and photographs of people (Gardner & Schoen, 1962). These tests have all evolved, it appears, out of attempts to test consistency of category width performances in varying situations.

Demand for portable equipment and the necessity of possible group testing have resulted in other forms and variations. Clayton and Jackson's version is a printed list of names of familiar items that are grouped with pencil and paper rather than with an actual sorting procedure. Clayton & Jackson (1961), Gardner & Schoen (1962), Tobin (1963) and Messick & Kogan (1963) have all used this instrument for obtaining measures of category width. Glixman (1965) has used a verbal form of Gardner's original object sort test.

Although all the tests just referred to have involved multidimensional sorting procedures, they have differed among themselves in at least one of several ways: perceptual form, meaning domain, familiarity of items, complexity and difficulty. While these differences have been useful in determining individual consistency across test forms and in extending testing to large numbers of subjects, the variations and their implications have seldom been recognized when test results have been discussed and compared. Yet, testing in the field of intelligence has told us that manner of presentation is important and the whole rationale for training of testers to administer individual intelligence tests is so that the normative data accumulated from a given form of presentation is also available to the individual tester, given the additional fact that the subject tested is, or would have been, a possible candidate for inclusion in the normative sample. This, in turn, draws attention to another disquieting trend in testing in the cognitive area.

Nearly all these studies have been carried out with adults, usually freshman college students. Yet, it is somewhat disturbing to note that the results have been used in the production of theories applicable to younger students and to children. Few studies have actually looked at categorizing style in younger populations and when they have, they have usually used adaptations of adult test forms. The children's form of the Pettigrew Scale has already been mentioned. A line drawing adaptation of Clayton and Jackson's verbal test has been designed for use with children (Kogan & Wallach, 1965) and it might be pointed out that the authors found that the low correlations of category width values between it and the children's version of the Pettigrew

Scale indicated that, for their population, the two indices were unrelated. Furthermore, no studies to our knowledge have clearly established the relative effects at different ages of these tests whether they be in adult or children's form.

It should be recognized that there is a place and function for the speculative and for the theorizing. Smith, Bruner & White (1958) noted this and pointed out that in the development of a science there is a strategy of discovery and a strategy of proof. It appears, however, that eventually the psychometric bases of the tests must be clearly established so as to proceed from speculation to fact; from discovery to proof. Recently, the apprehensions felt and the doubts experienced about discrepancies in test and procedures in the domain of cognitive style have been elegantly expressed by Klein, Barr and Wolitsky (1967)

> It would seem profitable to determine the distinctions, if any, among these concepts, to continue to explore their dimensionality and degree of generality, and to study their hierarchical organization. The use of different measures for similar concepts makes comparison of different studies difficult. Large scale multivariate research on cognitive controls and the further delineation of individual cognitive consistencies promises to be an important contribution to our knowledge of personality organization.
>
> (p. 515)

Some attempts have been made in the "further delineation of individual cognitive consistencies" by examining the labels individuals use to define the categories they form on the sorting tests (Gardner & Schoen, 1962; Kagan Moss & Sigel, 1963; Wallach & Kogan, 1965). The labels are believed to be indicative of the quality or level of response preferred by a subject. A review of the research on this aspect of categorizing behaviour indicates the highly exploratory and tentative nature of such studies.

2. Level of Response

Thus far the discussion has been limited to one dimension of categorizing style, category width. Just as style is but one index of categorizing behaviour, so too is category width but one aspect of style. In addition to a span or realm, each category has a content and it is of interest to study individual differences and characteristic modes of response in this function of the categorizing process. Multidimensional tests of categorizing behaviour, that is the sorting test, can be used for this purpose. For obvious reasons, the scale-type, unidimensional test cannot.

The significance of an individual's category content as indicated by the label he has placed on the category has not been widely investigated. There is not, as yet, conclusive evidence to support the contention that level of response which indicates the basis of category inclusion is an enduring and self-consistent aspect of categorizing style. One suspects that this is so, largely because of the difficulty in obtaining a reliable objective criterion against which the responses may be measured.

a. Measurement of level of response

All procedures used to assess the qualitative level of content of an individual's categories assume that the language used to label the categories reflects this level. Thus, the subject is asked to indicate verbally what the items he has grouped together hold in common. Several methods have been used to measure the quality or level of these responses. Kagan, Moss and Sigal (1960) employed an ego-centric versus stimulus-centered criterion to classify the content of categories formed on a figure sorting test. They concluded that children exhibited more

frequent ego-centered responses than adults. In a subsequent study (Kagan, Moss & Sigal, 1963) these same researchers used an active-analytical versus passive-relational classification to judge response level. Results were clouded by the use of widely different test forms for the adults and the children, but some evidence was presented to show that passive-relational responses indicated a developmentally low mode of functioning. Kennedy & Kates (1964) scored categories as adequate or inadequate. This classification, however, was found to be too closely tied to intellectual ability to be a valid indicator of cognitive style.

The most common method appears to be one which scores labels along a concrete-functional-abstract continuum, according to a scheme devised by Rapaport, Gill and Schafer (1946). In this procedure, an abstract response is considered to be one which takes cognizance of a common attribute of objects with other diverse properties. The functional classification is based on use and the concrete is based on location, belonging or simple association. An abstraction index based on the proportion of abstract responses to total responses has been used with limited success (Gardner, Messick & Jackson, 1960; Gardner & Schoen, 1962).

b. Level of Abstraction as a Dimension of Style

There is a substantial body of research which suggests that the level of abstraction at which a person chooses to function is a highly generalized trait, discernible in a wide range of behaviour (Vinacke, 1952). Weigl (1927) found that subjects characteristically and consistently exhibited either a concrete or a conceptual approach in sorting tasks. Goldstein and Scheerer's scoring scheme was based on an

abstract-concrete classification (Goldstein & Scheerer, 1941). The tripartite system of Rapaport (Rapaport, Gill & Schafer, 1946) also included an abstract as well as a concrete and functional definition. Hanfmann (1941) marshalled evidence to support a view that individuals consistently display a thinking (conceptual) or a perceptual (concrete) approach to a sorting task. Harvey, Hunt and Schroeder (1961) have based their studies on the assumption that an abstract attitude is a highly generalized characteristic.

Gardner & Schoen (1962), however, have presented evidence to suggest that the level of abstraction at which a person chooses to function on the object sort test is a relatively transient mode of response. They indicated that an adult's preferential level of abstraction as revealed by the labels he gave to the categories he had formed was consistent for short periods of a day or two but varied over longer periods such as two years. Nothing is known of the stability of preferred level of response in childhood.

In their study, Gardner & Schoen (1962) used three sorting tests to study levels of abstraction. These were Clayton & Jackson's (1961) word sorting test, a behaviour sorting test and a photo sorting test. It was found that concrete responses predominated in the paper and pencil test, that the number of abstract responses increased in the behaviour sorting test and that abstract responses predominated in the photo-sorting test. This seemed to indicate that the nature of the test had such a potent influence on the abstraction index that the preferred level of response could not be said to be a consistent, transituational mode of response, that it was not a true dimension of style. However, one may suspect that the difficulty of the tests, especially the photo-sorting test, may

have forced the use of the abstract responses. If such were the case, the tests were not true measures of categorizing style but were more a measure of capacity to abstract.

In summing up the evidence presented regarding the measurement of response level, it appears that no criterion has proved entirely satisfactory. It appears, too, that the relative value of the abstraction index as an indicator of a consistent and enduring categorizing style is open to debate. The whole issue merits further study with tests that are true measures of style and not capacity.

C. SUMMARY

A review of the evidence thus far presented suggests that there are self-consistent and enduring styles of categorizing when these styles are ordered along the category width dimension. This has been confirmed in the literature by a wide variety of experimental procedures carried out, for the most part, on adults of similar age. Conclusions are less clear about the stylistic characteristics of the labels used to define the categories, and many of the findings in this regard are quite tentative.

Similarly, the meaning and interpretations placed on the category width and response level have resulted in much lively controversy (Bruner & Tajfel, 1965a, 1965b; Gardner & Schoen, 1965). Some of the confusion and discord has been the result of the use of a wide variety of tests and scoring procedures with little attention being paid to the significance of the resulting variability that might have been produced. The almost exclusive use of a single age group has also placed limitation on interpretations and understandings of the observed phenomena.

Since the effects of age and test form appear to be of central concern to further study in this area, a review of the theory and research on age and test form effects as they relate to the categorization process is indicated.

CHAPTER III

A REVIEW OF THEORY AND RESEARCH RELATING TO EFFECTS OF AGE AND TEST FORM ON CATEGORIZING BEHAVIOUR

A. EFFECTS OF AGE

If one wishes to examine the stylistic aspects of cognitive functioning in subjects of varying ages to find out how differences among people in cognitive modes arise, the first step is to search for evidence to see if, in fact, significant age differences have been found to exist. However, few studies have been concerned with the development of cognitive styles in children; even fewer have focussed on the development of categorizing style as defined by the widths and the contents of the categories. This, perhaps, is not surprising. The chief concern of the cognitive psychologist has always been with allowing the subject to exhibit his behaviour and there has been less interest in explaining the process by which that behaviour has developed. Anderson (Anderson & Audubel, 1965) has summed up what has come to be called the cognitive approach to the study of higher mental processes in this way;

> Among cognitive psychologists there is more emphasis on what the organism is like now and less emphasis in actual practice on determining the controlling antecedent factors that account for the current status of the organism. Furthermore, at any point in time, the cognitive psychologist typically manipulates stimuli more to get the organism to exhibit its processes than in an effort to determine how and to what extent the organism can be explained and controlled in terms of its environment.
>
> (p. 396)

1. Research on Cognitive Styles in Children

It has been pointed out previously that most research in this field has been carried on with subjects of college age. Thus, very little is known of the formation, growth and stability of categorizing styles prior to adulthood.

Witkin and associates (Witkin et al., 1962) have undertaken developmental studies of "modes of articulation" in children of different ages. Thus, assuming that relative effects of tests are similar for all the relevant age levels, he could make comparisons and plot changes from age to age. It was found that for mode of field approach and articulation of body movement, there was a change, over time, within the group as a whole towards a more analytical approach. Results suggested that relatively more or less differentiation was a stable characteristic of a growing child. That is, children at age 10 who, relative to their group have a highly developed degree of differentiation, tended to have the same relative standing at age 17.

Santostephano & Paley (1966) carried out an investigation of the development of two cognitive controls, focussing vs. scanning and constricted vs. flexible, in children aged six years, nine years and twelve years. Results indicated that such cognitive styles did operate in children and their growth appeared to be consistent with developmental theory. To our knowledge no studies have been reported on the development of categorizing style defined by the span or width of the category and its content, although Gardner has indicated that he is at work on certain unspecified facets of this problem (Gardner, 1964).

Identification of certain aspects of style in categorizing in children has been made, however. Wallach & Kogan (1965) using a

picture form of Clayton & Jackson's sorting test and the children's Pettigrew Scale reported that fifth-grade children displayed identifiable styles in breadth of categorizing on both tasks. Wallach & Caron (1959) using a children's form of the Pettigrew Scale arrived at a similar conclusion with sixth-grade children. No between age comparisons can be made however, because of the divergent nature of the test samples and the fact that the Wallach & Kogan study was primarily designed to explore the creativity variable. Moreover, no comparisons can be made with results obtained from other studies using older populations and adult forms of these tests for there have been no long term studies to determine the validity and comparability of these tests for different age levels.

One longitudinal study, that was carried out at the Fels Institute (Kagan, Moss & Sigal, 1960; 1963), has explored age differences in categorizing styles defined by breadth of category and level of response. However, they used widely different measures for adults and children; this has made age comparisons impossible for the present. This study is still in progress and will eventually provide some psychometric basis for the use of its tests at varying age levels. Unfortunately, these tests are not widely used elsewhere.

The paucity of research directly relevant to the development of categorizing style in children makes it necessary to turn to developmental theory for possible <u>a priori</u> hypotheses on age differences in this particular area of cognitive functioning. Logically, age differences could be expected because of the well known influence of growth and experience.

2. Developmental Theory and Age Differences in Cognitive Functioning

The literature abounds in evidence to support a developmental theory of cognitive functioning (Flavell, 1964; Goldman & Levine, 1963; Inhelder & Piaget, 1958; Piaget, 1953; Scheerer, 1954; Vygotsky, 1962; Werner, 1961; Zaslow, 1959). According to this theory, development begins with a global phase, progresses through a period of individuation or differentiation and culminates in a synthesis or integration of the differentiated parts. Thus, through development, early forms are superseded by more effective forms.

Throughout many of these studies, the implication has been that differentiation is to be valued because, other things being equal, the differentiated person has more resources (more categories) for coping with his environment. But, what does a differentiated form of response imply as far as categorizing style is concerned?

a. Developmental theory and category width

The interpretation of number of categories formed on an object sorting test as a measure of conceptual differentiation (Gardner & Schoen, 1962) brings to focus the role of developmental theory in cognitive style. It seems to imply that those who sort many categories are exhibiting a high degree of differentiation. In the light of developmental theory this would suggest that those with narrow category widths are displaying a developmentally more advanced level than those who sort few categories, the broad categorizers. This could mean that one might expect children to have broader category widths than adults.

It is hazardous, however, to postulate a simple relationship between age and category width on the basis of the differentiation

hypothesis alone. Developmental theory as formulated by Werner (1962) and Vygotsky (1962) has implications for the integration of a system. Integration is a patterning of total structure (Witkin et al., 1962), a synthesizing of differentiated parts into a unified whole. There is, then, the possibility that there are two kinds of broad categorizers: those who group together or synthesize from detailed differentiation and those who experience diffusely and group together without differentiation. If such were the case, all narrow categorizers would be exhibiting a developmentally mature form of cognitive style, but broad categorizers could well be at either a developmentally low or a more advanced level. Much of this rationale was used by Tobin (1963) to study levels of cognitive style in a single age group, college freshmen. He found some evidence to support the view that such a dichotomy exists among broad categorizers. However, the homogeneous nature of his sample, being of similar age, and the same sex, may well have minimized other possible differences.

There is yet another possibility to be considered when age differences in categorizing breadth are hypothesized on the basis of developmental theory. It is entirely conceivable that one's preferential mode of functioning as revealed by a free sort test in no way reflects one's capabilities as indicated by one's developmental level. One may prefer to function at a level below that of which one is actually capable. This is particularly true of adults since they presumably have more levels from which to choose. It may be reasonable to expect, therefore, that developmental level is a potent determiner of preferential response in the younger, but not in the older, individual.

The interesting implications suggested for possible age differences in breadth of categorizing by developmental theory accentuates

the need for studies to determine the extent and direction of age differences in cognitive styles.

b. Developmental theory and level of response

Both Werner (1961) and Piaget (Flavell, 1963; Piaget, 1953) have expressed the view that the essential difference between adult and child intellect is denoted by the appearance of an abstract faculty in early adolescence. Thus, cognitive growth can be seen as a progressive development in the ability to abstract. Since level of response in this investigation is measured by an abstraction index, developmental theory may have implications for the study of age differences in response levels.

It is not known whether people preferentially operate at their optimum cognitive levels. Although there is some evidence to suggest that preferential level of response as determined by an abstraction index is not related to ability to abstract in adult females (Gardner & Schoen, 1962), nor in adult males (Tobin, 1963), the pattern may be quite different for younger children.

On the basis of developmental theory one could not expect a young child or pre-adolescent to reveal an habitually abstract cognitive mode as it is, theoretically, beyond his capability (Inhelder & Piaget, 1958). However, one might expect an adolescent to be able to operate abstractly since Piaget has indicated that ability to deal with abstract operations emerges at about 11 or 12 years of age (Flavell, 1963; Inhelder & Piaget, 1953). A study employing children sampled at the age when abstracting ability has just developed and compared with adults in whom there has been opportunity for experience and practice to affect the use of the abstract mode should be of interest.

In summary, evidence from developmental theory suggests that age differences in cognitive functioning exist. Age, whether conceived of in terms of maturation or experience, appears to be a possible potent factor in determining the categories (concepts) available and, perhaps, preferred for ordering an array of items. If preferential mode of response reflects in any way the availability of concepts or the individual's optimum level of cognitive development, one would expect to find age differences in cognitive style.

B. EFFECTS OF TEST FORM

It has been noted that investigations pertaining to cognitive style are rooted in the field of psychology that explores individual differences in ways of organizing and experiencing stimulus events. Throughout this area of study one theme is clear: We organize and structure an array of items not only as <u>they</u> are but also as <u>we</u> are. Thus, the main concern of investigations in this area has been with the effect of the individual as a mediator between stimulus and response. Yet, the assumption that the category formed is predictably a function of idiosyncracies in categorizing behaviour in nc way rules out the assumption that it is also dependent upon the nature of the stimuli being categorized. Virtually no studies have been concerned with preferential modes of categorizing behaviour from this view point. And it is recognized that when we become concerned about the possible effects of test forms on modes or styles of categorizing we are further blurring the boundaries between the cognitive and behaviorist territory, a recent trend noted by Ausubel (Anderson & Ausubel, 1965).

1. Studies Regarding the Effects of Test Form

The neo-Freudian theory from which much of the concept of cognitive style is derived suggests for many researchers that man's breadth of categorizing and, perhaps to a lesser extent, his level of response should be relatively unaffected by changes in the form of the stimuli to be categorized. Thus, studies in cognitive style have largely ignored the influence of stimulus material on categorizing style preferring to emphasize the individual's consistency rather than his variations across test forms. Correlations of category width values between tests have been positive but moderately low, usually in the neighborhood of 0.30-0.50 for sorting tasks (Fillenbaum, 1959; Gardner & Schoen, 1962). Thus, much of the variance is left unexplained.

Early studies on conceptualizing were greatly concerned with the differential effect of stimuli on the order and ease with which concepts were attained (Hull, 1920; Smoke, 1932; Heidbreder, 1946, 1949; Heidbreder & Overstreet, 1948). The experiments of Heidbreder are of particular interest. She presented evidence to indicate that there is a relatively consistent hierarchical dominance in which concepts are usually attained. Further studies (Dattman & Israel, 1951; Grant, 1951; Grant & Curran, 1953) revealed that the order of concepts was neither permanent nor constant but was subject to the influence of such situational factors as stimulus material, age and experience. All these studies were carried out using guided sorting procedures and it now seems crucial to ask if the same situation holds under free sorting techniques, that is, do a subject's preferential sortings, more especially category width and level of response, reflect changes in stimuli presentation?

a. Test form and category width

Although many test forms have been used in the measurement and study of category width, there is only one investigation known to us which directly measured the effects of different stimuli presentations on the breadth of categories formed. Glixman (1965) employed three series of statements, each from differing meaning domains, and required that the subjects categorize them under free conditions of sorting. The three meaning domains were ordered from peripherally to personally relevant: objects, war, self. It was found that a greater number of categories was used to sort objects than for the other, more personally relevant domains. He interpreted the results, in the light of information theory, to mean that by setting the number of categories to be used for the domain, a subject controls the amount of information with which he has to deal. It appears that one can choose broad categories for items related to self because one is relatively certain about the accuracy of one's information.

Apart from revealing that differences in tests are related to differences in preferred category widths, there is little that one can generalize from this one study. One can only wonder if other variations in test forms would be reflected in variations in breadth of categorizing. For example, if meaning were held constant and test form were varied, would category width differences be found?

b. Test form and response level

There is some direct evidence from the literature which points to a relationship between the form of the stimuli being categorized under free conditions and the level of response as determined by the

quality of the label placed on the category. The study of Gardner & Schoen (1962) which indicated that the nature of the test (whether it be because of form or difficulty) was a strong determiner of response level has been previously discussed.

Forgus and Fowler (1957) found the basic mode of response to be concrete unless test items were familiar, in which case an abstract mode predominated. Their study, although allowing personal preferences within certain limitations, was not strictly speaking a "free" procedure. Reed (1946) in a study similar to Forgus' in methods showed that the more complex the stimuli, the more likely was the response to be based on sensory or concrete cues.

Lundsteen & Michael (1966) who tapped a dimension of cognitive style in response level by using three forms of stimuli presentation - words, paragraphs and stories - which did not involve a sorting procedure found that with children, the more complex the material, the more abstract the mode of response. If this appears to contradict Reed's findings, it must be remembered that different procedures were involved, and that Reed's subjects were adults and Lundsteen's & Michael's were young children.

All this seems to suggest that response level, although predominantly concrete, may be somewhat affected by stimulus form in free sort procedures, especially if there is a variation in complexity or difficulty among the stimulus forms. This may have relevance for the question of whether there is an age differential in test form effects as whether or not a test form is considered to be difficult or complex may well depend on the age of the subject.

2. Developmental Theory and Test Form Effects

Research previously cited has suggested that whether or not test form differences are reflected in changes in categorizing behaviour conceptualized as style, might depend on the age of the individual. Developmental psychology tells us that there is increasing specificity of psychological traits with increasing age (Garrett, 1946). The relatively immature respond globally to different situations; they have a homogeneous, generalized response which comes into play regardless of the situation. Those at a higher developmental level, with a greater repertory of responses, respond specifically to each situation. On this basis one would expect that the hypothesized effects of changes in stimuli would depend upon the age of the subject.

Thus, in the light of developmental theory and studies in concept attainment, it appears that whether or not stimulus form (i.e., differences in test form) affects category width or level of response, may well depend on the age of the categorizer.

CHAPTER IV

THE EXPERIMENT

A. STATEMENT OF THE PROBLEM

A review of the available literature demonstrates the speculative nature and the intriguing prospects of the study of categorizing behaviour in terms of individual styles. The status of the research, it appears, is somewhat analogous to that of the investigations into measurement of intelligence in the early 1930's when there were many studies and many tests but insufficient data on any given population, either cross-sectional or longitudinal, to make cross-referencing possible. The almost exclusive use of college students as subjects and the growing proliferation of tests with which to assess consistent response modes in categorizing behaviour has given rise to certain doubts and questions. It is not enough to note and attempt to measure a phenomenon such as a characteristic tendency to use broad or narrow categories; the curious will always wonder how it came to develop in a particular individual. Because of the limited amount of evidence available from the few reported studies at varying ages or stages of development, the recurring question of how a consistent response mode known as style arises cannot be adequately answered.

To do this is not an easy or short term task. Given that a phenomenon is found in an adult, then one must demonstrate that phenomenon at increasingly younger ages. If the results are consistent with younger age levels then it is possible to extrapolate adult scores

downwards to any age level. However, if one reaches an age where the test becomes unsuitable, for example a printed word test for non-readers, one can only extrapolate downwards and to another form of test, provided similar results have been obtained under both test conditions at the older age. So far no one has demonstrated that age consistency exists, but most workers have acted as though it does.

Additionally, with the widespread use of increasingly different test forms - shapes, objects, pictures, words - it seems important to ask if the differences between tests are so insignificant as to allow them to be used interchangeably even at a single age level. Is an experimenter justified in drawing the same conclusions from a test given with material of one kind, such as real objects, from that given with another, such as printed words? It is quite possible that differences in stimulus form and presentation could produce measureable differences in the breadth of categories and level of response used in the sorting task. It is not known what these differences could be, if they follow a predictable pattern or, indeed, if they exist at all.

Thus, the question of possible age differences in categorizing style and the interesting challenge of mapping their growth over time in the individual is confounded by the already confused problem of the effects of test form variations on a single age group or on an individual. It is probable that the same tests cannot be used for all ages and it is possible that similar or identical tests could have diverse effects depending upon the age of the subject. Accordingly, it was decided to make a modest attempt to fill in some of the gaps in the existing knowledge.

This research, then, was designed to study, at two age levels, the effect of verbal, picture and object forms of a free-sort test on an individual's categorizing behaviour, when categorizing behaviour is defined as: (1) the number of categories formed, (2) the percentage of abstract labels placed on the categories thus defined. Because of recent controversy two additional measures, number of unsorted objects and the total category index are included for exploratory purposes.

Four broad questions are posed:

1) Does categorizing behaviour defined as number of categories formed and level of response change as a function of test form?

2) Does categorizing behaviour defined as number of categories formed and level of response change as a function of age?

3) Is there a test form and age interaction? That is, do various test forms differentially affect categorizing behaviour at varying ages?

4) Is there individual consistency in categorizing behaviour across object, picture and verbal forms of the sorting test? That is, are there significant inter-test correlations?

Before supplying a description of the procedures and methods of this experiment, it is necessary to explicitly underline the reasoning behind the experimental design. Although much of the rationale is provided in the literature already reviewed, further amplication is required since many of the tests and procedures used in this study are of quite recent origin and, hence, may be relatively unknown. Accordingly, a discussion relevant to the choice of subjects, tests, and testing and scoring techniques precedes the account of the actual experimental procedure.

B. RATIONALE BEHIND THE EXPERIMENTAL DESIGN

1. The Subjects

a. Sex

Many researchers have pursued the study of sex differences in cognitive functioning (Gardner et al., 1959; Glixman, 1965; Kogan & Wallach, 1964; Messick & Kogan, 1963; Pettigrew, 1958; Wallach & Caron, 1959; Wallach & Kogan, 1965). Findings have been somewhat ambiguous. For example, Wallach & Kogan (1965) and Gardner and associates (1959) found negligible sex differences in number of categories formed and number of objects left unsorted whilst other studies have revealed significant sex differences (Glixman, 1965; Messick & Kogan, 1963; Pettigrew, 1958). Lack of consistency in the findings prompted the decision to confine the sample used in this study to one sex.

b. Age

Little attention has been given to the developmental aspect of categorizing behaviour under free sorting conditions. Studies using a developmental approach have usually employed structured sorting techniques designed to measure conceptualizing capacity (Goldman & Levine, 1963; Reichard, Schneider & Rapaport, 1944). Those investigations concerned, as this one is, with tapping a preferential mode of categorizing have been carried out with subjects at a single age level, usually first year college students. To facilitate comparisons with previous studies and thus make interpretation of test results more meaningful, it was decided to include a group of first year college girls in this experiment. So that some indication of the developmental aspect of categorizing style could be determined, a sample of younger girls was also

chosen. These girls were grade six and seven students in their twelfth year as of April 1, 1966. This particular age level was selected because it was felt to be the youngest age at which one could be reasonably confident that the tests to be used would be, in fact, measures of preferential categorizing mode and the full range of responses would not necessarily be limited by a maturational inability to form abstract concepts. Rationale for the choice of this age level was based on Piaget's developmental theory of cognitive functioning (Inhelder and Piaget, 1958). According to Piaget, most children at the age of twelve would be in the final stage of development in the evolution of thought processes. They should be able to think abstractly and should already possess a large and varied repertory of concepts. This theory is further supported by other studies which have suggested that children of even much younger ages are capable of thinking abstractly (Reichard, Schneider & Rapaport, 1944).

2. The Test and Testing Techniques

a. The technique of free sorting

The test used in this study employs a free sorting technique as opposed to a structure technique. The implications of this free sort procedure may perhaps be best understood by a comparison of it with the structured or guided sorting test.

As has been previously noted, investigators have long used structured sorting techniques in the study of the thought processes. Among the more popular of the sorting tests has been the Goldstein-Scheerer Object Sorting Test (Rapaport, Gill & Schafer, 1946). This typical object sorting test is administered in two parts. In the first

phase the subject is given an item (e.g., a pencil) and is asked to pick out from a group of 33 common objects all the other items that belong with it. In the second part, groups of objects are presented to the Subject and he is asked to tell what each group has in common. Thus, in the first instance, the category is explicitly or implicitly defined and the Subject is simply required to establish appropriate category boundaries. In the second instance, the category itself is provided and the Subject is asked to label it.

The more recent free sort tests such as the one developed by Gardner (1953) differ from their predecessors such as the one just described both in what they purport to measure and in design.

In a free-sort test no predetermined structuring is provided. The subject is faced with an array of objects and is instructed to "put together into groups the objects which seem to belong together for some reason or another." (Gardner, 1953) Perhaps the Gardner Object Sort Test has been the most widely used test of this type. It includes 73 common objects which the Subject is required to group together in a way that is "most logical, most natural and most comfortable" to him. Therefore, unlike the earlier structured sorting tests, the free sort tests are not tests of capacity to think abstractly or to form concepts. Rather, they are tests of preferential conceptual style. The Subject is allowed to respond freely and there is no compulsion and no incentive for him to produce "best" or "right" groupings.

Advocates of the free sort method of studying categorizing or conceptualizing style point out that while structured object sorting tests may have important uses in diagnosing schizophrenia, brain-damage or other disorders of the thought process, the free sort is more suitable for the study of thought in normal subjects (Gardner, 1953).

They believe it more nearly approximates the cognitive tasks that face the individual from day to day (Gardner & Schoen, 1965). In a free sort test the subject must not only formulate concepts but also he must differentiate them one from the other.

b. Choice of test and test forms

The tests devised for use in this investigation are three item-form variations of the Clayton and Jackson sorting test (Clayton & Jackson, 1961). The test, which in turn was adapted from Gardner's object sort test (Gardner, 1953), consists of a list of names of 50 familiar objects chosen for their specificity and clarity of meaning. Instructions require the subject to group objects which seem to belong together by writing down the names of objects under categories labelled A, B, C, etc. Sloane, Gorlow & Jackson (1963) have indicated that the Clayton & Jackson test and the Gardner test both load heavily on the same factor.

The Clayton and Jackson sorting test, because it is readily portable, inexpensive and easily administered, is an excellent instrument for testing large numbers of people and for this reason would be useful for large scale developmental studies. However, it is not certain that the present unmodified verbal test form would be suitable at lower age levels and already item-form variations of this test have been designed for use with children (Wallach & Kogan, 1965). Thus, because of its widespread and increasing popularity for use in one form or another with varying age groups, and because of its potential for future use, the Clayton and Jackson test was considered to be a worthwhile test for study.

Three main item-form variations of the test are possible: verbal, picture, object. It should be noted that this hierarchy (verbal, picture, object) represents a progressive increase in the amount of information available to the sorter about the test items. For example, the verbal forms give no indication as to color or shape or material, the pictorial forms indicate shape but no color, and not necessarily material, and the object forms portray all three: color, shape, material. This appears to have implications for categorization as a means of information processing. Significant differences in categorizing behaviour among these test forms, if found, might have significance for both Gardner's and Bruner's theories of the categorizing process, and more especially of category width.

In brief, it would appear that a neo-Freudian interpretation of category width as a form of impulse expression suggests that differences in perceptual form of the test would have little effect on the preferential or habitual mode of response. An information theory explanation of category width might admit to significant differences in breadth of categorizing because of the variations in the amount of available information provided by the three forms. The verbal test gives few perceptual cues; pictures in black and white give more; objects, 3-dimensional and colored, give still more. It is quite probable that the risk involved in categorizing efficiently and parsimoniously varies tremendously among the three forms with risk probably being greatest where least information is given.

c. Test order effects

From the outset it appeared that the order in which the test forms were presented might differentially affect test scores. Since

it was impossible to determine just what these effects might be, the order of test presentation was systematically controlled so that test order effect would not contaminate measured effects of test form on categorizing behaviour (see Appendix B for test order and series).

Obviously when the same individual is administered a similar test three times, there is an unavoidable possibility that repetition will affect his behaviour. To minimize this, care was taken to administer each test to each individual at intervals of one week. On all second and third test situations, instructions were modified to remind the subject that this was not a memory test and that she was still free to sort in the manner which was most natural, comfortable and logical to her (Appendix A). It was, of course, not possible to eliminate all order effects, specifically those tied to sheer repetition, so variance due to order was statistically analyzed, but it must of necessity remain of secondary interest in this investigation.

3. Aspects of Categorizing Behaviour

There are various dimensions along which categorizing behaviour can be assessed. The free sort test yields measures of several such dimensions against which the effects of item form and age can be gauged. The dimensions chosen for study in this investigation were selected because of their measurability and their relative independence (Gardner & Schoen, 1962; Messick & Kogan, 1963) as well as for their importance in the conceptualizing process as indicated by the literature reviewed.

These dimensions are:

a) Number of categories defined (Messick & Kogan, 1963)

b) Per cent abstract labels, i.e., abstraction index (Gardner &

Schoen, 1962; Tobin, 1963)

c) Number of objects left unsorted (Messick & Kogan, 1963)

d) Total number of categories formed (Gardner & Schoen, 1962)

a. Number of Categories defined

When a subject is confronted with an array of diverse objects to be sorted into comfortable groupings, as he is in a free-sort test, he must decide how broad or how narrow a range of objects can be logically or naturally grouped together. Some categorizers produce few groupings and place a large number of objects under a single label, while others generate more categories and consign fewer exemplars to each group. Thus, the number of groupings reflects breadth of categorizing with those forming few groups being known as broad categorizers because the boundaries of their categories are relatively wide, and those forming many groups being known as narrow categorizers, the boundaries of their categories being more restricted.

The categorization process has often been conceived of as largely a problem in "breadth" of categories. This dimension of categorizing behaviour was first studied by Gardner (1953) who initially called it a measure of equivalence range, that is, the range of discriminable events subsumed under one label or rubric. Subsequently he re-defined it as "conceptual differentiation" (Gardner & Schoen, 1962). Whether viewed as equivalence range or conceptual differentiation, this aspect of categorizing behaviour is measured by the number of categories formed on a sorting task. It was decided to employ this measure in this study because breadth of category has been found to be one of the most consistent and persistent aspects of individual categorizing style (Gardner, 1953; Fillenbaum, 1959; Gardner, Jackson & Messick,

1960; Gardner & Long, 1961; Gardner & Schoen, 1962; Sloane, 1959).

b. Abstraction index

The method used in this investigation for interpreting the labels placed on the grouping is one which was derived from the work of Rapaport, Gill and Schafer (1946), modified by the subsequent studies of Gardner and Schoen (1962), and most explicitly formulated by Tobin (1963). According to this scheme, reasons given for any grouping can be placed into one of three rubrics: concrete, functional and abstract or conceptual. A conceptual label is one that represents true abstraction and therefore the definition is based on a common characteristic of otherwise diverse objects. The abstraction score used by Gardner and his associates consisted to a ratio of abstract definitions to the total number of categories defined (Gardner & Schoen, 1962).

Although some previous studies had indicated that the abstraction index may not be a fruitful method of exploring qualitative differences in labels placed on categories (Gardner & Schoen, 1962), it was decided to retain its use in this study. The long historical concern with abstraction on the part of researchers investigating developmental theories of cognitive functioning suggested that some consideration be given to use of the abstraction index in a study designed to examine the possible significance of age level on preferential categorizing behaviour.

c. Number of objects unsorted

Until recently, each unsorted object was defined in the literature as a category in itself. However, Messick and Kogan (1963) have indicated that number of unsorted objects is an independent aspect

of categorizing style, unrelated to number of categories defined. It was found that the two measures do not correlate similarly with other personality and cognitive measures. Although their investigation was not a definitive study because of the nature of the sample used, it was decided in this study to include unsorted objects as a separate measure since further valuable information as to the dimensions of this variable might be found.

d. Total categories

Much of the research prior to 1963 had been analyzed by considering that a category can consist of a single, unlabelled item. Thus, number of categories would be composed of both defined categories plus each, single item left in a classification of miscellaneous, or indeed, left unsorted. Although recent research seriously questions the validity of this view, the results of this investigation, initially at least, have been analyzed according to a total categories index as well as to a defined categories concept. This was done to facilitate comparisons with previous studies which used this index and to further comparisons between the two measures themselves.

C. THE PROCEDURE

From the foregoing it is seen that the experiment to study the effects of age and test form on stylistic tendencies in categorizing behaviour was confined to individuals of one sex (female) at two ages (12 and 18) who were presented, in a specified order, with each of three forms of sorting test (modified Clayton & Jackson). The measuring instrument was a free sorting test, i.e., unstructured and administered under free and relaxed conditions. The three forms were

object, picture and verbal. Results of the categorization were to be viewed in terms of number of defined categories sorted, per cent abstract labels placed on categories, number of objects left unsorted, and total number of categories formed. Fuller details of the procedures are provided below and in Appendix A.

1. Testing Procedure

Sixty girls participated in this study. Group 1, drawn from the sixth and seventh grades of three suburban schools, consisted of 30 girls who were in their twelfth year at the time of testing. The 30 subjects in Group 2 were first year college students enrolled in a Bachelor of Education program at Macdonald College. Their mean age was 18.3 years. Thirty-four girls in all volunteered to participate as members of Group 2 but four had to withdraw prior to testing because of test scheduling and transportation difficulties. It is interesting to note that since the entire enrollment in the first year Bachelor of Education program was 39 girls, our group of volunteers was almost a total sample.

Each subject in Group 1 was selected by the principal and staff of her school. The selection criterion was such that the subjects were deemed to be capable, in terms of intelligence and academic achievement, of undertaking a first year college course at the appropriate future date. Volunteers were then called for from among the girls thus selected. In actual fact, each girl who met the selection criterion sought to be, and was, included in the sample.

In this way an attempt was made to roughly match the two groups on the basis of intellectual ability (undefined) and inclusion status.

Testing of all 60 subjects was done by the writer during the months of March and April, 1966. Prior to testing the examiner met with the subjects to brief them on the purpose of the experiment and to assure them of anonymity. They were told that the experiment was designed to study thinking but that the tasks were not measures of intelligence or of how well people think because there was no "best way" of doing them. They were assured that the results would not appear on their records nor in any way affect their academic standings. All subjects were requested not to discuss their sortings with their classmates until the experiment had been completed.

It is believed that this preliminary meeting was successful in allaying any fears and anxiety on the part of the subjects and in exciting their interest and enthusiasm. During the test periods, without exception, the girls approached the sorting tasks eagerly and freely. They all appeared to accept the fact that there were no "right" or "wrong" ways to classify the items and, while some subjects were certainly more particular with their sortings than others, no anxiety or tension was discernible.

All subjects were administered each of the three forms of the sorting test in sequence, seven days apart. Order of test presentation was varied to provide for six series in each age group. Subjects were assigned at random to a series. Thus, there were five girls at each age level assigned to each series. The test sequence for each series appears in Appendix B, Table III.

The presentation of the 50 items on each of three test forms was standardized. The cards, pictures or objects were laid out on a large surface in the same pattern of 10 rows with five items each.

All items were visible to the subjects from the outset and appeared in the same order and position for each subject in all tests.

In the verbal test names of the items were printed in large type on 3x5 white cards, one card for each item. In this way, the original test was modified slightly to allow for an actual sorting procedure.

The picture test was composed of these 50 items represented in 5x7 glossy photographs of actual objects. All pictures were checked for clarity and ease of recognition on a group of 15 sixth grade girls who were not part of the test sample. It was assumed that if the items were recognizable by this group they would also be by the older sample.

The object form of the test included 3-dimensional models of the items. Each object was a miniature replica except for the dime and the aspirin. In these instances the actual item was used. This form was also pre-tested on an additional group of sixth grade girls.

The complete, specific instructions for the initial administration of the test are included in Appendix A. Modifications of instructions for the second and third test periods are also given. Each girl was seen individually and care was taken to provide a relaxed, permissive atmosphere so that individual preferences had the opportunity to operate fully.

It should be noted that, while every effort was made to standardize all aspects of test presentation for both age levels, one unavoidable difference existed. The younger group were tested at school during normal class time, although not in their normal classrooms. The time required to do the test was taken from the usual classroom periods. Most of the older group, however, were tested during the evening in an informal setting and time for the test was taken from

their leisure hours.

2. Scoring Procedure

When the subjects had indicated they were finished with their sortings, the groupings and the reasons for the groupings were recorded. Throughout the testing, notes were made by the examiner on the manner in which each subject went about the task and comments of those who verbalized their thought processes were noted. Following the entire test schedule, each protocol was scored by the examiner for:

1) number of defined categories as per Messick and Kogan (1963) (groupings of two or more items assigned to a common label)
2) abstraction index, or percentage of abstract labels used to define the categories sorted as per Gardner & Schoen (1962) and Tobin (1963)
3) number of items left unsorted (Messick & Kogan, 1963)
4) total categories sorted (each unsorted item is called a category and added to defined categories to provide a total categories index) (Gardner & Schoen, 1962)

Scoring for number of defined categories, number of unsorted items and total categories was an uncomplicated, objective procedure carried out in accordance with the previously discussed definitions of these dimensions. A further word is necessary, however, to amplify the scoring procedure to obtain the abstraction index.

a. Scoring procedure for abstraction index

Each subject's abstraction index for each test was calculated by obtaining the percentage of abstract labels relative to the total

number of labels used to define the categories. This measure was found by Gardner et al. (1960) to load 0.57 on a level of abstraction factor.

The criteria for deciding whether or not a label was abstract was based largely on Rapaport's rationale as modified by Tobin (Tobin, 1963). Abstract definitions were considered to be those which referred to essential, inferential properties of objects. Functional and concrete definitions were also scored but were not analyzed. Functional definitions were those relating to the use to which an object is put; concrete definitions were those referring to a single, usually external, property of an object.

All protocols were scored twice by the same rater with an interval of ten months between scoring. Agreement between the two was nearly 96%. Chief discrepancy was a decision on the second scoring to classify as concrete all categories defined on the basis of the materials of which the objects were made. This is a departure from Rapaport's system (Rapaport, Gill & Schafer, 1946), but one which seemed more logical in terms of the accepted definitions of concrete and abstract, since material is a readily observable, perceptual attribute. Some support is found for this view in the literature and was accepted by Tobin in his scoring procedure (Tobin, 1963). As an additional check, 20 protocols were drawn at random and scored by an independent rater. Comparison of the ratings of the two examiners yielded agreement which covered over 96% of the cases.

While it must be acknowledged that the decision as to whether or not a category is abstract is a subjective one, acceptance of Tobin's scoring instructions converts it to one which is virtually objective. Tobin's instructions are extremely explicit and abundant examples are

given. He has chosen these examples, it appears, from the labels actually given by individuals to define the categories sorted using the Clayton and Jackson test. Guidelines used in scoring the tests in this investigation are given in Appendix A. The reader who wishes a more comprehensive account is referred to Tobin (1963).

CHAPTER V

HYPOTHESES

Based on the review of the literature, it is possible to offer several predictions regarding the effects of age and test form on categorizing behaviour when that behaviour is defined in terms of category width and level of response.

In the light of developmental theory, the following hypotheses can be proposed:

1) There will be significant between age differences in the number of categories formed and in the level of response on all three forms of a free sort test.

a) Older subjects will sort more categories on the tests than younger subjects because they are at a more differentiated level of development.

b) Older subjects will define their categories with a greater percentage of abstract labels than younger subjects because they have greater facility for abstraction.

2) Younger subjects will perform more globally on all forms of the test. Since specificity of response is a function of age, the older subjects may be more specific in their performances so that the width of the category formed and the level of abstraction used may, for them, be a function of the material of the test. Older subjects will, therefore, be less consistent than younger subjects from one form of the test to another.

3) If differentiation is a function of age and the amount of information a function of the test form, older subjects will sort the objects into more categories if the form is real object rather than pictorial, and if pictorial rather than verbal. (The actual objects being concrete and tangible, and so communicating information as to texture, weight, color, etc., will have the effect of producing narrower categories or a greater number of categories will be required to provide for the use of all the information about them.)

On the basis of style theory, the following additional hypothesis is offered:

4) While all subjects will be consistent in category width across test forms, the older subjects will be more consistent than the younger subjects since the development of a cognitive style represents increasing control over action and should be more resistant to change at older age levels.

Because of the recent controversy between Bruner and Tajfel on the one hand and Gardner and Schoen on the other on the meaning and interpretation of category width,* it is also possible to set up one other hypothesis based upon categorization as a decision process with risk-taking implications.

5) Subjects will sort increasingly fewer categories if the tests are in the order verbal, picture, object; (the more information supplied, the more knowledge one has about the item and the more generous one can be in his category boundaries without incurring the risk of being wrong.)

---

*No hypothesis on the basis of style theory is made regarding level of response since its relevance to "style" has not, as yet, been fully determined.

It will be noted that Hypotheses 2 and 4 are in conflict and similarly Hypotheses 3 and 5 are in opposition. It is difficult to suggest which are the more plausible. Empirical evidence is scant and the hypotheses are *a priori*. However, the relatively explicit terms of developmental theory as compared to the amorphous nature of style theories, still in their infancy, favor those hypotheses based on the former, i.e., Hypothesis 2 rather than Hypothesis 4 and Hypothesis 3 rather than Hypothesis 5.

CHAPTER VI

RESULTS

Raw scores for all tests and measures are tabulated in Appendix B, Tables I and II and from these data appropriate totals and means were calculated. Significance of age, test form and test order differences was determined by pseudo-three way classification analyses of variance according to McNemar (1962). The Duncan Multiple Range Test (LeClerg, Leonard & Clark, 1962) was used where applicable to test significance of mean differences. To analyze data on response level, the per cent scores were transformed by arc sine conversions (LeClerg, Leonard & Clark, 1962). Correlations to determine consistency among individual scores on each test form were computed for the number of defined categories and the percentage of abstract responses. Defined level of significance in all cases was considered to be 0.05 per cent.

The results thus obtained are recorded under the following headings: (A) Age and test form effects, (B) Effects of age and test form on individual consistency, and (C) Order effects.

A. AGE AND TEST FORM EFFECTS

Means were computed for the aspects of categorizing behaviour being explored in this study and summarized on the basis of test form and age for purposes of general comparisons (Table 1).

Table 1. Mean Values of Scores Obtained by Two Age Groups on Three Forms of a Free Sort Test

	No. of defined categories	No. unsorted	Total categories	Percentage of abstract responses
Group 1 (Younger)				
Verbal	12.63	2.4	15.03	21.5
Picture	13.03	2.9	15.93	25.0
Object	12.33	2.37	14.70	24.3
Grand Mean	12.67	2.55	15.22	23.7
Group 2 (Older)				
Verbal	11.40	2.5	13.83	41.7
Picture	11.70	2.43	14.13	41.9
Object	10.47	2.8	13.27	41.7
Grand Mean	11.19	2.58	13.54	41.7

Statistical analyses for age and test form effects were carried out on number of defined categories, number of unsorted objects, total number of categories and percentage of abstract responses and, therefore, results will be recorded for each of these aspects individually.

1. Number of Categories Defined

The analysis indicated significant differences between the two age groups in number of categories defined (Table 2). Significant test form differences were also found. There was, however, no significant test form-age interaction effect.

Table 2. Analysis of Variance for Number of Categories Defined on Three Forms of a Free Sort Test by Two Age Groups

Source of Variance	df	ss	ms	F
Total	179	1,124.06		
Age	1	98.27	98.27	7.67**
Subjects within age	58	743.79	12.82	
Test	2	28.74	14.37	6.68**
Test x Age	2	3.84	1.92	0.89 N.S.
Residual	116	249.42	2.15	

**Significant at 0.01 level N.S. Not significant

These results pointed to the necessity of further examination of mean differences, both within and between age groups. The Duncan Multiple Range Test was applied to make all possible mean comparisons (Table 3). The shortest significant range values for test mean differences appear in Appendix B, Table VIII.

Table 3. Significant Mean Differences for Number of Categories Defined on Three Test Forms by Two Age Groups. Duncan Multiple Range Test

	Group 2 (Older)			Group 1 (Younger)		
Group Test	Object	Verbal	Picture	Object	Verbal	Picture
Mean	10.47	11.40	11.70	12.33	12.63	13.03
		N.S.	N.S.		N.S.	
			N.S.			N.S.

N.S. - non-significant differences indicated by underscoring

Data appearing in Table 3 clarify the significant age differences reported in Table 2. Older subjects sorted the items into fewer defined categories than the younger subjects on each of the three test forms, also the relative effects of test forms on number of categories defined were in the same hierarchy, object, verbal, picture, for both age groups. Subjects sorted fewest defined categories on the object form, most on the picture form. Not all differences between test forms, however, reached significance (Table 3).

Examination of Table 3 for test form effects within age groups reveals that the older subjects (Group 2) defined significantly fewer categories on the object form of the test than on the verbal or the picture forms. Therefore, for Group 2 the object form of the test was significantly different from all the others. There were no significant differences between the verbal and picture test forms.

In Group 1, which was composed of the younger subjects, there were fewer categories formed on the object sort test than on the picture test, but there were no significant differences between the object and verbal test forms. Similarly, there were no significant differences between the verbal and picture forms. Thus, in summary, for the younger group, the verbal test was not significantly different from either the verbal or object tests but the object test form was significantly different from the pictorial form.

2. Number of Unsorted Objects

"F" values presented in Table 4 indicated that neither the age of the subjects nor the form of the test were significant factors in determining the number of objects left unsorted.

Table 4. Analysis of Variance for Number of Unsorted Objects on Three Forms of a Free-Sort Test at Two Age Levels

Source of Variance	df	ss	ms	F
Total	179	1,176.2		
Age	1	0.02	0.02	.0014 N.S.
Subjects within age	58	816.18	14.07	
Test form	2	1.43	0.72	.238 N.S.
Test form x age	2	6.24	3.12	1.03 N.S.
Residual	116	352.31	3.04	

N.S. - non-significant

3. Total Categories

Analysis on the basis of a total categories index showed that there were significant test form differences in number of total categories sorted (Table 5). There were no significant differences between age groups in number of total categories formed.

Table 5. Analysis of Variance for Total Number of Categories Formed on Three Forms of a Free-Sort Test by Two Age Groups

Source of Variance	df	ss	ms	F
Total	179	2,504.99		
Age	1	95.32	95.32	3.03 N.S.
Subjects within age	58	1,821.66	31.41	
Test	2	31.34	15.67	3.28*
Test x age	2	2.90	1.45	0.37 N.S.
Residual	116	553.77	4.78	

N.S. - non-significant *significant 0.05 level

Further examination of test form differences appears in Table 6. Since the analysis of variance indicated no significant age differences, the totals for the two age groups on each test were combined on the assumption that for the purposes of this particular investigation the two groups could be considered samples from the same population. Calculations for the shortest significant range values appear in the Appendix (Appendix B, Table IX).

Table 6. Significant Mean Differences for Total Categories Formed on Three Forms of a Free-Sort Test. Duncan Multiple Range Test. Groups 2 and 1 Combined

Group	Object	Verbal	Picture
Mean	14.02	14.43	15.03
	N.S.		
		N.S.	

N.S. - non-significant mean difference indicated by underscoring

Data showed that significantly fewer categories were sorted on the object form of the test than on the picture form (Table 6). There were no significant differences in the effects of the verbal test as compared to the object test. Similarly there were no significant differences found between the verbal and picture forms of the test. Thus, the verbal test form was not significantly different from either the picture or the object forms, but the object form was significantly

different from the picture.

4. Percentage of Abstract Responses

Highly significant age differences were found in percentage of abstract labels used to define categories (Table 7). Test form had no significant effect on proportion of abstract responses.

Table 7. Analysis of Variance for Per Cent Abstract Responses on Three Forms of a Free-Sort Test at Two Age Levels (arc sine transformations)

Source of Variance	df	ss	ms	F
Total	179	40,582.5		
Age	1	5,633.6	5,633.6	18.45**
Subjects within age	58	17,706.22	305.27	
Test form	2	51.51	26.76	0.18 N.S.
Test form x age	2	41.87	20.93	0.14 N.S.
Residual	116	17,148.66	147.8	

**significant 0.01 level

N.S.-non-significant

Since the foregoing results justified a closer look at mean differences on the basis of age only, it was not necessary to use a Duncan Multiple Range Test. The mean percentage of abstract responses for the younger group across test forms is 23.7; the mean percentage of abstract responses for the older group across test forms is 41.7 (Table 1). Therefore, on the basis of information obtained in Table 7, it could be stated that the older group gave significantly more abstract

responses across all test forms than the younger group.

B. EFFECTS OF AGE AND TEST FORM ON INDIVIDUAL CONSISTENCY

Individual scores for the two main aspects of style studied in this investigation were correlated among test forms to determine consistency of response. Results are summarized in Tables 8 and 9.

Table 8. Correlations Between Scores on Test Forms for Number of Defined Categories

df = 28 t = 2.05

Test Comparisons	Group 1	Group 2
Verbal - Picture	.71	.46
Verbal - Object	.71	.62
Object - Picture	.83	.32 N.S.

N.S. - non-significant

A casual inspection of the data in Table 8 indicates that correlations between scores are higher for the younger group than for the older group. All correlations are significant and positive except for that between the object and picture test forms at the older age level which, while positive, fails to reach the defined level of significance (0.05). The highest correlation was obtained between these two tests, object and picture, at the younger age level. Highest correlations were obtained between scores on the object and picture test forms in

the younger group; comparisons of scores on these tests at the older level yielded the lowest (insignificant) correlation obtained.

All correlations between scores indicating level of abstraction are significant and positive (Table 9). Correlations at both age levels are highest between scores on verbal and picture test forms, lowest for those obtained between the object and all other test forms. This is true at both age levels.

Table 9. Correlations Between Scores on Test Forms for Percentage of Abstract Responses

df = 28 t = 2.05

Comparisons	Group 1 (Younger)	Group 2 (Older)
Verbal - Picture	.52	.60
Verbal - Object	.45	.40
Object - Picture	.45	.48

C. ORDER EFFECTS

From the raw data recorded in the Appendix means were computed on the basis of test order and are presented in Table 10.

Because of the experimental design, comparisons of test order were averaged over all tests and analyzed separately. Analyses of variance similar to those of Tables 2, 4, 5, and 7 were performed and appear in Appendix B (Tables IV, V, VI and VII). No simple order effects reached a defined level of significance. There was a significant

age order interaction noted for the total categories index (Appendix B, Table VI).

Table 10. Summary of Results on Order Effects. Mean Values of Scores Obtained by Two Age Groups on Three Consecutive Free-Sort Tests

	No. of defined categories	No. unsorted	Total categories	Per cent abstract responses
Group 1 (Younger)				
1st test	12.53	3.3	15.87	18.8
2nd test	12.67	2.2	15.83	23.7
3rd test	12.80	2.1	14.87	28.6
Grand Means	12.67	2.55	15.22	23.7
Group 2 (Older)				
1st test	10.77	2.5	13.30	44.6
2nd test	11.43	2.5	13.97	46.2
3rd test	11.37	2.7	14.03	39.1
Grand Means	11.19	2.58	13.54	40

CHAPTER VII

DISCUSSION OF RESULTS AND THEIR IMPLICATIONS

This experiment was designed primarily to consider the effect of age and test form on the kind and number of categories produced on a sorting task. It has been noted that both age and form of test significantly affected the number of defined categories sorted (Table 2) and that age, but not test form, affected the kind of labels to define the categories (Table 7). For the two ages sampled, however, there were no interactions between age and form of test (Tables 2 & 7). These same variables, age and test form, did not affect the number of unsorted categories (Table 4) and, when these unsorted categories were added to the number of defined categories to form the total categories index, the effect of age was no longer significant (Table 5).

A. RELATIONSHIP AMONG TOTAL CATEGORIES, DEFINED CATEGORIES AND UNSORTED OBJECTS

Analysis for the total categories index was included, it may be recalled, for comparison with defined categories as criterion of category width. It appears that age differences in categorizing breadth were masked when a total categories index was used (Table 5). This is interesting in view of the extensive use of the total categories index by Gardner and associates (Gardner, 1953; Gardner *et al.*, 1959; Gardner *et al.*, 1960; Gardner & Schoen, 1962). This group of workers has emphasized the consistency of styles in category width across test forms; it is possible that this observed self-consistency may be partially a

result of the use of the total categories index for measurement.

The fact that age differentially affected the number of defined categories of sorted objects and the number of objects unsorted supports the view that the two dimensions are, indeed, independent aspects of categorizing style (Messick & Kogan, 1963). There seems to be, therefore, little justification for using a total categories measure as it may well obscure actual differences in categorizing behaviour. In the following discussion of results only the defined categories criterion of category width will be used.

B. AGE EFFECTS

The broadly stated hypothesis (Hypothesis 1) that significant between-age differences would be found in the number of categories formed and in the level of abstraction used to label the category content was supported (Tables 2 & 7). However, while older subjects generated significantly more abstract responses than younger subjects (Tables 1 & 7), they did not sort more defined categories (Tables 1 & 2). This is of interest because developmental theory suggests that, although subjects become more differentiated with age, they are increasingly able to integrate or synthesize as well. This latter function is believed to be the more mature form of cognitive development (Werner, 1961). Since abstract definitions are considered to be more all-inclusive than non-abstract (Scheerer, 1954), and since older subjects gave a large percentage of abstract labels to their categories (Table 1), it appears that they may be synthesizing or integrating. However, this is a post hoc observation and further investigation would be required to confirm it.

Thus, taken as a whole, the effects of age appear to be quite consistent with developmental theory as formulated by Piaget (1953) and by Werner (1961). Further study at younger age levels and at intervening ages would be needed, however, before one could confidently state that growth of categorizing style follows predicted patterns of developmental theory. Additionally, longitudinal rather than cross-sectional methods of research might yield more accurate and useful data regarding the development of individual differences in categorizing style, though both are warranted. Further testing of the younger group as they mature could provide some of the additional information necessary for confirming or rejecting this notion.

C. TEST FORM EFFECTS

Significant test form differences in breadth of categorizing were found at both age levels (Tables 2 & 3). In the younger group significantly fewer categories were sorted on the object form of the test than on the picture form, but there were no significant differences between object and verbal forms. Similarly, there were no significant differences between scores on the verbal and picture tests. In short, for the younger group, only the object and picture test forms were significantly different from each other.

In the older group, the object form of the test produced significantly different results from both the verbal and the picture forms. Results on the picture test were not different from those on the verbal, however. Thus, effects of varying test forms appear to be more pronounced at the older than at the younger age level. Also, the range in the number of categories defined from object to picture test was 1.23

for the older group, compared with a range of 0.70 for the younger. This evidence is in partial support of the hypothesis from developmental theory (Hypothesis 2) which predicted that older subjects would respond more specifically to each test and that their sortings would be more reflective of the nature of the test than would the sortings of the younger group.

The results pertaining to the direction of the effects of test form on category width are a little difficult to interpret. Simply not enough is known about the nature of test form differences; one cannot state with any degree of certainty what information is transmitted to an individual by a particular item in any particular form. Hypotheses were based on the assumption that the object test supplied the most information about an item, the picture somewhat less, and the verbal least. Accordingly, it was predicted that the direction of the effects of differences in test form would also be in that order. It appears that this assumption may have been incorrect for the tests differentially affected category width in the following order: object, verbal, picture. Furthermore, it was the same for both ages (Table 1). It may be that category width is not a function of the amount of information available because subjects actually sorted fewest categories on the object form of the test (Table 3), the test form which was believed to give the most information about the items. There is, then, no evidence to support Hypothesis 3.

There is, however, some support for the hypothesis derived from the view of categorization as a decision process with risk-taking implications (Hypothesis 5). It states that subjects should sort fewest categories in the test form providing the most information about the items.

This proved to be true. However, they did not sort the fewest categories on the verbal form of the test, the form viewed as giving the least information about the item. Assuming that the amount of information supplied is a critical factor, and that the more information the test supplies the more generous one can be in category boundaries without incurring the risk of increased error, it would appear that the subjects got the most information from the object form of the test and the least from the picture form. If this were the case, the imagination may have supplied the information for the verbal form of the test. Tending to confirm this idea was the fact that many subjects, during their sortings, indicated that they liked the freedom of the verbal form of the test. It allowed them to supply their own information about the properties of the objects.

Although it is inviting to think of the results in this way, one cannot ignore the possibility that the order in which the various test forms were presented may have had some effect on the interpretation of the items. This could be particularly true in the case of the verbal test. Because order of test form presentation was controlled, not all subjects sorted the verbal form of the items first and so their visualization of these items may have been subject to prior influences by the object and picture forms. Further work with changes in the experimental design would be needed to determine the validity of this conjecture.

Then again, it is possible that subjects sorted the greatest number of groups on the picture test because they found it confusing and, therefore, the available information was, for them, indiscernible. It is interesting to note that, when subjects were asked which test they preferred, none of the older group chose the picture form and only two

of the younger group did so. Spontaneous comments during sorting indicated that at least some of the subjects found the pictures "difficult" to categorize. Because all pictures were taken at the same distance from the camera, the objects in them were not to scale. For example, the pencil appeared as large as the tree. This may have caused confusion and affected sortings. If line drawings had been used instead of photographs, or if different photographic techniques had been employed, results might have been different. Given the additional fact of the non-significant correlation between pictures and objects for the older group (Table 8), which is in contrast to previously cited research and theory, one is led to suspect that something about the picture test, itself, operated to produce the unpredicted results. Whether this was peculiar to our particular test or is generally true of photographic and other types of pictorial tests should be determined. It could have important educational connotations in view of the use made of pictures in textbooks for reading and other purposes.

There is yet another way of looking at the differences in object, picture and verbal forms of the same test. As well as differing in the amount of information they supply about an item, they can also be considered to vary along the concrete-abstract dimension. Although no hypotheses were made, it was considered possible that the level of response might reflect these differences in test form. For example, subjects might produce more abstract responses on the abstract (verbal) form of the test and more concrete responses on the concrete (object) form of the test. Results indicated that level of response was not affected by variations in test form at either age level (Table 7). A different experimental design with variations in time between testing periods might

indicate whether this was a result of a consistent style or whether it was an effect of learning transfer.

D. INDIVIDUAL CONSISTENCY

All inter-test correlations for category width values were significant with the exception of the object-picture comparison for Group 2 (Table 8). In addition, level of response scores correlated significantly between test forms (Table 9). This indicates a certain degree of self-consistency and thereby lends support to style theory. Other studies, previously discussed, have indicated that individuals exhibit stylistic tendencies in category widths. The present study, while confirming this, supports the notion that there may be stylistic components in level of response as well.

In general, correlations between test forms were higher for the younger group than for the older group (Table 8) and so predictions based on style theory broke down. It was hypothesized that the older group would be more consistent, that is, have higher inter-test correlations than the younger group (Hypothesis 4). This was not so (Table 8). On the other hand, in the light of developmental theory, it was expected that increasing age would bring increasing specificity thus producing lower correlations or greater variations among an individual's test scores (Hypothesis 2). Since the younger group did respond more globally to the three test forms, Hypothesis 2 is supported.

Correlations for level of response scores were generally lower than those for category width values, although they were still significant (Table 9). There did not appear to be any age differences in so far as self-consistency of response is concerned.

It is of further interest to note that highest inter-test correlations for the younger group on the category width dimension were obtained between the object and the picture forms of the test (.83). Yet, for the older group these tests produced a non-significant correlation (.32). The highest inter-test correlation for the older group was between the object and verbal forms (.62).

It is tempting to speculate about the possible meanings of these results. One might have expected all subjects to behave most consistently on the object-picture test forms as, intuitively, these tests appear to be most alike. Perhaps the older group, having greater experience with verbal materials, was more at ease with the verbal form than the younger group. Perhaps, too, the older group may have seen ambiguities in the picture test that were not apparent to the younger group. These speculations would have to be studied in much more detail before they could be confirmed or rejected.

E. AGE AND TEST FORM INTERACTION EFFECTS

It is of interest to note that for both age groups fewest groups were formed on the object test, more were produced on the verbal form, and most groups were sorted on the picture form (Table 3). Thus, no test-age interaction occurred (Table 2). There were no other recorded interactions between test and age among the other measured aspects of categorizing behaviour. However, an inspection of the correlations obtained between tests for category width values at both ages showed an interaction trend (Table 8). It appeared that younger subjects showed greatest consistency between the picture and object test forms, the forms in which the older subjects were least consistent. This

effect was not found, however, when the correlations of per cent abstract responses for each test form were compared (Table 9). Both age groups showed greatest consistency of response between verbal and picture forms. Thus, there is only scant evidence in this investigation to support the notion that varying the tests may have had differential effects at varying ages. However, it must be remembered that according to Piaget's theories, the two age groups studied were both in the final stage of cognitive development, and so, perhaps, few interactions could be expected. Further testing at widely varying ages would be needed before the supposition should be finally discarded.

F. ORDER EFFECTS

There was no direct, statistically measurable effect of test order on any aspect of categorization. However, there was a trend at both age levels to sort into more defined categories with each successive test (Table 10). This could indicate a learned increase in specificity. Also, the number of total categories tended to decrease with subsequent tests in the younger group but increase in the older group. This, in fact, was indicated by a significant age-order interaction for total categories (Appendix B, Table VI). Thus, although test order effects were not pronounced, and no inferences are drawn, the importance of repetitive actions on test effects for educational practice is of such consequence that further work seems justified in this area.

G. SUMMARY OF RESULTS IN TERMS OF HYPOTHESES

In review, if one turns to the hypotheses based on developmental theory, it was found:

1) Significant age differences were indicated both in number of categories formed and in level of response on all forms of the sorting test, as predicted.

a) Older subjects did not sort more categories on tests than younger subjects. This hypothesis is, therefore, rejected.

b) In terms of the abstraction index, it was predicted that older subjects would form more abstract categories than younger subjects. This proved to be so.

2) Regarding the number of defined categories produced, the younger children behaved more globally across test forms than did the older subjects. This was indicated by fewer significant differences between mean scores on varying tests and higher inter-test correlations at the younger age level. There were no discernible differences between the two age groups in consistency for level of response. Generally speaking, then, older subjects could be considered to be less consistent than younger subjects and so this hypothesis is largely supported.

3) It was predicted that the number of categories formed would be higher for the object than for the pictorial and for the pictorial than for the verbal forms of the test. This was not so. The object form gave fewest categories while the pictorial form produced the greatest number. Thus, increased information appeared to result in broader categories.

For the hypothesis based on style theory it was found:

4) All inter-test correlations were significant excepting that obtained for category width values between the object and picture forms in the older group. However, by inspection it could be seen that inter-test correlations were not higher for older individuals than for the

younger. Thus, although the first part of the hypothesis which predicted self-consistency for all subjects is supported, the second part predicting increasing consistency with increasing age is not.

For the hypothesis from information theory which views categorization as a decision process, it was found:

5) As predicted, subjects sorted fewer categories for the object test as compared to the other test forms. This difference was significant in each instance in the case of the older group, and for the object as compared to the picture form for the younger group, although subjects did not sort the greatest number of categories on the verbal test as hypothesized. There was no significant difference between verbal and picture tests in number of categories defined at either age level. Thus, this hypothesis is generally supported.

As a final comment, it should be noted that the hypotheses from developmental theory regarding age effects and from information theory regarding test form effects were largely supported. The hypothesis from developmental theory regarding test form effects was rejected. It was not possible to either retain or discard the hypothesis from style theory.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

This study was designed primarily to determine the effects of age and test form on certain stylistic aspects of categorizing behaviour. The subjects were composed of two groups of girls aged 12 and 18 respectively. These two groups were crudely matched for general ability and inclusion status. Three test forms of the Clayton & Jackson free sort test were administered to each girl at each age level. These tests were the original verbal form, slightly modified (by placing the name of each item on a separate card to allow for a sorting procedure), an object form, composed of models or replicas of the items, and a picture form, composed of photographs of these items.

Under free and relaxed conditions subjects were asked to sort these items into "most natural, most comfortable and most logical" categories. Each subject sorted all three sets of test material in a prearranged sequence a week apart. Test order effects were reduced to a minimum by having five subjects at each age sort at each of the six possible order combinations.

Number of defined categories and percentage of abstract labels were recorded for each subject and each test. Analyses of variance were carried out for age, test form and order effects. Correlations for defined categories and per cent abstract responses were computed at each age level for all possible test form combinations. Number of total categories and number of unsorted items were also scored and

analyzed for exploratory purposes.

In reviewing the outcome of this investigation as a whole, several salient points should be emphasized. First of all, results indicated the existence of significant age differences in the measured aspects of categorizing behaviour, category width and level of response, with older subjects sorting fewer and more abstract categories. Thus, there were positive indications that developmental theory may prove most useful in predicting the growth patterns of personal styles in such behaviour. There was some evidence also to suggest that test forms differentially affected category widths, especially at the older age level. The direction of the test form effects on category width levels appeared to be similar regardless of age but, for the age levels studied, there might have been a test-age interaction effect in terms of individual consistency. Other trends at other ages can only be hypothesized.

As with many initial studies, this experiment gave rise to many new questions and uncovered many unforeseen problems. The picture test did not produce expected results in several instances and thus, in future, closer attention may need to be directed to this test form. Since our version was obtained by photography from a standard distance, it did not reproduce the objects to scale size. This may have been a factor contributing to the interesting findings. On the other hand, such results may have been a function of photograph tests in general. Since some investigators, in the past, have elected to use line drawings rather than photographs in pictorial test forms, it would be interesting to utilize line drawings as a fourth test form in subsequent studies. A comparison of the two pictorial tests might prove

valuable.

In the results obtained one cannot be sure that transfer effects from one form of the test to another may not have produced, or helped to produce, some of the negative results: For example, one notes the absence of significant test form effects on level of response. Since each subject did not do each test in every possible order, and no subject repeated any one test form, there could be no meaningful test form-test order analysis. Thus, the possibility of an unrecognized test form-test order interaction exists. Consequently, further studies must try to separate this effect from the development of a categorizing style, if indeed it can be separated. At any rate, the interval between testings must be varied in some systematic way.

Thus, while much more work is needed to plot the developmental sequence of stylistic tendencies in categorizing behaviour, the results of this investigation suggest that it would be a worthwhile endeavour. To accomplish this, however, there is still the need for adequate, psychometrically based tests for measuring the various dimensions of categorizing behaviour at varying ages. All this may take considerable time, but results should prove to be of value in the further understanding of individual differences in cognitive functioning.

BIBLIOGRAPHY

Anderson, R. C. and Ausubel, D. P. Readings in the psychology of cognition. New York: Holt, Rinehart and Winston, 1965.

Berkowitz, L. The judgmental process in personality functioning. Psychol. Rev., 1960, 67, 130-142.

Bruner, J. S., Goodnow, Jacqueline J., and Austin, G. A. A study of thinking. New York: Wiley, 1956.

Bruner, J. S., Postman, L., and Rodrigues, J. Expectation and the perception of color. Amer. J. Psychol., 1951, 64, 216-227.

Bruner, J. S. and Rodrigues, J. Some determinants of apparent size. J. abnorm. soc. Psychol., 1953, 48, 17-24.

Bruner, J. S. and Tajfel, H. Cognitive risk and environmental change. J. abnorm. soc. Psychol., 1961, 62, 231-241.

Bruner, J. S. and Tajfel, H. Width of category and concept differentiation: A note on some comments by Gardner and Schoen. J. pers. soc. Psychol., 1965a, 2, 261-264.

Bruner, J. S. and Tajfel, H. A rejoinder. J. pers. soc. Psychol., 1965b, 2, 267-268.

Clayton, Martha B. and Jackson, D. N. Equivalence range, acquiescence and overgeneralization. Educ. psychol. Measmt., 1961, 21, 371-382.

Dattman, P. E. and Israel, H. E. The order of dominance among conceptual capacities: An experimental test of Heidbreder's hypothesis. J. Psychol., 1951, 31, 147-160.

Fenichel, O. The psychoanalytic theory of neurosis. New York: Norton, 1945.

Fillenbaum, S. Some stylistic aspects of categorizing behaviour. J. Pers., 1959, 27, 187-195.

Flavell, J. J. The developmental psychology of Jean Piaget. New York: Van Nostrand, 1963.

Forgus, R. H. Perception, the basic process in development. New York: McGraw-Hill, 1966.

Forgus, R. H. and Fowler, H. The order of dominance in concept attainment as affected by experience. J. Psychol., 1957, 44, 105-108.

Freeman, F. S. Theory and practice of psychological testing. New York: Holt, 1951.

Freud, Anna. The ego and the mechanisms of defense. (orig. published London: Hogarth Press, 1937) New York: International Universities, 1946.

Freud, S. (1925). The ego and the id. In R. M. Hutchins (Ed.) The Major Works of Sigmund Freud. Chicago: Encyclopaedia Britannica, Inc., 1952, 697-717.

Gardner, R. W. Cognitive styles and categorizing behaviour. J. Pers., 1953, 22, 214-233.

_____. The development of cognitive structure. In Constance Scheerer (Ed.), Cognition: theory, research, promise. New York: Harper and Row, 1964, pp. 147-171.

Gardner, R. W. and Long, R. I. The stability of cognitive controls. J. abnorm. soc. Psychol., 1960, 1961, 485-487.

Gardner, R. W. and Schoen, R. A. Differentiation and abstraction in concept formation. Psychol. Monogr., 1962, 76, No. 41 (Whole no. 560).

_____. Reply to a note by Bruner and Tajfel. J. pers. soc. Psychol., 1965, 2, 264-267.

Garrett, H. A. The developmental theory of intelligence. Amer. Psychologist, 1946, 1, 372-378.

Glixman, A. F. Categorizing behavior as a function of meaning domain. J. pers. soc. Psychol., 1965, 2, 370-377.

Goldman, A. E. and Levine, M. A developmental study of object sorting. Child Develpm., 1963, 34, 649-666.

Goldstein, K. and Gelb, A. Zur Psychologie des optischen Wahrnehmunger und Erkennungs-vorange. Z. Ges. Neurol. Psychiat., 1918, 41. As summarized in Forgus, R. H., Perception the basic process in development. New York: McGraw-Hill, 1966.

Goldstein, K. and Scheerer, M. Abstract and concrete behavior: An experimental study with special tests. Psychol. Monogr., 1941, 53, No. 2 (Whole No. 239).

Grant, D. A. Perceptual vs. analytical responses in the number concept of a Weigl-type card sorting test. J. exp. Psychol., 1951, 41, 23-29.

Grant, D. A. and Curran, J. F. Relative difficulty of number, form and color concepts of a Weigl-type problem using unsystematic number cards. J. exp. Psychol., 1953, 43, 408-413.

Guilford, J. P. The structure of intellect. Psychol. Bull., 1956, 33, 267-293.

Hanfmann, E. A study of personal patterns in an intellectual performance. Char. and Pers., 1941, 9, 315-325.

Hanfmann, E. and Kasanin, J. A method for the study of concept formation. J. Psychol., 1937, 3, 521-540.

Harvey, O. J., Hunt, D. E. and Schroeder, H. M. Conceptual systems and personality organization. New York: John Wiley & Sons, 1961.

Heidbreder, Edna. The attainment of concepts. I. Methodology and terminology. J. genet. Psychol., 1946, 35, 173-189.

_____. Conceptual achievements during card sorting. J. Psychol., 1949, 27, 263-309.

Heidbreder, Edna and Overstreet, P. V. Critical features and contexts. J. Psychol., 1948, 25, 299-239.

Hull, C. L. Quantitative aspects of the evolution of concepts. Psychol. Monogr., 1920 (No. 123).

Inhelder, Barbel and Piaget, J. The growth of logical thinking from childhood to adolescence. New York: Basic Books, 1958.

Kagan, J., Moss, H. A. and Sigel, I. E. Conceptual style and the use of affect labels. Merrill-Palmer Quart., 1960, 6, 261-268.

_____. Psychological significance of styles of conceptualization. In J. C. Wright & J. Kagan (Eds.) Basic cognitive processes in children. Monogr. Soc. Res. Child Develpm., 1963, 28, No. 2 (Serial No. 86), pp. 73-112.

Kennedy, K. and Kates, S. L. Conceptual sorting and personality adjustment in children. J. abnorm. soc. Psychol., 1964, 68, 211-214.

Klein, G. S. A clinical perspective for personality search. J. abnorm. soc. Psychol., 1949, 44, 42-49.

_____. Need and regulation. In M. R. Jones (Ed.) Nebraska symposium on motivation. Lincoln, Neb.: Univ. of Nebraska Press, 1954, pp. 224-274.

_____. Cognitive control and motivation. In G. Lindzey (Ed.) Assessment of human motives. New York: Rinehart, 1958, pp. 87-115.

Klein, G. S., Barr, H. L. and Wolitsky, D. L. Personality. In P. R. Farnsworth, O. McNemar, and Q. McNemar (Eds.) Annual Rev. of Psychol. Palo Alto, Calif.: Annual Reviews, Inc., 1967, pp. 467-560.

Kogan, N. and Walloch, M. A. Risk-taking, a study in cognition and personality. New York: Holt, Rinehart and Winston, 1964.

LeClerg, E. L., Leonard, W. H. and Clark, A. G. Field plot technique. Minneapolis: Burgess Publ. Co., 1962.

Lundsteen, S. W. and Michael, W. B. Validation of three tests of cognitive style in verbalization for third and sixth grades. Educ. and psychol. Measmt., 26, 449-461.

Messick, S. and Damarin, F. Cognitive styles and memory for faces. J. abnorm. soc. Psychol., 1964, 69, 313-318.

Messick, S. and Fritsky, N. Dimensions of analytic attitude. J. Pers., 1963, 31, 346-370.

Messick, S. and Kogan, N. Differentiation and compartmentalization in object sorting measures of categorizing style. Percept. mot. Skills., 1963, 16, 47-51.

_____. Category width and quantitative attitude. Percept. mot. Skills, 1965, 20, 493-497.

Piaget, J. The origin of intelligence in the child. London: Routledge and Kegan Paul, 1953.

Pettigrew, T. F. The measurement and correlates of category width as a cognitive variable. J. Pers., 1958, 26, 532-544.

Rapaport, D. (Ed.) Organization and pathology of thought. New York: Columbia U. Press, 1951.

Rapaport, D. Cognitive structures. In Contemporary approaches to cognition. Cambridge, Mass.: Harvard Univ. Press, 1957, pp. 157-200.

Rapaport, D., Gill, M. and Schafer, R. Diagnostic psychological testing. Chicago: The Year Book Publishers, Inc., 1946, Vol. I & II.

Reed, H. B. IV. The influence of the complexity of the stimuli. J. exp. Psychol., 1946, 36, 252-261.

Reichard, S., Schneider, M. and Rapaport, D. The development of concept formation in children. Amer. J. Orthopsychiat., 1944, 14, 156-161.

Rogers, C. R. A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.) Psychology: A study of a science. Vol. 3, Formulations of the person and the social context. New York: McGraw-Hill, 1959, pp. 184-256.

_____. The actualizing tendency in relation to "motives" and to consciousness. In M. R. Jones (Ed.) Nebraska symposium on motivation. Lincoln, Nebraska: Univ. of Nebraska Press, 1963.

Santostephano, S. and Paley, Evelyn. Development of cognitive controls in children. Child Develpm., 1964, 35, 939-949.

Scheerer, M. Personality functioning and cognitive psychology. J. Pers., 1953, 22, 1-16.

_____. Cognitive theory. In Gardner Lindzey (Ed.) Handbook of social psychology. Cambridge, Mass.: Addison-Wesley, 1954, ch. 3.

Scheerer, Constance (Ed.). Cognition: Theory, research, promise. New York: Harper & Row, 1964.

Sherif, M. and Hovland, C. I. Judgmental phenomena and scales of attitude measurements: Placement of items with individual choice of number of categories. J. abnorm. soc. Psychol., 1953, 48, 135-141.

Sloane, H. N. A construct validation of equivalence range. Unpublished doctoral dissertation, Pennsylvania State U., 1959.

Sloane, H. N., Gorlow, L. and Jackson, D. N. Cognitive styles in equivalence range. Percept. mot. Skills, 1963, 16, 389-404.

Smith, M. B., Bruner, J. S. and White, R. W. Opinions and personality. New York: Science Editions, John Wiley & Sons, 1964.

Smoke, K. L. An objective study of concept formation. Psychol. Monogr., 1932, No. 191.

Tajfel, H., Richardson, A. and Everstine, L. Individual consistencies in categorizing: A study of judgmental behaviour. J. Pers., 1964a, 32, 90-108.

_____. Individual consistencies in conditions of risk-taking. J. Pers., 1964b, 32, 550-565.

Thurstone, L. L. Psychological implications of factor analysis. Amer. Psychologist, 1948, 3, 402-408.

Tobin, S. A. Categorization styles as a function of category width and level of abstraction. Unpublished doctoral dissertation, Univ. of Michigan, 1963.

Van de Geer, J. P. and Jaspars, J. F. M. Cognitive functions. In P. Farnsworth, O. McNemar and Q. McNemar (Eds.) Annual Rev. of Psychol. Palo Alto, Calif.: Annual Reviews, Inc., 1966.

Vinacke, W. E. The psychology of thinking. New York: McGraw-Hill, 1952.

Vygotsky, L. S. Thought and language. New York: Wiley, 1962.

Wallach, M. A. Research on children's thinking. In H. W. Stevenson (ed.) Child psychology: The sixty-second yearbook of the NSSE. Chicago: Univ. Chicago Press, 1963, pp. 236-276.

Wallach, M. A. and Caron, A. J. Attribute criteriality and sex-linked conservatism as determinants of psychological similarity. J. abnorm. soc. Psychol., 1959, 59, 43-50.

Wallach, M. A. and Kogan, N. Modes of thinking in young children. New York: Holt, Rinehart, Winston, 1965.

Weigl, E. On the psychology of so-called processes of abstraction (trans. by M. J. Rioch; orig. publ. 1927). J. abnorm. soc. Psychol., 1941, 36, 3-33.

Werner, H. Comparative psychology of mental development (orig. publ. Chicago Follett, 1948). New York: Science Editions, Inc. John Wiley & Sons, 1961.

Weschler, D. Cognitive, conative and non-intellective intelligence. Amer. Psychologist, 1950, 5, 78-83.

Witkin, H. A. Origins of cognitive style. In Constance Scheerer (Ed.) Cognition: Theory, research, promise. New York: Harper & Row, 1964, pp. 172-205.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R. and Karp, S. A. Psychological differentiation. New York: John Wiley & Sons, 1962.

Witkin, H. A., Lewis, Helen B., Hertzman, M., Machouer, Karen, Meissner, Pearl B., and Wapner, S. Personality through perception. New York: Harper, 1954.

Zaslow, R. W. A study of concept formation in normals, mental defectives and brain damaged adults. Genet. Psychol. Monogr., 1961, 63, 279-338.

APPENDIX A

THE TEST: DIRECTIONS FOR ADMINISTERING AND SCORING

The Clayton and Jackson Sorting Test

lamp post	hat	towel
television set	golf club	a watch
rock	picture	tree
hammer	wallet	postage stamp
ruler	spoon	chair
gloves	telephone	band aid
sweater	pillow	milk bottle cap
aspirin	rug	window shade
canoe	screw driver	a bolt
coat	measuring cup	shoe
pencil	cigarette	daffodil
book-end	radio	a baseball
lamp	a match	handkerchief
rope	dictionary	comb
a tire	a pot	toboggan
pistol	nail scissors	dime
arrow	shovel	

Instructions For Giving the Tests

I. Instructions for the first administration of the test to each individual.

1. First test, verbal form.

"First of all, I want you to know that there is no one right way of doing what I am going to ask you to do. Everyone does it in his own way. You should do it in the way that seems most natural, most logical and most comfortable to you. HERE ARE THE NAMES OF SOME FAMILIAR OBJECTS. All you are asked to do is to put together into groups the objects which seem to you to belong together. Later, when you are finished I will ask you the reason you put the objects together the way you did.

You may use as many or as few groups as you like, - just so long as the objects in each group belong together for a particular reason. If, after you have thought about all the objects, a few of them don't seem to belong with any of the others, you may put each of those objects into groups by themselves.

Please sort all the items. If you have any questions, please ask me."

2. First test, object form.

For, "Here are the names of some familiar objects," substitute: "HERE ARE SOME OBJECTS WITH WHICH YOU ARE FAMILIAR."

3. First test, picture form.

For, "Here are the names of some familiar objects," substitute: "HERE ARE SOME PICTURES OF FAMILIAR OBJECTS. FOR EXAMPLE, THIS IS A LAMP-POST, A BOOK-END, A ROCK, A WINDOW-SHADE."

II. Modification for the second and third administrations of the test for each individual.

For "First of all, I want you to know that there is no one right way of doing what I am going to ask you to do. Everyone does it in his own way. You should do it in the way that seems most natural, most logical and most comfortable to you," substitute: "FIRST OF ALL, PLEASE UNDERSTAND THAT THIS STUDY IS NOT CONCERNED WITH ABILITY TO REMEMBER. SINCE THERE IS NO ONE WAY TO DO THIS TASK, YOU SHOULD DO IT IN THE WAY THAT IS MOST LOGICAL, MOST NATURAL AND MOST COMFORTABLE TO YOU TODAY".

Instructions For Scoring Category Labels

The tripartite scheme of Rapaport (Rapaport, Gill & Schafer, 1946) was used in scoring every subject's preferred level of abstraction for each category and so each label was given a concrete, functional or abstract classification. However, the index for analysis was the percentage of abstract responses used relative to the total number of responses as per Gardner & Schoen (1962). Many of the examples and definitions given in this section are taken from Tobin's directions for scoring which were adapted from Rapaport's system (Tobin, 1963). The general criteria used was based on internal versus external properties. Generally, the more a category describes the external properties of things as revealed by the five senses, the closer it is to developmentally early, concrete thought. The more it refers to inherent, essential properties of objects, then the more abstract it is considered to be. Unlike Rapaport, who appeared to be somewhat inconsistent by classifying "made of" categories as abstract, Tobin scored such categories as concrete. This seemed more logical in view of the defining criterion (internal vs. external properties) and was adopted in this study. Thus, "made of paper", "made of plastic", etc. were considered to be concrete and non-abstract. The reader is referred to Tobin's dissertation (Tobin, 1963) and to Diagnostic Psychological Testing (Rapaport, Gill & Schafer, 1946) for a more comprehensive treatment of scoring procedures. A brief summary follows.

I. General Scoring Directions

A. The language used.

All labels are scored according to the language used to describe

them. Sometimes subjects may form groups that could be abstract concepts, but the label they use to define them is phrased in a concrete or functional way. For example, the same group of objects could be described:

a) found in a sports store (concrete)

b) used in sports (functional)

c) sports equipment (abstract)

Rule: Always score according to language used.

B. Compound Categories.

1. Sometimes a single group of objects is given more than one concept by the subject:

a) tools used to repair things (abstract and functional)

b) used for cooking in a kitchen (functional and concrete)

Rule: Score <u>both</u> concepts according to language and count each concept as a label when computing the total number of labels.

2. Sometimes more than one group of objects is given a single rubric by the subject:

a) children's clothes and toys

b) tools and sports equipment

Rule: Since two categories are actually involved, divide the grouping into the two and score both as concrete since this is a personalized affective response.

II. Specific Scoring Instructions

A. Concrete categories.

The following criteria are used to define a concrete category:

1. Affective responses

Examples:- I like them; dangerous; bad for people.

Note: not good for building things. This is a functional response.

2. Parts of objects

Examples: They all have handles; they all have legs.

3. Belonging

Examples: Owned by athletes; belong to a school girl.

4. Location

Examples: Found in the kitchen; found in a woman's purse.

Note: used in the bathroom is partially concrete so is scored as such and is also scored as functional.

5. Origins

Examples: made by J. C. Higgins; come from the ground; come from trees; get them from a store.

6. An object used to do something to or with another

Examples: light a cigarette with a match; tie the rope to the rock; take the tire off with the hammer.

7. Properties of words

Examples: all begin with "d"; all have four letters.

8. Size and shape

Examples: all large; all round; all square corners; all small and round.

9. Symbolic meanings

In this grouping the subject re-interprets items.

Examples: screw-driver, pistol, arrow - "for screwing"
lamp and lamp-post "Mutt and Jeff"

10. Fabulated

In this type of response one attribute of an object is taken as the starting point from which a story unfolds and includes other objects in the group.

Example: (rock, rope, hammer) This rock is a mountain and I am going mountain climbing with the rope. I'll need a hammer too.

11. Syncretistic

These responses are vague and overgeneralized. They could include many other objects not in the given sorting.

Examples: all manufactured; all exist.

12. References to "made of".

All references to what an item is made of are included in this classification, even relatively sophisticated responses such as "made of fabric".

B. Functional Categories.

1. Uses to which things can be put

Examples: used for grooming; used for recreation.

2. Things that objects do

Examples: they grow; they run; they tell time.
(radio and t.v.)- get news from them
(hammer and rock)- hit people with them

3. Responses beginning with "for"

Examples: for cooking; for sports; for bathing.

C. Abstract Categories.

1. Complex class names

Examples: tools; sports equipment; bathroom

items; kitchen utensils; wearing apparel.

2. One-word responses with abstract conceptual connotations.

   Examples: sports; nature; furniture.

   Note: One word categories with concrete and functional connotations are scored accordingly, e.g. "used" - functional; "house" - concrete.

3. All responses beginning with "associated with", "pertaining to", or "have to do with", provided the second part is abstract.

   Examples: associated with sports; have to do with transportation; pertaining to communications

   Note: pertaining to a tool shop, however, is concrete associated with building things, - functional.

4. Miscellaneous

   "man-made" - is abstract because made is used in the abstract sense.

APPENDIX B

RAW DATA

TABLES I - IX

Table I. Raw Scores, All Measures, By Form of Test: Group 1

Subject	Number Defined Categories			Number Unsorted			Total Categories			Per cent Abstract Responses		
	Verb.	Pic.	Obj.	Verb.	Pic.	Obj.	Verb.	Pic.	Obj.	Verb.	Pic.	Obj.
A 1	10	12	12	2	3	3	12	15	15	0	15	21
2	10	10	10	7	4	0	17	14	10	0	20	10
3	12	13	13	10	2	2	22	15	15	15	19	7
4	14	13	11	1	0	2	15	13	13	43	38	82
5	13	13	12	8	9	2	21	22	14	0	8	33
B 1	14	13	12	0	4	4	14	17	16	21	8	17
2	12	12	12	0	1	1	12	13	13	17	31	25
3	11	11	8	4	2	2	15	13	10	27	9	13
4	17	18	18	0	0	0	17	18	18	47	50	42
5	14	18	15	1	2	1	15	20	16	50	29	38
C 1	13	17	12	0	2	3	13	19	15	14	35	17
2	18	18	18	6	6	6	24	24	24	47	52	30
3	13	13	13	2	2	2	15	15	15	61	92	50
4	13	13	11	2	0	0	15	13	11	21	38	18
5	10	10	12	0	4	6	10	14	18	40	10	13
D 1	7	9	7	0	1	1	7	10	8	25	20	29
2	11	11	12	1	0	0	12	11	12	17	0	13
3	15	13	14	1	3	1	16	16	15	6	23	14
4	15	12	12	3	14	7	18	26	19	13	8	25
5	13	15	14	2	1	2	15	16	16	50	22	33
E 1	12	13	13	1	1	0	13	14	13	42	61	31
2	13	16	15	0	0	1	13	16	16	8	6	6
3	12	10	10	1	0	1	13	10	11	25	10	20
4	7	12	12	1	0	0	8	12	12	13	25	0
5	14	12	11	2	6	2	16	18	13	27	27	25
M 1	18	16	17	4	5	2	22	21	19	26	24	18
2	13	12	9	5	9	12	18	21	21	15	23	33
3	12	13	12	3	3	4	15	16	16	23	31	50
4	12	12	12	0	0	2	12	12	14	25	50	25
5	11	11	11	5	3	2	16	14	13	18	27	42

Table II. Raw Scores, All Measures, By Form of Test: Group 2

Subject	Number Defined Categories			Number Unsorted			Total Categories			Per cent Abstract Responses		
	Verb.	Pic.	Obj.	Verb.	Pic.	Obj.	Verb	Pic.	Obj.	Verb	Pic.	Obj.
F 1	9	9	8	3	2	1	12	11	9	27	10	0
2	10	11	10	4	2	4	14	13	14	28	55	27
3	12	16	13	1	3	3	13	19	16	40	52	53
4	12	13	14	6	8	8	18	21	22	69	69	86
5	10	12	11	1	1	1	11	13	12	36	42	46
G 1	14	14	12	3	1	2	17	15	14	47	19	18
2	10	13	12	5	6	5	15	19	17	46	43	50
3	15	9	14	1	0	1	16	9	15	60	56	71
4	16	14	13	2	9	7	18	23	20	26	25	33
5	10	8	9	1	0	0	11	8	9	39	33	30
H 1	9	11	10	2	0	0	9	11	12	67	36	54
2	12	11	11	0	1	0	12	12	11	33	50	42
3	12	13	12	0	0	4	12	13	16	67	54	75
4	9	9	9	1	1	2	10	10	11	88	89	89
5	12	13	11	6	5	6	18	18	17	58	54	45
I 1	13	9	9	1	1	1	14	10	10	0	33	78
2	10	13	11	1	2	2	11	15	13	64	57	64
3	14	14	14	4	0	3	18	14	17	70	64	64
4	10	12	10	1	1	1	11	13	11	10	15	30
5	14	10	12	2	2	1	16	12	13	22	10	9
J 1	13	15	10	2	4	4	15	19	14	15	13	42
2	10	12	7	3	2	2	13	14	9	37	50	12
3	13	12	9	2	1	1	15	13	10	69	92	14
4	11	16	6	2	3	2	13	19	8	50	41	50
5	6	6	6	4	3	5	10	9	11	84	33	67
K 1	9	9	9	1	1	0	10	10	9	22	22	22
2	12	12	10	8	3	6	20	15	16	54	46	27
3	10	10	10	1	0	0	11	10	10	40	25	40
4	12	11	13	3	8	8	15	19	21	0	33	22
5	13	14	9	4	3	4	17	17	13	43	50	27

Table III. Test Sequence For Each Treatment Series

Series		Test Order		
Group 1	Group 2	1	2	3
A	F	verbal	picture	object
B	H	picture	object	verbal
C	I	object	verbal	picture
D	K	picture	verbal	object
E	G	verbal	object	picture
M	J	object	picture	verbal

Table IV. Analysis of Variance For Defined Categories, Groups 1 and 2, Age and Order Effects

Source of Variance	df	ss	ms	F
Total	179	1124.06		
Age	1	98.27	98.27	7.67**
Subjects within age	58	743.79	12.82	
Order	2	6.96	3.48	1.47 N.S.
Order x age	2	2.20	1.10	.46 N.S.
Residual	115	272.82	2.37	

** - significant at 0.01 level
N.S. - not significant

Table V. Analysis of Variance For Unsorted Items, Groups 1 and 2, Age and Order Effects

Source of Variance	df	ss	ms	F
Total	179	1176.2		
Age	1	0.02	.02	.0014 N.S.
Subjects within age	58	816.18	14.07	
Test order	2	12.10	6.05	2.12 N.S.
Test order x age	2	15.75	7.88	2.76 N.S.
Residual	116	332.15	2.86	

N.S. - not significant

Table VI. Analysis of Variance for Total Categories, Groups 1 and 2, Age and Order Effects

Source of Variance	df	ss	ms	F
Total	179	2504.99		
Age	1	95.32	95.32	3.03 N.S.
Subjects within age	58	1821.66	31.41	
Order	2	.84	.42	.11 N.S.
Order x age	2	123.96	61.98	15.36**
Residual	116	463.21	3.99	

** - significant at 0.01 level
N.S. - not significant

Table VII. Analysis of Variance for Per cent Abstract Responses
Groups 1 and 2, Age and Order Effects

Source of Variance	df	ss	ms	F
Total	179	40582.35		
Age	1	5633.6	5633.6	18.45**
Subjects within age	58	17706.22	305.27	
Test order	2	264.31	132.16	.93 N.S.
Test order x age	2	524.50	262.25	1.85 N.S.
Residual	116	16453.72	141.84	

** - significant at 0.01 level
N.S. - not significant

Table VIII. Shortest significant range values for test means on defined categories*

P	2	3	4	5	6
5%	0.75	0.79	0.82	0.84	0.85
1%	0.99	1.03	1.07	1.09	1.10

* df = 116 S.E. Mean =0.268

Table IX. Shortest significant range values for test means on total categories*

P	2	3
5%	.781	.823
1%	1.03	1.08

* df = 119 S.E. Mean = 0.279