MORRISSETTE, Mary Pauline, 1938-
THE USE OF CATEGORIES BY BRIGHT, NORMAL, AND SUBNORMAL PRE-ADOLESCENT GIRLS ON THE PIKUNAS GRAPHOSCOPIC SCALE AND THE STEPHENS' CATEGORIZATION TASKS.

The University of Oklahoma, Ph.D., 1967
Psychology, clinical

University Microfilms, Inc., Ann Arbor, Michigan

THE UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE USE OF CATEGORIES BY BRIGHT, NORMAL, AND SUBNORMAL PRE-ADOLESCENT GIRLS ON THE PIKUNAS GRAPHOSCOPIC SCALE AND THE STEPHENS' CATEGORIZATION TASKS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

MARY P. MORRISSETTE

Norman, Oklahoma

1967

THE USE OF CATEGORIES FOR BRIGHT, NORMAL, AND SUBNORMAL PRE-ADOLESCENT GIRLS ON THE PIKUNAS GRAPHOSCOPIC SCALE AND THE STEPHENS' CATEGORIZATION TASKS

APPROVED BY

P. T. Teska

W. B. Ragan

Mary Clare Petty

O. J. Rupiper

DISSERTATION COMMITTEE

ACKNOWLEDGMENTS

With a deep sense of gratitude the writer wishes to express her indebtedness to Dr. P. T. Teska for his enlightening lectures, for his confidence in and guidance of the writer as a doctoral candidate, and for his interest and assistance in the conducting and reporting of this study.

Appreciation is extended to the other members of my doctoral committee, Dr. William Ragan, Dr. Mary C. Petty, and Dr. Omer J. Rupiper for their constructive criticisms and encouragement at various stages of this study. Dr. Rupiper and Dr. Charles Bridges were most helpful with their suggestions on statistical analysis.

Appreciation is also extended to Mr. Lester M. Reed, Superintendent of the Norman, Oklahoma, public school system, and to Dr. Larry Hayes and the Board of Education, Oklahoma City, Oklahoma, for their cooperation in this research and for the use of their schools and students. A very special thanks is given to the principals and teachers and students of the schools in which the testing was done.

The author also expresses appreciation to Dr. Wyatt Stephens who devised the Stephens' categorization tasks and to Dr. Justin Pikunas who devised the Pikunas Graphoscopic Scale, for their assistance and permission to use their testing instruments.

The author cannot adequately express in words her appreciation and indebtedness to her mother, and younger sisters, Denise, Doris, and Dianne, for their confidence, love, understanding, and encouragement through her college and graduate studies.

Many thanks are also expressed to my very special and close friend, Dutch Hassell, who assisted me in the testing and provided many helpful suggestions in the reporting of the study.

In putting together a dissertation of this type there are many unseen hands whose skills are as important and necessary as that of the author's. They are the ones who originally contributed research, interest, hard work, and ideas, thereby developing the area concerned that much more. It is because of their generosity and dedication that this paper is possible. To these individuals, I express sincere appreciation.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

THE USE OF CATEGORIES BY BRIGHT, NORMAL, AND SUBNORMAL PRE-ADOLESCENT GIRLS ON THE PIKUNAS GRAPHOSCOPIC SCALE AND THE STEPHENS' CATEGORIZATION TASKS

CHAPTER I

INTRODUCTION

We are concerned not only with the role of categorization in the thinking of subnormal and normal children, but also with how it relates to and forms a part of the thinking of superior children. The plea for research in the area of mental retardation by the late President John F. Kennedy and the subsequent federal aid for making this research possible is finally being carried out. Research for the gifted child has been neglected, too, (Newland, 1941) for we know no more about the thinking patterns of the gifted child than those of the retarded one. To date, virtually no studies, experimental or comparative, concerned with the use and stock of categories have been carried out with retarded and bright children. It would therefore appear that study in this area of mental functioning is more than justified at this time. We need greatly to start exploring this phase to provide a basis for or to eradicate such assumptions as the alleged simplicity of thinking in the retarded child (Doll, 1941) and the complexity of thinking in the gifted child. Specific differences between the thinking patterns of the mentally retarded on the continuum of retardation as well as between them and children classified as normal or superior have, as yet, to be identified with any degree of clarity. Empirically, it is suggested that the retardates' capacities are quantitatively different, that he learns less and learns it more slowly. We can also pose the very logical hypothesis and provide some evidence in its support that the mentally retarded child learns in a qualitatively different way in that his impaired functions cause his perceptions of the environment to be limited and different. On the other hand, the gifted child has been traditionally viewed as possessing quantitatively and qualitatively greater abilities such as learning, assimilating, perceiving, evaluating, and functioning. It is from these observations that we have found a necessity for devising special techniques of instruction and for setting up special classes for our gifted and retarded children. Although these observations have been in large part descriptive and clinical, more recent comparative and experimental studies seem to offer hope for better communication between investigators and more continuity in the procedures, results, and conclusions of studies. Only with this approach and an avoidance of a tenacious clinging to traditional assumptions will progress in research be possible for these two extremes of the population.

The rationale underlying the present study is as follows:

1. Psychological literature permits us to define a category as a schema for evaluating impinging stimulus objects and events. A category represents a varying definitiveness and breadth along some specifiable dimension, e.g., <u>hot-cold</u>, <u>good-bad</u>, as abstracted from the experience of objects in one's environs. Once evolved, a category

serves as a psychological yardstick in terms of which stimuli are compared and gauged. It is a kind of experiential filter through which objects are screened and evaluated on their way from sensory reception to ultimate response evocation.

2. Categories, in their interrelatedness, serve the critical function of providing a system of ordering by means of which the environment is broken down and organized and differentiated and integrated into its many facets. It is in this capacity that categories provide the medium through which one establishes and maintains ties with one's environs. It is through the milieu of categorization that one can orient himself meaningfully in relation to time, space, and other objects and dimensions of the psychological world. It is on this same basis that one's self-identity and existence are articulated and maintained. Threats to such ties lead either to efforts at restoring them by reorientation and organization or to breakdown and ultimate destruction of self.

3. Some writers suggest that the normal or average individual functions as indicated in Steps 1 and 2. To show that one does not derive the same information from various categories of one's environs, or to take much more or much less time to respond to a category, indicates that one differs from the normal or average person.

4. As categorization is defined and explained in Steps 1 and 2 above, it seems logical to arrive at the premise that because retardates seem to have difficulty in everyday functioning and because the bright child displays relative ease in everyday functioning, it is probably a result of their differences in stock and use of categories

as compared to the normal child that they meet with their environs in dissimilar manners suggesting an inner difference in thinking, hence performance.

5. Most studies related to concept-formation or categorization have made use of male subjects. We therefore felt that the use of female Ss would give us added insight into this phenomenon, particularly as it applies to feminine thinking. In addition, the author believed that both testing measures would reveal interesting differences in the stock and use of categories that the male Ss had previously displayed. Furthermore, the findings derived from female Ss would suggest practical educational procedures and practices that could be unique to them.

6. Pre-adolescent Ss were used because some experiments indicate that children develop conceptual abilities better and more quickly at and after nine years of age. There were also practical advantages in testing that age group.

7. The categorization process could be explored through both the Stephens' tasks and the Pikunas Graphoscopic Scale, since both these tools were concerned with performance on conceptual categories and had demonstrated this potential.

Review of the Literature

General Literature

The area of "categorization" in psychological literature reveals a need for much sophisticated research (Kvaraceus, 1958; Zigler, 1962). Categorization, concept formation and ability, abstraction, generalization, or whatever other main title this area appears under, has been investigated, but often with inadequate procedures or in ways yielding distorted and over-stated conclusions. Yet, concepts or categories, as they will be referred to here, are most closely aligned with thinking (Bruner, Goodnow & Austin, 1956; Vinacke, 1952; Brown, 1958; Church, 1961), hence learning and mastery of life, making more stringent research a must.

The basic phenomenon involved in categorization allows the individual to maintain contact with his environment. It is a matter of responding to the many and varied stimuli of one's environs through evaluation and classification. The categories develop in time from percepts, memories, images, and the re-organization of an individual's experiences in a problem-solving manner. Several noted psychologists including Helson (1947), Thurstone (1927a) and Razran (1949) anchor categorization in perceptual theory viewing it as resulting from response behavior to stimuli that orders the world about us. Razran (1949) states that the human discrimination continuum has more steps than the animal. Categorization may be similar or different or have several shades of similarity and dissimilarity. Discrimination becomes more efficient and refined as the organism grows with gross approach-avoidance behavior no longer acting in an all-or-none manner but as a complex many-stepped response based on various properties of objects or events. Basically, it is a discriminal process which refines itself as the organism matures and consequently brings in added elements when the higher mental processes are involved. Bruner, Goodnow, and Austin (1956) define two levels of categorization, the perceptual and the conceptual level--"One of the principal differences between the two forms of categorization--the perceptual on the one hand and the conceptual on the other--is the immediacy of the attributes by which their fitness to a category is determined. In the perceptual sense, the relevant attributes are more immediately given by which we judge the categorical identity of an object at least in simple perceptual situations. At the other end, the attainment of knowledge about the attributes that are relevant may require a difficult strategy of search. . . ."

To Bruner, Goodnow, and Austin:

1. Categorization reduces the complexity of the environment.
2. Categorization is the means by which the objects of the world about us are identified.
3. Categorization reduces the necessity of constant learning.
4. Categorization permits the ordering and relating of classes of events (1956, p. 13).

It comes to mind, then, that categorization is what was frequently discussed by our philosophers as "generalization"--indeed, an old topic. Further, categorization is exactly what traditional psychophysics investigated by comparing experiences in different forms. Psychophysicists such as Weber and Woodsworth (Boring, 1957; Peters, 1953) were concerned with scaling experiences on a continuum according to intensity and magnitude. Psychophysics probed the relationships between physical reality and the corresponding experienced dimensions. It was the stimulus dimension that was most important to traditional psychophysics so that the presentation of a controlled series of stimuli with the quantified corresponding responses was standard. Being based on sensory experiences, these judgments were much the same for most people even though such judgments are a complex outcome of behavior based on stimulus traits, risk-taking, values, motives, pay-offs, probabilities of occurrence, expectations of the person and physical limits of the receptor mechanism. Today, psychophysics seeks to identify variables that affect behavior of a person in his judgments, actions, and responses.

It becomes obvious from the review of literature, then, that though we have investigated and isolated variables such as age and socio-economic level associated with categorization, we have, in fact, either falsified or have failed to clarify the several assumptions made for or against bright, normal, and subnormal individuals. We still view the subnormal and bright individual in an almost mystical manner with both falling at opposite ends of a continuum which we are not even sure exists.

The above information gives us a general introduction to the scientific setting and analysis of "categorization" in a psychological framework. What follows serves to provide continuity in viewing "categorization" as part of the learning process, hence its role in mental development. Categorization is only one facet of mental development when viewed as conceptualization. It is, however, a most important facet where mental subnormality or superiority is involved.

It seems proper to view mental development as an increasing capacity to discriminate the world of objects and events. This capacity is often referred to as the intellect. Pikunas says that "intellect is manifested in three interrelative ways: concept formation, symbolic or abstract reasoning, and judgment." He defines these as follows:

Concept formation is the mental process of abstraction, comparison, and generalization, which results in the creation of an image that represents some essential traits or qualities of the particular object. Words and speech are symbolic and observable expressions of mental images or concepts.

Symbolic reasoning is a mental operation in the course of which relationships between two or more words or concepts are discerned and understood. Evaluation and classification of concepts, beliefs, opinions, judgments, hypotheses, and theories are a further achievement of symbolic reasoning. Judgment is the natural effect of reasoning by the use of concepts and known facts or relationships (1959, p. 112).

Pikunas then concludes that there are four stages of intellectual maturation which start with periods of questioning.

The first fifteen to seventeen months of life are considered as the preconceptual level. The infant may already be able to associate symbols with the objects they represent. However, this does not substantiate the abstraction of at least the few essential qualities necessary for a conceptual image.

The ability to develop concepts becomes obvious during the first period of questioning. Questions are at first related to the names of various subjects and objects of the child's environment. In the latter part of the second year, when he becomes curious to know the names of distinguishable phenomena in his surroundings, the child asks "What is it?" The meaningful use of newly acquired names or titles indicates that a child forms definite mental images and is able to use verbal symbols to represent them. Thus the conceptual level is introduced by curiosity about names, or nominal realism, as Jean Piaget calls it. Concept formation precedes reasoning and judgment. At approximately the twenty-month level, an infant forms his first simple judgments. . . .

The second period of questioning introduces the so-called precausal level of reasoning. The child now begins to realize that things serve certain purposes, and he wants to know these purposes. What for? How? Why? These are the usual questions of purposeful thinking at five or six years of age.

The third period of questioning evidences interest in causal relationships. Thus, from eight or nine years of age, the child's curiosity is directed to all types of changes and occurrences. He wants to know about causes and effects. Causal thinking represents the scientific dimension of recognition. It proceeds from concrete to abstract and metaphysical relations.

> In the fourth period of questioning, the child discovers the essential or philosophical dimension of existence. Excessive preoccupation with thinking and reasoning at thirteen or fourteen years of age leads to the emergence of many existential, essential, and spiritual problems. The individual acquires an abstract, theoretical, and very critical attitude, accompanied by much skepticism, doubt, disbelief of authority. Curiosity and rational sentiments lead the individual to investigate many sources until he finds proofs to justify his assumptions. The beginnings of <u>self-answering</u> characterize the period of puberty and adolescence.
>
> Achievement or performance due to practical application of intellectual capacities is referred to as intelligence. The psychologist, therefore, tests intelligence by requiring the subject to perform certain tasks, the efficient performance of which presupposes the presence and functioning of intellect. Intellect enables man to meet and solve natural and artificial problems of an individual, social, environmental, or cultural type. Its contribution to adjustment is of capital value (1957, pp. 112-113).

Thus, we have a summary of the mental development of the child which is in agreement with the general literature (Carmichael, 1954).

Both Pikunas's and Piaget's approaches to explanations of mental development stem from a philosophical approach toward thinking, which is as it should be; still, we have progressed far enough in experimentation with conceptualization to be able to make very clear attempts at discovering the categories or concepts children acquire at different ages and how these operate in the child.

Experimental Literature

Overall, the literature on categorization is little in amount and frequently poor in quality with confounding variables running rampant in several of these studies. We will see from the review below that there is no research which compares normal, subnormal, and bright children on repertory and use of categories except for the two studies (Stephens, 1964; Van Osdol, 1964) which made use of the Stephens' categorization tasks. Studies exist, however, that do contribute indirectly to this issue.

These are:

1. Studies on concept development in the normal child, e.g., Piaget's classical studies.
2. Studies associated with variables related to concept formation. These variables are age, training or experience, socio-economic level, and vocabulary and performance on general mental ability tests.
3. Psychological studies comparing children of different intellectual levels.

Studies of Concept Development in the Normal Child

Piaget has done extensive work on concept development in the normal child. His works (1929, 1930) lead us to think that the process of abstraction proceeds uniformly along developmental lines in normal persons from concrete and private conceptualizations of the very young child to the abstract behavior of the older child, the latter being a function of language and communication skills, as well as a function of environmental experience. Investigators both support and criticize Piaget. Nagy (1948) supported his findings by analyzing children's theories of death. Three stages were discovered: the denial of death which is equivalent to Piaget's animism (all is alive for the child between ages three to five); the personification of death which is the artificialism Piaget described in children five to nine years of age; and death viewed as an inevitable process for nine-year olds and older.

The latter is Piaget's "realism." Gribsby (1932) studied the concepts of time, space, cause, part-whole relations, discordance, and number in children between ages two years-eight months and six years-four months. She concluded that although there were gradations, these were more related to mental age than to chronological age. She was in agreement with Piaget on the causal relations category. Deutsche (1937), however, disagreed in large part with Piaget and Gribsby's findings on causal relations after testing out 732 subjects (Ss) between the ages of eight and sixteen. She concluded that there was no evidence that children's reasoning developed by stages. Oakes (1947) also studied causal relations in seventy-seven children between ages four years-ten months and seven years-four months. He was in agreement with Deutsche. The experiments by Welch and Long (1940, 1942) also reflected disagreement with Piaget on age-stages in concept-formation. They were forced to conclude that children's concepts change with increasing age, but in a gradual progression toward greater understanding rather than in clear-cut stages. That the concepts of children differ from the concepts of adults in degree rather than in kind is due mainly to the amount of experience and knowledge members of each group have acquired respectively.

Studies Associated with Variables Related to Concept Formation

The author will now review the literature associated with variables related to concept formation. These are age, training or experience, socio-economic level, and vocabulary and performance on general mental ability tests.

Age. Welch and Long (1942, 1939) have done notable investigative work on conceptualization in children. Their several studies indicate a growth from the simple to the more complex levels for conceptualizing. They refer to first-hierarchy concepts and second-hierarchy concepts which are levels of abstractness. Welch and Long have overwhelming evidence showing that children eight years and older can at least deal with the lower hierarchical levels.

A study by Reichard, Schneider, and Rapaport (1944) on conceptualization confirmed the results of Piaget. They used the Weigl Color-Form Sorting tests (successes measured by the ability to shift from one category to another) on 234 normal white children, ages four to fourteen. The results showed that increased age provided for a better ability to group objects together and gave rise to better abstract explanations of the groupings. Inadequate responses decreased with age increase. They found that the five and six-year olds gave a concretistic performance which was based on the non-essential features of objects. Children up to eight or ten years gave a functional performance where classification was based on use or value, etc.; and finally, the older children reflected a conceptual level wherein they classified on the basis of abstract features.

Colby and Robertson (1942) investigated color-form abstraction among children of ages three and one-half to nine and one-half. They concluded, upon retest after a lapse of a year, that there was a main developmental trend which showed a shift from color toward form dominance as means of classification. Welch (1939, 1940) supported this finding though he said that form was dominant below age three with color dominating between ages three and six and a return of form dominance after age six.

The concept of time appears to develop over a period of several years. Ames (1946) and Friedman (1944a, 1944b) indicated that the four-year old recognizes morning and afternoon, the five-year old knows what day it is, and the eleven-year old knows the time system. Other time concepts develop later, most children reaching a full understanding of time by ages fourteen-fifteen.

Several studies on magnitude or size (Thrum, 1935; Hicks and Stewart, 1930; Welch, 1939) show that the simple concepts of size can be acquired as young as fourteen months, but that the concept of middle-size relations is not common under age five. It does appear, however, that children between ages three and five can learn the concept of intermediacy and apply it in a particular situation depending on their MA level.

Long (1940) has shown that the concept of roundness is present as early as three years of age. Gellerman (1933) and Munn and Steining (1931) have shown that the concept of triangularity developed between the ages of fifteen months and two years of age. The concept of contradictory relations was investigated by Dixon in 1949, who found that the noting of contradiction becomes apparent by ages three and four at the same time as the concepts "big" and "little" are developed.

Cause-effect relations were studied by Lacey and Dellenbach (1939). They concluded that this concept was not grasped until eight or nine years of age.

Social concepts in children were studied by Ordan (1945). He found that recognition of these concepts increases from year to year, but that it is not until twelve-thirteen years of age that children recognize social problems with any degree of understanding.

Socio-economic Status. It appears that socio-economic level has a very low relation to scores on concept tests according to the findings of Deutsche (1937) and Ordan (1945).

Vocabulary and Performance on General Mental Ability Tests. According to Deutsche (1937) and Ordan (1945), there is a very obvious relationship between vocabulary and concept scores. Piaget (1926) long ago stressed the interrelationship between language and the thought processes of the child, and even before Piaget, Terman (1916) reported a correlation of .91 between mental age and the vocabulary test in reference to the 1916 Stanford-Binet Scale. Terman and Merrill (1937) again reported a high correlation between the vocabulary test on the Revised Stanford-Binet Scale and on the mental age rating on the scale as a whole. These correlations ranged between .65 and .91 for the different age groups.

Both Wechsler (1944) and Rapaport (1945) have suggested that subtest response patterns on the Wechsler scales were indicative of various thought processes of the subject, e.g., schizophrenics show a high vocabulary score when compared to other subtest scores (Myers and Gifford, 1943); mental defectives show a higher performance IQ than a verbal IQ. Scatter analysis on the Binet scales has also supposedly displayed various thought processes. Feifel (1949) did an expansive study on the vocabulary responses of 370 normal and abnormal Ss ranging from age fifteen through eighty. He found that normal Ss used synonyms in defining the Stanford-Binet vocabulary words while abnormal Ss used description, inferior explanations, illustration, demonstration, and repetition.

Study after study on scatter analysis of both the Binet and Wechsler scales has proved to be of little diagnostic value (Anastasi, 1954; Cronbach, 1960). These scores do not reflect specific modes of thinking, because these scales were designed to prevent any factor except "general ability" from influencing scores to any measurable degree. The subtests are factorially complex; therefore, by their very nature, these subtests are incapable of further breakdown for detailed predictive ability.

Training or Experience. Deutsche (1937), Oakes (1947), and Ordan (1945) suggest that concepts and use of concepts differ among children as much as because of their various experiential backgrounds as because of their intellectual levels. Both the studies of Welch and Long (1943) and Gellerman (1933) show that children can be taught different concepts at very young ages.

In summary, we must conclude that the studies that contribute to the rationale of this study yield, at best, only some information on the number of categories children possess and can use.

The studies apparently rely upon an investigation of factors that are either philosophical, i.e., realism, causation, reasoning, time, etc. (Piaget's influence), or perceptual in nature, i.e., form, size, etc. These factors are most important to conceptual learning, but it is now imperative that we be concerned with the more common categories used by children of different mental abilities in everyday intellectual pursuits.

Psychological Studies Comparing Children of Different Intellectual Levels

Several studies can be brought together under this heading to lend some indirect information on "categorization" and patterns of thinking.

Werner and Strauss (1941; 1940-41) have made several investigations which try to distinguish thinking patterns and facets in familial and brain-injured retardates. It appears that the brain-injured retardates use a most incoherent procedure to copy a mosaic pattern of marbles from a marble board, while the familial retardate uses a well-rounded, continuous, and global manner. These investigators (Werner and Strauss, 1940-41) also had normal, familially retarded, and brain-injured retarded children repeat melodic tunes that had been played on the piano. They found that the familially retarded children made errors similar to those made by the normal children in repeating the tunes and that both these groups tended to simplify and make the tunes more homogeneous and less articulate when the patterns were difficult. They tended toward a global, round, and homogeneous approach. The brain-injured, however, made errors that were rarely made by subjects from the other two groups, and their reproductions of the tunes were strange and unrelated patterns reflecting no synthesis. This study also showed that brain-injured retardates, being more susceptible to the distractions of a background in trying to reproduce figures, have a greater difficulty than the familials with figure-ground relations.

The work of Lewin (1935) and Kounin (1941) has led us to assume that the conceptual behavior of the subnormal is rigid. In fact, the approach to concept formation in subnormals has traditionally been viewed as concrete and rigid while the normal individual supposedly uses an abstract and flexible approach. Goldstein and Scheerer (1941) reflect this approach in their development of sorting tests which were used mainly to study the concrete and abstract abilities of pathological cases.

Bolles (1937), Weigl (1941), Hanfmann and Kasamin (1937), and Iscoe and Giller (1959-1960) also approached the study of conceptual behavior as a concrete versus an abstract dichotomy. All of these studies have contributed to the issue of conceptual behavior in general, but only minimally to "categorization," especially since the above studies were mainly concerned with adult performance and pathological thinking.

Osborn (1960-1961) presented familially retarded, brain-injured retarded, and normal individuals (ages ten-thirty) with an altered Bausfield procedure (presentation of words at random with Ss automatically reproducing these in groups or clusters or categories implying a basic organizational tendency) replacing the words with pictures. Osborn found no significant differences between the brain-injured and familial retardates, and both these groups recalled and organized the pictures conceptually as adequately as the normal Ss. Nevertheless, the retardates, did show qualitative differences in the way in which they arrived at their total scores indicating poor functioning related to inappropriate learning habits.

Griffith, Sptiz and Lipman (1959) made a study on the relationship between concept formation and verbal mediators. They administered an abstraction task in which normal and retarded Ss had to discover a similarity among three words, e.g., bird, airplane, kite; night, cave, closet. In another session, the Ss were presented again with the stimulus words and were asked to define these to see how many would make use of the abstraction used with the original triads of words. It was found that the percentage of abstractions attained increased as the number of words defined in common with a possible abstraction increased. Retarded individuals and the normal seven-year olds were not very successful at concept attainment unless they had defined at least two words in terms of an acceptable abstraction. The nine-year old normal Ss were quite successful at concept attainment even when they defined only one word in terms of an abstraction. It seems, then, that there is a continuous quality between age and ability to conceptualize.

With two exceptions, all research dealt with comparative studies of differences in abstract and concrete behavior. The two exceptions were studies carried out by Bensberg (1958) and Martin and Blum (1961) and represent the only attempts made to study the process of concept formation in mental defectives as a group.

Bensberg (1958) focused attention on the mediation hypothesis and the concept of habit strength in male defectives. Seeking to evaluate the effects of variations in the dimension of similarity (form or color) and in the number of pretraining trials upon subsequent performance in a task in which only one dimension--form--was relevant, he used sixty male defectives with a mean IQ of 47.01 and a mean CA of 20.31. They were assigned at random to five groups of twelve Ss each. The pretraining materials consisted of stimuli varying in color and form, and the task was to establish a response--button pressing--to either form or color as a relevant dimension. Group I (color) and Group III (form) were carried to a criterion of one errorless presentation of the stimulus series; Groups II (color) and IV (form) were given additional trials. Group V, the control group, performed a neutral task. Color and form tasks were equated for difficulty level. In the second part of the experiment, a paired-associate task was used to test for generalization of pretraining responses in which form was the only relevant dimension. The colors and forms of the generalization task were different from those used in pretraining in that the forms of the generalization task were paired with nonsense syllables, and all groups performed the same task.

Bensberg anticipated that transfer would be negative for the groups pretrained on color and positive for the groups pretrained on form, while overlearning would increase the predicted effect. Statistical analysis indicated that the groups did not differ significantly on the pretraining task and that all predictions made for performance on the transfer task were confirmed except one. The group given additional pretraining on form did not show a significantly higher level of performance on the transfer task than the form group carried to a criterion.

Martin and Blum (1961) employed the dimension of oddity in a comparative study of conceptual learning and generalization in normal children, familial defectives, and mongoloids, with groups further differentiated according to sex. Stimulus objects were presented three at a time to all Ss, the odd item being varied in size, color, form, or spatial orientation. Two training series were run with correction in which size was the odd dimension, followed by a series of eight tests of generalization. The sequence of odd dimensions in the test series was size, form, spatial orientation, color, spatial orientation, form, and size. Each test was presented for six noncorrected trials, with candy and social approval as reinforcers. The sketchy descriptive data presented indicated that the training task was relatively easy for all groups to master although there was some slight suggestion of differences in favor of normal and familial Ss. The measure of generalization for all the tests was a mean number of correct responses on the first trial. An overall measure of learning for the eight trials was also computed in terms of the number of correct responses on trials two to six.

An analysis of variance of the generalization data revealed that group differences alone were noteworthy. After adjustment by covariance for group differences in MA, however, the differences associated with diagnostic category lost significance, while sex differences and the sex by group interaction became expressive. The results for the generalization series further suggested that in normal and familial Ss and in mongoloid boys, there was a tendency for orientation tests to be most difficult and for form tests to be easiest. Form was the most difficult and color easiest for the mongoloid girls.

On the overall measure of learning the only meaningful difference found was between groups. This difference remained significant even after group differences were adjusted for MA. The adjusted group means of the normal and familial Ss did not differ, but the mean of the mongoloid group was lower than either. The order of difficulty of tests for the familial and normal Ss was from most to least difficult: orientation, color, size, and form. No consistent pattern emerged for the mongoloids with the exception that the orientation tests tended to be most difficult for both males and females. The similarity between the generalization and learning measures in pattern of difficulty should be noted.

Although the studies presented in the literature section do not yield all the knowledge and understanding of human behavior that we could have received, let us remember that those who participated in its production put in much effort and that, in fact, we have quite a stock of information. Linking of this information and a common interpretation would yield added understanding.

Summary

A review of the literature reveals that very little research is directly concerned with the categorization process though some studies on concept-formation, conceptualization, and generalization give indirect evidence related to it. A large proportion of the studies reflect improper research with absence of control experimentation, failure to define the problem and key terms, obvious interacting variables, and confounding influences. Furthermore, the absence of a uniform vocabulary in psychological research promotes poor communication, makes our investigative efforts fruitless, and permits society to continue to operate under falsifications and entrenched routines. The research for both bright and subnormal children has not justified our present routines of dealing with them, nor has it greatly enlightened our understanding of their mental processes. Still, notable studies have been conducted in this area.

For example, it appears that Stephens (1963) developed an excellent approach for studying a facet of the thinking process through examination of categorizing abilities and differences of normal and subnormal boys. This approach is based on the theory that an individual uses many conceptual categories to interpret stimuli and lend meaning to his experiential world. It follows that if a person possesses very few categories and uses these ineffectively because they are poorly developed, he will view and interpret stimuli and experiences in a different manner than the average person. On the other hand, the individual who reflects a very broad range of categories and extra ability and sensitivity for noting associative cues leading to effective category use might be viewed as unlike the average person, but in a superior way. The present study is concerned with all three levels of functioning.

It is felt that further analysis of categorization abilities in a comparative study as explored through performance on Stephens' tasks and on the PGS with drawings as the mode of communication will prove valuable to increase understanding of thinking patterns and intellective traits of children. Furthermore, additional analysis should help to provide a more realistic base or displace some current educational practices and encourage more research.

CHAPTER II

STATEMENT OF THE PROBLEM

Introduction

The problem of this study was to compare the performance of subnormal, normal, and bright subjects on two testing instruments--the Pikunas Graphoscopic Scale (PGS) and Stephens' categorization tasks--in an effort to determine the relationship between these two instruments, and more specifically, to study the use of conceptual categories by these subjects of different intellectual capacities on the two testing tools.

This study was concerned with determining whether or not a significant correlation existed between the abilities of the different groups to make greater or less successful use of test categories on Stephens' categorization tasks and to employ more or fewer content categories in response to the semi-structured perceptual cues on the PGS: that is, if bright subjects can make successful use of a greater number of test categories on Stephens' tasks than do subnormal subjects (Ss), so will they be able to make use of a greater number of content categories than do subnormals in response to the perceptual cues on the PGS. A positive correlation here would indicate a continuous function between the response of a subject to an overt stimulus which elicits conceptual categories and to a semi-structured cue which elicits conceptual categories. The positive relationship would also seem to indicate that children are quite able to utilize their own experiential framework despite the degree to which a stimulus is structured if the conceptual category is developed and associated or known within.

In the use of Stephens' tasks, determining to what degree the different groups were successful in identifying members of categories after these were named by the examiner and ascertaining the possible variances between the different groups in speed of obtaining correct responses were the concerns of the author.

Interest in comparing the groups on quantity and quality of content based on different categorical responses--for example, the number of original versus popular responses used by each group and the number of content categories used by each group on the PGS was warranted. It was thought that these groups would also show differences on type of content depending on age; human content would be drawn possibly with greater frequency with IQ increase and age increase.

Hypotheses to Be Tested

With respect to Stephens' categorization tasks and the Pikunas Graphoscopic Scale, the following hypotheses were formulated:

1. There is no statistically significant relationship between the number of successes attained by each group of subnormal, normal, and bright subjects on Stephens' unstructured categorization tasks and the number of different content categories these respective groups employ in response to the perceptual cues on the PGS.

2. There is no statistically significant relationship between the number of successes attained by each group of subnormal, normal, and bright subjects on Stephens' naming tasks and the number of different content categories these respective groups employ in response to the perceptual cues on the PGS.

Hypotheses specifically related to Stephens' categorization tasks are as follows:

3. There is no statistically significant difference between bright and normal, normal and subnormal, and bright and subnormal groups of subjects in the number of correct responses they achieve in response to Stephens' unstructured categorization tasks.
4. There is no statistically significant difference between the bright and normal, normal and subnormal, and bright and subnormal groups of subjects in the number of correct category names specified when members of each group are required to name categories they made use of on Stephens' unstructured categorization tasks.
5. There is no statistically significant difference between the bright and normal, normal and subnormal, and bright and subnormal groups of subjects in the number of correct responses they achieve in response to Stephens' structured categorization tasks.

Hypotheses related to time differences on Stephens' categorization tasks are as follows:

6. There is no statistically significant difference between the bright and normal, normal and subnormal, and bright and subnormal groups in their total times for responses to the unstructured categorization tasks.

7. There is no statistically significant difference between the bright and normal, normal and subnormal, and bright and subnormal groups in their total times for responses to the category naming tasks.

8. There is no statistically significant difference between the bright and normal, normal and subnormal, and bright and subnormal groups in their total times for responses to the structured categorization tasks.

Hypotheses related to the PGS are as follows:

9. On the basis of different frequencies of categorical responses (Human, Animal, etc.), bright, normal, and subnormal children will not produce content that differs from one another on the PGS.

9a. Original responses on the PGS will be produced with greater frequency with IQ increase.

An original response is a drawing that has a separate theme or concept which adequately incorporates the stimuli. Similar responses appearing on subsequent frames are not original.

9b. Bright, normal, and subnormal children differ in the number of common or popular responses they produce on the PGS.

A popular response refers to the expectancy of certain drawings for each stimuli as defined on pp. 5-7 of the Pikunas Graphoscopic Scale Manual (1959).

9c. Children of different IQ levels do not differ significantly on the number of content categories they employ in responding to the PGS.

9d. Bright girls do not produce more human responses than normal and subnormal girls on the PGS.

9e. Animal responses on the PGS will not be produced with greater frequency with IQ increase.

10. On the basis of different frequencies of categorical responses (Human, Animal, etc.), children of different ages will not differ on type of content produced on the PGS.

10a. Human content will not be produced with greater frequency with increasing age.

10b. House content will not be drawn with diminishing frequency with age increase.

CHAPTER III

PROCEDURE OF THE STUDY

Introduction

The purpose of this study was to determine if there were significant statistical differences in the performance of bright, normal, and subnormal children utilizing a variety of conceptual categories under different conditions. The study sought answers to assumptions previously made about the intellective traits of children of different intellectual abilities. Is it, in fact, true that bright children have a much broader range of categories to operate with than subnormal children? How do normal and bright children compare in their abilities to use categories? How quickly do each of these groups respond in employing the categories under different conditions? Do bright children show a significantly greater variety of content score than do both normal and subnormal children? Are bright children much more original in responding to the perceptual cues? Do subnormal children have a more difficult time drawing upon their conceptual categories and making meaningful use of these? Are normal children in the group most apt to give the common or popular responses to perceptual cues?

Answers to these questions will open a path of suggestions relevant to teaching practices for children falling within these intellectual ranges in addition to providing us with more factual information concerning the intellective traits of children of different intellectual abilities.

The Instruments

This study made use of two tools of research which have been published: The Stephens' categorization tasks, an experimental instrument devised by Wyatt Stephens (1963), and the Pikunas Graphoscopic Scale (PGS) (Appendix A), a multi-dimensional projective test of personality with limited standardization.

Stephens' categorication tasks consisted of a series of twenty-seven cards, each eight inches by eighteen inches, on which were located seven different pictures. Each test card represented one of the twenty-seven categories. These categories are listed below:

Sample: Size
Sample: Form
1. Color
2. Number
3. Detail
4. Orientation in space
5. Heat
6. Clothing
7. Fruits versus vegetables
8. Flying versus non-flying objects
9. Containers versus non-containers
10. Tools versus non-tools
11. Cutting versus non-cutting equipment
12. Sex differences in children
13. Age differences in men
14. Sex differences in adults
15. Happy versus sad children
16. Ugly versus pretty women
17. Land vehicles versus airborne or amphibious vehicles
18. Land animals versus airborne or amphibious animals
19. Young boys versus other living things
20. Clothing made from animal products versus other wearing apparel

21. Footwear versus other clothing
22. Furniture versus other household objects
23. Cooking equipment versus other household objects
24. Male versus female wearing apparel
25. Even number of dots versus odd numbers of dots

On each test card there were seven randomly ordered figures or pictures, four of which represented the category and three of which were incorrect responses in terms of the category which was being tested. In this study, as in Stephens' study, the subject performed three tasks.

> First, he was required to independently decide upon the appropriate category for each card. Second, he was required to provide a name for each of the categories he had used as a basis for arriving at his responses. Third, he was required to find the items on each card which represented the correct category, after the name of that category had been specified by the examiner (Stephens, 1963, pp. 29-30).

There was no need to carry out a pilot study with this tool since its author had already demonstrated its discriminating ability in his own publication (Stephens, 1964, pp. 311-315).

The second tool of research employed in this study was the Pikunas Graphoscopic Scale (PGS), a semi-structured projective drawing test designed primarily for use with children and adolescents. It consisted of ten drawing plates on one sheet of heavy paper (approximately 15" x 19"). Eight of these ten plates have semi-structured perceptual cues, one of these ten being fairly structured with another compensating for it in the sense that it is a free drawing answering the question "What is the thing you like to draw most often?" Subjects are also asked to give a written answer to the question "What else would you like to draw?" Appendix A is a copy of the PGS. The reproduction differs from the original in size and in color of the cues. The PGS grew out of Gestalt-orientated psychological research and it relies on drawing as a means of self-expression and communication.

In being tested, subjects observe the given perceptual cues and form associations and develop concepts for drawing.

The Pilot Study

The pilot study was concerned with the responses of bright, normal, and subnormal children solely on the PGS since the Stephens study had already adequately demonstrated the efficiency of his instrument in discriminatory ability.

The main purposes of the pilot study were (1) to give support (by observing the raw scores) to the assumption that children of different intellectual potential will produce content that differs in response to the semi-structured perceptual cues on the PGS (because of their various experiential framework); (2) to give support to the further assumptions that bright, normal, and subnormal children will differ in number of popular and original responses and in variety of content in responding to the PGS cues; (3) to give support to the assumption that boys and girls differ in the type of content they produce in response to the PGS cues, e.g., boys produce more factory-made objects whereas girls produce more human figures; and (4) to reveal any mechanical problems that might exist in the administration and scoring of the PGS.

Three groups of ten children each were tested in the pilot study. There were five boys and five girls in each group. All of the intelligence quotients listed below had been achieved by the children in no more than one to one-and-one-half years prior to this testing. The intelligence quotients for the normal, high average, and bright groups were from the California Test of Mental Maturity (CTMM). The intelligence quotients for the subnormal children had been obtained on the Stanford-Binet or the Wechsler Intelligence Scale for Children (WISC). Group I, the bright group, showed an IQ range of 126-140. Group II, the high average group, showed an IQ range of 116-122. Group III, the normal group, showed an IQ range of 93-110. Group IV, the subnormal group, showed an IQ range of 64-75 (educable range); these boys and girls were enrolled in a special class and were all ten to thirteen years of age. The subjects from Groups I, II, and III, all from the same school, were enrolled in fifth and sixth grade regular classrooms and were all ten to twelve years of age.

All these children were checked out prior to testing to insure control of variables that might influence or mask the desired results. All the children were Caucasian, Protestant, free from physical handicaps, of the upper-middle class, and seemed free of serious emotional problems as judged by their school records and behavior. These factors were determined with the aid of the school principal and the teachers of the children.

Preliminary procedure for testing was as follows:

All subjects were instructed by their classroom teachers to bring their five basic color crayons (black, red, green, blue, and yellow) to the testing room. The ten children from each group were tested together in three different test periods.

The seating plan in the testing room had been re-arranged so that the subjects were safe distances from one another to insure no copying. Once seated, each subject was given a PGS blank, a pencil, and a small box in which they could put their crayons. They were told that this task had no bearing on their grades but that it was important for them to have and use their own ideas. They were told to raise their hand if they had questions during the testing period and that the examiner would go to their desk to answer their questions. They were told to ask for any extra crayons or pencils if needed and of the importance of silence during the session.

The procedure for testing was as follows:

1. The subjects were instructed to write their name, address, sex, age, grade, and date of birth on their PGS blank.
2. Then, the following instructions were given:

> This is a drawing task. (The examiner shows the test blank). Look at these different squares that have different marks on them. In the first one there's a child and a dog. Try to finish it. Then look at each of these marks (point). They make you think of something to draw. You can draw anything you like, and start at any space you want, but I want you to number the squares in order as you draw in each one (demonstrate). Try to make these marks (point again to some of them) a part of your drawings. Use any of the five crayons you have brought with you. After you have completed the drawings which should include the square at the bottom (#10), raise your hand, and I will nod "yes" for you to come up to my desk with your box of crayons and your drawing test.--Remember now, draw as well as you can.

Once at the examiner's desk, the subject was given further instructions as follows:

> 3. Well, now write down the name of each thing you drew in each space and then answer in writing the question: "What else would you like to draw?"

The subject was aided in writing and spelling if necessary. When the subject had completed the test, the examiner expressed her appreciation to the subject and told him he could return to his classroom. A record of time was kept on each subject, and a specially constructed sheet was used to score each subject's responses.

Three of the Pikunas Graphoscopic Scales were not evaluated because they were incomplete or reflected severe emotional disturbance. Two of these scales were from subnormal subjects and the other from a bright subject. This fact in itself reflected the importance of using strict measures in evaluating good adjustment in screening subjects for the study. When these responses were included, the differences between the bright and subnormal subjects were simply maximized in this case.

The response categories were scored as Human (H), Animal (A), House (Hs), Factory-made object (Fobj.), or Miscellaneous (Misc.) content.

The tables presented below give raw scores or raw score means or averages. Raw score means or averages are used when the number of subjects in a group are unequal. Averages are a means of equating the results.

Observation of the data presented in Table 1 below shows support of the first assumption of the pilot study and that part of the second assumption concerned with variety of content. The first assumption stated that children of different intellectual potential would produce content that differs in response to the semi-structured perceptual cues on the PGS.

Table 2 presented following shows support for the second assumption which is concerned with number of popular responses, and the number of content categories used in responding to the PGS cues.

The author will not present a table or figure on the frequency of original responses though gross observation clearly indicated

Table 1

Comparison of Bright, High Average, Normal, and Subnormal Groups on Frequency of Types of Content Categories Produced on the Pikunas Graphoscopic Scale
(N = 27)

	Raw Score Averages on Content Categories				
	Human	Animal	House	Fobj.	Misc.
Bright (3 girls - 3 boys)	2.0	0.8	0.8	2.8	2.5
High Average (6 girls - 4 boys)	2.7	0.9	0.5	2.6	2.4
Normal (3 girls - 4 boys)	1.5	2.3	0.4	2.6	2.3
Subnormal (2 girls - 4 boys)	1.3	0.3	1.0	4.8	1.8

Table 2

Comparison of Bright, High Average, Normal, and Subnormal Groups on Number of Content Categories and Popular Responses Employed on the PGS
(N = 27)

	Groups			
	Bright (n=6)	High Av. (n=10)	Normal (n=7)	Subn. (n=4)
Raw Score Averages On No. of Content Categories	6.0	5.5	5.0	4.2
Raw Score Averages On No. of Popular Responses	3.2	3.1	3.4	2.5

the support of the assumption that bright subjects displayed a greater number of original responses than did the subnormal subjects.

Table 3 presented below shows support for the third assumption which is concerned with type of content produced by subjects of different sexes.

Table 3

Comparison of Girls and Boys of Different Intellectual Abilities on Types of Content Produced on the PGS
(N = 27)

	Raw Score Averages									
	Girls					Boys				
	H	A	Hs	Fobj.	Misc.	H	A	Hs	Fobj.	Misc.
Bright (3 girls - 3 boys)	1.7	1.7	0.0	3.7	2.0	2.3	0.0	0.7	3.0	3.0
High Average (6 girls - 4 boys)	2.7	1.2	0.5	2.2	2.5	2.8	0.3	0.3	3.3	2.2
Normal (3 girls - 4 boys)	2.0	0.0	0.0	3.7	3.3	1.0	3.3	0.8	2.5	1.5
Subnormal (2 girls - 2 boys)	1.5	0.5	0.5	4.0	2.5	1.0	0.0	1.5	5.5	1.0
Mean of Averages (14 girls-13 boys)	2.1	1.1	0.5	2.6	2.6	1.8	1.1	0.7	3.3	2.0

Table 4 further indicates sex differences of subjects of different intellectual ability as related to the frequency of their popular responses and content categories.

The mechanical procedure of the administration revealed the necessity for more individualized instruction and attention for the subnormal subjects.

Table 4

Comparison of Girls and Boys of Different Intellectual Abilities on Frequency of Popular Responses and Content Categories on the PGS
(N = 27)

	No. of Content Categories		No. of Popular Responses	
	Raw Score Averages			
	Girls	Boys	Girls	Boys
Bright (3 girls - 3 boys)	6.3	5.7	3.7	2.7
High Average (6 girls - 4 boys)	6.0	4.8	3.0	3.3
Normal (3 girls - 4 boys)	5.0	5.0	4.3	2.8
Subnormal (2 girls - 2 boys)	5.0	3.5	1.5	3.5
Mean of Averages (14 girls-13 boys)	5.7	3.2	4.9	3.0

It became clear that the data sheet needed revision to include a further breakdown of the content categories especially under the "Miscellaneous (Misc.)" category along with wider spacing for personal data and response totals.

The pilot study also reflected time differences in the completion of the PGS as shown in minute units in Table 5 below.

Table 5

Comparison of Bright, High Average, Normal, and Subnormal Groups on Time for Completion of the PGS
(N = 27)

	Minutes to Completion for Each Subject by Group			
	Bright	High Aver.	Normal	Subnormal
	59	40	59	29
	45	67	41	29
	44	50	30	44
	26	48	24	30
	24	25	51	
	25	22	54	
		46	59	
		49		
		49		
		43		
Totals	223	439	318	132
Means	37 min.	44 min.	45 min.	33 min.

Gross observation of the pilot study data revealed sex differences related to colors used, and the order of completion of the frames. It also appeared that children of similar intellectual abilities tended to use a similar order of completion of the PGS frames.

In summary, one can conclude that examination of the results of the pilot study indicated sufficiently great trends and raw score differences between groups to warrant further research on these hypotheses.

The Sample

Subjects included in the present study were ninety girls from the Oklahoma City and Norman, Oklahoma, public school system. The schools from which these subjects were selected fell within the same socio-economic level as determined from a chart graphed at the County Building in Norman and at the Capitol Building in Oklahoma City, Oklahoma. These charts have shown that the location of a home and some other establishments are one of the best indicators of socio-economic levels. In addition, the addresses of the subjects were checked against this chart to insure their being from a lower-middle to middle-middle socio-economic level.

All subjects were Caucasian, Protestant, and had always attended Oklahoma schools.

Each subject was screened prior to testing to insure good vision, good hearing, absence of physical handicaps and of emotional disturbance. Subjects who evidenced impairments of the above nature were excluded from the sample.

The examiner determined the presence of defects by school record information, observation, and testing. The examiner ascertained the presence or absence of physical handicaps and hearing difficulty by observing and conversing with the child. Visual acuity was checked through school record information, observation, and by testing the child on her ability to identify objects, words, and colors.

The Goodenough Draw-a-Person Test (Goodenough, 1926) was administered to all the subjects in order to eliminate those evidencing emotional disturbance or pathology and to assess an IQ as well as verify the obtained California Test of Mental Maturity (CTMM), Stanford-Binet (S-B), or Weschler Intelligence Scale for Children (WISC) IQ for the respective subjects. These obtained IQs had been achieved within the past one to one and one-half years.

The IQ ranges for the subjects were as follows: Group I consisted of thirty bright girls with an IQ range of 125 and above as measured on the CTMM and verified on the Goodenough Draw-a-Person Test; Group II consisted of thirty normal girls with an IQ range of 90 to 110 as measured on the CTMM and verified on the Goodenough; and Group III which consisted of thirty subnormal girls with an IQ range of 55 to 75 as measured on the S-B or the WISC and verified on the Goodenough. The IQ range gaps between the groups were purposely set to maximize the differences in the results.

All the subjects were nine, ten, or eleven years of age and in the fourth, fifth, or sixth grade, or in a special class. Each of the three groups was composed of ten, nine-year old girls; ten, ten-year old girls; and ten, eleven-year old girls. This grouping provided for a balanced design with an "n" of 10 per cell which provided for clearcut results (rather than indications of trends) in analysis of variance designs as was used here. This type of grouping also provided for more ease in calculating the statistics. Further, this particular chronological age range was used because the literature indicated that children developed conceptual abilities better and more quickly at and after nine years of age. There were also practical advantages in testing this age group. The subjects were cooperative, could follow instructions readily, were able to write, and had a recent IQ on their school record at this age. It was of special interest to this author to study conceptual abilities in young girls since much of the research already carried out had been done with boys.

In summary, the ninety girls used as subjects in this study had the following characteristics:

1. All the subjects were girls attending schools in the Oklahoma public school system. The thirty normal and the thirty bright subjects were enrolled in regular classes in the Norman, Oklahoma, public school system whereas the thirty subnormal subjects were enrolled in special classes in the Oklahoma City public school system.
2. All subjects were from lower-middle to middle-middle socio-economic level families.
3. All subjects were free of visual or hearing difficulties, physical handicaps, or severe emotional problems.
4. All subjects were Caucasian, Protestant, and had always attended Oklahoma schools.
5. All the subjects were from nine to eleven years of age. Each of the three groups of bright, normal, and subnormal subjects were made up of ten, nine-year old girls; ten, ten-year old girls; and ten, eleven-year old girls.
6. Group I was made up of thirty bright girls and showed an IQ range of 125 and above.
7. Group II was made up of thirty normal girls and showed an IQ range of 90 to 110.
8. Group III was made up of thirty subnormal girls and showed an IQ range of 55 to 75.

Administration of the Tests

There were two sections to the procedure since both a group and an individual instrument were used. The group instrument, the Pikunas Graphoscopic Scale, was administered first. The individual instrument was Stephens' categorization tasks.

The PGS was administered to six to fifteen Ss at a time. The subnormals were tested in the smaller groups because of their need for individual instruction and attention. The subjects were told to report to the testing room at a particular pre-arranged time. The room was large and physically arranged so that there were safe distances between the seats to prevent copying. One or two proctors, depending on the size of the group being tested, aided the examiner.

The following procedure was used in the administration of the PGS:

1. The subjects were each given a lead pencil, a box of color pencils with black, blue, green, red and yellow colors, a PGS blank and a sheet of unlined, blank paper.

2. The subjects were asked to complete the Goodenough Draw-a-Person Test by drawing a person on the blank sheet of paper given to them. These were completed, signed, and turned in to the examiner.

3. The subjects were instructed that this task had no bearing on their grades, that it was important for them to have and use their own ideas, and to be silent during the testing session. They were told to raise their hand if they had questions, and that the examiner would go to their desk to answer their questions. They were also told that extra lead and color pencils were available.

4. The subjects were instructed to write their name, address, school, grade, teacher, age, sex, and birth date on their PGS blank.

5. Then, the following instructions were given:

> "This is a drawing task. (The examiner shows the test blank). Look at these different squares that have different marks on them. In the first one there's a child and a dog. Try to finish it. Then look at each of these marks (point). They make you think of something to draw. You can draw anything you like, and start at any space you want, but I want you to number the squares in order as you draw in each one (demonstrate). Try to make these marks (point again to some of them) a part of your drawings. Use any of the five color pencils I have given you. After you have completed the drawings which should include the square at the bottom (#10), raise your hand, and I will nod 'yes' for you to come up to my desk with your box of color pencils and your drawing test.--Remember now, draw as well as you can."

6. Once at the examiner's desk, the subject was given further instructions as follows:

> "Well, now write down the name of each thing you drew in each space and then answer in writing the question: 'What else would you like to draw?'"

The subject was aided in writing and spelling if necessary. The instructions under Step 5 above were given with blackboard demonstrations when necessary, and the instructions were given individually when the subject failed to grasp them when administered to the group. This inability was most common for the subnormal Ss.

Once the test was completed, the examiner expressed her appreciation to the subject, and told her she could return to her classroom.

A time record was kept on each subject, covering that period from when the examiner said "go" to the whole group until the subject brought her completed drawing test to the examiner's desk.

Appendix B is a record sheet for data collected on the PGS.

For the administration of the Stephens' categorization tasks, each subject was tested individually in a quiet, well-lighted room. The subject was seated across the table from the examiner and administered the Stephens' categorization tasks with a modified procedure version from the Stephens (1963, pp. 36-37) and Van Osdol (1964, pp. 43-44) studies. The following procedure was used:

1. The examiner placed four, small, red identical vehicles in a row in front of the subject and said, "I want you to look at these little cars and I will show you what we are going to do. This will help you to understand what I want you to do. These cars are all alike. They are all red. They are shaped alike. In real life they all have a driver, an engine, and four black wheels. So you can see they are all alike."

   Three red cars were then removed and three different vehicles were aligned next to the remaining red car. The subject was then asked, "Are these all alike?" The subject usually answered by saying <u>no</u> or by shaking her head negatively. Sometimes one would volunteer to point out the differences. The examiner then stated, "No, these are not alike. They are different colors--red, blue, yellow, green--and they are shaped differently; but even though they don't look alike they still are alike in some ways. They each have an engine, a driver, and four wheels. They can be driven down the street. They burn gasoline. So, you can see that things can be alike in many ways; they may look alike; they may smell alike in many ways; they may feel alike; or they may do things alike."

   The cars were removed and the subject was given the four black plastic checkers, and the examiner said, "Now, I'll show you what we are going to do. I have some pictures on these cards of lots of things" (Van Osdol, 1964, p. 43). "On each card some of the things go together because they are most alike. We're going to look at each card, and put the checkers on the things which are most alike. I'll show you what I mean with the first two cards."

2. The examiner presented each sample card, and aided the subject, when necessary, in the correct solution, each time verbalizing the correct category following correct placement of the checkers.

3. The examiner then presented the first test card, saying: "Let's do this one. Which of these are most alike?" The subject's response and the time required to reach it were recorded. Then the examiner asked: "How or why are those most alike?" The subject's responses and the time she required to reach them were recorded.

4. The instructions presented in item 3 just above were given for each of the twenty-five items.

5. After the unstructured administration was completed, each card was presented again in the structured condition, wherein the examiner structured the situation by specifying the category which the subject should employ. Each card was placed before the subject and she was asked: "Which ones are the same color?", etc., naming the category for each card, until all cards had been tried by the subject. A time response was also recorded on this phase of the testing (Stephens, 1963, pp. 36-37).

After the completion of these three different categorization tasks, the subject was complimented on her effort, thanked, and permitted to return to her classroom.

Appendix C is a record sheet for data collected on the Stephens' categorization tasks.

Obtained Data

Data related to the personal history of each subject was obtained as well as test data. For screening purposes, it was mandatory to know the name, address, sex, age, grade, school, teacher, race, religion, socio-economic level, birth date, previous mental ability test scores, information related to adequacy of vision, hearing, adjustment, and the absence of physical handicaps.

The test data included a human figure drawing (Goodenough Test) from each subject and a PGS protocol with a time response. It included the subject's responses to Stephens' unstructured categorization tasks, the naming responses, the responses to the structured categorization tasks, and the time required to arrive at each of these responses.

The procedures described above enabled the examiner to evaluate her test data from the bright, normal, and subnormal subjects on their conceptualizing abilities and differences.

CHAPTER IV

ANALYSIS OF THE DATA

The general purpose of this study was to compare and obtain information on the categorization abilities of normal, subnormal, and bright, pre-adolescent girls. Two instruments were used to test the assumptions made in Chapter II: the Pikunas Graphoscopic Scale and the Stephens' categorization tasks.

Specifically, the examiner gathered information making comparisons between normal and subnormal, normal and bright, and subnormal and bright groups of Ss possible. These comparisons were all related to and dependent upon ability to use categories that had been observed to be of importance in everyday intellectual functioning (Stephens, 1963). The examiner was concerned with the possible relationship that could exist between the functioning of the aforementioned groups on Stephens' tasks and the PGS. It appeared (from the pilot study) that the collection of data on these IQ groups from the two instruments could show a relationship between successful identification and use of categories on the Stephens' tasks and the number of content categories employed in responding to the PGS cues. In other words, it was assumed that the bright Ss would have the greater number of successes on the Stephens' unstructured and naming categorization tasks and that these Ss would also make use of the most types of content in comparison to normal and subnormal Ss in responding to the PGS cues. The normal Ss would be next in frequency of successes on the Stephens' tasks and variety of content employed on the PGS, and the subnormal Ss would be the least successful on the Stephens' tasks along with the least number of kinds of content produced on the PGS.

The next three hypotheses were concerned with the performance of the subnormal and normal, normal and bright, and bright and subnormal groups on the Stephens' unstructured categorization tasks (determining the correct category for organizing pictures of items), the naming tasks (giving the correct name to the category just successfully employed in the unstructured situation), and the structured categorization tasks (finding of the correct pictures identifying the category just named by the examiner). It was assumed that the groups, as listed immediately above, would differ significantly from one another on these three tasks.

The next three hypotheses were concerned with total time differences by the various IQ groups on the three Stephens' tasks.

The remainder of the assumptions were concerned with the content categories produced by the different IQ or age groups on the PGS. These will be discussed specifically in the latter part of the chapter.

It was proposed that the data collected on the categorization tasks and on the PGS would help explain the traditional assumption of simplified thinking patterns of the subnormals versus the complex thinking patterns of bright individuals. It was logical to assume that the child who was least able to succeed on the Stephens' tasks and that these Ss would also make use of the most types of content in comparison to normal and subnormal Ss in responding to the PGS cues. The normal Ss would be next in frequency of successes on the Stephens' tasks and variety of content employed on the PGS, and the subnormal Ss would be the least successful on the Stephens' tasks along with the least number of kinds of content produced on the PGS.

The next three hypotheses were concerned with the performance of the subnormal and normal, normal and bright, and bright and subnormal groups on the Stephens' unstructured categorization tasks (determining the correct category for organizing pictures of items), the naming tasks (giving the correct name to the category just successfully employed in the unstructured situation), and the structured categorization tasks (finding of the correct pictures identifying the category just named by the examiner). It was assumed that the groups, as listed immediately above, would differ significantly from one another on these three tasks.

The next three hypotheses were concerned with total time differences by the various IQ groups on the three Stephens' tasks.

The remainder of the assumptions were concerned with the content categories produced by the different IQ or age groups on the PGS. These will be discussed specifically in the latter part of the chapter.

It was proposed that the data collected on the categorization tasks and on the PGS would help explain the traditional assumption of simplified thinking patterns of the subnormals versus the complex thinking patterns of bright individuals. It was logical to assume that the child who was least able to succeed on the Stephens' tasks and responded with a very narrow range of content categories on the PGS would be the least able to interpret and evaluate his experiential world effectively, and that conversely, the more successful on Stephens' tasks and the greater the variety of content produced on the PGS, the more effectively would the individual be able to function and interrelate with his environment.

Specifically, Stephens' unstructured categorization tasks reflect a level of effectiveness in applying categories independently in one's experiential framework. Stephens' naming tasks are an index of the extent to which a category is developed and delineated. Stephens' structured categorization tasks serve as an index of one's stock or repertory of categories. The content produced on the PGS is reflective of one's stock of categories and one's ability to draw upon this repertory of categories and make effective use of the categories in a meaningful whole as stimulated by a semi-structured cue. In other words, the child taking the PGS must possess a stock of categories, observe the perceptual cues, and form the appropriate associations to develop the concepts for drawing. On Stephens' tasks, the child acts similarly since he must possess a stock of categories (as indicated by the structured categorization tasks), be able to draw upon these for determining the appropriate category for organizing pictures of items by use of picture association cues (as indicated by the unstructured categorization task), and the category employed must be developed and delineated enough such that the subject makes meaningful use of it in his experiential world (as indicated by Stephens' naming tasks). In verbalizing the response category name, the child gives external proof of the development of the category. Examination of the types of errors made in naming the categories, the quality and nature of the S's verbalizing in arriving at the appropriate category name, etc., can serve as clues to differences in the thinking apparatus of bright, normal, and subnormal children. In addition, the PGS gives added clues to differences between individuals in observation of stimulus-cue incorporation; lightness, continuity and completeness; relation to standards; popularity of contents; disproportion; recurrences; disorientation, lack of elaboration and continuity; heaviness of the drawings; specified movement; use of third dimension; symbolic and abstract content, etc. Only some of these clues will be discussed, however, since they are not part of the objectives of the main study but might be suggestive for further research.

Three groups of pre-adolescent girls were included in this study. Each group consisted of ten, nine-year olds; ten, ten-year olds; and ten, eleven-year olds. The Ss were all from the lower-middle to middle-middle socio-economic level homes. They were Caucasian and Protestant and had always attended Oklahoma schools. They were free of gross physical handicaps, had good hearing and vision, and were free of any severe psychopathology. The normal and bright girls were all in the fourth, fifth, or sixth grade in a regular classroom. The subnormal girls were enrolled in the special education classes. The bright group (Group I) had an IQ range of 125 and above as measured on the California Test of Mental Maturity (CTMM) and verified on the Goodenough Draw-a-Person Test. The normal group (Group II) had an IQ range of 90 to 110 as measured on the CTMM and verified on the Goodenough, and the subnormal group (Group III) had an IQ range of 55 to 75 as measured on the Stanford-Binet or WISC and verified on the Goodenough. These three groups were similar except for their measured intellective levels.

Statistical Treatment

The Pearson product-moment correlations were used to evaluate Hypotheses 1 and 2 (Walker and Lev, 1953, p. 234). A Pearson correlation value is a measure of the linear relationship between a specific set of paired X and Y scores. In this case, the X scores are the number of content categories used by members of an IQ group on the PGS and the Y scores are the number of correct responses achieved by these members on either the unstructured or naming category tasks. Degree of freedom for table reference is equal to N - 2 where N stands for the number of pairs of scores. Our correlations use 30 pairs of scores so require 28 degrees of freedom which show a significant value of .361 for the .05 level and a value of .463 for the .01 level of confidence. Whenever r is equal to or greater in absolute size than the appropriate significant value, regardless of whether r is negative or positive, we can conclude that it is significant at the level of confidence we are using.

Analyses of variance were used to evaluate Hypotheses 3 through 8 and 9d. The computational measures and formula used to arrive at the final expression of the significance or insignificance of our F-values are found in Chapter 9 of Walker and Lev's <u>Statistical Inference</u> (1953) book. The analysis of variance technique is simply a method which provides an objective criterion for deciding whether the variability between groups is large enough in comparison with the variability within groups to justify the inference that the means of the populations from which the different groups were drawn are not all the same. The manner of tabular report for the analysis of variance results on the several hypotheses follow the usual procedure of giving the source of variance, sum of squares, degrees of freedom (df), mean square (MS), F-value, and probability (p). F-values require two df values because the degree of significance of any obtained F depends upon two factors, the number of groups and the number of Ss within the groups. Most of the tables presented below make use of 2df for the MS in the numerator because three groups are being compared and 87df for the mean square in the denominator because 90 Ss are being used. The null hypotheses were rejected when results yielded F-values that were 3.10 or larger for the .05 level of confidence and 4.85 or larger for the .01 level of confidence with 2 and 87df.

Hypotheses 9, 9c, and 10 were evaluated by means of t-tests. This statistical procedure is described fully under the result section for Hypothesis 9 on pages 73-75.

Linear trend analyses were used to evaluate Hypotheses 9a, 9b, 9e, 10a, and 10b. The precise procedure for the analyses are described by Edwards' (1960) in Chapter 14 of Experimental Design in Psychological Research. The manner of tabular report for linear trend analyses is as described above for the analyses of variance.

The several hypotheses are evaluated in order. The data analyses showing the acceptance or rejection of the hypotheses at the .05 or .01 levels of significance are presented in tabular form and in the text of this chapter.

Relationship of Performance on the PGS and on Stephens' Unstructured Categorization Tasks

Pearson product-moment correlations between frequency of successes on the Stephens' unstructured categorization tasks and the number of different content categories used on the PGS were carried out for bright, normal, and subnormal groups of subjects.

Hypothesis 1 assumed that there would be no relationship between the number of successes attained by each group of subnormal, normal, and bright Ss on Stephens' unstructured categorization tasks and the number of different content categories these respective groups employed in response to the perceptual cues on the PGS. Table 6 below shows that there is a significant relationship between the performance of the subnormals on Stephens' unstructured categorization tasks and their performance on the PGS. In other words, it is true that subnormal subjects used fewer categories in responding to the perceptual cues on the PGS just as they used fewer categories successfully on Stephens' unstructured categorization tasks. This correlation, however, did not hold true for the performance of the bright and normal Ss on these two instruments. Hypothesis 1 was rejected only in part since only the coefficient of correlation, r, for the subnormal Ss, which was .402, was significant at the .05 level of confidence with 28df. The r value for the normal Ss was .013 and for the bright Ss it was .060, neither of which was significant at the .05 level with 28df. These results are summarized in Table 6 below.

Table 6

Correlation of Number of Content Categories Used on the PGS and the Number of Correct Responses Obtained on Stephens' Unstructured Categorization Tasks
(n = 30)

	Bright		Normal		Subnormal	
	Oontent Cat. Used	Correct Responses	Content Cat. Used	Correct Responses	Content Cat. Used	Correct Responses
	X	Y	X	Y	X	Y
1.	6	17	7	17	7	16
2.	6	16	8	14	6	10
3.	4	17	5	19	5	12
4.	4	20	6	17	6	16
5.	4	22	7	15	4	12
6.	4	21	6	18	2	7
7.	8	17	6	18	5	18
8.	7	17	5	16	2	14
9.	7	20	5	13	2	9
10.	5	22	7	15	6	10
11.	7	14	6	16	1	9
12.	7	17	6	10	2	14
13.	7	17	7	14	4	7
14.	8	17	6	17	5	5
15.	6	20	7	13	5	10
16.	7	18	7	19	3	14
17.	8	20	8	15	3	14
18.	5	14	4	13	6	15
19.	5	20	5	19	2	7
20.	6	15	7	15	5	9
21.	6	19	6	15	6	20
22.	4	21	6	16	5	12
23.	3	22	7	13	7	15
24.	5	22	6	14	5	6
25.	6	16	6	16	5	13
26.	6	16	5	15	2	6
27.	5	17	8	20	4	16
28.	5	20	6	20	6	15
29.	5	18	7	16	2	9
30.	5	16	7	16	6	11

Bright	Normal	Subnormal
$\Sigma X = 171$ $\Sigma Y = 548$	$\Sigma X = 189$ $\Sigma Y = 474$	$\Sigma X = 129$ $\Sigma Y = 351$
$\Sigma X^2 = 1027$ $\Sigma Y^2 = 1018$	$\Sigma X^2 = 1219$ $\Sigma Y^2 = 7648$	$\Sigma X^2 = 645$ $\Sigma Y^2 = 4541$
$2\Sigma XY = 6164$	$2\Sigma XY = 5974$	$2\Sigma XY = 3178$
$\Sigma XY = 3082$	$\Sigma XY = 2987$	$\Sigma XY = 1589$
$r = +.060$	$r = +.013$	$r = +.402$
$p > .05$	$p > .05$	$p < .05$

Relationship of Performance on the PGS and on Stephens' Naming Tasks

Hypothesis 2 assumed that there was no relationship between frequency of successes on Stephens' naming tasks by subnormal, normal, and bright subjects and the number of different content categories these respective groups employed in response to the perceptual cues on the PGS. Pearson product-moment correlations demonstrated that there was a relationship between the performance on Stephens' naming tasks and performance on the PGS. There was a positive correlation for the subnormal subjects and the bright subjects on these two tools, so that Hypothesis 2 was rejected in part as was Hypothesis 1. The r for the subnormals was .469, making it possible to reject Hypothesis 2 at the .01 level with 28 df. The r for the brights was .430, making it possible to reject Hypothesis 2 at the .05 level. The r for the normals was .188 which is not significant at the .05 level and thereby makes Hypothesis 2 accepted for that IQ group. These results are summarized in Table 7 below.

Results on Stephens' Categorization Tasks

The results for Hypotheses 3, 4, and 5 on the performance of bright, normal, and subnormal Ss on Stephens' categorization tasks and Hypotheses 6, 7, and 8 on time comparisons by these groups on Stephens' tasks will be presented in the next section of this chapter.

Table 7

Correlation of Number of Content Categories Used on the PGS and the Number of Correct Responses Obtained on Stephens' Naming Tasks
(n = 30)

	Bright		Normal		Subnormal	
	Content Cat. Used	Correct Responses	Content Cat. Used	Correct Responses	Content Cat. Used	Correct Responses
	X	Y	X	Y	X	Y
1.	6	14	7	14	7	15
2.	6	14	8	14	6	8
3.	4	17	5	18	5	10
4.	4	18	6	15	6	12
5.	4	20	7	13	4	9
6.	4	18	6	17	2	5
7.	8	16	6	17	5	5
8.	7	15	5	14	2	10
9.	7	19	5	11	2	7
10.	5	21	7	12	6	8
11.	7	14	6	14	1	5
12.	7	16	6	6	2	11
13.	7	14	7	10	4	5
14.	8	15	6	13	5	5
15.	6	20	7	11	5	8
16.	7	16	7	14	3	11
17.	8	17	8	13	3	12
18.	5	13	4	13	6	11
19.	5	18	5	16	2	6
20.	6	12	7	13	5	6
21.	6	17	6	13	6	18
22.	4	21	6	12	5	11
23.	3	21	7	9	7	13
24.	5	21	6	14	5	6
25.	6	14	6	16	5	10
26.	6	15	5	14	2	4
27.	5	14	8	16	4	12
28.	5	19	6	18	6	12
29.	5	15	7	16	2	3
30.	5	13	7	13	6	6
	$\Sigma X = 171$	$\Sigma Y = 497$	$\Sigma X = 189$	$\Sigma Y = 404$	$\Sigma X = 129$	$\Sigma Y = 264$
	$\Sigma X^2 = 1027$	$\Sigma Y^2 = 8451$	$\Sigma X^2 = 1219$	$\Sigma Y^2 = 5642$	$\Sigma X^2 = 645$	$\Sigma Y^2 = 2694$
	$2\Sigma XY = 5574$		$2\Sigma XY = 5062$		$2\Sigma XY = 2442$	
	$\Sigma XY = 2787$		$\Sigma XY = 2531$		$\Sigma XY = 1221$	
	$r = +.430$		$r = -.188$		$r = +.469$	
	$p < .05$		$p > .05$		$p < .01$	

An Overall Evaluation of Hypotheses 3, 4, and 5

The author deemed it necessary to analyze Hypotheses 3, 4, and 5 as a group in a two-way classification, 3 x 3 analysis of variance for repeated measures as well as individually by analyses of variance for simple effects. These hypotheses assumed that the bright, normal, and subnormal groups of subjects would not differ significantly from one another on Stephens' unstructured tasks, his naming tasks, and the structured tasks. The raw score data making this analysis possible is seen in Table 8 below.

Table 8

Total Numbers of Correct Responses on Stephens' Unstructured, Structured, and Naming Categorization Tasks by Bright, Normal, and Subnormal Girls
(n = 30)

	Categorization Tasks		
	Unstructured	Structured	Naming
Bright	548	721	497
Normal	474	677	409
Subnormal	351	587	264

The results of the analysis of variance are presented in Table 9. The F-value of 59.03 for the Intelligence Level is significant at the .01 level with 2, 87df. This result means that the bright, normal, and subnormal Ss differ significantly from one another in their abilities to achieve correct responses on the PGS. The F-value 245.04 for categorization tasks is significant at the .01 level with 2, 87df. This finding means that the unstructured, structured, and naming tasks differ significantly from one another so that we can say that they are not equally able of eliciting correct responses. Simply speaking, the means of both of these variables are not equal and therefore permit us to reject null Hypotheses 3, 4, and 5.

Table 9

Analysis of Variance for Data on Stephens' Categorization Tasks by Bright, Normal, and Subnormal Girls

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	1809.99	2	904.99	59.03**
(T - 1): Categorization Task	3999.92	2	1999.60	245.04**
I(R - 1): Error Term	1333.78	87	15.33	
(I - 1)(T - 1): Intelligence Level x Categorization Task (Interaction)	84.06	4	21.01	2.50*
I(R - 1)(T - 1): Error Term x Task	1420.02	174	8.16	
Total	8648	269		

** = Significant at .01 level
* = Significant at .05 level

The hypothesis that the interaction of these two variables (Intelligence Level x Categorization Task) is zero is rejected at the .05 level with an F-value of 2.50 and 4, 87df. This fact would force us to conclude that the differential effects of the experimental conditions are not due to one another independently, but rather to the interaction effects of the variables. It should be clear that this interactive component needs further investigation since interaction is generally a result of something above and beyond the factors involved.

The author is not convinced of the possible meaningfulness of this interaction effect since it is significant only at the .05 level. Until further investigation proves different, it seems cautious and wise to retain the null hypothesis of no significant interaction between the two variables. This decision appears justified since the F-value was not significant at the .01 level and the F-values concerned were most significant.

Performance of Bright, Normal, and Subnormal Subjects on Stephens' Unstructured Categorization Tasks

Hypothesis 3 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of Ss would not differ significantly from one another in the number of correct responses they achieved in response to Stephens' unstructured categorization tasks. This hypothesis was rejected at the .01 level of significance with an F-value of 21.53 with 2, 87df as determined by an analysis of variance test for simple effects (Walker and Lev, 1953, ch. 9; Edwards, 1960, ch. 12). The data analysis for Hypothesis 3 is presented in Table 10 below. The raw score totals of number of correct responses obtained on the unstructured tasks were 548 for the bright Ss, 474 for the normal Ss, and 351 for the subnormal Ss as seen in Table 11. The maximum possible raw score total for each vs.

Table 10

Analysis of Variance for Bright, Normal, and Subnormal Subjects on Stephens' Unstructured Categorization Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	660.16	2	330.08	21.53**
I(R - 1): Error Term		87	15.33	

** = Significant at .01 level

Table 11

Raw Scores, Means, and Medians Found for Brights, Normals, and Subnormals on Number of Correct Responses To Stephens' Unstructured Categorization Tasks (n = 30)

	Number of Correct Responses on Unstructured Tasks		
	Brights	Normals	Subnormals
1.	17	17	16
2.	16	14	10
3.	17	19	12
4.	20	17	16
5.	22	15	12
6.	21	18	7
7.	17	18	18
8.	17	16	14
9.	20	13	9
10.	22	15	10
11.	14	16	9
12.	17	10	14
13.	17	14	7
14.	17	17	5
15.	20	13	10
16.	18	19	14
17.	20	15	14
18.	14	13	15
19.	20	19	7
20.	15	15	9
21.	19	15	20
22.	21	16	12
23.	22	13	15
24.	22	14	6
25.	16	16	13
26.	16	15	6
27.	17	20	16
28.	20	20	15
29.	18	16	9
30.	16	16	11
Totals	548	474	351
Means	18.27	15.80	11.70
Medians	17.50	16.17	12.33

group was 750. Table 11 also shows the measures of central tendency for each of these sets of group scores. The means (average number of correct responses) were 18.27 for the brights, 15.80 for the normals, and 11.70 for the subnormals. The medians were 17.50 for the brights, 16.17 for the normals, and 12.33 for the subnormals. The modes (most frequent scores) were 17 for the brights, 15 and 16 (bimodal) for the normals, and 9 and 14 (bimodal) for the subnormals.

Performance of Bright, Normal, and Subnormal Subjects on Stephens' Naming Tasks

Hypothesis 4 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of Ss would not differ significantly from one another in the number of correct responses they achieved in response to Stephens' naming tasks. This hypothesis was rejected at the .01 level of significance with an F-value of 30.09 with 2, 87df as determined by an analysis of variance for simple effects. The data analysis for Hypothesis 4 is presented in Table 12 below. The raw score totals for number of correct naming responses

Table 12

Analysis of Variance for Bright, Normal, and Subnormal Subjects on Stephens' Naming Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	922.86	2	461.43	30.09**
I(R - 1): Error Term		87	15.33	

** Significant at .01 level

Table 13

Raw Scores, Means, and Medians Found for Brights, Normals and Subnormals on Number of Correct Responses Obtained in Naming The Completed Unstructured Tasks
(n = 30)

	No. of Correct Responses Obtained on the Naming Tasks		
	Brights	Normals	Subnormals
1.	14	14	15
2.	14	14	8
3.	17	18	10
4.	18	15	12
5.	20	13	9
6.	18	17	5
7.	16	17	5
8.	15	14	10
9.	19	11	7
10.	21	12	8
11.	14	14	5
12.	16	6	11
13.	14	10	5
14.	15	13	5
15.	20	11	8
16.	16	14	11
17.	17	13	12
18.	13	13	11
19.	18	16	6
20.	12	13	6
21.	17	13	18
22.	21	12	11
23.	21	9	13
24.	21	14	6
25.	14	16	10
26.	15	14	4
27.	14	16	12
28.	19	18	12
29.	15	16	3
30.	13	13	6
Totals	497	409	264
Means	16.57	13.57	9.63
Medians	16.67	14.14	8.50

achieved by the three groups were as follows: 497 by the bright Ss, 409 by the normal Ss, and 264 by the subnormal Ss as seen in Table 13. The maximum possible raw score total for each group was 750. Table 13 also shows the measures of central tendency for each of these set of group scores. The means were 16.57 for the brights, 13.57 for the normals, and 9.63 for the subnormals. The medians were 16.67 for the brights, 14.14 for the normals, and 8.50 for the subnormals. The modes were 14 for the brights, 13 and 14 (bimodal) for the normals, and 5 for the subnormals.

Performance of Bright, Normal, and Subnormal Subjects on Stephens' Structured Categorization Tasks

Hypothesis 5 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of Ss would not differ significantly from one another in the number of correct responses they achieved in response to Stephens' structured tasks. This hypothesis was rejected at the .01 level of significance with an F-value of 10.14 with 2, 87df as determined by an analysis of variance for simple effects. The data analysis for Hypothesis 5 is presented in Table 14 below. The

Table 14

Analysis of Variance for Bright, Normal, and Subnormal Subjects on Stephens' Structured Categorization Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	311.03	2	155.51	10.14**
I(R - 1): Error Term		87	15.33	

** Significant at .01 level

Table 15

Raw Scores, Means, and Medians Found for Brights, Normals, and Subnormals on Number of Correct Responses Obtained on Stephens' Structured Categorization Tasks (n = 30)

	Number of Correct Responses Obtained on the Structured Tasks		
	Brights	Normals	Subnormals
1.	25	24	22
2.	22	23	18
3.	25	24	21
4.	23	24	21
5.	24	21	21
6.	23	23	17
7.	23	24	23
8.	23	22	21
9.	25	20	17
10.	24	22	16
11.	25	21	18
12.	24	22	21
13.	25	22	16
14.	25	23	19
15.	24	20	18
16.	24	21	14
17.	20	23	22
18.	25	21	20
19.	25	22	18
20.	25	22	22
21.	24	20	23
22.	25	24	20
23.	24	25	24
24.	25	24	21
25.	24	24	23
26.	22	24	11
27.	25	24	21
28.	25	23	20
29.	25	22	19
30.	23	23	20
Totals	721	677	587
Means	24.20	22.57	19.07
Medians	24.88	23.17	20.75

raw score totals for the number of correct responses to the structured categorization tasks were 721 by the bright Ss, 677 by the normal Ss, and 587 by the subnormal Ss as seen in Table 15. The maximum possible raw score total for each group was 750. Table 15 also shows the measures of central tendency for each of these sets of group scores. The means were 24.20 for the brights, 22.57 for the normals, and 19.07 for the subnormals. The medians were 24.88 for the brights, 23.17 for the normals, and 20.75 for the subnormals. The modes were 25 for the brights, 24 for the normals, and 21 for the subnormals.

Total Time Differences between Bright, Normal and Subnormal Subjects on Stephens' Unstructured Categorization Tasks

Hypothesis 6 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of Ss would not differ significantly from one another on their total times for responses to Stephens' unstructured categorization tasks. This hypothesis was accepted since an F-value of .325 with 2, 87df was not significant at the .05 level as determined by a 1 x 3 analysis of variance test for simple effects. The total times in seconds for completion of Stephens' unstructured categorization tasks for each subject is given in Table 16 below. This table also presents the total times in seconds and in minutes for each group for the unstructured tasks as well as the mean time scores in seconds and in minutes for each group. These times are as follows: 13,667 seconds, or 227.8 minutes for total task time with a mean of 455.6 seconds, or 7.6 minutes for the brights; 12,850 seconds,

Table 16

Raw Scores in Seconds and Means for Total Times in Seconds and Minutes for Bright, Normal, and Subnormal Subjects on Stephens' Unstructured Categorization Tasks (n = 30)

	Brights	Normals	Subnormals
1.	248	516	910
2.	316	406	431
3.	472	343	342
4.	322	349	342
5.	326	499	425
6.	486	305	588
7.	425	403	553
8.	445	546	360
9.	471	256	365
10.	317	382	692
11.	618	402	259
12.	726	564	407
13.	477	416	422
14.	594	534	659
15.	492	305	498
16.	666	383	456
17.	372	433	417
18.	467	565	464
19.	494	372	413
20.	589	547	364
21.	582	437	425
22.	835	378	658
23.	318	283	591
24.	435	247	311
25.	300	428	386
26.	550	501	342
27.	331	426	363
28.	329	353	378
29.	334	826	375
30.	330	445	237
Total	13,667 seconds or 227.8 min.	12,850 seconds or 214.2 min.	13,433 sec. or 233.9 min.
Means	455.6 seconds or 7.6 min.	428.3 seconds or 7.1 min.	447.8 sec. or 7.5 min.

or 214.2 minutes for total task time with a mean of 428.3 seconds, or 7.1 minutes for the normal subjects; and 13,433 seconds, or 233.9 minutes for total task time with a mean of 447.8 seconds, or 7.5 minutes for subnormal subjects. The data analysis for Hypothesis 6 is presented in Table 17 below.

Table 17

Analysis of Variance on Total Time for Stephens' Unstructured Categorization Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	11,801.49	2	5900.74	0.33 n.s.
I(R - 1): Error Term	1,579,535.40	87	18155.51	

n.s. = Not significant

Total Time Differences Between Bright, Normal, and Subnormal Subjects on Stephens' Naming Tasks

Hypothesis 7 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of subjects would not differ significantly from one another on their total times for responses to Stephens' naming tasks. This hypothesis was rejected at the .01 level of significance with a 13.67 F-value with 2, 87df as determined by a 1 x 3 analysis of variance test for simple effects. The total times in seconds for completion of Stephens' naming tasks for each S are given in Table 18 below.

Table 18

Raw Scores in Seconds and Means for Total Times in Seconds and Minutes for Bright, Normal, and Subnormal Subjects on Stephens' Naming Tasks
(n = 30)

	Brights	Normals	Subnormals
1.	74	112	273
2.	90	114	207
3.	78	136	129
4.	69	121	148
5.	71	236	182
6.	93	141	297
7.	124	238	127
8.	141	252	138
9.	71	102	111
10.	74	214	100
11.	134	149	79
12.	110	309	103
13.	117	134	98
14.	146	250	113
15.	114	181	93
16.	117	249	105
17.	92	122	83
18.	118	190	146
19.	101	153	122
20.	88	180	161
21.	106	138	77
22.	110	129	74
23.	108	132	99
24.	88	147	110
25.	162	113	86
26.	146	102	96
27.	80	122	92
28.	78	125	108
29.	79	226	118
30.	92	134	97
Totals	3,071 seconds or 51.2 min.	4,951 seconds or 82.5 min.	3,772 seconds or 62.9 min.
Means	102.4 seconds or 1.7 min.	165.0 seconds or 2.6 min.	125.7 seconds or 2.1 min.

This table also presents the total times in seconds and in minutes for each group for the naming tasks as well as the mean time scores in seconds and in minutes for each group. These are 3,071 seconds, or 51.2 minutes for total task time-with a mean of 102.4 seconds, or 1.7 minutes for the bright subjects; 4,951 seconds, or 82.5 minutes for total task time with a mean of 165.0 seconds, or 2.6 minutes for the normal Ss; and 3,772 seconds, or 62.9 minutes for total task time with a mean of 125.7 seconds, or 2.1 minutes for subnormal Ss. The data analysis for Hypothesis 7 is presented in Table 19 below.

Table 19

Analysis of Variance on Total Time for Stephens' Naming Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	60,176.63	2	30,088.01	13.67**
I(R - 1): Error Term	191,505.80	87	2,201.21	

** = Significant at .01 level

Total Time Differences Between Bright, Normal, and Subnormal Subjects on Stephens' Structured Categorization Tasks

Hypothesis 8 assumed that bright and normal, normal and subnormal, and bright and subnormal groups of subjects would not differ significantly from one another on their total times for responses to Stephens' structured categorization tasks. The hypothesis was rejected at the .01 level of significance with a 21.98 F-value with 2, 87df as determined by a I x 3 analysis of variance test for simple effects in Table 20 below. The total times in seconds for completion of Stephens' structured tasks for each S is presented in Table 21. This table also gives the total times in seconds and in minutes for each group for the structured tasks as well as the mean time scores in seconds and in minutes for each group. These are 4,314 seconds, or 71.9 minutes for total task time with a mean of 143.8 seconds, or 2.4 minutes for the bright Ss; 5,519 seconds, or 92.0 minutes for total task time with a mean of 184.0 seconds, or 3.1 minutes for the normal Ss; and 6,818 seconds, or 113.6 minutes for total task time with a mean of 227.3 seconds, or 3.8 minutes for subnormal Ss.

Table 20

Analysis of Variance on Total Time for Stephens' Structured Categorization Tasks

Source of Variance	Sum of Squares	df	Mean Square	F
(I - 1): Intelligence Level	104,549.35	2	52,274.67	21.98**
I(R - 1): Error Term	206,913.64	87	2,378.30	

** Significant at .01 level

Table 21

Raw Scores in Seconds and Means for Total Times in Seconds and Minutes for Bright, Normal, and Subnormal Subjects on Stephens' Structured Categorization Tasks (n = 30)

	Brights	Normals	Subnormals
1.	108	158	336
2.	175	162	186
3.	112	155	238
4.	138	138	150
5.	100	166	252
6.	141	155	347
7.	162	165	232
8.	262	259	207
9.	113	153	156
10.	101	253	294
11.	249	188	159
12.	157	230	198
13.	156	208	168
14.	146	230	193
15.	129	214	206
16.	134	166	268
17.	117	162	165
18.	147	222	326
19.	89	171	260
20.	165	223	211
21.	174	147	225
22.	145	144	265
23.	122	168	192
24.	123	160	184
25.	120	197	216
26.	195	167	270
27.	119	167	195
28.	144	120	186
29.	118	286	394
30.	153	185	139
Totals	4,314 seconds or 71.9 min.	5,519 seconds or 92.0 min.	6,818 seconds or 113.6 min.
Means	143.8 seconds or 2.4 min.	184.0 seconds or 3.1 min.	227.3 seconds or 3.8 min.

Results on the Pikunas Graphoscopic Scale

The remainder of the Hypotheses, 9 through 10b, are concerned exclusively with the second tool used in this study, the Pikunas Graphoscopic Scale.

Comparison of Bright, Normal, and Subnormal Subjects on Content Produced on the PGS

Hypothesis 9 assumed that bright, normal, and subnormal subjects would not differ significantly on the content they produced on the PGS. This hypothesis was rejected only in part. Twenty-seven t-tests comparing each group with the two others on nine different types of the most common responses elicited on the PGS were carried out as summarized in Table 22.

F-values were computed in the following fashion to determine equal or unequal variances. First, the variances (s^2) of the two groups being compared were found by the following formula:

$$s^2 = \frac{N\Sigma X^2 - (\Sigma X)^2}{N(N - 1)}$$

and then the Variance Test was applied in the following manner:

$$\text{Variance Test} = \frac{\text{Largest } s^2}{\text{Smallest } s^2} = \text{F-value}$$

If the groups compared on a particular type content, e.g., human, showed equal variances as determined by an insignificant F-value at or beyond the .05 level, the following t-test formula for equal variances was applied:

Table 22

Comparison of Bright, Normal, and Subnormal Girls on Types of Content Produced on the PGS

Type of Response	Groups Compared	F-value	t-value
Human	Brights vs. S-N	1.460	1.440
	Brights vs. Normals	1.310	2.417**
	Normals vs. S-N	1.114	0.850
Animal	Brights vs. S-N	13.658**	3.731**
	Brights vs. Normals	1.729	1.918
	Normals vs. S-N	7.900	0.706
House	Brights vs. S-N	1.643	0.541
	Brights vs. Normals	1.185	0.196
	Normals vs. S-N	1.386	0.696
Fobj.	Brights vs. S-N	2.197*	0.294
	Brights vs. Normals	2.088*	1.796
	Normals vs. S-N	1.052	1.276
Plant	Brights vs. S-N	2.697**	2.020*
	Brights vs. Normals	1.420	1.525
	Normals vs. S-N	1.899*	0.377
Clothing	Brights vs. S-N	6.828**	2.070
	Brights vs. Normals	1.773	0.665
	Normals vs. S-N	2.109**	2.395*
Food	Brights vs. S-N	2.625**	1.515
	Brights vs. Normals	2.146*	1.626
	Normals vs. S-N	1.223	0.000
Nature	Brights vs. S-N	1.133	0.287
	Brights vs. Normals	1.728	1.210
	Normals vs. S-N	1.947*	1.471
Miscellaneous	Brights vs. S-N	9.100**	3.130**
	Brights vs. Normals	1.576	4.383**
	Normals vs. S-N	5.774**	1.615

** Significant at .01 level
* Significant at .05 level

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s^2 \frac{N_1 + N_2}{N_1 N_2}}}$$

with the formula for the s^2 in the denominator being

$$s^2 = \frac{(N_1 - 1)s_1{}^2 + (N_2 - 1)s_2{}^2}{N_1 + N_2 - 2} .$$

If the variances of the two groups compared yielded a significant F-value, the following t-test formula was used:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1{}^2}{N_1} + \frac{s_2{}^2}{N_2}}} .$$

The t-test values were considered significant when the probability was larger than the previously set level of significance which was the .05 level with 58 df in this case.

As Table 22 shows, bright and normal Ss differed significantly on human content giving a t-value of 2.417 which is significant at the .01 level of significance with 58df. Bright and subnormal Ss differed significantly on animal content giving a t-value of 3.731 which is significant at the 1% level with 58df. Bright and subnormal subjects also differed significantly on plant content giving a t-value of 2.020 which is significant at the .05 level with 58df. Both brights and subnormals and normals and subnormals differed significantly on clothing content yielding t-values significant at the .05 level with 58df.

Bright and subnormal Ss and bright and normal Ss differed significantly on miscellaneous content yielding t-values significant at the .01 level with 58df.

Comparison of Bright, Normal, and Subnormal Subjects on Production of Original Responses on the PGS

Hypothesis 9a assumed that there would be an increase in original responses on the PGS with IQ increase. This hypothesis was analyzed by means of a linear trend as displayed in Table 23 and it was accepted showing an F-value of 85.64 significant at the .01 level with 1, 87df. The raw score totals for the original responses were 175 for the brights, 129 for the normals, and 67 for the subnormal Ss.

Table 23

Linear Trend Analysis Data Showing Increase of Original Responses with IQ Increase

Source of Variance	Sum of Squares	df	Mean Square	F
Linear Trend	194.40	1	194.40	85.64**
Error Term	197.84	87	2.27	

** Significant at .01 level

Comparison of Bright, Normal, and Subnormal Subjects on Production of Popular Responses on the PGS

Hypothesis 9b assumed that bright, normal, and subnormal children would differ significantly in the number of popular responses they produced on the PGS. A linear trend analysis, as summarized in Table 24 below, reveals that these children of different IQ groups do differ significantly on popular responses produced on the PGS as shown by an F-value of 4.22 significant at the .05 level with 1, 87df, and also, that the popular response increased with IQ increase. The raw score totals for the popular responses were 116 for the brights, 94 for the normals, and 64 for the subnormals.

Table 24

Linear Trend Analysis Data Showing Increase of Popular Responses on the PGS with IQ Increase

Source of Variation	Sum of Squares	df	Mean Square	F
Linear Trend	45.06	1	45.06	4.22*
Error Term	929.47	87	10.68	

* Significant at .05 level

Comparison of Bright, Normal, and Subnormal Subjects on Number of Content Categories Employed on the PGS

Hypothesis 9c assumed that children of different IQ levels did not differ significantly on the number of content categories they employed in responding to the PGS. The data for this hypothesis was analyzed by means of three t-tests between the various groups as seen in Table 25.

Table 25

Data Analysis Revealing Differences in Number of Content Categories used on the PGS by Bright, Normal, And Subnormal Subjects

Groups Compared	F-value	t-value
Brights vs. Subnormals	1.726	3.475**
Brights vs. Normals	1.849	1.981
Normals vs. Subnormals	3.192**	5.417**

** Significant at .01 level

It is clear that Hypothesis 9c must be rejected in part since a comparison of the brights with the subnormals reveals a t-value of 3.475 which is significant at the .01 level of confidence with 58df. A comparison of the normals with the subnormals also reveals a t-value of 5.417 which is significant at the .01 level with 58df thus indicating that children of different IQ levels employ different numbers of content categories in responding to the PGS.

Comparison of Bright, Normal, and Subnormal Subjects on Production of Human Responses on the PGS

Hypothesis 9d assumed that bright girls did not produce more human responses than normal and subnormal girls on the PGS. An analysis of variance using the number of human responses as the dependent measure reveals that this hypothesis must be rejected since the F-value is 4.61 which is significant at the .05 level with 1, 87df. Thus, it is true that bright girls produce more human responses than normal and subnormal girls on the PGS. The analysis is summarized in Table 26.

Table 26

Analysis of Variance Contrasting Brights versus the Normals and Subnormals on Human Responses on the PGS

Source of Variation	Sum of Squares	df	Mean Square	F
Brights versus Normals and Subnormals	9.79	1	9.79	4.61*
Error Term	184.64	87	2.12	

* Significant at .05 level

Comparison of Bright, Normal, and Subnormal Subjects on Production of Animal Responses on the PGS

Hypothesis 9e assumed that animal responses would not be produced with greater frequency with IQ increase. A linear trend analysis as summarized in Table 27 below reveals an F-value of 9.69, which is significant at the .01 level with 1, 87df, and thereby permits us to reject the hypothesis and conclude that animal responses on the PGS are produced with greater frequency with IQ increase. The raw score totals for production of animal responses were 47 for the brights, 28 for the normals, and 23 for the subnormals.

Table 27

Linear Trend Analysis Data Showing Increase of Animal Responses on the PGS with IQ Increase

Source of Variation	Sum of Squares	df	Mean Square	F
Linear Trend	9.60	1	9.60	9.69**
Error Term	86.88	87	0.99	

** Significant at .01 level

Comparison of Nine, Ten, and Eleven-Year Old Girls On Content Produced on the PGS

Hypothesis 10 assumed that girls of different ages would not differ on the type of content they produced on the PGS. This hypothesis was rejected in part. Twenty-seven t-tests comparing nine, ten, and eleven-year olds with one another as groups on nine different types of the most common responses elicited on the PGS were carried out as summarized in Table 28.

The approach to this analysis was identical to that for Hypothesis 9; hence, it is unnecessary to repeat the procedure here. The F-values were calculated to determine whether the variances of the two groups being compared were equal or unequal, thus determining the t-test formula to be used.

Table 28 reveals a t-value of 3.361 for nine versus eleven-year olds on human content; a t-value of 6.557 for nine versus ten-year olds; and a t-value of 3.448 for ten versus eleven-year olds. These values are all significant at the .01 level of confidence with 58df. These same comparisons on factory-made object (Fobj) content also reveal t-values which are significant at the .01 level with 58df. The nine vs. eleven-year olds show a t-value of 3.448 on factory-made object content. The nine versus ten-year olds show a t-value of 2.913, and the ten versus eleven-year olds show a t-value of 6.140 for factory-made object content. The only other t-value found in this series is that of nature content when the nine and eleven-year olds are contrasted. This t-value was 2.167 which is significant at the .05 level with 58df. It is the eleven-year old group that shows the greater use of the nature content. It is clear, then, that Hypothesis 10 is rejected in part since these three age groups revealed differences on human, factory-made object, and nature responses.

Table 28

Comparison of 9, 10, and 11-Year Old Girls on Types of Content Produced on the PGS

Type of Response	Groups Compared	F-value	t-value
Human	9 x 11	1.232	3.361**
	9 x 10	0.226	6.557**
	10 x 11	1.136	3.448**
Animal	9 x 11	1.389	0.176
	9 x 10	1.389	0.176
	10 x 11	1.000	0.000
House	9 x 11	1.081	1.272
	9 x 10	1.084	0.565
	10 x 11	1.171	0.740
Fobj.	9 x 11	1.463	3.448**
	9 x 10	1.069	2.913**
	10 x 11	1.564	6.140**
Plant	9 x 11	1.261	0.170
	9 x 10	1.098	0.874
	10 x 11	1.149	0.652
Clothing	9 x 11	1.026	0.181
	9 x 10	1.527	0.409
	10 x 11	1.567	0.606
Food	9 x 11	1.231	0.000
	9 x 10	1.616	0.794
	10 x 11	1.989*	0.741
Nature	9 x 11	4.469**	2.167*
	9 x 10	6.391**	1.600
	10 x 11	1.430	0.224
Miscellaneous	9 x 11	2.610**	1.988
	9 x 10	1.815	1.321
	10 x 11	1.438	0.730

** Significant at .01 level
* Significant at .05 level

Comparison of Nine, Ten, and Eleven-Year Old Girls On Production of Human Content on the PGS

Hypothesis 10a proposed that human content on the PGS would not be produced with greater frequency with age increase. A linear trend analysis summarized in Table 29 indicates that we must accept this proposal since its F-value is 1.12 which is not significant at the .05 level with 1, 87df. A quadratic trend analysis also summarized in this table reveals an F-value of 3.36 which is not significant at the .05 level with 1, 87df. The raw score totals for production of human responses were 65 for the nine-year old girls, 89 for the ten-year old girls and 77 for the eleven-year old girls.

Table 29

Linear and Quadratic Trend Analysis on Human Content on the PGS for 9, 10, and 11 Year Old Girls

Source of Variance	Sum of Squares	df	Mean Square	F
Linear Trend	2.4	1	2.4	1.12 n.s.
Quadratic Trend	7.2	1	7.2	3.36 n.s.
Error Term	186.5	87	2.14	

n.s. = Not significant

Comparison of Nine, Ten, and Eleven-Year Old Girls on
House Content on the PGS

Hypothesis 10b proposed that house content would not decrease with age increase and, as seen in Table 30 below, this hypothesis is accepted since its linear trend analysis reveals an F-value of 1.71 which is not significant at the .05 level with 1, 87df. The raw score totals for production of house content were 21 for the nine-year olds, 18 for the ten-year olds, and 14 for the eleven-year old girls. The raw score totals are diminutive with age increase, but not enough to yield a significant trend.

Table 30

Linear Trend Analysis on House Content on the PGS
For 9, 10, and 11 Year Old Girls

Source of Variation	Sum of Squares	df	Mean Square	F
Linear Trend	0.82	1	0.82	1.71 n.s.
Error Term	41.97	87	0.48	

n.s. = Not significant

The present chapter presented the results of an investigation with bright, normal, and subnormal pre-adolescent girls concerning their categorization abilities.

CHAPTER V

SUMMARY AND CONCLUSIONS

The general purpose of the present study was to shed more light on the thinking patterns of children of different intellectual capacities. More specifically, the author was concerned with the categorization abilities of pre-adolescent girls of different IQs and how categorization relates to intellectual functioning. Categorizing abilities were assessed by means of two testing instruments: the Pikunas Graphoscopic Scale (PGS) and the Stephens' categorization tasks. The Stephens' instrument had formerly displayed its discriminating ability for normal and subnormal boys (Stephens, 1963). The PGS is a drawing task based on the ability to draw upon one's conceptual categories and employ these meaningfully in response to semi-structured perceptual cues. The PGS had not been used for exploring categorization abilities as such, but both these tools rely on means of self-expression and communication and were therefore readily viewed as related and capable of reflecting various facets of categorizing ability. The types of content produced on the PGS reflect categories and consequently inner functioning (Pikunas, 1959) as do Stephens' tasks. On the Stephens' categorization tasks, bright, normal, and retarded Ss were required to perform unstructured (S required to use categories to organize pictures of items), naming (S required to verbalize the category name just employed), and structured (S required to identify the pictures representing the category after the examiner has named the category) categorization tasks. The degree of success on these tasks by the Ss from the different IQ levels was reflective of their modes of intellectual operation. The author proceeded from the assumption that children with different IQs would display varying abilities in speed of responding, in the number of categories they would use correctly, in the stock of categories they possessed, in being able to make meaningful use of the categories, and in verbalizing the categories employed.

The Stephens' categorization tasks make use of twenty-seven categories common to everyday functioning. Each category is represented by seven randomly arranged pictures on a rectangular card. Four of these pictures depict the category having common properties. The other three pictures have no common associative cues with the category concerned. The subject is responsible for identifying the four pictures representing the category being tested.

Three groups of Ss were compared in the present study. Each group was comprised of ten, nine-year olds; ten, ten-year olds; and ten, eleven-year old girls giving us thirty Ss per group. Each group reflected an IQ level. The bright group showed IQs of 125 and above on the California Test of Mental Maturity (CTMM); the normal group showed IQs of 90 to 110 on the CTMM; and the subnormal group showed IQs of 55 to 75 as assessed from the WISC and Stanford-Binet. These IQs were verified on the Goodenough. The Ss from the bright and normal groups were all from the fourth, fifth, or sixth grade and were enrolled in a regular classroom. The subnormals were enrolled in special education classes. All the girls were from low-middle to middle-middle socio-economic level homes and showed no visual or hearing difficulties, no gross physical handicap, and no severe psychopathological disturbance.

Relationship of Performance on the Stephens' Instrument and on the Pikunas Graphoscopic Scale

The first two hypotheses were concerned with the relationship between the Stephens instrument and the PGS. Hypothesis 1 assumed that there would be no relationship between the number of successes attained by each group of subnormal, normal, and bright Ss on Stephens' unstructured categorization tasks and the number of different content categories these respective groups employed in response to the perceptual cues on the PGS. This hypothesis held true for both the bright and normal Ss since the coefficients of correlation for these two groups were insignificant at the .05 level. The subnormal Ss showed a significant r value at the .05 level so that we can reject Hypothesis 1 in part and conclude that for subnormals, at least, there was a linear relationship between their degree of success on Stephens' unstructured categorization tasks and the number of content categories they employ on the PGS.

Hypothesis 2 assumed that there would be no relationship between the number of successes attained by each group of subnormal, normal, and bright Ss on Stephens' naming tasks and the number of different content categories these respective groups employed in response to the perceptual cues on the PGS. This hypothesis held true for the normal Ss since their correlation coefficient value, was not significant at the .05 level. In contrast, the hypothesis was rejected for bright and subnormal Ss since their pairs of scores obtained from the two tools reflected a linear relationship significant at the .05 level. We concluded, therefore, that at least for bright and subnormal Ss, there was a positive relationship between the number of successes they attained on Stephens' naming tasks and the number of content categories they made use of on the PGS. The results for Hypothesis 2 permitted us to say that there was a clear tendency for subnormal Ss to succeed on fewer naming categories and, correspondingly, to use fewer categories to respond to the PGS just as bright Ss succeeded on more naming categories and, correspondingly, used more categories in responding to the PGS cues.

The apparent linear relationships for bright and subnormal Ss on performance on the Stephens' naming tasks and on the PGS appear stronger than those found between the PGS and performance on Stephens' unstructured tasks. This indication probably meant that Stephens' naming tasks were more closely aligned with performance on the PGS than were the Stephens' unstructured tasks; that is, the mental operations needed for verbalizing the unstructured category (which is the naming task) required development and delineation of that category which was precisely what a subject needed and made use of to perform on the PGS. These two performances (success on naming tasks and use of content categories on PGS) were more equivalent in difficulty and in what they required in terms of intellectual processing.

Performance of Bright, Normal, and Subnormal Subjects on Stephens' Unstructured, Structured, and Naming Tasks

An overall evaluation of Hypotheses 3, 4, and 5 was carried out. Hypotheses 3, 4, and 5 stated that there were no significant differences between bright, normal, and subnormal groups of subjects on the number of correct responses they achieved on Stephens' unstructured, structured, and naming tasks. A two-way classification, 3 x 3 analysis of variance, permitted us to reject these hypotheses at the .01 level. We concluded, then, that both the intelligence level variables (independent) and the categorization task variables (dependent) did not show means that were estimates of a common population. In other words, bright, normal, and subnormal Ss differed significantly from one another on the number of correct responses they attained on Stephens' unstructured, structured, and naming tasks.

Interaction between the intelligence level variable and the categorization task variable was significant at the .05 level, but the writer felt justified in requiring a .01 level of significance for concluding that the two variables did show significant interaction. The F-values for the variables were highly significant and the F-value for interaction was not significant at the .01 level; it was therefore concluded that the interaction was insignificant. This indication made it very clear, however, that the interaction effect must be investigated further through more meaningful comparisons among the groups or the like.

Individual analyses for Hypotheses 3, 4, and 5 were also carried out. Hypothesis 3 stated that bright, normal, and subnormal groups of subjects would not differ significantly from one another in the number of correct responses they achieved on Stephens' unstructured tasks. The hypothesis was rejected at the .01 level, and we can conclude that bright, normal, and subnormal groups of subjects did differ significantly from one another on the number of correct responses they achieved on Stephens' unstructured tasks. A look at the raw score totals in Table 11 reveals that the bright Ss were the most successful in number of correct responses to the unstructured tasks, the normals were medially successful, and the subnormals were the least successful. Since it was with unstructured tasks that we were concerned, and the unstructured tasks required independent application of an appropriate category (Stephens, 1963), we concluded from our data that bright Ss were most able to apply appropriate categories independently, normals were medially able, and subnormals were the least able. In fact, this conclusion suggests that subnormal Ss are at quite a disadvantage in being able to lend meaning to their world of experiences through categorization when compared to bright and normal Ss. At the opposite end of the continuum, we can see that the bright Ss have quite an advantage over normal Ss in being able to lend meaning to their experiential world through categorization. The lack of ability of subnormals to employ categories meaningfully undoubtedly leaves many environmental stimuli free-floating and meaningless and therefore lost to them, or it may be that they attempt to fit the stimuli into an already narrowed range of operations. It is undoubtedly perception of this phenomenon that led Kounin (1941) and Lewin (1935) to postulate and expand upon their theory of "rigidity" of behavior in subnormals and "flexibility" of behavior in normals. In concluding that subnormals cannot independently apply categories as well as do brights and normals, we are saying that their intellectualizing is not as effective as that of brights and normals, that subnormals cannot make use of associative cues in their environs as effectively, and therefore that they cannot order and classify their experiences as well as others so that the experiential world is more incomprehensible and chaotic to them. We can say that the subnormals intake of information is narrowed because he cannot make the necessary mental associations that would give the information meaning and permit intake.

Hypothesis 4 stated that bright, normal, and subnormal groups of Ss would not differ significantly from one another in the number of correct responses they achieved on Stephens' naming tasks. The hypothesis was rejected at the .01 level, permitting us to conclude that bright, normal, and subnormal groups of Ss do differ significantly from one another in the number of correct responses they achieved on Stephens' naming tasks. The raw score totals showed that the brights were most successful, the normals medially successful, and the subnormals least successful in correctly verbalizing the name of the unstructured tasks. It would seem logical to assume that if a subject can point out the correct pictures representing a category being tested (unstructured task), he should be able to verbalize the name of the category (naming task). Stephens' tasks show that this conclusion is not sound for any of the groups in this study as seen from their raw score totals on unstructured and naming tasks. The bright Ss showed 548 successes on the unstructured tasks and 497 successes on the naming tasks, which is a discrepancy of 51. The normal Ss showed 474 successes on the unstructured tasks and 409 successes on the naming tasks, which is a discrepancy of 65. The subnormal Ss showed 351 successes on the unstructured tasks and 264 successes on the naming tasks which leaves a discrepancy of 87. It seems that to be able to name a category requires more than just possession of it. It demands development and delineation of that category. Again, we conclude that because the subnormal shows the largest discrepancy and the least number of successes on the naming tasks, he has the least understanding of the categories he possesses. This fact means that his categories are not as well developed and delineated as those possessed by normals and brights. This conclusion is especially relevant to the poor functioning the subnormal displays in an academic setting where specific conceptual categories are used. The normal Ss show an equivalent inferiority to the bright Ss as the subnormal Ss did to the normal Ss.

Stephens (1963) pointed out that his subnormals showed a greater discrepancy than the normals in their successes on the unstructured and naming tasks. Nevertheless, he found that even the normals showed that their understanding of the categories (naming task) did not keep pace with their performance on the unstructured tasks where only use of the categories and not specific naming was mandatory. It is interesting to note that not only did this study support what Stephens found, but also did it show that with a higher IQ level group (brights with IQs of 125 and above) added to the picture, there was as much of a difference seen between the normals and brights as there was between normals and subnormals. This determination permitted us to say that there was as much of a gap between bright and normal Ss as there was between normal and subnormal Ss in ability to name a category one possessed. This finding gave some support to Stephens' proposition that there may be "a sequence of category development in which general ability to use categories is followed by the specific ability to name them" (Stephens, 1963, pp. 80-81). We might note, too, that Stephens' study used boys while this study used girls yet this same discrepancy was found.

Hypothesis 5 stated that bright, normal, and subnormal groups of Ss would not differ significantly from one another in the number of correct responses they achieved on Stephens' structured categorization tasks. The hypothesis was rejected at the .01 level, permitting us to conclude that bright, normal, and subnormal Ss did differ significantly from one another in the number of correct responses they achieved on Stephens' structured tasks. The raw score totals showed that the brights were most successful, the normals medially successful, and the subnormals least successful on Stephens' structured tasks.

Obtaining correct responses on the structured tasks was much less difficult than obtaining correct responses on both the unstructured and naming tasks. For success on the structured tasks, a subject had merely to identify the correct pictures representing the test category after the category name was given by the examiner. In other words, a subject did not have to make an independent decision requiring appropriate use of a category nor did he have to name the category. The facility of the structured tasks in comparison to the unstructured and naming tasks is seen in the mean differences between the tasks by the groups. The means (average number of correct responses) for the unstructured tasks were 18.27 for the bright Ss, 15.80 for the normal Ss, and 11.70 for the subnormal Ss. The means for the naming tasks were 16.57 for the brights, 13.57 for the normal Ss, and 9.63 for the subnormal Ss. The means for the structured tasks were 24.20 for the bright Ss, 22.57 for the normal Ss, and 19.07 for the subnormal Ss. The means for the structured tasks are the highest for all three groups which reflects their lesser level of complexity. If a category was present at any functional level, the structured tasks appeared sensitive enough to detect it and thus indicate the repertory of categories one possessed.

Rejection of Hypothesis 5 left us to conclude that subnormal Ss possessed far fewer categories than did the normal and bright Ss. The subnormal Ss made ninety successes fewer than the normal Ss while the normal Ss made forty-four successes fewer than the bright Ss on the naming tasks. A narrow range of categories indicated the extent of difficulty one might have in interpreting one's experiences and in conceptual manipulations despite the presence of strong associative cues suggesting appropriate category use. The broad range of categories that the bright Ss displayed indicates the extent of meaningfulness they can give incoming stimuli and their experiential world.

Time Comparisons for Bright, Normal, and Subnormal Subjects on Stephens' Categorization Tasks

Hypotheses 6, 7, and 8 were concerned with total time differences by bright, normal, and subnormal subjects on Stephens' unstructured,

structured, and naming tasks. The time differences were analyzed by analyses of variance for each type task and the respective null hypotheses for the tasks were rejected or accepted at the .01 or .05 level of significance. Total time for these hypotheses means the summation of the seconds each subject from a group took to complete each task under a certain condition; that is, the number of seconds it took bright Ss to complete all the naming tasks. Incorrect or correct responses alone were not taken into consideration because the Ss were being compared as a group, not as individuals or on each category.

Hypothesis 6 assumed that bright, normal, and subnormal groups of Ss would not differ significantly from one another on their total times for responses to Stephens' unstructured categorization tasks. The hypothesis was accepted at the .05 level, permitting us to conclude that the bright, normal, and subnormal groups of subjects did not differ significantly from one another on the total number of seconds they took to respond to Stephens' unstructured tasks. This finding did not support Stephens' result (1963, p. 83) of a clear difference between normals and subnormals in their times for completion of the unstructured tasks. We must keep in mind, however, that Stephens was comparing only normals and subnormals, he was using boys, and his statistical mode of comparing was based on mean times for each category, not on total time differences. The writer feels that the addition of the performance of the bright group of Ss may have helped to mask the differences between the normals and subnormals in the statistical analysis though this is said with hesitation since it is quite obvious to the examiner that the bright Ss did, in fact, take the most time to complete the unstructured tasks. These Ss used a calm, collected, and calculated approach, thereby taking more time but achieving many more correct responses than normals and subnormals. In general, the bright Ss did not use a trial and error approach to arrive at the correct responses as did the subnormal Ss, and, less frequently, the normals. The bright Ss seemed to figure the appropriate category mentally while still holding the checkers and then quickly placed them on the correct items while the normals and subnormals seemed to shuffle the checkers around on different items while thinking.

A look at the total time means for the unstructured tasks revealed interesting differences. On the unstructured tasks, the bright Ss showed a total time mean of 7.6 minutes, the normal Ss showed a total time mean of 7.1 minutes, and the subnormal Ss showed a total time mean of 7.5 minutes. It appeared that the bright Ss take more time than the normal Ss, but in so doing achieved significantly more correct responses. The subnormal Ss took almost as much time to respond as the bright Ss, but for different reasons--they used trial and error a great deal, were more distractible, doubted their choice of answers, and often looked up at the examiner for approval.

Hypothesis 7 assumed that bright, normal, and subnormal groups of subjects would not differ significantly from one another on their total times for responses to Stephens' naming tasks. The hypothesis was rejected at the .01 level and permitted us to conclude that bright, normal, and subnormal groups of Ss did differ significantly from one another on the total number of seconds they took to respond to the naming tasks. The finding supported Stephens' result (1963, p. 83). Calculation of the total time means revealed that it took the bright Ss an average of 1.7 minutes, the normal Ss 2.6 minutes, and the subnormals 2.1 minutes to complete the naming tasks. The quicker responding of the bright Ss was in part due to the extra time they took to achieve successfully on the unstructured tasks; that is, they had figured out the appropriate category on the unstructured tasks and could therefore verbalize it very quickly on the naming tasks. The normal Ss showed much continuity in performing on the unstructured and then on the naming tasks so that it was difficult for the examiner to know exactly when to stop timing for the unstructured task and start timing for the naming task. The normals seemed to perform and think simultaneously and then verbalize before having completed the unstructured task. This indication left room for timing error on the part of the examiner. The bright subject was more methodic in that she figured the response, acted quickly, and then verbalized without showing the overlap of functioning that the normals displayed. The subnormal showed more erratic performance in that she appeared to respond as quickly as the normal if she felt certain she had the correct category, but then seemed to give up very readily when uncertain of it. The latter probably accounts for the lesser total time average the subnormal shows when compared to the normal.

Hypothesis 8 assumed that bright, normal, and subnormal groups of Ss would not differ significantly from one another on their total times for responses to Stephens' structured categorization tasks. The hypothesis was rejected at the .01 level, permitting us to conclude that bright, normal, and subnormal groups of Ss did differ significantly from one another on the total number of seconds they took to respond to the structured tasks. The finding supported Stephens' result (1963, p. 84). The total time means showed that it took the bright Ss an average of 2.4 minutes, the normals 3.1 minutes, and the subnormals 3.8 minutes to complete the structured tasks showing us the direction of the differences. In taking the most time, the subnormal showed that she had the most difficulty in responding to categories even after the examiner had stated the category sought for. It also suggested that the subnormal took relatively more time to pick out the associative cues behind the correct items for successful category identification.

The data on Hypotheses 3 through 8 led us to conclude that bright, normal, and subnormal Ss differed significantly from one another on the number of successes they made on Stephens' unstructured, structured, and naming tasks. The bright Ss showed the most successes for all three categorization tasks. The normal Ss showed more successes than the subnormals, but fewer than the bright Ss on all three types of tasks. The subnormals showed the least number of successes on the tasks. From this information we can say that bright Ss have the largest stock of categories and can use these categories most effectively since their categories were very well developed and these Ss were most sensitive to the associative cues making appropriate category use possible. It seemed that there was as much of a discrepancy between brights and normals as there was between normals and subnormals in performance on all the tasks and therefore in categorization ability and functioning. The subnormal Ss showed the least number of categories in their repertory, had the most difficulty in using the categories effectively and had the least ability to note the associative cues needed for appropriate category use implying relative lack of category development and delineation.

The time comparisons led us to conclude that the subnormal girls took the most time to complete the structured categorization tasks while the bright girls took the least amount of time, a conclusion that reflected the latter group's possession of well-delineated categories and extra sensitivity to cues making correct category identification facile. The normal girls took the most time to complete the naming tasks. This factor seemed susceptible to error because the normal girl typically overlapped her performance on the unstructured tasks with the verbalizing of the response which is the naming task. The subnormal girls took less time than the normal girls because they tended to give up more readily when the category appeared too difficult for successful completion and they succeeded on fewer naming tasks. Because of their initial approach, the bright girls took the longest time for the unstructured tasks and the least time for the naming tasks. They figured out the category, placed the checkers promptly, and then verbalized the response. They appeared to sacrifice time initially in order to proceed methodically and achieve more correct responses on both the unstructured and naming tasks finally. They used an organizational approach to solve the category problem posed.

Content Produced by Bright, Normal, and Subnormal Girls on the Pikunas Graphoscopic Scale

The remaining hypotheses were concerned with the performance of bright, normal, and subnormal groups of nine, ten, and eleven-year old girls on the PGS. The types of content produced on this drawing task was representative of the types of categories girls have in their repertory and can draw upon to make meaningful use of when stimulated by a semi-structured cue. The process of intellectualizing to produce a category here was shown to be aligned with that involved in the Stephens' unstructured and naming tasks by means of significant correlations found for Hypotheses 1 and 2. Both Stephens' tool and the PGS showed significant differences between bright, normal, and subnormal girls on the number of categories they possessed and on the effectiveness with which they could use the categories.

Hypothesis 9 assumed that bright, normal, and subnormal Ss would not differ significantly from one another on the types of content they produced on the PGS. The nine most common content categories from the three groups were compared by t-tests. The hypothesis was rejected because of the following findings: It was found that bright and normal Ss differed significantly only on the amount of human and miscellaneous content they produced. The bright Ss used human content most frequently while the normal Ss used the miscellaneous category most frequently. The bright and subnormal differed significantly on the animal, clothing, and plant content produced with the bright Ss using these categories most often. These two groups also differed on the miscellaneous category with the subnormal Ss using it the most. The normal and subnormal Ss also differed on the clothing category with the normal girls producing it more frequently.

The above data permitted us to conclude that the girls from the different IQ levels equated for age did differ on the types of categories they possessed and drew upon to interpret and lend meaning to stimuli. The widest differences were found between the two extremes of the IQ level continuum though intermediary differences were found, too. The greater amount of human and clothing content produced by the bright Ss is probably due to the earlier social awakening and greater maturity of these Ss since they had been equated for age. The clothing content produced points to sex role awareness and acceptance.

In relation to Hypothesis 9, it was interesting to note the categories most frequently used by the girls at the different intelligence levels. A descending order of frequency of content categories produced on the PGS by bright girls was as follows: human, animal, factory-made objects, miscellaneous, plant, house, clothing, nature, and food. The descending order of frequency of content categories by normal girls was as follows: human, factory-made objects, miscellaneous, animal, house, clothing, plant, food, and nature. The descending order of frequency of content categories by subnormal girls was as follows: human, miscellaneous, factory-made objects, animal, house, plant, food, nature, and clothing.

Hypothesis 9a assumed that there would be an increase in original responses on the PGS with IQ increase. A linear trend analysis with a significant F-value at the .01 level permitted us to conclude that girls produced more original responses with increasing IQs. The raw score totals showed that the bright girls produced 175 original responses, the normals produced 129, and the subnormals produced 67. These results supported Pikunas's (1963, pp. 47-48) contention that children with low mental endowment reflect a lack of originality on this scale. The results also supported the innumerable literature sources (DeHaan and Havighurst, 1961) that state that, relatively speaking, bright subjects are most original and creative.

Hypothesis 9b assumed that bright, normal, and subnormal children would differ significantly in the number of popular responses they produced on the PGS. A significant F-value at the .05 level derived from a linear trend analysis permitted us to conclude that children of different IQ level groups did differ in the number of popular responses they produced on the PGS and that the number of popular responses increased with IQ increase. The raw score data showed 116 popular responses for the bright girls, 94 for the normals, and 64 for the subnormals. The examiner had expected the normal Ss to show the most number of popular or common responses as based on the notion that the average or normal child would reflect the most conformity of behavior. This would have been displayed psychologically by a more similar repertory of categories and therefore more similar interpretation of the PGS cues. In addition, the fact that the bright Ss would display more originality would negate the possibility of more popular responses from them. This result was not found, however, and it left the writer to speculate on the reason for the higher production of popular responses by the bright girls. The answer may be in the scoring procedure chosen or, more likely, it may be a reflection of the population chosen. The bright and normal girls were from Norman, Oklahoma, which is a university town with a high percentage of the children finally attending college. The pre-adolescent girls, especially the bright ones, know at an early age that they will attend college and seek to imitate dress and manners of the average college coed. The conformity is extreme and, similarly, the behavior. It is possible to assume, then, that the bright girls would tend to interpret cues in a similar manner in that their behavioral functioning is most influenced by conformity.

Hypothesis 9c assumed that children of different IQ levels did not differ significantly on the number of content categories they employed in responding to the PGS. Significant t-values at the .01 level permitted us to reject the hypothesis. The bright and subnormal groups of subjects differed significantly with the bright Ss showing the larger number of content categories employed. The subnormal and normal groups of Ss differed significantly with the normal Ss showing the larger number of content categories employed. The t-value for the normals versus brights was not significant. The results here showed that bright and normal Ss have a much broader range of categories than subnormals as judged from the greater variety of content categories they employed to respond to the PGS cues. The result supported Pikunas's (1963, p. 47) contention that the mentally subnormal use a low number of content categories.

Hypothesis 9d stated that bright girls did not produce more human responses than normal and subnormal girls on the PGS. The hypothesis was rejected at the .05 level permitting us to conclude that brights do produce more human responses than normal and subnormal Ss on the PGS. This determination is explained in terms of the greater maturity, earlier identification, and social consciousness of the bright girls in this age range.

Hypothesis 9e assumed that animal responses would not be produced with greater frequency with IQ increase. A significant F-value at the .05 level derived from a linear trend analysis permitted us to reject the hypothesis and conclude that animal responses were produced with greater frequency with IQ increase. The bright Ss showed the most frequent use of the animal category, the subnormals the least frequent use of it, and the normals the medial use of it. Pikunas and Carberry (1960) report of animal drawings on the PGS and explain this as related to an increase in instinctual urges concomitant with puberty and due to a rise of interest in pets. In attributing greater maturity to bright Ss, we can readily see that the usual thirteen-year old level performance might be lowered to the ten to eleven-year old level, and cause the greater number of animal responses to be seen here for the brights. It is also interesting to note the number of girls, especially the brighter ones, who did draw a horse (usually in Frame 10) and expressed a desire for a horse or other pet during the testing session. This again may be in part due to a reflection of regional sectionalism since the data is being collected in a southwestern town where the cowboy heritage is still relatively dominant and the geographic location lends itself to use of horses.

Based upon the data gathered concerning the performance of bright, normal, and subnormal girls on the PGS, the following conclusions were warranted. Bright, normal, and subnormal pre-adolescent girls did differ on the types of content categories they employed to respond to the PGS cues. The bright Ss appeared to use the human and clothing categories the most, reflecting an earlier social consciousness and greater maturity. In contrast, the subnormals, and the normals to a certain extent, reflected a more concrete performance, employing the factory-made objects and miscellaneous categories more frequently. Animal responses were produced with greater frequency with IQ increase. The bright Ss used far more original responses than subnormal Ss, but the normal Ss showed as much superiority over the subnormals as the bright Ss did over the normal Ss. Surprisingly, the bright Ss also used the most popular or common responses where the examiner had expected the normal Ss to do so. The bright Ss and the normal Ss both used a greater variety of content categories than the subnormals to respond to the PGS cues, though a similar difference between bright and normal Ss was not found.

Content Produced by Nine, Ten, and Eleven-Year Old Girls on the Pikunas Graphoscopic Scale

Hypothesis 10 stated that girls of different ages would not differ significantly on the type of content they produced on the PGS. The nine most common content categories from the three groups were compared by t-tests. The hypothesis was rejected because of the following findings. The nine and eleven, nine and ten, and ten and eleven-year old groups were compared on the number of times they employed human content as a category. These results were significant at the .01 level regardless of which groups were being compared. The ten-year old Ss showed the most frequent use of the human category and the nine-year olds used it least. Both linear and quadratic trend analyses used to test Hypothesis 10a, which assumed that human content would not be produced with greater frequency with age increase, proved non-significant. We therefore accepted the hypothesis. The fact that the ten-year old girls gave the most human content out of the three age levels used seems to stand in contrast with what Pikunas and Carberry (1960, p. 3) found in that both their nine and eleven-year old Ss (boys and girls) used human content more frequently. We must keep in mind, however, that Pikunas and Carberry were using both sexes and children of average intellectual ability. This study used only girls, and of special importance is the fact that each group was comprised of ten bright girls, ten normals, and ten subnormals. It might well be that the IQ variable is confounding the results.

In refutation of Hypothesis 10, results on the factory-made objects category comparing nine and ten-year olds, nine and eleven-year olds, and ten and eleven-year olds were significant at the .01 level. The raw scores reveal that the eleven-year olds used this category most frequently, the nine-year olds used this category medially, and the ten-year olds used it least often. Pikunas and Carberry (1963, p. 3) showed that the ten-year olds used the factory-made object category most, the eleven-year olds used it medially, and the subnormals used it least.

As indicated immediately above, however, the differences are probably due to our population sample differences.

In relation to Hypothesis 10, it was interesting to note the categories most frequently used by the girls at different age levels. A descending order of frequency of content categories used by nine-year old girls was: human, miscellaneous, factory-made objects, animal, house, plant, food, clothing, and nature. The ten-year old girls showed the following descending order of frequency: human, miscellaneous, factory-made objects, animal, plant and house (difference of 1), nature, food, and clothing (latter two were equal). The eleven-year old girls showed the following descending order of frequency of content categories used: human, factory-made objects, miscellaneous, animal, plant and house (difference of 1), food and clothing (used equally), and nature.

Generally speaking, the most commonly used content categories in this study by age were the same as those found by Pikunas and Carberry (1960) even though the present study overlooks the use of the miscellaneous category which is justified when one views the category as non-existing if every category produced was taken as an entity. Pikunas and Carberry (1960) found that human content was the most frequently used category at all their age levels (five through fifteen). McCarty (1924) studied the interests and abilities of children through drawing and found this to be so. It was true for the present study, too, when the data was analyzed according to age and according to IQ level.

Factory-made objects was the second most commonly used category, the animal category was third in rank, and house and plant categories were fourth and fifth. This was true when the Ss were compared by age and on the basis of intelligence level.

Hypothesis 10b assumed that house content would not decrease with age increase. A non-significant F-value derived from a linear trend analysis led us to accept the hypothesis.

The comparisons of the pre-adolescent girls by age appeared to yield indications of differences on the types of categories they possessed and used most frequently in everyday functioning; but these indications were generally not large enough to cause a significant difference to show as determined by powerful statistical tests like t-tests, linear trend analyses, and analyses of variance. In other words, a year-to-year comparison is too small. It seems, however, that differences would become manifest if the groups compared reflected age clusters at the various important developmental stages.

Summary of the Conclusions

Analysis of data collected in this study permits us several conclusions on the categorizing abilities of bright, normal, and subnormal pre-adolescent girls with similar age and background characteristics:

1. There is a positive linear relationship between the number of successes subnormal girls achieve on Stephens' categorization tasks and the number of content categories they employ to respond to the PGS cues. The bright girls show a similar relationship.

The normal girls show more variation in response, and thus, the relationship was not true for them.

2. Bright girls are the most successful when compared to normal and subnormal girls in identifying the correct category as a result of noting appropriate associative cues for the category being tested (independent use of the category). They are most successful at verbalizing their correct responses reflecting good category development and delineation, and they display the broadest range of categories.
3. The subnormal girls are least successful when compared to normal and bright girls in identifying the correct category as a result of noting appropriate associative cues for the category being tested. They are least successful at verbalizing their correct responses reflecting poorer category development and delineation, and they display the narrowest range of categories.
4. On almost every comparison made in this study, the normal girls' performance and achievements are as inferior to the bright girls as they are superior to the subnormal girls.
5. Bright, normal, and subnormal girls of similar age and background reflect a lack of understanding of some of the categories they use though this gap seems greatest in the subnormal girls and smallest in the bright girls.
6. Time comparisons reveal that bright girls, when compared to normal and subnormal girls, take the most time to respond to the unstructured categorization tasks and, subsequently, take the least time to respond to the naming and structured tasks achieving the

largest number of correct categories on all three types of categorization tasks.

7. Time comparisons reveal that the subnormal girls when compared to normal and bright girls take the most time to respond to the structured tasks and take less time than normal girls to respond to the unstructured tasks because they give up easily and succeed less frequently. The normal girls take more time because they expend more effort and succeed better.
8. Bright, normal, and subnormal girls differ from one another on the types of categories they employ to respond to semi-structured cues reflecting differences in their stock of categories and how they interpret environmental stimuli.
9. Nine, ten, and eleven-year old girls differ from one another on the types of categories they employ to respond to semi-structured cues reflecting differences in their stock of categories and how they interpret environmental stimuli. The differences between the girls are more clear-cut when compared by IQ level than by age when age comparisons are on a year-to-year basis.
10. Original responses to semi-structured cues are produced with more frequency with IQ increase.
11. The human content category is the most common response displayed by pre-adolescent girls regardless of their age or intelligence level.
12. Bright girls make use of the human, animal, and clothing categories more frequently than normal and subnormal girls in response to the PGS

cues. The subnormal girls show a more concrete performance, employing the factory-made objects category most.

Implications and Recommendations

It was implied that the findings of this study would help explain or negate the assumptions of simplicity versus complexity of thinking in the respective subnormal or bright child. As judged from the performance displayed and the conclusions drawn from this study, it seems that psychologists are justified, at least in some of their theories and understandings, where mental processes of the lowly mentally endowed and the gifted are concerned. This study cannot, in final analysis, say that the concrete versus abstract theoretical approaches to the study of thinking should be abandoned since it is obvious from this study that the subnormal child is more concretistic while the bright uses more abstraction with the normal falling in between. The tangible seems to mean more to the subnormal as a result of the fewer thinking dimensions he owns, the fewer categories he can employ independently, and the narrower range of categories he has in stock. Whatever and however the subnormal and bright have and use within their minds certainly causes dissimilar overt performance in terms of how they interpret stimuli and lend meaning to their experiential worlds. The normal child almost always seems to achieve to a level that falls between the bright and subnormal. He seems to display some of the intellective traits of both the bright and subnormal though he has an approach all his own that reflects more persistence and effort than either of the other groups. It is no wonder that literature constantly views the subnormal, normal, and bright on a continuum since their performances usually orient them as such. There are qualitative differences, but our tests of subjects' performances must be quantified for statistical analyses and therefore force us to speak habitually in terms of a continuum.

The findings here can only add to the view of simplified thinking in subnormals and a complex mode of thinking in brights. The subnormal does not note the associative cues for correct category identification as readily or at all when he does not already have it in his repertory. Furthermore, even when he does recognize the associative cues, it takes him longer to do so and longer to actually identify the category. His psychic as well as his motor coordination seems slower.

We must pay clear attention at any time in evaluating the results here to keep in mind that the subnormal population was homogeneous in that the IQs of the Ss were similar; but of extreme importance is the fact that the population was not homogeneous in that both familial and brain-injured retardates were used--and, who knows what other subdivisions of these Ss based on physiology would be necessary to get truly meaningful results? This determination is evidenced since previous testing reveals interesting differences among the performances of various types of retardates. The familial retardate usually behaves like the normal but achieves less when results are quantified. The mongoloid reflects various patterns. The brain-injured retardate is usually impulsive, gives up readily, and also achieves poorly. It is hoped that this is the last study that has failed to take the heterogeniety of a subnormal population into consideration. Certainly, investigation of the categorization process as described in the procedure of this study in a subnormal population alone where heterogeniety of the Ss was controlled would be most interesting and could yield information for instructional practices. The same research with the gifted child could prove valuable.

Although the present study does not discount the value of the abstract versus concrete and the simple versus complex approaches to subnormality and giftedness, the writer feels that Stephens' categorization appraoch, as was used here, has a great deal to offer through further research. It seems that categorizing ability, which has category use and repertory as a core, could be used to explore intellectual processing. Categorization is a wholistic concept with a dual meaning in that it represents what one does now with what one learned or acquired from the past. Investigations with categorization could explore sex differences, socio-economic level differences, age differences, and developmental level differences. The types of associative cues needed for learning categories could be studied. Categorization task items could be studied for validity and reliability as measures. Categorization tests could be devised that would have discriminating powers and predictive validity. The process of category development in terms of human growth could be studied. How the categorization process relates to other intellectual functioning and dimensions could be explored.

How categories relate to one another and how they function in intellectualizing could be studied. In fact, a whole psychological approach to the study of the thinking dimension in man could emerge from further research along this line.

As for educational practices, it seems quite apparent from findings here that repetition as a means of learning is indeed slow and ought not be used extensively at all. It would seem best to teach by means of categories already known or to teach categories by means of discovery of associative cues and then further learning from the newly known categories. The associative cues could be ideas, visual and auditory stimuli, kinesthetic, or a blend of two or more. Learning by association seems far superior to rote learning for children from any intelligence levels. Rote learning requires no intellectual manipulation and can therefore be very frustrating. If a child has no knowledge of a category, or, for that matter, any type of information, a repetitious performance to acquire it is meaningless and perplexing.

It cannot be concluded from this experiment that we should discontinue the traditional concrete to abstract appraoch of teaching the subnormal. In fact, if a conclusion is to be made, this study would support the traditional practice since the subnormal does seem more comfortable and capable with the tangible or concrete as opposed to the abstract. Why, then, should not the concrete be the starting point? Still, it may be that added structure of learning tasks and more meaningful associations would substitute for the usual concrete to abstract method of instruction. An important facet seems to be that of using clear-cut and few dimensions to instruct the subnormal, as seems necessary for instruction in the younger child. Conversely, several dimensions (this is enrichment really) should be used and can be used in teaching the bright child to maximize the intake of his information, provide more knowledge, and challenge and interest in the learning process. The bright child, like the subnormal and normal child, needs well-organized and methodic instruction. One important difference, however, is that the bright child seems more able to organize information--to order and classify the various stimuli and give them meaning--so that if material is not presented systematically, he is still at an advantage since he has a better organizational principle even though he, too, suffers information loss from an absence of systematic instruction especially at the younger age levels.

Further research on categorization could give us new insights and information on the types of categories children possess and use most. It could tell us the categories that are most meaningful to the various subject areas and even how to develop these needed categories.

These statements are just a few of the possibilities and promises of further research on categorization.

REFERENCES

Anastasi, Anne. Psychological testing. New York: MacMillan., 1954.

Ames, L. B. The development of the sense of time in the young child. J. genet. Psychol., 1946, 68, 97-125.

Bensberg, G. Concept learning in mental defectives as a function of appropriate and inappropriate "attention sets." J. educ. Psychol., 1958, 49, 137-143.

Bolles, M. M. The basis of pertinence: A study of test performance of aments, dements, and normal children of the same mental age. Arch. Psychol., N. Y., 1937, No. 212.

Boring, Edwin G. A history of experimental psychology. New York: Appleton-Century-Crofts, Inc., 1957.

Brown, Roger. Words and things. Glencoe, Illinois: The Free Press, 1958.

Bruner, Jerome S., Goodnow, Jacqueline J., & Austin, G. A. A study of thinking. New York: John Wiley and Sons, Inc., 1956.

Carmichael, Leonard, (Ed.) Manual of child psychology. New York: John Wiley and Sons, Inc., 1954.

Church, Joseph. Language and the discovery of reality. New York: Random House, 1961.

Colby, M. G., and Robertson, Janis G. Genetic studies in abstraction. Jnl. comp. physiol. Psychol., 1942, 33, 385-401.

Cronbach, Lee J. Essentials of psychological testing. New York: Harper & Brothers, Publishers, 1960. Pp. 157-213.

DeHann, Robert F. and Havighurst, Robert J. Educating gifted children. Chicago: University of Chicago Press, 1961.

Deutsche, J. M. The development of children's concepts of causal relations. Inst. Child Welf. Monogr. Series. Minneapolis: Univer. Minnesota Press, 1937.

Dixon, J. C. Concept formation and emergence of contradictory relations. J. exp. Psychol., 1949, 39, 144-149.

Doll, Edgar A. The essentials of an inclusive concept of mental deficiency. Am. J. ment. Defic., 1941, 46, 214-219.

Edwards, Allen L. Experimental design in psychological research. New York: Holt, Rinehart & Winston, 1960.

Feifel, Harold. Qualitative differences in the vocabulary responses of normals and subnormals. Genet. Psychol. Monogr., 1949, 39, 151-204.

Friedman, K. C. Time concepts of elementary school children. Elem. School J., 1944, 44, 337-342. (a)

_______. Time concepts of junior and senior high school pupils and of adults. Sch. Rev., 1944, 52, 233-238. (b)

Gellerman, L. W. II. Form vs. background. J. genet. Psychol., 1933, 42, 28-50.

Goldstein, K., and Scherrer, M. Abstract and concrete behavior; an experimental study with special tests. Psychol. Monogr., 1941, No. 239.

Goodenough, F. L. Measurement of intelligence by drawings. New York: World Book Co., 1926.

Griffith, Belver C., Spitz, H. H., and Lipman, R. S. Verbal mediation and concept formation in retarded and normal subjects. J. exp. Psychol., 1959, 58, 247-251.

Grigsby, O. J. An experimental study of development of concepts of relationship in preschoool children as evidenced by their expressive ability. Jnl. exp. Educ., 1932, 1, 144-162.

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

Helson, H. Adaptation-level as a frame of reference for prediction of psycho-physical data. Amer. J. Psychol., 1947, 60, 1-29.

Hicks, J. A., and Stewart, F. D. The learning of abstract concepts of size. Child Developm., 1930, 1, 195-203.

Iscoe, Ira, and Giller, Donald. Areas of concept formation in the mentally retarded. Am. J. ment. Defic., 1959, 63, 112-116.

Kounin, J. S. Experimental studies of rigidity. I. The measurement of rigidity in normal and feeble-minded persons. Charac. & Pers., 1941, 9, 251-272.

Kvarceus, William C. Research in special education: Its status and function. J. except. Child., 1958, 24, 249-254.

Lacey, J. I., and Dallenbach, K. M. Acquisition by children of the cause-effect relationship. Amer. J. Psychol., 1939, 52, 103-110.

Lewin, Kurt. A dynamic theory of personality: Selected papers. Trans. D. K. Adams and K. E. Zener. New York: McGraw Hill, 1935.

Long, L. Conceptual relationships in children. The concept of roundness. J. genet. Psychol., 1940, 57, 289-315.

Long, L. and Welch, L. Influence of levels of abstractness on reasoning ability. J. Psychol., 1942, 13, 41-59.

Martin, W. E., and Blum, A. Interest generalization and learning in mentally normal and subnormal children. J. comp. physiol. Psychol., 1961, 54, 28-32.

McCarty, S. A. Children's drawings. Baltimore: Williams and Wilkins, 1924.

Munn, N. L., and Steining, B. R. The relative efficacy of form and background in a child's discrimination of visual patterns. J. genet. Psychol., 1931, 39, 73-90.

Myers, C. R., and Gifford, Elizabeth. Rescoring the Stanford-Binet. Bulletin of the Canadian Psychol. Asso., I (April, 1941, No. 29).

Nagy, M. The child's theories concerning death. J. genet. Psychol., 1948, 73, 3-27.

Newland, T. Ernest. A critique of research on the gifted. J. except. Child., 1941, 29, 391-398.

Oakes, M. E. Children's explanation of natural phenomena. Teachers College, Columbia Univ., Contrib. to Educ., 1947, No. 926.

Ordan, H. Social concepts and the child mind. New York: King's Crown Press, 1945.

Osbern, W. J. Associative clustering in organic familial retardates. Amer. J. ment. Defic., 1960, 65, 351-357.

Peters, R. S. (Ed.) Brett's history of psychology. New York: MacMillan, 1953.

Piaget, J. The child's conception of physical causality. New York: Harcourt, Brace, 1930.

_______. The child's conception of the world. New York: Harcourt, Brace, 1929.

_______. The language and thought of the child. New York: Harcourt, Brace, 1926.

Pikunas, Justin. Fundamental child psychology. Milwaukee, Wis.: Bruce Pub. Co., 1957. Pp. 112-113.

_______. The Pikunas Graphoscopic Scale manual. Detroit, Mich.: University of Detroit, 1959.

Pikunas, Justin, and Carberry, Hugh. Developmental levels in the content of children's drawings. Detroit: University of Detroit, 1960.

Rapaport, D. Diagnostic psychological testing. Chicago: Yearbook Publishers, 1945. Vol. 1.

Razran, G. H. S. Stimulus generalization of conditioned responses. Psychol. Bull., 1949, 46, 337-365.

Reichard, Suzanne, Schneider, Marion, & Rapaport, David. The development of concept formation in children. Amer. J. Orthopsychiat., 1944, 14, 156-162.

Stephens, Wyatt. A comparison of normal and subnormal boys on tasks requiring the use of selected categories. Published doctoral dissertation, Univer. of Oklahoma, 1963.

_______. A comparison of the performance of normal and subnormal boys on structured categorization tasks. J. except. Child., 1964, 30, 311-315.

Terman, L. M. The measurement of intelligence. Boston: Houghton Mifflin, 1916.

Terman, L. M. and Merrill, M. A. Measuring intelligence. Boston: Houghton Mifflin, 1937.

Thrum, M. E. The development of concepts of magnitude. Child Developm., 1935, 6, 120-140.

Thurstone, L. L. A law of comparative judgment. Psychol. Rev., 1927a, 34, 273-286.

Van Osdol, William R. A comparison of normal and subnormal subjects using visual and auditory structured categorization tasks. Unpublished doctoral dissertation, Univer. of Oklahoma, 1964.

Vinacke, W. Edgar. The psychology of thinking. New York: McGraw-Hill Book Co., 1952.

Walker, Helen M., and Lev, Joseph. Statistical inference. New York: Henry Holt and Co., 1953. Chs. 9 and 10.

Wechsler, D. Measurement of adult intelligence. (2nd ed.) Baltimore: The Williams & Wilkins Co., 1944.

Weigl, E. On the psychology of so-called processes of abstraction (Trans. Margeret J. Rioch). J. abnorm. soc. Psychol., 1941, 36, 3-33.

Welch, L. A preliminary investigation of some aspects of the hierarchical development of concepts. J. genet. Psychol., 1940, 22, 359-378.

Welch, L. The development of size discrimination between the ages of 12 and 40 months. J. genet. Psychol., 1939, 55, 243-268.

_______. The span of generalization below the two-year age level. J. genet. Psychol., 1939, 55, 269-297.

Welch, L., & Long, L. Comparison of the reasoning ability of two age groups. J. genet. Psychol., 1943, 62, 63-76.

Welch, L., & Long, L. The higher structured phases of concept formation of children. J. Psychol., 1940, 9, 59-95.

Werner, Heinz, & Strauss, Alfred A. Causal factors in low performance. Amer. J. ment. Defic., 1940-41, 45, 213-218.

Werner, Heinz, & Strauss, Alfred, A. Pathology of figure-ground relation in the child. J. abnorm. soc. Psychol., 1941, 36, 236-248.

Zigler, Edward. An overview of research in learning, motivation, and perception. J. except. Child., 1962, 28, 445-448.

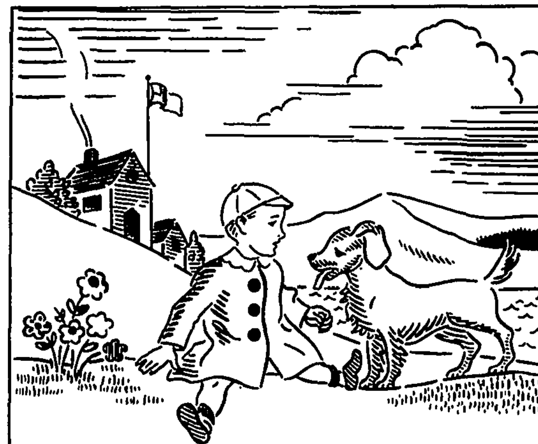
APPENDIX A

THE PIKUNAS GRAPHOSCOPIC SCALE

Name ______________________

____

____



1

2

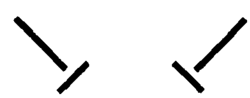
3

4

5

6

7



8

9

What is the thing you like to draw most often?

Draw it in the empty space on the right. →

What else would you like to draw?

10

PGS Third Printing 1961

APPENDIX B

DATA SHEET FOR THE PIKUNAS GRAPHOSCOPIC SCALE

Name: ______________________ IQ Level: ______

Address: ______________________ Age: ________

Date tested: __________________ Time: _____ to _____ = _____min.

School: _______________________ No. of Popular Responses: ______

Teacher: ______________________ No. of Original Responses: ______

Grade: ________________________ No. of Content Categories: ______

DOB: __________________________ No. of Correct Responses:

IQ: ___________________________ Unstructured tasks: ________

Test: __________________ Naming tasks: ___________

Date: __________________ Structured tasks: __________

Socio-economic level: __________

Order of Completion	Frame No.	Content Categories								
		Human	Animal	House	Fobj.	Plant	Cloth.	Food	Nature	Misc.
	1									
	2									
	3									
	4									
	5									
	6									
	7									
	8									
	9									
	10									
	11									
Totals										

APPENDIX C

DATA SHEET FOR STEPHENS' CATEGORIZATION TASKS

Name: ______________________ IQ Level: ________ Age: ________

Cat.	Cor.	Unstructured	Time	Cor.	Naming	Time	Cor.	Structured	Time
1									
2									
3									
4									
5									
6									
7									
8									
9									
10									
11									
12									
13									
14									
15									
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
Total									

School: ______________ Goodenough: ________ Results: No. correct
Obtained IQ: ______________ Visual acuity: ________ Unstruc: ________
Date tested: ______________ Color vision: ________ Naming: ________
C. A.: ___ M.A.: ___ Grade: __ Hearing: ________ Struc.: ________