HURT, Jr., Maure, 1934-
RELATIONSHIPS OF THE MODE OF
CATEGORIZATION STYLE TO ACHIEVEMENT
ON SELECTED INTELLECTUAL MEASURES.

University of Arizona, Ph.D., 1969
Education, psychology

University Microfilms, Inc., Ann Arbor, Michigan
RELATIONSHIPS OF THE MODE OF CATEGORIZATION STYLE TO ACHIEVEMENT ON SELECTED INTELLECTUAL MEASURES

by

Maure Hurt, Jr.

A Dissertation Submitted to the Faculty of the DEPARTMENT OF EDUCATIONAL PSYCHOLOGY
In Partial Fulfillment of the Requirements For the Degree of DOCTOR OF PHILOSOPHY In the Graduate College
THE UNIVERSITY OF ARIZONA

1969
I hereby recommend that this dissertation prepared under my direction by MAURE HURT JR. entitled RELATIONSHIPS OF THE MODE OF CATEGORIZATION STYLE TO ACHIEVEMENT ON SELECTED INTELLECTUAL MEASURES be accepted as fulfilling the dissertation requirement of the degree of Doctor of Philosophy after inspection of the final copy of the dissertation, the following members of the Final Examination Committee concur in its approval and recommend its acceptance:*

This approval and acceptance is contingent on the candidate's adequate performance and defense of this dissertation at the final oral examination. The inclusion of this sheet bound into the library copy of the dissertation is evidence of satisfactory performance at the final examination.
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: [Signature]

[Name]
ACKNOWLEDGMENTS

The author is deeply indebted to Dr. Marie M. Hughes, who provided guidance throughout the writing of this dissertation and without whose material assistance this study could not have been completed. In addition the author would like to thank the other members of his committee, Dr. Jerald Reece, Dr. Robert Kraner, Dr. John Bergan, and Dr. Bruce Howard.

Appreciation is expressed to the staff of the Arizona Center for Early Childhood Education who assisted in gathering and scoring the data.

Finally, the author would like to express his thanks to his wife Linde for her support and encouragement during the course of writing this dissertation.
# TABLE OF CONTENTS

LIST OF TABLES ........................................ vi
ABSTRACT ........................................... viii

CHAPTER

I. THE PROBLEM, HYPOTHESES, AND THE DEFINITIONS OF TERMS USED ................. 1
   The Problem ........................................ 1
   Statement of the Problem .......................... 2
   Importance of the Study............................ 2
   Theoretical Basis of the Study ..................... 3
   Limitations of the Study ........................... 8
   Definitions of Terms Used ........................... 9
   Standardized Primary Instruments .................. 9
   Intellectual Tasks ................................... 10
   Hypotheses to be Tested ............................ 10
   Organization of the Dissertation .................... 13

II. REVIEW OF THE LITERATURE ......................... 16
   Literature of Categorization ....................... 16
   Psychometric Testing and the Culturally Deprived .... 27

III. SAMPLE, MATERIALS, AND PROCEDURES ............. 39
   Sample ............................................. 39
   Instruments and Materials ........................... 41
      Standardized Primary Instruments ................. 41
   Transformations .................................... 50
   Time Schedule of Testing ............................ 50

IV. RESULTS ........................................... 52
   Results on Preferred Mode of Categorization .......... 52
      Types of Preferred Mode of Categorization .......... 52
      Individual Mode Intercorrelations .................. 53
      Combined Mode Inter-relationships .................. 57
      Sex Differences in Grouping Results ................ 58
TABLE OF CONTENTS--Continued

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results Concerning Hypothesis One</td>
<td>60</td>
</tr>
<tr>
<td>Results Concerning Hypothesis Two</td>
<td>60</td>
</tr>
<tr>
<td>Factor Analysis of First Year Tests</td>
<td>60</td>
</tr>
<tr>
<td>Results Concerning Hypothesis Three</td>
<td>64</td>
</tr>
<tr>
<td>Results Concerning Hypothesis Four</td>
<td>66</td>
</tr>
<tr>
<td>Canonical Correlation of Categorization and Metropolitan Achievement Tests</td>
<td>67</td>
</tr>
<tr>
<td>Relationship of Standardized Primary Instruments to Metropolitan Achievement Tests</td>
<td>67</td>
</tr>
</tbody>
</table>

V. DISCUSSION, SUMMARY, AND CONCLUSIONS ...................................... 71

Discussion ......................................................................................... 71
  Hypothesis One ........................................................................... 71
  Hypothesis Two ........................................................................... 72
  Hypothesis Three ........................................................................ 73
  Hypothesis Four ........................................................................... 74
Summary .............................................................................................. 76
Conclusions ......................................................................................... 79
  Material Dependency .................................................................. 80
  Process and Product Divergency ............................................. 81
  Language Bias ............................................................................. 82
  Recommendations for Future Research ...................................... 82
  Psychometrizing Categorization .............................................. 83

APPENDIX A. SCORING RULES FOR CATEGORIZATION RESPONSES .................... 85

APPENDIX B. PROCEDURES AND SCORING FOR HUGHES CAR DISCRIMINATION TEST ......................................................................................... 90

APPENDIX C. FIGURES SUPPLIED FOR INTELLECTUAL TASKS: GROUPING ........ 93

APPENDIX D. SAMPLE PROTOCOL FROM INTELLECTUAL TASKS .................... 94

REFERENCES ....................................................................................... 97
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Relationships of Conceptual Style and Selected Variables from the Pederson and Wender (1968) Study</td>
<td>26</td>
</tr>
<tr>
<td>2. Reliability Coefficients from the Robinson and Hanson (1968) Study</td>
<td>33</td>
</tr>
<tr>
<td>4. Age of Sample</td>
<td>40</td>
</tr>
<tr>
<td>5. Test-Retest Correlations for Hughes Car Discrimination Task</td>
<td>45</td>
</tr>
<tr>
<td>6. Scores Derived from Intellectual Tasks: Grouping</td>
<td>48</td>
</tr>
<tr>
<td>7. Inter-Scorer Reliability for Intellectual Tasks: Grouping</td>
<td>49</td>
</tr>
<tr>
<td>8. Testing Schedule</td>
<td>51</td>
</tr>
<tr>
<td>9. Inter-relationships of Descriptive Preferred Mode Types</td>
<td>54</td>
</tr>
<tr>
<td>10. Inter-relationships of Categorical-Inferential Preferred Mode Types</td>
<td>55</td>
</tr>
<tr>
<td>11. Inter-relationships of Relational Preferred Mode Types</td>
<td>56</td>
</tr>
<tr>
<td>12. Inter-relationships of Thematic Preferred Mode Types</td>
<td>56</td>
</tr>
<tr>
<td>13. Relationships of Frequencies of Groups in the Four Categories</td>
<td>58</td>
</tr>
<tr>
<td>14. Analysis of Variance Summary Standardized Primary Instruments with Preferred Mode of Categorization as the Independent Variable</td>
<td>61</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>15. Factor Loadings for First Year Standardized Primary Instruments</td>
<td>63</td>
</tr>
<tr>
<td>16. Analysis of Variance Summary Results of Hughes Car Discrimination Task Using Flexibility as the Independent Variable</td>
<td>65</td>
</tr>
<tr>
<td>17. Canonical Correlation: Preferred Mode of Categorization and MAT Reading and Word Knowledge</td>
<td>68</td>
</tr>
<tr>
<td>18. Canonical Correlation: First Year Standardized Primary Instruments and MAT Reading and Word Knowledge</td>
<td>69</td>
</tr>
</tbody>
</table>
ABSTRACT

This study is an investigation of the relationships between the preferred mode of categorization of an individual and the scores that that individual achieves on selected achievement, vocabulary, intelligence, and discrimination tests.

Scott (1962) reported that mode of categorization correlated with achievement on elementary science achievement tests and Yeatts (1969) reports similar findings with academic achievement in other areas as well. Sigel (1965) states that mode of categorization as a part of cognitive style could affect the responses given on the commonly used standardized tests.

Four hypotheses were tested. The first hypothesis stated that analyses of variance using the preferred mode of categorization as the independent variable would show differences significant at the .05 level, with achievement on selected standardized tests. The second hypothesis was similar to the first, except that factor scores, resulting from a factor analysis of the selected standardized tests, replaced the test scores as the dependent variables. The third hypothesis stated that degree of flexibility as evidenced by the use of more than one mode of categorization would, when used as the independent variable, differentiate
between groups on the scores achieved on the selected standardized tests. The fourth hypothesis stated that conceptual style as represented by utilization of modes of categorization in the first grade would significantly predict reading achievement at the third grade level.

The sample was composed of 87 randomly selected, Mexican-American children entering the first grade in schools in poverty areas in Tucson, Arizona. All subjects met at least one of the following criteria: (1) Spanish surname or (2) spoke Spanish in the home according to school records.

The subjects were tested with the Metropolitan Readiness Tests, the Van Alstyne Picture Vocabulary Test, and the Goodenough-Harris Drawing Test in the first grade and the Metropolitan Achievement Test: Reading at the end of third grade. In addition to these tests, the Intellectual Tasks: Grouping was given in first grade as well as the Hughes Car Discrimination Task. Both Tasks were developed at the Arizona Center for Early Childhood Education.

The Intellectual Tasks: Grouping is an object sorting task utilizing 44 rubber figures of fathers, mothers, and children, service workers such as firemen, policemen, nurses, and doctors, plus domestic and wild animals. The subjects were asked to group the figures into groups that "belonged" together and then give the reason for belonging in that group. The transcript of the
verbal responses given was scored using a modified version of the system presented by Wallach and Kogan (1965). Four modes of categorization were used: (1) descriptive based on similarities of objective physical attributes, (2) categorical-inferential based on representation of the grouping concept by all members of the group, (3) relational based on an expressed relationship of figures within a group, and (4) thematic based on a story or theme involving the figures in the group.

The Hughes Car Discrimination Task was intended to estimate the ability of the subject to discriminate to a relatively fine degree between model cars presented in pairs. The scoring was based on the number of differences the subject was able to discriminate.

The first, second, and fourth hypotheses were not supported by the results. The third hypothesis, concerning flexibility, was upheld in the analysis of the Hughes Car Discrimination Task in the Sedan and Total scores with the group having the highest degree of flexibility, as evidenced by the use of all four modes of categorization, achieving the highest scores in discrimination.

Several factors may account for the results: (1) the duplication of figures in the object sorting task may have introduced a set response, (2) object sorting may be subject to a material dependency phenomena which influences the results of a given type of materials, and (3) the
product in the form of achievement on the standardized tests may be restricted in this type of sample and appear to be the same, but the process, categorization, may be different.
CHAPTER I

THE PROBLEM, HYPOTHESES, AND THE DEFINITIONS OF TERMS USED

In the past decade there has been an increasing interest in the function of cognitive style in learning ability and academic achievement. Scott's (1962) work in cognitive style seems to indicate that certain manifestations of cognitive style, such as preferred mode of categorization seem to have a discriminating effect when achievement was measured in the area of science in upper elementary grades.

Sigel (1965) indicates that mode of categorization or perceptual organization has a relationship to the responses that an individual gives for some items on the Wechsler Intelligence Scale for Children, moreover, he contends that the phenomena has a greater potential for influence among the children of cultural minority groups, and in a wide variety of standardized tests.

The Problem

This section contains the statement of the problem, its importance, the theoretical base, and the limitations.
Statement of the Problem

It was the purpose of this study to examine the relationship of selected aptitude, achievement, vocabulary, and discrimination tests, and the factors therein with responses representing a subject's behavior on a categorization task. In addition the relationship of reading achievement to categorization was of especial interest.

Importance of the Study

Individualization of instruction has long been a goal of educators, but in order to accomplish this goal it is necessary to know how to individualize instruction so as to match the instruction with the instructed. If one follows the reasoning of Bruner, Goodnow, and Austin (1956), Bruner, Olver, Greenfield, and collaborators (1966), or Piaget (Inhelder and Piaget, 1964) it is clearly evident that they consider learning as a function of cognitive structure and categorization in particular. Individual preference for a given mode of categorization (Sigel, 1965) may have influence upon the amount of knowledge gained in the academic environment and the skills mastered as measured by performance on the various standardized intelligence and achievement tests used in the public schools. Clarification of this influence or relationship will enhance the ability of the curriculum builder to take
advantage of the abilities of the individual student implied in his preferential mode of categorization.

Theoretical Basis of the Study

Guilford (1967) provides a structure of the intellect and Piaget (Inhelder and Piaget, 1964) provides a process of development of that intellect.

Cognitive style and categorization fit neatly into the structure of intellect as hypothesized by Guilford. The structure of intellect as formulated by Guilford is composed of three dimensions. The first dimension is composed of operations which encompass evaluation, convergent production, divergent production, memory, and cognition. The second dimension of the three-dimensional model is concerned with products, products in the form of units, classes, relations, systems, transformations, and implications. The final dimension of the cubical model is that of content. Content includes figural, symbolic, semantic, and behavioral content. Categorization at its various levels of sophistication can be related to a large number of the hundred-and-twenty cells in the model, but fits specifically into the cells defined by cognition on the operations plane, classes and relations on the product surface of the cube, and figural, symbolic, and semantic on the content dimension. Guilford does not specifically state the ontogeny of each of the cells in his model, but he notes that some of the
cells seem to develop later than others (Guilford, 1967, p. 417) depending on the use that the individual has for that particular combination of vectors. The Figural Classification Test that has been used by Guilford (1967, p. 30) to identify and mark the cell of cognition of figural classes in his factor analytic studies of the structure of intellect model bears a strong resemblance to the sorting tasks used by Wallach and Kogan (1965), Kagan, Moss, and Sigel (1963), and others (Yeatts, 1969) working in the area of cognitive style.

Piaget (Flavell, 1963) complements Guilford in that he postulates the epistemological history of classificatory cognition. For Piaget (Inhelder and Piaget, 1964) there are three stages in the developmental process which is the cognitive basis of categorization. Piaget, as reported by Flavell (1963, p. 303) sees the first stage occurring around the age of three in most children. Characteristic of this stage is the construction of figural collections, categories based on shifting criteria. An illustration of this first stage is given by Inhelder and Piaget (1964) as follows:

MAR (2:11). The instructions are identical. He begins with a pile of circles, then aligns a number of squares, continuing with semi-circles and circles.

A row of jumbled elements is: "A train, ch, ch, ch!" (p. 33).

It appeared to the observer that the categorization was logical in the example given above until the verbal
explanation was given by the child. In the young child, Piaget's view is that the boundaries of classification are not fixed, but shift between logical class and infralogical whole. In the preceding example the logical class would be the groups with the common geometrical shape while the shift to the infralogical whole takes place with the identification by the child of the grouping as a "train."

The second stage in Piaget's developmental theory is that typified by non-figural collections, according to Flavell (1963), between the ages of five-and-one-half and eight approximately. At this stage classificatory behavior is characterized by groups based on similarity of attributes with the potential of sub-class if required by the situation. At this stage though, the child cannot keep in mind at all times the sine qua non of the third stage the ability to classify under the concept of inclusion. The third stage then for Piaget, is arrived at by the child when categorization can be performed using the concept of inclusion. This concept defines the operation of being aware that if a sub-class is a part of a class it does not exhaust the class. In mathematical terms \( A + A' = B \) and \( B = A + A' \), that is to say that the reversibility aspect must be present to substantiate the presence of the inclusion concept. The child must be able to recognize that \( A \) and \( A' \) can be classes unto themselves and simultaneously be members of class \( B \). At this time Piaget terms the
developmental stage that of concrete-operational cognition in his developmental theory (Flavell, 1963, p. 306).

Hanfmann and Kasanin (1942) have designated several stages of categorization while investigating the sorting task behavior of adults. They also note the writings of Vigotsky in his 1932 book, *Thinking and Speech*, in which he described the initial classificatory behavior of the child which Hanfmann and Kasanin found in some of their schizophrenic subjects.

These authors proposed a model which in the primary stage bears a marked resemblance to Piaget's first stage. For Hanfmann and Kasanin this stage is characterized by intra-category relationships which are vulnerable to shifting by the subject. For example with colored blocks the first block of a group may be a red square, the next block selected on a red criterion, a red circle, but the third member of the group, a blue circle, is selected on the circle criterion. The connecting link in the first two items selected being color, but the second and third selected items based on form, yet all included in the same group. This is similar to the example given by Inhelder and Piaget (1964) of the four-year-old with sequential groups, but who labeled the completed group as a train.

The intermediate level is based on both classification and pseudo-classifications wherein the rules may be established by the subject, but are not binding in that
exceptions are allowed or even chance may enter into the classificatory behavior.

The highest level of categorization occurs when the relationship between members of a class is governed by a general concept therefore the class has a hierarchical organization in which hypotheses can be formulated and tested logically in the process of defining a category. Again this system has a high degree of commonality with the concept of inclusion put forth by Piaget in his stage three.

It is interesting to consider that this classification model was developed by Hanfmann and Kasanin while working with adults, both normal and those with mental disease. They did not find any differences in classificatory behavior between males and females, but noted that there was a significant difference between the abilities of college educated subjects and those with a high school or less education.

Bruner (Bruner et al., 1956, 1966) is concerned with the categorizing behavior of human beings, but more in the formation of new categories, than in the development of the ability in the mental growth of the child. Bruner's work fits into the higher levels hypothesized by both Piaget (Inhelder and Piaget, 1964) and Goldstein and Scheerer (1941). The types of categorization, the use of conjunctive concepts, disjunctive concepts, or relational
concepts are foremost in categorizing behavior investigated by Bruner (Bruner et al., 1956, p. 41).

Conjunctive categorization is based on joint positive attributes, whereas disjunctive categorization shares the joint attributes feature, but in positive-negative combinations. Relational categorization is based upon a "specificable relationship between defining attributes [Bruner et al., 1956, p. 43]."

Thus one sees a theoretical basis accommodating categorization with theories for its development in children with sequential stages of increasing complexity of classificatory behavior, and hypothesized strategies for the creation and conformation of categories (Bruner et al., 1956; Guilford, 1967; Hanfmann and Kasanin, 1942; Kagan et al., 1963; Inhelder and Piaget, 1964). In summation, Bruner (Bruner et al., 1956) states unequivocally that "virtually all cognitive activity involves and is dependent on the process of categorization [p. 246]."

Limitations of the Study

This study uses a sample that is intended to be representative of the Mexican-American cultural minority. The schools from which the sample comes serve an area that is considered a poverty area under the provisions of Public Law 89-10.
Age is considered a second limitation, the group of children in the sample range from six years old to seven years old at the inception of the study.

A third limitation is the nature of the instrument measuring preferred mode of categorization. Reliability is inferred by the number of groupings that a subject forms in a given mode, but a formal measure such as a test-retest reliability coefficient is not available on this instrument at the present stage of development.

A final limitation of the study is the definition of Mexican-American which is taken to mean in this study those children whose surname is Spanish in origin or who are reported by the classroom teacher as speaking Spanish in the home.

Definitions of Terms Used

The following definitions are meant to clarify certain concepts and terms as used in this study. The definitions of the grouping categories will be developed in the review of the literature and so are omitted at this point.

Standardized Primary Instruments

This term is used to indicate the selected standardized instruments used in this study for the assessment of intelligence and achievement. The following measures are subsumed under this definition: (1) the
Goodenough-Harris Drawing Test; (2) the Metropolitan Achievement Tests, Form A, Primary; (3) the Metropolitan Readiness Tests; and (4) the Van Alstyne Picture Vocabulary Test. Although not a standardized test the Hughes Car Discrimination Test is included in this definition.

Intellectual Tasks

This term is used to denote the set of tasks from which categorization scores are derived, specifically the Grouping section of the Intellectual Tasks.

Hypotheses to be Tested

The hypotheses are formulated on two levels. The first level is the direct test score level where the intelligence, achievement, and vocabulary answers are represented by standard scores based on the tables given in the appropriate test manuals for the age or grade level of the subject. At the second level the factor scores of the combined intelligence, achievement, and vocabulary tests will be tested. All hypotheses will be tested at the .05 level of significance.

The hypotheses are as follows:

1. The preferred mode of categorization as evidenced by highest frequency of use of a given mode, when used as the independent variable, will show significant differences in intelligence, achievement, and vocabulary measures.
2. The preferred mode of categorization, when used as the independent variable, will significantly differentiate between groups on the factor scores of the first year standardized primary instruments.

3. Flexibility as evidenced by the use of more than one mode of categorization, when used as an independent variable, will significantly differentiate between scores achieved on the standardized primary instruments.

4. Conceptual style as represented by utilization of modes of categorization in the first grade will significantly predict reading achievement at the third grade level.

The hypotheses are based on the writings of Sigel (1965) in which he states that mode of categorization affects intelligence scores. Given the degree of correlation found between the various intelligence tests and the achievement tests used in the public schools (Anastasi, 1968) the results of the selected achievement tests used in this study should be influenced as well. This first hypothesis was intended to test whether or not a given mode, or modes, of categorization is superior to any other mode, or modes, at this stage in the development of the child.
Nunnally (1967) points out that the factor structure for a given test will vary with the population tested. The population in this study deviated quite markedly from the norming population on the standardized tests and it was believed that the factor structure might deviate as well. Therefore, the variance of a given subtest could possibly represent something entirely different in terms of constructs measured than the test designer had intended. Factor analysis, using a principal axis approach and varimax rotations, will allow the subtest variances to be distributed into common factors.

The third hypothesis concerning flexibility was the logical result of the preceding two hypotheses. If one may introduce the concept of a categorization mode-linked item in a test where that item can be more easily answered correctly if the examinee's categorization mode is complementary to the concomitant categorization mode-link of the question. With heterogeneity of mode-linked questions, the relationships with a test or subtest may be obscured hence the factor score design of hypothesis two. Given that the construct of the mode-linked question is valid, then it follows that the subject who can operate in more than one mode of categorization has a better cognitive structure with which to deal with the problems presented in intellectual tests as well as in life as a whole. This hypothesis was limited in its testing by allowing free response
to the categorizing tasks so that subjects who categorized in only one mode were not required to categorize in another. Therefore, their ability in this particular is not known. Flexibility in this case was based on spontaneous flexibility rather than imposed.

The fourth hypothesis was formulated to determine whether or not a given pattern of preferred mode of categorization in combination with the ability to use other modes would seem to indicate a greater facility in achievement on the Metropolitan Achievement Test: Word Knowledge and Reading.

Organization of the Dissertation

This study is organized into five chapters. The first chapter introduces the study including the hypotheses to be tested, the theoretical basis of the study, and the definitions of terms to be used.

Chapter II is a review of the literature relevant to categorization, discrimination, and the use of standardized tests with the culturally deprived.

The third chapter deals with the methods of the study. The procedures are described, the sample selection process explained, and the materials used in the investigation are presented. The procedures section discusses the standardized tests used in measuring achievement and intellectual level, with a short description of the administrative
procedures and the tests themselves. In addition the newly developed tests and tasks are described, such as the Intellectual Tasks and the Hughes Car Discrimination Task. Finally the statistical treatment of the data is noted and the research design described.

The fourth chapter is divided into five sections. The first section is concerned with the results dealing with the intra-relationships of the modes of categorization. In addition this section presents the results of the comparisons of preferred mode of categorization with sex of the subject as the independent variable.

The second section presents the results obtained from the statistical analysis of hypothesis one. The third section is concerned with the statistical results of hypothesis two. Hypothesis three statistical results are contained in section four and the analysis of hypothesis four is reported in section five.

The fifth and last chapter is concerned with the discussion and conclusions drawn from the results of the study, the implications of the results, and recommendations for further research.

The information contained in the appendices is the result of the efforts of the research workers at the Arizona Center for Early Childhood Education in developing an instrument to measure the cognitive repertoire of young children. Work was begun in January of 1966 by Marie M.
Hughes and her associates. The result of this work is the Intellectual Tasks Battery. The manual of directions for administering and scoring this battery was the source for the four appendices.
CHAPTER II

REVIEW OF THE LITERATURE

The literature pertaining to this study was drawn from two areas of research, the first area was in the field of cognition and classificatory behavior, the second from the body of research in the area of tests and measurement's related to achievement in reading, especially when employed in the measurement of the culturally deprived and minority groups.

Literature of Categorization

Categorization and classificatory behavior have been examined from several viewpoints. One viewpoint is that of the researcher interested in the ontogeny of classificatory behavior such as Piaget (Inhelder and Piaget, 1964) who centered his attention upon the various stages of development of cognitive categorization. Other have been interested in the formation of categories in adults, both normal (Bruner et al., 1956) and abnormal (Goldstein and Scheerer, 1941; Hanfmann and Kasanin, 1942). The attention of another group of investigators has focused upon the qualities of the categories formed, such as Kagan, Moss, and Sigel (1960, 1963) and Gardner (1953). The relationships of the types of categorization forms yet another area.
where research has been done (Sloane, Gorlow, and Jackson, 1963). Cross-cultural studies concerning categorization have been rather rare in the literature, but Price-Williams (1962) reports on some African cross-cultural work and the Bruner group at Harvard University have done some experiments using a rural-urban continuum (Bruner et al., 1956).

Goldstein and Scheerer (1941), working primarily with mentally-ill adults found that their subjects were unable to shift from concrete to abstract criteria for grouping tasks dealing with the groupings of colored blocks. In their experiments with mentally-deficient adult subjects the grouping task was performed on a situational criterion and the subjects appeared to be unable to group on the criterion of form nor the criterion of color. These authors also noted that in some of their previous experiments with young children, the rationale for integration into groups tended to be on a thematic basis.

More recent investigations indicate that brain damage also incapacitates the adult to some degree in his classificatory behavior. Milner (in Warren and Akert, 1964, pp. 208-209) reported that some patients with impairment of the bilateral frontal lobes of the brain were unable to function on the Wisconsin Card Sorting Test (WCST). One of these patients who was unable to perform the categorizing operations of the WCST later earned a
master's degree in accounting. No report was given as to any later success with the same test by the accountant.

The categorization model of Hanfmann and Kasanin (1942) has already been described in Chapter I. The research that led to the establishment of this model indicated that college educated, normal adult subjects (professional staff and graduate students at a mental hospital) were able to perform better on categorization tasks than those subjects with less than a college education (non-professional staff and attendants), who in turn were able to function at a higher level than schizophrenics with a college level education. At the bottom of the categorization ability hierarchy were those schizophrenics who had a high school education or less.

Those researchers working with mentally-abnormal subjects have been primarily concerned with inabilities rather than abilities and the instruments were organized as problem solving tasks with the subject confronted with the task of discovering the proper solution. Classification problems were created and the subjects asked to discover the predetermined solution.

This contrasts with the research done concerning the development of the child where the solution was not predetermined, but left to the discretion of the subject with the experimenter accepting the solution given and attempting to uncover the underlying cognitive concept.
Rapaport, Gill, and Schafer (1945, p. 391) came to the conclusion that concept formation is a basic thought process and in the following years researchers, influenced in part by Rapaport, began investigating the dimensions of concept formation in the development of the child. Dimensions such as equivalence range and mode of categorization emerged from these investigations.

Equivalence range is a concept developed from categorizing behavior which refers to the degree of attribute divergency of stimuli accepted by the subject in a given category in a categorizing task. Gardner's (1953) work indicated that equivalence range for an individual had a positive correlation across five tests dealing with object sorting, size constancy, and brightness of illumination. His pilot study gave a subjective indication that categorization may be dependent upon time available for the task at hand.

Subsequently in a later study, Gardner, Jackson, and Messick (1960) found no relationship between size constancy and object sorting although Gardner in his earlier study had.

Fillenbaum (1959) reported a coarseness-fineness continuum in categorizing behavior in reference to category size and criterial aspects. This would appear to be related to the concept of category breadth or the degree
of tolerance that an individual has for deviations from a standard yet warranting inclusion in a given category.

Gardner et al. (1960) found that the retinal size of objects tended to relate to grouping of similar sized objects. In this study the authors reported that the number of groups formed by the female subjects used in the study had virtually no relationship to mathematical aptitude, but correlated negatively with the results obtained on the Full Range Picture Vocabulary Test. The authors stated in conclusion that, "... the major theoretical contribution of the present study derives from the clear evidence that intellectual abilities and cognitive controls are not isolated aspects of cognitive organization but mutually related [p. 123]."

The findings above relating to mathematical aptitudes differ from that of Pettigrew (1958, pp. 532-544) who found that with a sample comprised of both males and females there was a correlation between category width and the Quantitative score from the American College Entrance Examination \( r = .26 \). Loomis and Moskowitz (1958) compared tolerance of ambiguity and equivalence range and found a positive correlation on this comparison as well.

Messick and Kogan (1965) factor analyzed a battery of tests consisting of Pettigrew's Estimation Questionnaire and three versions of a quantitative aptitude test. The first version of the test had widely diverging response
choices, the second had narrowly diverging response choices, and the last version required the subject to supply the answers in an open-ended format. The factor analysis of the Pettigrew instrument yielded two factors, the first factor was not related to any of the aptitude test forms, but the second had a common loading with the broad response form of the aptitude test. This suggested to the authors that those subjects who had a tendency for broad range categorization were able to eliminate those answers that did not appear to be within the realm of possibility and reject those more rapidly than those subjects who demonstrated a narrow categorization preference. The advantage for the broad categorizers disappeared with the other two versions of the test. It occurs to the writer that this might possibly account for the divergence in findings between Pettigrew (1958) with a positive correlation between breadth of category and the quantitative section of the ACE and Gardner et al. (1960) with a chance relationship between the same two variables. If the test used in the latter study was one of the narrow breadth type one would not expect a relationship, while the instrument used in Pettigrew's study may have a relatively broader response format thereby yielding a positive correlation.

Cross-cultural studies have been done by Price-Williams (1962) who found, when using materials indigenous
to the area, that there were no differences in size of category between children of the same Tiv tribal group, half of whom had attended a local school and half had not. Indigenous materials were defined as those materials which the children would come in contact with outside the environment of the school. (Two black beetles and a fish were among the materials chosen.) A parallel test was constructed using school materials and in this instance there was a significant difference between groups in their categorizing behavior. The average age was approximately nine years, but due to the lack of records for the non-school sample this was only an estimate.

Scholnick, Osler, and Katzenellenbogen (1968) investigated concept formation across two grades with middle and lower class subjects. After a training period the experimental groups showed no IQ effect nor did social class have a significant effect, although with the non-trained control groups both of these variables were significant.

Greenfield et al. (in Bruner et al., 1966) discussed a series of studies using rural and urban populations in Mexico and Senegal. Another study comparing Eskimos and whites in Alaska was presented by Reich in the same publication. These findings were compared on a combined basis and the most impressive difference in the rural-urban dichotomy was that of degree of abstraction. The rural
children tended to categorize using concrete criteria, but the urban children as the age of the group increased used more and more abstract criteria for categorization. The authors commented, "But what is most striking in the extension of this difference is when we compare schooled and unschooled children. Schooling appears to be the most powerful factor we have found in the stimulation of abstraction [Bruner et al., 1966, p. 315]."

Reichard, Schneider, and Rapaport (1944) found the same type of concrete-abstract hierarchy noted in the previous study in the acquisition of concepts, beginning with the concrete attributes of a concept and progressing through the functional attributes and culminating in the categorization of the acquired concept into a general class and conceptual purpose.

Chase (1969) investigated the concrete-abstract continuum and its relationship to reading by relating a story to a fourth grade sample ($N = 65$) and a sixth grade sample ($N = 70$) and asking the subjects to explain the meaning of the story. Those students who explained the meaning in concrete terms were compared with those who used abstract ideas. A crossbreak analysis of reading ability with grade level and concrete-abstract used as partitioning variables yielded a highly significant difference ($>.001$) with the better readers using more abstraction in the story
meaning explanation. It was also found that more sixth
graders used abstraction than did fourth graders.

Kagan et al. (1960, 1963) developed the concept of
two general types of categorization in children, *stimulus*
centered and *egocentric*. Both of these general types have
three formal conceptual categories that have been identi-
fied by the authors. The first category is that of *descrip-
tive*, based on similarity of object physical attributes
within a grouping. The second categorization mode is that
of *categorical-inferential* characterized by each item of a
grouping formed by the subject exemplifying the categoriza-
tional criterion, but not hinging upon a physical attribute
or direct reference to the items in the grouping. The
third categorizational mode is that of *relational*. The
criterial attribute of this mode of categorization is that
of functional relationship within the grouping. No single
item is exemplary of the grouping criterion.

Wallach and Kogan (1965) modified this system and
used it in their study of the modes of thinking in young
children so as to accommodate the object sorting tasks they
employed. The results indicated that the males differed
from the females in the percentage of descriptive groups
formed, but did not vary significantly in the other two
categories used, inferential and relational. In addition
to a higher percentage of descriptive groups the males were
found to have a significantly broader mean band width as
well. When creativity and intelligence were used as partitioning variables for females, Wallach and Kogan found no significant differences in any of the modes of categorization. The analysis of the results for males yielded somewhat confusing results which the authors feel may be an artifact of the scoring system which does not take into account the quality of the categorization criterion. The analysis of the descriptive percentage scores was non-significant on either the creativity variable or the intelligence variable. The same was true of the inferential percentage score although the intelligence variable approached significance. The analysis of variance for relational responses was the only analysis of categorization mode that yielded a significant difference, that being the interaction at the .02 level. In that particular case the high intelligence-low creativity cell and the low intelligence-low creativity cell had approximately equal means. The authors indicate that these results may be caused by a modal patterning that contains some elements of commonality between groups, thus accounting for the similarity on some measures, but divergency on others.

Pederson and Wender (1968) used Sigel's Object Sorting Task and correlated the results with some other measures among which was the WISC. Table 1 reflects the results that were obtained with 30 subjects observed at 2-1/2 years of age on the behavioral scales and at a mean
Table 1. Relationships of Conceptual Style and Selected Variables from the Pederson and Wender (1968) Study

<table>
<thead>
<tr>
<th>Categorizational mode</th>
<th>Relational</th>
<th>Inferential</th>
<th>Descriptive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variables</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Intelligence)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WISC Performance</td>
<td>-.25</td>
<td>.35</td>
<td>.15</td>
</tr>
<tr>
<td>WISC Verbal</td>
<td>.07</td>
<td>.07</td>
<td>-.20</td>
</tr>
<tr>
<td>(Behavioral)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical Contact</td>
<td>.46</td>
<td>-.34</td>
<td>-.17</td>
</tr>
<tr>
<td>Attention Seeking</td>
<td>.11</td>
<td>-.38</td>
<td>.14</td>
</tr>
<tr>
<td>Orality</td>
<td>.48</td>
<td>-.35</td>
<td>-.18</td>
</tr>
</tbody>
</table>

The subjects in the study had been observed in a nursery at the younger age level and the behavioral scores recorded at that time. Physical contact was defined as seeking physical contact with an adult, attention seeking was recorded when the subject sought non-needed help or recognition and orality indicated the amount of licking, drooling, or thumb-sucking the subject engaged in.

Hughes (Hughes and Coxon, 1968) developed a categorization task which was used by Smart (1969) with Mexican-American children of lower and middle socioeconomic
levels. She found that the categorization on the tasks did not differ significantly in cross-class comparisons. In the instrument used, the definitions of categories were based on that of Wallach and Kogan (1965) which in turn originated from the classification system of Kagan et al. (1960, 1963). In the Intellectual Tasks: Grouping (Hughes and Coxon, 1968) these categories were adapted to a figure sorting task and extended to a fourth category, thematic, as well. The categories developed are utilized in this study and are defined in Appendix A.

In summary, the literature concerning categorization is still far from complete. Several studies have indicated that aspects of categorization are related to achievement and intelligence to a significant degree, but the role of categorization in learning the primary school skills is certainly not clear at the present time. Different dimensions of categorization have been tentatively identified in various studies such as the concrete-abstraction dimension, the modes of categorization, and equivalence range. On a developmental level much of the research is on a theoretical basis and to a lesser degree experimental.

**Psychometric Testing and the Culturally Deprived**

A large body of literature including the appearance of several journals has been formed in the area of
psychological testing since the turn of the century. World War I brought a surge to the testing movement as millions of men entering the armed forces were classified through the use of the Army Alpha (Thorndike and Hagen, 1961). At the same time the psychometricians became aware of the shortcomings of the verbal intelligence test when used with illiterates or those whose native tongue was other than English. The Army Beta was a response to this problem, but the problem loomed larger in the period after World War II. Intelligence tests were used to place students in homogeneous groups for instructional purposes or as admittance standards to special classes, both advanced and retarded. The "culture-free" test came into being, such as the Davis-Eells, then later the culture-free aspect was altered to "culture-fair" when further research indicated that the culture-free tests tended to correlate with socioeconomic status and thereby exhibited the same "sins" in discriminating against the culturally deprived individual as did the more traditional tests. Some of the same underlying factors such as the common experience of all individuals in the social environment that influence the traditional standardized tests influence the culture-free type as well (Thorndike and Hagen, 1961).

At the same time the concept of intelligence became so inextricably confounded with the concept of ultimate intellectual capacity of the individual in the mind of the
unsophisticated that there was a strong opposition to the use of any standardized test. Several books appeared (Gross, 1962) attacking psychological testing and New York City abolished the use of intelligence tests for placement in the public school system. Professionals in the field were concerned as well (Fishman et al., 1964).

Investigations concerned with the use of standardized tests and the culturally disadvantaged revealed that properly used and interpreted, the intelligence and achievement results gained from testing the culturally deprived can give the educator and the psychologist valuable information.

The picture vocabulary tests have been found to estimate verbal IQ to a significant level. Anderson and Flax (1968) found that there is a significant correlation ($r = .38$ for six-year-olds, .59 for seven-year-olds, and .41 for eight-year-olds) between the Peabody Picture Vocabulary Test and the WISC. This degree of relationship state the authors, is about the same as the Performance and Verbal sections of the WISC correlate with each other.

The Metropolitan Readiness Tests (MRT) and the earlier version the Metropolitan Reading Readiness Tests (MRRT) have been frequently used in the assessment of the readiness of the child to begin school and for the curriculum of the first grade. Bagford (1968) reported that the MRRT given at the first grade level correlated with the
Lorge-Thorndike Intelligence Test at sixth grade level at the .50 level for the Verbal portion and .41 for the Performance part. This same approximate level of correlation was observed when the MRRT was correlated with the Iowa Tests of Basic Skills in the fourth through sixth grades.

When teachers estimated the level of readiness of their students after readiness training in kindergarten on a scale of five, as reported in a study by Koppman and LaPray (1969) the quintiserial correlations were .47 for Copying, .67 for Letter Knowledge, but only .02 for Word Knowledge, all subtests of the MRT.

Sipay (1964) conducted a study to find out at what level of functioning the standardized tests actually measured. He used, among others, the Metropolitan Achievement Tests. The author devised a paragraph reading test based on the Scott, Foresman series with nine levels of difficulty. Two hundred-and-two fourth grade students became the sample and three criterial levels of reading achievement were established. Level one was defined as correctly pronouncing 96+% of the words in the paragraph with comprehension at the 60% level, level two required 90+% pronunciation with the same level of comprehension, and the frustration level occurred when the student could not pronounce more than 10% of the words and comprehension fell to 50%. The MAT gave a mean grade level difference of
.79 at level one, .11 at level two, and -1.15 at the frustration level. This indicated to Sipay that the MAT gives approximately the same grade level as level two in paragraph reading, but is significantly different if the criteria of levels one or three are used.

Kingston (1962) used the Metropolitan Reading Readiness Tests (1950 version) to predict third and fourth grade achievement for 272 subjects and found that the MRRT subtests varied in predictive validity from .3 to .6. He reported that Numbers and Matching subtests predict better than do other parts of the MRRT with overall achievement on the Stanford Achievement Tests. The author indicated that Numbers seemed to be more heavily weighted on mental ability and Matching on perceptual ability. Weintraub (1967) supported this finding when he reported that Numbers on the MRT was the best predictor of reading achievement among the subtests and was almost the equal of the combined total in prediction of reading at the end of first grade.

There is very little doubt that the various readiness tests do predict reading for a portion of the population, but the question has arisen as to this prediction with children coming from culturally deprived areas. Anastasi (1968) points out that since the school environment is heavily loaded with the same factors that the intelligence and aptitude tests measure the predictive ability is quite real for all groups culturally deprived
and otherwise on a group basis. Karp and Sigel (1965) reviewed the literature and supported Anastasi's viewpoint when they indicated that there seems to be a handicap in absolute terms, but since the criterion is usually based on a cognitive structure similar to the one measured by the aptitude test the predictive validity is much the same for the culturally deprived child as it is for the middle class Anglo-American child.

Hanson and Robinson (1967) measured Negro and white children entering school and as second and third graders. They found that the Negroes entered first grade at a lower average level of readiness than the whites and proceeded to fall farther and farther behind as they progressed through the first three years of school. The same two authors (Robinson and Hanson, 1968) found that the Goodenough-Harris Draw-a-Man Test, the MRT, and the MAT:Primary I had reliabilities approximating the norming groups when used with Negro students during the first year of school. See Table 2 for the coefficients.

Mitchell (1967) found that the MRT has about the same validity coefficients for Mexican-Americans as it does for Anglo-American students. The mean for the cultural minorities in Mitchell's study was about a standard deviation below the mean for the Anglo-American sample and the same was true for the criterion measure, the Stanford Achievement Tests, 1963 Revision, Primary I, Form X. The
Table 2. Reliability Coefficients from the Robinson and Hanson (1968) Study

<table>
<thead>
<tr>
<th>Test</th>
<th>1st Grade</th>
<th>2nd Grade</th>
<th>3rd Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draw-a-Man</td>
<td>.86*</td>
<td>.91*</td>
<td>.81*</td>
</tr>
<tr>
<td>MRT</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT: Primary I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Discrimination</td>
<td>.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT: Primary II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Discrimination</td>
<td>.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT: Elementary Form</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Discrimination</td>
<td>.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Test-Retest, all others are KR 20 reliability coefficients.*
mean age of the sample was approximately 6.5 years old at the time of the initial testing. In an earlier study Mitchell (1962) found the same type of results using the MRRT and the MAT:1959 Form as the criterion. The combined correlation matrix for both studies is given in Table 3 and indicates that the tests are about as effective for the Anglo-American population as they are for the Negro and other cultural minorities in terms of prediction of those particular criterial factors that are measured in the MAT:1959 Form.


<table>
<thead>
<tr>
<th>Criterion Measure</th>
<th>Validity Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford Achievement Test</td>
<td></td>
</tr>
<tr>
<td>Paragraph Meaning</td>
<td>.56</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>.59</td>
</tr>
<tr>
<td>Word Study Skills</td>
<td>.54</td>
</tr>
<tr>
<td>Spelling</td>
<td>.54</td>
</tr>
<tr>
<td>Metropolitan Achievement Tests</td>
<td></td>
</tr>
<tr>
<td>Word Knowledge</td>
<td>.56</td>
</tr>
<tr>
<td>Word Discrimination</td>
<td>.56</td>
</tr>
<tr>
<td>Reading</td>
<td>.51</td>
</tr>
<tr>
<td>Average of the three MAT scores</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>.58</td>
</tr>
</tbody>
</table>
Arnold (1969) found that the reliability of the MAT:Primary II to be .95 when using Cronbach's Alpha measure. The subtest results were: (1) $r = .87$ for Word Knowledge, (2) $r = .89$ for Word Discrimination, and (3) $r = .89$ for Reading. The sample was composed of 210 Mexican-American children in the third grade.

The Goodenough-Harris Draw-a-Man Test has been used to predict achievement in the primary grades as well. Panther (1967) reports that the Draw-a-Man Test correlated with the Reading subtest of the MAT:Primary I at the .34 level ($p = .05$) at the end of first grade. Shipp and Louden (1964) administered the Draw-a-Man Test to 56 first grade boys and 59 girls in the same grade and found that the results correlated with achievement at the end of the year as measured by the Gray, Votaw, and Rogers Primary Achievement Tests total score at the .51 level for the entire sample. After separating the sexes, the correlation for the same two variables gave a figure of .44 for the girls and .50 for the boys. This group had a mean score of 113 on the IQ score of the Draw-a-Man Test.

Even though the literature indicates that the reliability and validity for the instruments used in this study are satisfactory for the Mexican-American sample there are still shortcomings in the use of these instruments. Darcy (1953, 1963) reviewed the literature dealing with language in the use of standardized tests in some
detail. She noted that when socioeconomic level is held constant there is still a difference in the verbal results, but that performance tends to be quite similar on English language tests.

The factor of speed enters into the cultural bias problem as well. Knapp (1960) tested a group of bilingual Mexican visa applicants (adults) and a monolingual group of job applicants on two formats of the Cattell Culture Free Test, one speeded and the other non-speeded and found that the discrepancy between the two groups, although favoring the monolinguals on each form was significantly less with the non-speeded version. Socioeconomic class was not controlled in that the Mexicans were applying for visas to do farm labor while the Anglo-Americans were applying for skilled aircraft industry jobs.

There is an alternative that suggests itself in the testing of bilinguals and that is to use the "native" language, usually not the language used in the schools, as the testing language. Keston and Jiminez (1954) tested fifty bilingual children in elementary school with both an English and Spanish version of the Stanford-Binet and reported that the mean on the English version was 86.0 contrasted with the Spanish version mean of 71.8, a significant difference. The standard deviations were 15.3 and 9.9 respectively indicating that the Spanish version did not have the variability of the English one and indicating
to the authors that the range of language used in the home is not equal to that used in the school as far as diversity is concerned. When the two versions were correlated with grade point average the resulting correlations were .11 for the Spanish version while the English version correlated at the .62 level, or about the same as for a monolingual middle-class group.

Jensen (1961) found that the intelligence level of a group of Anglo-Americans as measured by the California Test of Mental Maturity was related to the learning abilities demonstrated on a series of tests which Jensen devised to measure immediate recall, serial learning, and paired associates learning using familiar and abstract objects. For the Mexican-Americans on the contrary, the CTMM did not differentiate between rapid and slow learners and the verbal factor measured tended to be equated for the entire Mexican-American group with the variance arising from the performance scores.

In summary, the research in the area of testing subjects who are bilingual or culturally deprived, or both, indicates that the test results are usually biased against the deviant groups in terms of level of attainment. Conversely the results derived from these tests tend to correlate positively and significantly with the criterial learning tasks posed in the school environment that have a
verbal content that is similar in general to the tests themselves, and thus the tests can serve as predictors despite the departure from norm in achievement level. This assumes of course that the fallacy of considering the results of a test to be a measure of the ultimate intellectual capacity of the individual is guarded against at all times.

This study is concerned with the relationships of cognitive abilities to achievement in the upper primary levels. The cognitive abilities have been described in the literature reviewed on a traditional readiness basis, stemming from perceptual skills, problem solving abilities, plus the amount of knowledge of entry level academic skills such as recognition of alphabetic symbols. In addition readiness as defined by developmental level in conceptual abilities was discussed and reviewed in the work of such individuals as Piaget, Bruner, and Kagan and his associates. Research studies reporting the validity of the standardized primary instruments have shown that there is a considerable degree of predictive validity for the reading criterion. Less clear is the predictive validity of preferred mode of categorization, yet several studies suggest that this area of cognitive functioning has an effect on learning ability therefore the two areas are of mutual interest in the review of the literature and analysis of the experimental results.
CHAPTER III
SAMPLE, MATERIALS, AND PROCEDURES

The sample was formed by randomly selecting a group of Mexican-American children from entering first grade classes in elementary schools located in poverty areas. The sample selected was then administered selected intelligence, achievement, and discrimination tests as well as an instrument developed by Hughes (Hughes and Coxon, 1968) at the Arizona Center for Early Childhood Education. This instrument was designed to determine the preferred mode of categorization of the subject.

Sample

The sample used in this study was selected by using a table of random numbers and assigning identification numbers to each child in the population who fulfilled the following criteria: (1) Mexican-American ethnic background as defined by Spanish surname or by Spanish being the language of the home based on school records, (2) attendance at a school qualified to receive poverty funds under the provisions of PL 89-10, and (3) initial first grade experience in the 1966-1967 academic year.

Those students whose identification numbers were matched by the randomly drawn numbers were included in the sample. The sample used in this study was part of a larger
sample who were studied under a cooperative project between Tucson School District #1 and The University of Arizona. The original group in the study consisted of 87 students, 40 females and 47 males. As one might expect the sample was somewhat attenuated in the final testing period during the spring of the third year, but it is believed that the subjects lost from the sample were not lost due to any factor that would introduce a systematic bias in the results. Therefore, the sample was still considered to be random for the purposes of the study.

The average age of the sample was six years, two months as of the first of September, 1966, when the study began. Table 4 indicates that there was not a significant difference in age between males and females in the sample.

Table 4. Age of Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean Age in Months</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>74.60</td>
<td>3.89</td>
</tr>
<tr>
<td>Females</td>
<td>74.17</td>
<td>4.09</td>
</tr>
<tr>
<td>Total group</td>
<td>74.39</td>
<td>3.99</td>
</tr>
</tbody>
</table>
Instruments and Materials

The intellectual tests used in this study are among those commonly used in the public schools at the lower primary level. It was believed by the writer that in specific instances the normal type of administration procedures would not result in the optimum performance of the child, therefore certain deviations from the administrative procedures given in the manuals were instituted. These deviations are noted in the description of the administrative procedures given for each test.

Standardized Primary Instruments

These are tests which are published by commercial firms and with norms presented based on a national sample. There is one exception to this statement, the Hughes Car Discrimination Task was developed at the Arizona Center for Early Childhood Education and norms have not yet been developed. Inter-group comparisons use the scoring system developed by the Center.

Van Alstyne Picture Vocabulary Test. This 60 question picture vocabulary test has a relatively large number of items that are of a difficulty level appropriate for the sample in this study. It also has been translated by personnel of the Center for use in comparative studies into the Spanish language. The test is conducted by the examiner reading a stimulus word and the subject responding
by pointing to one of four pictures which is thought to be the most closely associated with the stimulus word.

The manual (Van Alstyne, 1961) gives a split-half reliability coefficient of .71 in the age range represented in the sample in this study. Validity was determined by using the Lorge-Thorndike Intelligence Test at the appropriate level and resulted in a validity coefficient of .69.

**Metropolitan Readiness Tests.** The Metropolitan Readiness Tests form a battery of tests designed to measure "the extent to which school beginners have developed in the several skills and abilities that contribute to readiness for first-grade instruction [Hildreth, Griffiths, and McGauvran, 1965, p. 3]."

It was believed that although the subjects in the sample were not necessarily immature or retarded, assessment would be more effective if the group size was less than the recommended fifteen and so with this test the subjects were tested in groups of five.

The Manual of Directions: Metropolitan Readiness Tests (Hildreth et al., 1965) gives the following descriptions for the six tests making up the battery:

**Test 1. Word Meaning,** a 16-item picture vocabulary test. The pupil selects from three pictures the one that illustrates the word the examiner names.

**Test 2. Listening,** a 16-item test of ability to comprehend phrases and sentences instead of
individual words. The pupil selects from three pictures the one which portrays a situation or event the examiner describes briefly.

Test 3. Matching, a 14-item test of visual perception involving the recognition of similarities. The pupil marks one of three pictures which matches a given picture.

Test 4. Alphabet, a 16-item test of ability to recognize lower case letters of the alphabet. The pupil marks one of three pictures which matches a given picture.

Test 5. Numbers, a 26-item test of number knowledge.

Test 6. Copying, a 14-item test which measures a combination of visual perception and motor control (p. 3).

The reliability reported for this battery was based on three separate samples and the averaged reliability figures based on odd-even coefficients corrected by the Spearman-Brown formula ranged from .63 to .93 with the exception of Listening which had a rather low reliability at .39.

The predictive validity of the Metropolitan Readiness Tests was given as median correlations of several studies, using the Metropolitan Achievement Tests:Primary I, Form A as the criterion. The correlations with Word Knowledge of the latter test was .67 and .65 for Reading.

Goodenough-Harris Drawing Test. This test was administered to the sample in the classroom environment with the entire class tested concurrently. The only deviation from the published instructions (Harris, 1963) was
that the directions were read over a tape recorder in both English and Spanish. The reliability was reported for this test by Harris (1963) as a range from .60 to .70 on a test-retest over a period of three months and the validity as .45 when using the Stanford Binet as the criterion measure with a group of kindergarteners.

Hughes Car Discrimination Task. This task was devised by Hughes (Hughes and Coxon, 1968) as an instrument to measure the ability of young children to discriminate between similar, but not identical objects. It measures both gross discrimination and fine discrimination, but only the fine discrimination function is used in this study. It is an individually administered instrument. The task materials consist of five model cars approximately four inches long and painted silver. Two of the cars are identical sedans, a third is a sports car, a fourth a convertible, and the last another type of sedan. The models are highly detailed and the subject is allowed to manipulate them during the course of the task administration and inspect them in any manner in which he deems appropriate.

The two identical sedans are initially presented and the subject asked to decide if they are the same or different. A second pair of cars consisting of the sports car and the convertible are then presented. If, in
response to the question of similarity of the two, the subject states that they are different he is then asked to indicate all the differences, but if the subject indicates that they are the same the questioning ceases. The latter occurrence was rather rare in this sample. The request to indicate all the differences is repeated twice and all responses are recorded. The same procedure is used with the third pair of cars composed of the two unlike sedans.

These tasks yielded discrimination frequency scores, uniqueness scores, and saliency scores for the first and second sets of cars, plus a total combined score for each of the types of scores. If one considers these to be parallel tests the following reliability coefficients can be shown (Table 5). Appendix B gives the rules and the derivations of the scores in this task.

Table 5. Test-Retest Correlations for Hughes Car Discrimination Task

<table>
<thead>
<tr>
<th>Sedan</th>
<th>Discrimination</th>
<th>Uniqueness</th>
<th>Saliency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convertible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
<td>.55</td>
<td>.46</td>
<td>.52</td>
</tr>
<tr>
<td>Uniqueness</td>
<td>.53</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
<td>.36</td>
<td></td>
</tr>
</tbody>
</table>

N = 72.
**Intellectual Tasks: Grouping.** This type of measure was used by Wallach and Kogan (1965) and modified by Hughes (Hughes and Coxon, 1968) to investigate the categorizing behavior of young children, using high interest materials in an open-ended format. The materials used in this task are composed of rubber figures of jungle animals, farm animals, and human figures. Appendix C has a list of all the figures used. The human figures are of service workers such as postmen, doctors, nurses, and firemen, or potential family group members such as babies, pre-adolescents, as well as adult males and females. The skin tone of half of the human figures is somewhat darker than that of the other half, providing a potential grouping on a skin tone criterion.

The task is individually administered by an examiner and a recorder. The verbal responses of the subject are recorded together with the physical grouping produced. The task consists of two parts, Labeling and Grouping. The Labeling section is intended to familiarize the subjects with the materials and build rapport and at the same time provide a verbal means of identifying any figure that was not previously known. The score is the number of figures the subject is able to identify, judged on the criteria: (1) the response is the name, but not the functional definition, and (2) if at least one of a given type group is correctly identified then credit is given for
that type of figure. The maximum score is 27 because of duplication of figures rather than 44, the total number of figures.

The second part of the task is Grouping. The subject is asked to organize the figures into groups with the instruction, "Show me what things go together." After the subject completes the grouping the examiner then asks "Why do these things go together?" The responses are recorded and the figures returned to their original position at one end of the table, in a random mass. The complete task requires the subject to repeat the grouping procedure in a "different way," but this part was not used in this study.

Several schemes of categorization are reported in the literature, but the model proposed by Kagan, Moss, and Sigel (1960) was selected for this study because of the number of studies using both this model and achievement scores in their design. (See Appendix A for scoring rules and Table 6 for types of scores.)

The inter-scorer reliability was established for the Intellectual Tasks: Grouping by selecting on a random basis, fifteen protocols. These protocols were duplicated and given to the scorers who were to eventually score all protocols used in this study. After the scorers had scored their set of protocols independently, the results were correlated using a Spearman Rho correlation for the
Table 6. Scores Derived from Intellectual Tasks: Grouping

<table>
<thead>
<tr>
<th>Name of Score</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeling</td>
<td>No. of figures labeled</td>
</tr>
<tr>
<td>Total correct items</td>
<td>No. of figures used correctly in all categories</td>
</tr>
<tr>
<td>Flexibility</td>
<td>No. of categories used</td>
</tr>
<tr>
<td>Descriptive groupings</td>
<td>No. of groups in that category</td>
</tr>
<tr>
<td>Categorical-Inferential groupings</td>
<td>No. of groups in that category</td>
</tr>
<tr>
<td>Relational groupings</td>
<td>No. of groups in that category</td>
</tr>
<tr>
<td>Thematic groupings</td>
<td>No. of groups in that category</td>
</tr>
<tr>
<td>Error groupings</td>
<td>No. of groups that are in error according to scoring rule</td>
</tr>
<tr>
<td>Descriptive item frequency</td>
<td>No. if items in that mode</td>
</tr>
<tr>
<td>Categorical-Inferential items frequency</td>
<td>No. of items in that mode</td>
</tr>
<tr>
<td>Relational item frequency</td>
<td>No. of items in that mode</td>
</tr>
<tr>
<td>Thematic item frequency</td>
<td>No. of items in that mode</td>
</tr>
</tbody>
</table>
classification of categorization mode. The results are presented in Table 7.

Table 7. Inter-Scorer Reliability for Intellectual Tasks: Grouping

<table>
<thead>
<tr>
<th>Category</th>
<th>Reliability Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>.99</td>
</tr>
<tr>
<td>Categorical-Inferential</td>
<td>.87</td>
</tr>
<tr>
<td>Relational</td>
<td>.99</td>
</tr>
<tr>
<td>Thematic</td>
<td>.79</td>
</tr>
</tbody>
</table>

Due to the procedure used in the Intellectual Tasks: Grouping the conventional means of determining reliability are not appropriate, but it can be inferred from other studies that this type of measure has a certain degree of reliability. Sloane et al. (1963), using an object sorting test reported a reliability of .75. Gardner (1953) determined that the reliabilities on the object sorting tests in his study were .55 and .53. Both of these studies utilized object sorting tests as does the present one.
Transformations

As one might expect with a deviant sample, the scores on most of the standardized primary measures were skewed in a positive direction, therefore in order to more nearly satisfy the assumptions of the parametric statistics used, it was necessary to effect some type of transformation. According to Winer (1962) and Nunnally (1967) the form of variation from normal exhibited by the test results is most frequently countered by using a logarithmic transformation. Winer also suggests using a coding system consisting of adding one point to all scores so as to have greater than zero values for each valid score. Since this aids in the technical aspects of computer analysis as well this suggestion was followed.

Time Schedule of Testing

The study took place over a three year period, extending from first grade to third grade. Table 8 gives the testing schedule and the tests involved.
Table 8. Testing Schedule

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Val Alstyne Pic. Vocab.</td>
<td>Fall</td>
<td></td>
</tr>
<tr>
<td>Metropolitan Readiness</td>
<td>Fall</td>
<td></td>
</tr>
<tr>
<td>Intellectual Tasks: Gp.</td>
<td>Spring</td>
<td></td>
</tr>
<tr>
<td>Hughes Car Discrimination</td>
<td>Spring</td>
<td></td>
</tr>
<tr>
<td>Goodenough-Harris Drawing</td>
<td>Fall</td>
<td></td>
</tr>
<tr>
<td>Metropolitan Achievement</td>
<td></td>
<td>Spring</td>
</tr>
</tbody>
</table>
CHAPTER IV

RESULTS

This chapter of the study is organized into five sections. The first is concerned with results and the four hypotheses are discussed in the succeeding four sections.

Results on Preferred Mode of Categorization

This section deals with the results concerning the concept of preferred mode of categorization and variations thereof. Also included in this section are the results of the comparisons of preferred mode of categorization and sex of the subject.

Types of Preferred Mode of Categorization

The concept of "preferred mode of categorization" may be defined by several criteria. In this study two types of preferred mode of categorization are used. The first type is based on the criterion which states that the preferred mode of categorization is that mode of categorization which the individual uses in forming the greatest number of groups. The second type of preferred mode of categorization has as its criterion the number of items grouped under a given mode of categorization. The mode of categorization that has the largest number of
items under its aegis is considered the preferred mode of
categorization.

In order to examine the relationships of types of
preferred mode of categorization other than the two used
in the study, two other types of preferred mode of
categorization were developed. These additional types
were based upon z-score transformations (McNemar, 1962) of
the frequencies in each category. The first of these addi­tional types was developed by using the frequency of items
in each mode as scores and calculating the mean and
standard deviation of each mode then converting each
individual's set of scores into z-scores and using the
criterion of the largest z-score for each individual as
indicating that individual's preferred mode of categoriza­tion. The second additional type used the same method but
instead of number of items in calculating z-score the
number of groups in each category was used. The preferred
mode was determined by using the same criterion as above.

Individual Mode Intercorrelations

The correlation matrices of the preferred mode
scores defined by the frequency criterion and the z-score
criterion are given in order of descriptive, categorical­
inferential, relational, and thematic.

Descriptive mode intercorrelations. Table 9 indi­
cates that the various possible types of preferred mode of
Table 9. Inter-relationships\textsuperscript{a} of Descriptive Preferred Mode Types

<table>
<thead>
<tr>
<th>Variable</th>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. z-score (items)</td>
<td>1.00</td>
<td>.95</td>
<td>.92</td>
<td>.93</td>
</tr>
<tr>
<td>2. z-score (groups)</td>
<td>.00</td>
<td>.87</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>3. Frequency (items)</td>
<td></td>
<td>1.00</td>
<td>.90</td>
<td></td>
</tr>
<tr>
<td>4. Frequency (groups)</td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}Phi Coefficients.

categorization for the descriptive mode are quite similar with an average correlation of .91 after converting the individual correlations to Fisher's Z, averaging and reconverting to the average correlation coefficient (McNemar, 1962).

**Categorical-inferential mode intercorrelations.**
The categorical-inferential preferred mode of categorization results are in Table 10. It is indicated that in this mode the types are less related than in the preceding mode. The average of the correlations for all types of preferred mode of categorization is .79, when subjected to the averaging procedure described previously.

**Relational mode intercorrelations.** The types of relational mode preference intercorrelations are given in
Table 10. Inter-relationships\textsuperscript{a} of Categorical-Inferential Preferred Mode Types

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. z-score (items)</td>
<td>1.00</td>
<td>.82</td>
<td>.72</td>
<td>.68</td>
</tr>
<tr>
<td>2. z-score (groups)</td>
<td>1.00</td>
<td>.66</td>
<td>.63</td>
<td></td>
</tr>
<tr>
<td>3. Frequency (items)</td>
<td></td>
<td>1.00</td>
<td>.91</td>
<td></td>
</tr>
<tr>
<td>4. Frequency (groups)</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Phi Coefficients.

Table 11. The average correlation for all types of preferred mode of categorization is .74 and not significantly different from that of the categorical-inferential average correlation.

**Thematic mode intercorrelations.** The thematic mode preference types were analyzed in a similar manner and the results given in Table 12. The average of the various types of preferred mode in the thematic category is somewhat lower than the preceding three comparable figures for the other modes at .67.

**Summary.** The average correlations for categorical-inferential, relational, and thematic are not significantly different; however, the descriptive average correlation is significantly different from the highest of the other three
Table 11. Inter-relationships\(^a\) of Relational Preferred Mode Types

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. z-score (items)</td>
<td>1.00</td>
<td>.88</td>
<td>.54</td>
<td>.58</td>
</tr>
<tr>
<td>2. z-score (groups)</td>
<td>1.00</td>
<td>.62</td>
<td>.58</td>
<td></td>
</tr>
<tr>
<td>3. Frequency (items)</td>
<td></td>
<td>1.00</td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>4. Frequency (groups)</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

\(^a\)Phi Coefficients.

Table 12. Inter-relationships\(^a\) of Thematic Preferred Mode Types

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. z-score (items)</td>
<td>1.00</td>
<td>.70</td>
<td>.78</td>
<td>.61</td>
</tr>
<tr>
<td>2. z-score (groups)</td>
<td>1.00</td>
<td>.54</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>3. Frequency (items)</td>
<td></td>
<td>1.00</td>
<td></td>
<td>.79</td>
</tr>
<tr>
<td>4. Frequency (groups)</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

\(^a\)Phi Coefficients.
Traditionally number of groups has been the metric used in studies of cognitive style (Kagan et al., 1960, 1963; Wallach and Kogan, 1965) and preferred mode of categorization; therefore, the decision was made to use frequency of groups and frequency of items as the criteria for preferred mode of categorization. It also facilitates comparison with other studies when the same metric is used. The item frequency criterion was added to provide a slightly different perspective for the same situation.

Combined Mode Inter-relationships

The frequency counts for each subject of number of groups in a given mode were correlated to determine the inter-relationships. Table 13 presents these data. The measures are not independent in this analysis and necessarily must yield some negative correlations if some are positive. However, it is interesting to note that when a subject used the descriptive mode as a categorizing style there was a significant decrease in the use of other modes as indicated by the -.52 correlation, contrasted with this result are the relationships of the other three modes which indicate that there was not necessarily a significant decrease in use of other modes if a mode other than descriptive was used.

Wallach and Kogan (1965) did a similar study presenting the results with the sexes separated. They
found that for males the correlation was \(-.70\) for the
descriptive and categorical-inferential modes, \(-.29\) for the
descriptive and relational pair, with values of \(-.37\) and
\(-.26\) for the females which are quite similar to Table 13
which combines both the males and females of this study.
The two studies differ on the categorical-inferential and
relational correlations with \(-.26\) and \(-.29\) for males and
females respectively on the Wallach and Kogan study, but
the comparable figure is \(-.09\) for the present study.

Sex Differences in Grouping Results

Analyses of variance were performed using sex as
the independent variable and the scores from the grouping
tasks as dependent variables in order to determine if the
sex of the subject might possibly act as a differentiating factor in grouping performance.

None of the analyses were significant except the one dealing with the number of items in the relational category. Only 42 of the subjects used this category, 21 males and 21 females with the females using significantly more items ($p = .05$) in relational groupings than did the males. This was also reflected in the tendency of the females to use more relational groups than did the males although this tendency was not significant.

As a further check of this relationship a Chi Square test was performed using the preferred mode of categorization (group frequency criterion) as the dependent variable and sex as the independent variable. The Chi Square value was 1.642 indicating a non-significant difference. When a similar analysis was done using the preferred mode of categorization based on the criterion of number of items the Chi Square value was 4.66, still less than significant at the .05 level.

In summary, this section has dealt with the interrelationships of the categorization scores based on item and group frequency counts, and the possibility of difference between sexes. One difference was significant and that was the number of items grouped under the relational category.
Results Concerning Hypothesis One

The first hypothesis states that preferred mode, when used as the independent variable, will show significant differences in intelligence, achievement, and vocabulary measures. In order to test this hypothesis the sample was divided by preferred mode of categorization of the type based on the most groups criterion. The groups resulting from this division were 32, 45, 6, and 4 in number for descriptive, categorical-inferential, relational, and thematic respectively. The results for these analyses are given in Table 14.

The results of the analysis of the MRT Word Meaning and Listening are significant, but later analysis indicates these tests have very low reliability with this group of subjects and therefore this significance is not accepted without further verification by other tests.

Results Concerning Hypothesis Two

Hypothesis two states that preferred mode of categorization, when used as the independent variable, will significantly differentiate between groups on the factor scores of the combined first year standardized primary instruments.

Factor Analysis of First Year Tests

In order to test this hypothesis the standardized primary instruments which were administered the first year
<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>MS&lt;sup&gt;a&lt;/sup&gt;</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRT Word Meaning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>.0566</td>
<td>3.43*</td>
</tr>
<tr>
<td>Within Groups</td>
<td>69</td>
<td>.0174</td>
<td></td>
</tr>
<tr>
<td>MRT Listening</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>.0353</td>
<td>3.06*</td>
</tr>
<tr>
<td>Within Groups</td>
<td>69</td>
<td>.0115</td>
<td></td>
</tr>
<tr>
<td>MRT Numbers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>.0651</td>
<td>2.30</td>
</tr>
<tr>
<td>Within Groups</td>
<td>69</td>
<td>.0283</td>
<td></td>
</tr>
<tr>
<td>MRT Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>.0329</td>
<td>2.37</td>
</tr>
<tr>
<td>Within Groups</td>
<td>69</td>
<td>.0139</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Logarithmic Transformation.

*<sup>p</sup> < .05.
of the study were subjected to factor analysis. The number of factors inherent in these tests was not hypothesized, rather an eigenvalue cutoff point of 1.00 was established in accordance with Nunnally's (1967) recommendations. This analysis resulted in the factor structure given in Table 15. Four factors emerged from the varimax rotation of the factor matrix.

The first factor is loaded heavily with the Sedan scores from the Hughes Car Discrimination Task and lightly with the Convertible scores from the same task, plus the Copying test from the MRT. The second factor appears to be a general readiness or intellectual level judged by the loading from the MRT and the Goodenough-Harris Drawing Test. The third factor loads heavily with the Convertible section of the Hughes Car Discrimination Task and the Alphabet test of the MRT. Factor four seems to be a vocabulary factor with loading coming from the Van Alstyne Picture Vocabulary Test and Word Meaning of the MRT Matching from MRT also loads on this factor, indicating perhaps a perceptual element as well.

Analysis of variance for factor scores. The analysis of variance upon the factor scores using preferred mode of categorization as the independent variable did not yield any significant differences. The closest approach to the stated .05 level of significance was on factor two with
Table 15. Factor Loadings for First Year Standardized Primary Instruments

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor Loadings</th>
<th></th>
<th></th>
<th></th>
<th>h²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor</td>
<td>1.</td>
<td>2.</td>
<td>3.</td>
<td>4.</td>
</tr>
<tr>
<td>MRT Word Meaning</td>
<td>.16</td>
<td>-.14</td>
<td>.10</td>
<td>.83</td>
<td>.74</td>
</tr>
<tr>
<td>MRT Listening</td>
<td>-.01</td>
<td>.66</td>
<td>-.07</td>
<td>.17</td>
<td>.47</td>
</tr>
<tr>
<td>MRT Matching</td>
<td>.02</td>
<td>.52</td>
<td>-.21</td>
<td>.45</td>
<td>.53</td>
</tr>
<tr>
<td>MRT Alphabet</td>
<td>-.23</td>
<td>.50</td>
<td>-.43</td>
<td>.17</td>
<td>.52</td>
</tr>
<tr>
<td>MRT Numbers</td>
<td>-.11</td>
<td>.62</td>
<td>-.38</td>
<td>.18</td>
<td>.57</td>
</tr>
<tr>
<td>MRT Copying</td>
<td>.34</td>
<td>.76</td>
<td>-.15</td>
<td>-.11</td>
<td>.74</td>
</tr>
<tr>
<td>Van Alstyne PVT</td>
<td>-.04</td>
<td>.39</td>
<td>-.31</td>
<td>.62</td>
<td>.65</td>
</tr>
<tr>
<td>Convert. Discrim.</td>
<td>.28</td>
<td>.06</td>
<td>-.90</td>
<td>.06</td>
<td>.89</td>
</tr>
<tr>
<td>Convert. Uniq.</td>
<td>.24</td>
<td>.07</td>
<td>-.85</td>
<td>.07</td>
<td>.80</td>
</tr>
<tr>
<td>Convert. Saliency</td>
<td>.29</td>
<td>-.01</td>
<td>-.78</td>
<td>-.01</td>
<td>.69</td>
</tr>
<tr>
<td>Sedan Discrim.</td>
<td>.95</td>
<td>.02</td>
<td>-.21</td>
<td>.08</td>
<td>.95</td>
</tr>
<tr>
<td>Sedan Uniq.</td>
<td>.90</td>
<td>.02</td>
<td>-.18</td>
<td>.09</td>
<td>.85</td>
</tr>
<tr>
<td>Sedan Saliency</td>
<td>.87</td>
<td>.08</td>
<td>-.25</td>
<td>.04</td>
<td>.83</td>
</tr>
<tr>
<td>G-H Drawing (man)</td>
<td>.11</td>
<td>.83</td>
<td>.13</td>
<td>-.12</td>
<td>.74</td>
</tr>
<tr>
<td>G-H Drawing (woman)</td>
<td>.08</td>
<td>.86</td>
<td>.19</td>
<td>-.03</td>
<td>.78</td>
</tr>
<tr>
<td>Percentage Variance</td>
<td>19.5</td>
<td>23.7</td>
<td>18.9</td>
<td>9.5</td>
<td></td>
</tr>
</tbody>
</table>
a .14 probability of chance. The means of the independent groups indicated that the subjects who used relational as the preferred mode of categorization tended to have the highest factor scores.

Inasmuch as the results of the analysis of variance did not yield any significant differences the second hypothesis is rejected at the .05 level of significance.

**Results Concerning Hypothesis Three**

It was hypothesized that degree of flexibility, defined as the number of modes of categorization that a subject used on the Intellectual Tasks: Grouping, would facilitate solving of the tasks presented in the standardized primary instruments. The greater ability the subject has in terms of flexibility in use of mode of categorization the greater number of mode-linked questions he will be able to answer was the hypothesized outcome. In order to test this hypothesis, four groups were formed, the first group used only one mode the next group two, the third group used three modes, and the fourth group used all four modes of categorization. An analysis of variance of the standardized primary instruments indicated that the only significant differences occurred with the Hughes Car Discrimination Tasks. For comparative purposes the Convertible section results are included although they did not reach the stated significance level (see Table 16).
Table 16. Analysis of Variance Summary Results of Hughes Car Discrimination Task Using Flexibility as the Independent Variable

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convertible Section</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>157.90</td>
<td>1.78</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>88.60</td>
<td></td>
</tr>
<tr>
<td>Uniqueness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>1144.81</td>
<td>2.05</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>559.60</td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>13.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Sedan Section</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>226.15</td>
<td>5.24**</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>43.17</td>
<td></td>
</tr>
<tr>
<td>Uniqueness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>1328.82</td>
<td>5.95**</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>223.46</td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>18.85</td>
<td>2.81*</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>6.70</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>577.41</td>
<td>2.96*</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>194.81</td>
<td></td>
</tr>
<tr>
<td>Uniqueness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>3745.82</td>
<td>3.45*</td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>1084.46</td>
<td></td>
</tr>
<tr>
<td>Saliency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between Groups</td>
<td>3</td>
<td>28.68</td>
<td></td>
</tr>
<tr>
<td>Within Groups</td>
<td>83</td>
<td>30.47</td>
<td>0.94</td>
</tr>
</tbody>
</table>

*p < .05.

**p < .01.
This difference between flexibility levels might be attributed to a higher degree of perseverance for the highly flexible subject because of the Sedan portion of the task following the Convertible section in the task format. Perhaps a better interpretation would be that the more flexible subject is better able to discriminate a larger number of differences in a pair of cars that have a greater degree of similarity and is not affected by a set in the discrimination of differences, but can change from one type of discrimination to another whereas the subject with a less flexible cognitive style exhausts the range of possibilities sooner and hence the lower scores.

**Results Concerning Hypothesis Four**

Hypothesis four stated that the conceptual style as represented by performance on the Intellectual Tasks: Grouping would predict later reading achievement at a significant level.

In order to test this hypothesis and compare the results with the more traditional predictive instruments, two canonical correlations were performed using the computer program by Veldman (1967, p. 286). In canonical correlation one set of variables is used to predict another set. The relationship is optimized by varying the weights of the component variables in each set so as to achieve the highest degree of prediction between the two sets. These
weights are then correlated with the original measures so as to allow interpretation similar to that used with the loading in factor analysis which is mathematically similar to canonical correlation, but uses a symmetrical matrix rather than the possibly asymmetric matrix of canonical correlation.

Canonical Correlation of Categorization and Metropolitan Achievement Tests

In the initial analysis the dichotomized results of the categorization tasks for preferred mode based on number of items in a given mode were used as one set of variables, while two tests of the Metropolitan Achievement Tests, Word Knowledge and Reading were used in the opposing set. Table 17 presents these results with the canonical loading for each of the variables in each set and the canonical correlation.

Relationship of Standardized Primary Instruments to Metropolitan Achievement Tests

For comparative purposes a canonical correlation was also computed for the standardized primary instruments given the first year of the study and the reading measures used in the previous canonical correlation. The results of this correlation are given in Table 18.

The results of the canonical correlation between the first year standardized primary instruments and the Reading and Word Knowledge tests of the MAT indicate a high
Table 17. Canonical Correlation: Preferred Mode of Categorization and MAT Reading and Word Knowledge

<table>
<thead>
<tr>
<th>Set A Variables</th>
<th>Loadings</th>
<th>Set B Variables</th>
<th>Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive (Gp)</td>
<td>.37</td>
<td>MAT Word Know.</td>
<td>-.43</td>
</tr>
<tr>
<td>Cat.-Infer. (Gp)</td>
<td>-.04</td>
<td>MAT Reading</td>
<td>.44</td>
</tr>
<tr>
<td>Relational (Gp)</td>
<td>-.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thematic (Gp)</td>
<td>-.85</td>
<td>The correlation between the sets is .31 which is not significant (p = .77).</td>
<td></td>
</tr>
<tr>
<td>Descriptive (item)</td>
<td>.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cat.-Infer. (item)</td>
<td>.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relational (item)</td>
<td>.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thematic (item)</td>
<td>-.68</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 18. Canonical Correlation: First Year Standardized Primary Instruments and MAT Reading and Word Knowledge

<table>
<thead>
<tr>
<th>Set A Variables</th>
<th>Loadings</th>
<th>Set B Variables</th>
<th>Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRT Word Meaning</td>
<td>.21</td>
<td>MAT Word Know.</td>
<td>.98</td>
</tr>
<tr>
<td>MRT Listening</td>
<td>.42</td>
<td>MAT Reading</td>
<td>.78</td>
</tr>
<tr>
<td>MRT Matching</td>
<td>.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRT Alphabet</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRT Numbers</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRT Copying</td>
<td>.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRT Total</td>
<td>.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Alstyne PVT</td>
<td>.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-H Drawing (man)</td>
<td>.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-H Drawing (woman)</td>
<td>.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Only those variables having a loading of .2+ were included in the table, this criterion was not met by the Hughes Car Discrimination Task scores.*
degree of relationship. The canonical correlation is .74 which is significant beyond the .01 level. The Van Alstyne also contributes to this relationship along with the Goodenough-Harris Drawing Test at about the same level as the MRT subtests with the exceptions of Alphabet and Numbers, which are higher.
CHAPTER V

DISCUSSION, SUMMARY, AND CONCLUSIONS

This chapter is concerned with the discussion of the results of the study, a summation of the study as a whole, and the conclusions drawn from the results of this study.

Discussion

The discussion section of this chapter deals with the results of each of the hypotheses in turn.

Hypothesis One

The preferred mode of categorization is evidenced by the highest frequency of use of a given mode, when used as an independent variable, will show significant differences in intelligence, achievement, and vocabulary measures.

The analysis of variance procedure used to test this hypothesis indicated significant differences for the Metropolitan Readiness Tests Word Meaning and Listening tests. The reliability of these two tests was later investigated and was found to be rather low with this group of subjects; therefore, even though the results are significant, to conclude that the hypothesis is upheld is tentative without further support from the other measures.

71
This support is not forthcoming at the .05 level of significance, although the Numbers test and the total score of the Metropolitan Readiness Tests approach significance. Those subjects using the relational mode tend to do better on the tests mentioned above with the exception of Word Meaning.

In view of the evidence from this study, although there is a tendency toward it, Hypothesis One is not supported and must be rejected at the .05 level of significance.

Hypothesis Two

The preferred mode of categorization, when used as an independent variable, will significantly differentiate between groups on the factor scores of the first year standardized primary instruments. A factor analysis of these scores yielded a four factor structure representing two discrimination factors, one general readiness or academic intelligence, and one vocabulary factor. The factor scores for each individual were computed and then subjected to a one-way analysis of variance using preferred mode of categorization as the independent variable. The F-ratios for all factors were below the critical value needed for significance.

Hypothesis Two is not supported by the results and must be rejected at the .05 level.
Hypothesis Three

The division of the sample by degree of flexibility as evidenced by the number of modes of categorization used, will show a significant difference in scores achieved on the standardized primary instruments.

This hypothesis was tested through the use of an analysis of variance design. The results showed that there were no significant differences except in the Sedan section of the Hughes Car Discrimination Task. In this comparison there were strong differences, indicating that those subjects who used all four modes of categorization scored significantly higher on Discrimination, Uniqueness, and Saliency. Because of the contribution of this section to the total scores on the task the differences held up for those scores as well, with the exception of Saliency.

The materials used in the Sedan section of the Hughes Car Discrimination Task consist of two model cars which are more similar in appearance than are the model convertibles of the Convertible section. As a result the actual frequency scores on the Sedan section are lower and the distinctions made are more subtle than on the Convertible section. The most attractive explanation of the findings for Hypothesis Three is that the subject who is more flexible in his cognitive organization has a greater ability to discriminate and this ability is demonstrated in the more difficult of the two discrimination
tasks. The difference in the two discrimination tasks is substantiated by the factor analysis results which indicate that each task has its own factor.

A model of this explanation would ideally have an ascending magnitude of means as the degree of flexibility becomes greater. This assumes equal values for all modes of categorization which may or may not be true. The possibility exists that a given combination of types of categorization may be conducive to performance on the Hughes Car Discrimination Task and when one is able to combine all types then the given combination must be present whereas with any less than totality the combination may not be present. An extension of this study which included another pair of cars requiring even greater discrimination might clarify the problem.

For the Hughes Car Discrimination Task the third hypothesis is supported at the .05 level of significance, but for the remainder of the standardized primary instruments the hypothesis must be rejected.

Hypothesis Four

The fourth hypothesis states that conceptual style as represented by utilization of modes of categorization in the first grade, will significantly predict reading achievement at the third grade level.
This hypothesis was tested by using the subject's preferred mode scores of the types based on frequency counts of items and groups. When these scores were related to the scores achieved on the Reading and Word Knowledge tests of the Metropolitan Achievement Tests through a canonical correlation the relationship did not approach significance ($p = .77$).

The highly significant results ($p = .003$) of the canonical correlation between the first year standardized primary instruments and the Metropolitan Achievement Tests are not unexpected. Mitchell's (1967) study indicated that the predictive relationship of the Metropolitan Readiness Tests and the Metropolitan Achievement Tests. Kingston (1962) found that the Metropolitan Reading Readiness Tests (a predecessor of the Metropolitan Readiness Tests) predicted academic achievement as measured on the Stanford Achievement Tests. Panther (1967) found that the Goodenough-Harris Draw-a-Man Test predicted achievement on the Reading test of the Metropolitan Achievement Tests.

In view of the supporting research it would have been rather odd if this group of tests had not predicted achievement in the Reading and Word Knowledge tests of the Metropolitan Achievement Tests.

The canonical correlation between the preferred mode of categorization and the Reading and Word Knowledge tests was not significant, although the literature
indicates that some positive relationships have been found (Yeatts, 1969; Pederson and Wender, 1968) between preferred mode of categorization and academic achievement or intelligence, the results are plausible. Two primary conditions contribute to this plausibility: (1) the interval of time between the predictor task and the criterion measure, and (2) the variation in forms of the predictor.

The prediction found in previous studies, although significant, was not high and for the most part the time interval was shorter than in this study. The increased time interval could be expected to attenuate the type of relationship that is stage-dependent such as the one under investigation.

There is a possibility that the duplication of figures was conducive to a set being induced toward a "pairing" type of categorization which would introduce an additional factor into the process of categorization.

**Summary**

In summary, this study was designed to determine the relationships of modes of categorization to a selected group of measures used in the primary levels of the elementary school.

The sample was composed of 87 Mexican-American children attending elementary schools in the areas
qualifying for poverty funds under the provisions of Public Law 89-10, in Tucson School District #1.

Bruner et al. (1956, 1966) indicated that the individual organizes the environment through categorization and learning is closely related to this cognitive structure derived from the categorization process. Piaget (Inhelder and Piaget, 1964; Flavell, 1963) has developed the "assimilation-accommodation" model which has elements of categorizational behavior. At the same time empirical research has shown categorizing behavior in object sorting to demonstrate a preferred mode of categorization (Wallach and Kogan, 1965). In addition Sigel (1965) has shown that mode of categorization could affect the results of certain questions on intelligence tests.

The subjects were tested in the first year of school with the Van Alstyne Picture Vocabulary Test, the Goodenough-Harris Drawing Test, and the Metropolitan Readiness Tests. In addition to these standardized tests measuring achievement, vocabulary, and intelligence, the Hughes Car Discrimination Task, developed by the Arizona Center for Early Childhood Education was used to measure discrimination ability. The preferred mode of categorization was determined by using an object sorting task, the Grouping section of the Intellectual Tasks, also developed by the Center.
The hypotheses were designed to answer the general question of whether or not a given mode of categorization would relate to achievement on the standardized primary instruments. This question was based on a review of the literature which indicated that the way an individual organizes his environment affects his learning ability. The first hypothesis dealt with the differentiating ability of preferred mode of categorization with respect to the standardized primary instruments. The second hypothesis was concerned with the same factors emerging from a factor analysis of the standardized primary instruments, using preferred mode of categorization as the independent variable. Flexibility in the use of more than one mode of categorization was the area of investigation in the third hypothesis. The question of mode-linked questions was raised with the assumption that the greater the degree of flexibility the better equipped the subject was to deal with the problems posed by the academic environment in general and the standardized primary instruments in particular. The fourth hypothesis was designed to answer the question concerning the relationship of preferred mode of categorization to achievement in reading as measured by the Reading and Word Knowledge tests of the Metropolitan Achievement Tests.

The results indicate that there is very little, if any, relationship between preferred mode of categorization
and the standardized primary instruments. There are no significant relationships or differences between the Metropolitan Readiness Tests, Van Alstyne Picture Vocabulary Test, Goodenough-Harris Drawing test, or Metropolitan Achievement Tests when preferred mode of categorization is used as an independent variable or correlated. None of the above standardized primary instruments relate to flexibility in number of modes used as measured by the Intellectual Tasks: Grouping. There is a significant difference in number of discriminations and uniqueness score which have shared variance, when flexibility in the use of different modes is used as the independent variable in an analysis of variance. Those subjects who used all four modes of categorization were significantly higher in terms of the mean number of discriminations mode on the Sedan section of the Hughes Car Discrimination Task.

Conclusions

The results of this study indicate that, as measured by the Intellectual Tasks: Grouping, the categorizing behavior of the sample was not related to their achievement on a series of standardized primary instruments. This result appears somewhat counter to the theoretical positions of Piaget (Flavell, 1963) and Bruner et al. (1956, 1966) where categorization as a factor in the
cognitive structure and learning ability of the individual is of central importance.

The assumption made then, was if there are individual differences in categorizing behavior this will be evidenced in the learning abilities of the individual as shown by achievement in the academic environment. The results of this study would indicate that perhaps that assumption was incorrect. Consideration of this empirical-theoretical inconsistency suggested two possibilities for reconciliation. The first is the question of material dependency of categorization behavior, the second the potential divergence between process and product.

Material Dependency

The question of material dependency of categorizational behavior has not been answered satisfactorily at the present time. In the more traditional object sorting tasks the materials are widely varied in nature thereby skirting the problem. On the other hand the physical nature of the figures is quite similar in the Intellectual Tasks: Grouping, all anthropomorphic in nature. Neither of the two types of object sorting tests answers the material dependency question in that one has a wide variety which tends to confound any material dependent effects and the other has but a narrow range in variety and offers no comparisons.
If the factor of material dependency enters into preferred mode of categorization there may be a generalized preferred mode of categorization that will be shown over a wide variety of materials, but that will deviate markedly on a more or less uniform class of stimulus objects. This generalized preference for a mode of categorization could well be related to achievement, but the degree of relationship is attenuated when demonstrated preferred mode is distorted by a material dependency factor.

Process and Product Divergency

Categorization is essentially a process. It is a process for the organization of the cognitive structure of the mind for learning, assimilation, problem solving. A part of process is revealed in the behavior of the subject exhibited in the performance of the Intellectual Tasks: Grouping. In that no specified problem is presented with a right or a wrong answer the subject is left to choose his own way. That way may be in deference to the assumed wish of the examiner or the free choice of the subject, depending on the subject.

On the other hand the standardized primary instruments pose definite problems as do the requirements of the early elementary academic environment. The student meets these problems and answers, more or less, to the best of his ability. The same is true in the academic environment,
to say that one has a preferred cognitive style does not mean that others cannot be employed for specific tasks and in addition that one cognitive style is better.

Language Bias

In the studies by Yeatts (1969), Kagan et al. (1960, 1963), and Wallach and Kogan (1965) the language bias was not discussed, but judging from the similarity of experimental means to published norms the samples did not deviate in the direction one would expect with a sample that was suspected of a systematic bias. It is probable that the standardized primary instruments would test somewhat different factors with subjects of a group not typical of the norming group and the relationship of categorizing style and these instruments may be altered as well. The same is true of the categorizing measure.

Recommendations for Future Research

In order to do this these problems must be resolved:

1. The flexibility of the individual in categorizing behavior must be ascertained in a way that gives an estimate of the total repertoire available to him not just the preferred mode of categorization.

2. The phenomena of categorization must be better understood in terms of material dependency. An answer must be found for the question of how different levels of concepts are categorized in the
continuum from concrete to abstract and degrees of complexity at each level.

3. An estimate of constancy of behavior of categorization both short term and over extended periods is urgently needed through the use of parallel instruments and longitudinal studies.

When these questions can be successfully answered then the field opens up for psychometrization of categorization.

Psychometrizing Categorization

The thoughts of Flavell (1963, p. 417) concerning the "psychometrizing" of Piaget's intellectual tasks and inferentially the area of categorization opens wider vistas for the field of psychometric testing. At the present time the standardized instruments used in the schools have a relatively large amount of shared variance. They are used to predict performance on other tests of the same type. The prediction in this case is quite high, but this situation does not hold up when the criterion is outside the realm of academic achievement as the broad range of tests in industrial psychology bears evidence.

Hunt (1961) developed the concept of environmental factors, from Piagetian orientation, and their relationship to intelligence. The developmental levels hypothesized by Piaget are incorporated into this analysis which leads one
to conclude that these stages should be amenable to measurement provided the instruments could be developed. Categorization holds promise as an instrument for this measurement. Bruner (Bruner et al., 1966) states that categorization is a necessary part of cognitive structure used by the individual in his total experiences and therefore extends beyond the relatively specialized academic environment. If and when this particular type of cognitive function can be isolated and measured then the potential for prediction can be reached, but more importantly the possibility for learning of this ability can be developed.
APPENDIX A

SCORING RULES FOR CATEGORIZATION RESPONSES

The scoring procedures have been developed by the researchers at the Arizona Center for Early Childhood Education (Hughes and Coxon, 1968). In this study only a portion of the Intellectual Tasks were used and therefore the scoring procedures were modified accordingly.

Labeling

Labeling is defined as the use of academically common names or labels for the rubber figures in the task. An academically common label is one that is likely to be used and considered correct in the school environment. The label must be the name of the task item, not a functional definition. If one figure of a given type is correctly labeled the group is considered correctly labeled. The maximum score for this section of the Intellectual Tasks is twenty-seven, the number of different types of figures in the task.

Categorization Mode

The groups that the subject makes are recorded in terms of figural content. The groups must conform to the following rules before being scored for mode of categorization.
1. A grouping had to have at least two figures in it to be scored a group.

2. Only the groups created on the first trial were scored.

3. Only the first verbal response for each grouping was scored. Therefore, only one mode of categorization was allowed for each grouping.

4. Flexibility was based on the number of categories used in the first trial only.

5. The figures in a given grouping had to be correctly identified as defined by the labeling criteria to be scored.

Those groupings that met the above criteria were then scored as to mode of categorization using the following criteria.

Descriptive Mode

The descriptive mode is that mode under which the organization of the figures is on the basis of similarities of objective physical attributes. The figures are grouped because they share some physical property such as identical size, color, shape, or physical position. Examples of this type of grouping are: "Because they are all/both little." "These have black spots." "They both have the same clothes on."
Categorical-Inferential

The organization of figures in this category does not rest on a single abstracted physical property shared by the figures, but rather represent groupings that take account of the figures as whole entities. Any figure in the grouping can be considered an instance of the conceptual label. This type of categorization may be tested by applying these two criteria: (1) can the grouping be identified by a single conceptual label, and (2) are all items in the grouping instances of that label. Examples of this category are: "Because they are all boys." "They are all jungle animals." "All these people are doctors."

Relational

This type of organizational behavior is characterized by grouping on the basis of relationship of figures within the group. Relational groupings include: (1) family relationships such as sisters, fathers, mothers, brothers, uncles, etc.; (2) functional relationships such as boss-employee, teacher-pupil, nurse-patient, animal-caretaker; and (3) physical and positional relationships such as big vs. little, young vs. old, tall vs. short. Relationships on a functional basis may include those groupings attributed shared activities such as: "They play together." "They live together."
Thematic

The organization in this mode is based on a story or incident or theme involving the figures in the grouping. The verbalization of a thematic grouping may range from a simple sentence to an elaborate story. Examples of thematic grouping are: "This is a zoo where people are watching the animals."

Errors in Grouping

A grouping was considered in error if the verbal reason given for that particular grouping was not possible in view of the physical grouping. Examples of this: If a dog and a calf are grouped together and the subject says that they are "sheep." In this case the verbal description of the categorization is taken as the criterion of grouping and if the physical grouping does not conform to the criterion stated it is considered an error. An error is also assessed if the figures are grouped on a basis of contrast such as the lion and the giraffe, with the stated reason being that the lion ate meat and the giraffe ate leaves.

Levels Within the Categorical-Inferential Mode

This is a system of grading the levels of operation within the categorical-inferential mode. Level I is defined by the criterion of functional definition of a group of like items, for example "Because they bring us
mail." Level II is defined by the criterion of verbal label of like items. For example, "They are both policemen." Level III is defined by the criterion of unlike items grouped with a functional verbal reason. For example, "Because they all live in the jungle." Level IV must have unlike figures grouped together under a nominal verbal classification such as, "They are all jungle animals."

Scores Derived

The following scores were derived for the categorization tasks:

1. Number of figures labeled.
2. Number of figures used in correct groupings.
3. Number of different categories.
4. Number of groups in the descriptive category.
5. Number of groups in the categorical-inferential category.
6. Number of groups in the relational category.
7. Number of groups in the thematic category.
8. Number of error groupings.
9. Number of figures used in descriptive groups.
10. Number of figures used in categorical-inferential groups.
11. Number of figures used in relational groups.
12. Number of figures used in thematic groups.
This test is intended to determine the discrimination ability of culturally deprived children using high interest materials of considerable detail in structure.

**Procedures**

The material consists of five cars, three sedans two of which are identical, and two convertibles, one a sports car the other a conventional convertible.

The examiner places the two duplicate sedans in front of the child and asks, "Are these two cars alike or different?" This is intended as a training question to establish rapport with the child and to make the next pair of sedans exhibit greater contrast. Next the examiner places the two convertibles in front of the child and asks, "Are these two cars alike or different?" If the subject replies that they are alike this trial is terminated, but if the subject determines that they are different then the examiner asks, "Show me some of the things that are different." The responses of the subject are recorded on a tally sheet. When the response rate tends to slow the examiner again asks the subject to "Show me some more things that
are different." This procedure is followed by a third request for more differences. After the responses generated by the third request have finished the convertibles are removed and the two sedans presented. The same procedure is followed with the three requests for differences and the tally made of each response. This completes the test.

**Scoring**

The scoring for the Hughes Car Discrimination Test is in the process of development which is an ancillary function of this study. As a result the standards used in this are unique to the group represented herein.

**Frequency of Discrimination**

This score is based on the total number of items that the subject specifically indicates as being different.

**Uniqueness of Discrimination**

This score is based on the relative rarity of response compared with the rest of the subjects in the study. Each discrimination is assigned a value that is dependent on the following criteria:
<table>
<thead>
<tr>
<th>Score</th>
<th>Commonality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The top 20% of the discriminations ranked on frequency of choice.</td>
</tr>
<tr>
<td>2</td>
<td>The next 20%.</td>
</tr>
<tr>
<td>3</td>
<td>The next 20%.</td>
</tr>
<tr>
<td>4</td>
<td>The next 20%.</td>
</tr>
<tr>
<td>5</td>
<td>The bottom 20%.</td>
</tr>
</tbody>
</table>

The total uniqueness score is the sum of the uniqueness ratings.

Saliency Score

This score is based on the top 20% of the items chosen on the first trial. This score was intended to represent the most readily apparent contrasting features between the two test pairs of cars.

The three scores were determined for both the sedans and the convertibles and also combined into a total score for analysis.
### FIGURES SUPPLIED FOR INTELLECTUAL TASKS: GROUPING

<table>
<thead>
<tr>
<th>Number</th>
<th>Description of figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Babies in sitting position one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Boys, preadolescents, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Girls, preadolescents, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Men, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Women, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Doctors, one white, one brown, differentiated from men by black bag in one hand.</td>
</tr>
<tr>
<td>2</td>
<td>Firemen, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Mailmen, with pouches, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Nurses in uniform, one white, one brown.</td>
</tr>
<tr>
<td>2</td>
<td>Policemen in uniform, one white, one brown.</td>
</tr>
<tr>
<td>1</td>
<td>Dog similar to a collie.</td>
</tr>
<tr>
<td>1</td>
<td>Bull.</td>
</tr>
<tr>
<td>1</td>
<td>Cow.</td>
</tr>
<tr>
<td>1</td>
<td>Calf.</td>
</tr>
<tr>
<td>5</td>
<td>Chickens, one is a rooster.</td>
</tr>
<tr>
<td>1</td>
<td>Goat.</td>
</tr>
<tr>
<td>1</td>
<td>Colt.</td>
</tr>
<tr>
<td>1</td>
<td>Horse.</td>
</tr>
<tr>
<td>3</td>
<td>Pigs.</td>
</tr>
<tr>
<td>2</td>
<td>Piglets.</td>
</tr>
<tr>
<td>1</td>
<td>Lamb.</td>
</tr>
<tr>
<td>1</td>
<td>Sheep.</td>
</tr>
<tr>
<td>1</td>
<td>Elephant.</td>
</tr>
<tr>
<td>1</td>
<td>Giraffe.</td>
</tr>
<tr>
<td>1</td>
<td>Lion.</td>
</tr>
<tr>
<td>1</td>
<td>Zebra.</td>
</tr>
<tr>
<td>1</td>
<td>Rhinoceros, in some trials a hippo was substituted.</td>
</tr>
</tbody>
</table>
APPENDIX D

SAMPLE PROTOCOL FROM INTELLECTUAL TASKS

The following is a reproduction of the protocol used to determine the preferred categorization style of subjects in the study. Not all the information available from the protocol was used in the study because it was believed that the preferred mode of categorization would be the most clearly apparent from the initial grouping.

Labeling

Directions

Place all animals and people at one end of the table and tell subject:

"Let's spread these out so we can see them better. These toys are for us to work with. You can pick them up, and you can move around the table to work better."

After the subject manipulates the toys awhile and seems confident in the situation, say:

"Let's try to name these toys. As you name them, put them over here so Mr. (s) . . . (recorder) can see them."

If child hesitates, pick up a familiar toy and hand it to him, saying:

"This is a . . . ."

Subject should be handling and placing the rest of the toys as he names them.

94
Other phrases the examiner may use to elicit labeling are:

"What about this one?"
"And this is a . . . ?"
"What do you call this?"

If the child does not label a toy after two questionings by the examiner, the examiner names it for him, saying:

"We call this a . . . ."

All the toys are then placed together on one end of the table and the Grouping task begins.

Scoring

The score for labeling is the total number of types of figures named correctly.

Grouping

Materials: See Appendix C

Procedures:

1. Place toys together on table.

2. Ask, "Show me what things go together." (pause) "Make some piles of the things that go together."

3. When child completes first grouping, point to the separate piles and ask, "Why do these things go together?"

4. After recording responses, ask, "Can you put them together in another way?"

5. Point to the new groupings and ask, "Why do these things go together?"

6. After recording responses, ask, "Can you put them together in still another way?"

Scoring:

Indicate each grouping by placing a numeral (1, 2, or 3) on the blank before the appropriate category on the checklist to indicate first, second, or third
classification. If no appropriate category appears on the checklist, indicate groupings with brief descriptive statement. Record child's reasons for each grouping. Under "Why" record DK if child states he does not know. Record NR if child gives no response to question "Why?"
REFERENCES


97


Hanson, E., and Robinson, H. A. Reading readiness and achievement of primary grade children of different socio-economic strata. The Reading Teacher, 1967, 21, 52-56.


