THE VALIDITY OF THE DRAW A PERSON: A QUANTITATIVE SCORING SYSTEM FOR KINDERGARTEN CHILDREN

DISSERTATION

Presented in Partial Fulfillment of the requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Suzanne H. Gottling, B.A., M.A.

* * * * *

The Ohio State University
1990

Dissertation Committee
J. A. Naglieri
J. L. Genshaft
J. L. Collins

Approved by

Adviser
College of Education
ACKNOWLEDGEMENT

This document represents the conclusion of several years of hard work. Although the author admits to a percentage of the work, she wants it known that many people, large and small, worked hard too. First, her adviser, Jack A. Naglieri, provided a model of intelligent inquiry which made this project seem important and interesting. Second, her husband, Jim, provided a model of patience, encouragement, and sharing which made this project bearable. Third, the kindergarten children and their teachers at Graham Road School in Reynoldsburg and West Franklin in the Southwestern District provided a model of willing cooperation and made this project possible.
VITA

March 20, 1934 ........................................... Born - Allentown, Pennsylvania

1956 .......................................................... B.A., Wilson College, Chambersburg, Pennsylvania

1970-1975 .................................................. Teacher, Columbus School for Girls, Bexley, Ohio

1975-1984 .................................................. Director Lower School, Columbus School for Girls, Bexley, Ohio

1984-1985 .................................................. Graduate Research Assistant, Math Anxiety Project, Department of Educational Services and Research, The Ohio State University, Columbus, Ohio

1985 .......................................................... M.A., College of Education, The Ohio State University, Columbus, Ohio

1985-1986 .................................................. Intern School Psychologist, Reynoldsburg City Schools, Reynoldsburg, Ohio

1986-1989 .................................................. Graduate Research Assistant, Department of Educational Services and Research, The Ohio State University, Columbus, Ohio

1989-Present .............................................. School Psychologist, Marburn Academy, Columbus, Ohio
PUBLICATIONS


FIELDS OF STUDY

Major Field: Education

Studies in Psychoeducational Assessment: Professor Jack A. Naglieri
Studies in Psychodiagnostic Assessment: Professor Henry Leland
Studies in School Psychology: Professors Collins and Genshaft
Studies in Human Development: Professors Werar, Gibbs, and Blythe
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................ ii

VITA...................................................................................................................................... iii

LIST OF TABLES ................................................................................................................... vii

LIST OF FIGURES ................................................................................................................ ix

CHAPTER | PAGE
--- | ---
I. INTRODUCTION .............................................................................................................. 1

Background ......................................................................................................................... 1
Validity of the DAP ............................................................................................................. 4

II. REVIEW OF THE LITERATURE .................................................................................... 13

Human Figure Drawings: From Darwin to Goodenough .............................................. 14
Construction of the Goodenough Draw A Man Test ....................................................... 22
  Reliability ........................................................................................................................ 24
  Validity ............................................................................................................................. 24
  Summary ......................................................................................................................... 25
Harris' Revision of Goodenough's Scale ........................................................................... 26
  Reliability ........................................................................................................................ 27
  Validity as a Test of Intelligence .................................................................................... 28
  Summary ......................................................................................................................... 30
Draw A Person: A Quantitative Scoring System (DAP-Q) ........................................... 31
  Reliability ........................................................................................................................ 32
  Validity ............................................................................................................................. 33
General Developmental Theory ...................................................................................... 33
The DAP and Models of Intelligence .............................................................................. 37
  The DAP and Possible Links to the Das-Luria Cognitive Processing Theory .......... 44
Motor Development ......................................................................................................... 49
Personality ......................................................................................................................... 52
The DAP and Kindergarten Children ............................................................................. 58
III. METHOD ......................................................................................................................... 61
   Subjects ......................................................................................................................... 61
   Procedures ..................................................................................................................... 62
   Experimental Tasks ....................................................................................................... 63
      Draw A Person ........................................................................................................... 63
      Bracken Basic Concept Scale ................................................................................... 65
      Motor Scale, McCarthy Scales of Children’s Abilities ............................................. 66
      Cognitive Assessment System: Experimental Version ............................................. 68
   Teacher Version of the Child Behavior Checklist ....................................................... 71
   Research Questions and Data Analysis ....................................................................... 73

IV. RESULTS ...................................................................................................................... 79
   PASS Tasks ..................................................................................................................... 79
   Standard Scores ........................................................................................................... 82
   Correlations of DAP TT with McCarthy Motor, BBCS, and CBCL Variables ............ 85
   The PASS Model ......................................................................................................... 90
   Factor Analysis of PASS Tasks ................................................................................... 92
   Correlations of DAP TT with Planning, Attention, Simultaneous and Successive Processing Scores ......................................................... 96
   Prediction of DAP TT Scores using Regression ......................................................... 97
   Gender Differences ..................................................................................................... 107
   Additional Analyses .................................................................................................... 111

V. DISCUSSION ............................................................................................................... 119
   The DAP and PASS Processing .................................................................................. 120
   The DAP and Conceptual Understanding ................................................................... 122
   The DAP and Motor Functioning ............................................................................... 125
   The DAP TT and the Teacher Version, Child Behavior Checklist ................................ 128
   The Influence of Gender ............................................................................................. 129
   Age Effects .................................................................................................................. 131
   Future Research ......................................................................................................... 131
   Conclusion ................................................................................................................. 134

REFERENCES .................................................................................................................. 136

APPENDIX A  PARENT NOTIFICATION LETTER ......................................................... 156

vi
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Description of sample</td>
<td>62</td>
</tr>
<tr>
<td>2. Raw Score Means and Standard Deviations of PASS Variables</td>
<td>80</td>
</tr>
<tr>
<td>3. Means and SDs of the Four Components of the PASS Model</td>
<td>82</td>
</tr>
<tr>
<td>4. Means and SDs for All Variables Which Have Been Nationally Standardized</td>
<td>83</td>
</tr>
<tr>
<td>5. Intercorrelations of Variables</td>
<td>87</td>
</tr>
<tr>
<td>6. Intercorrelations of Variables with Age Partialled Out</td>
<td>89</td>
</tr>
<tr>
<td>7. Intercorrelations of PASS Tasks</td>
<td>91</td>
</tr>
<tr>
<td>8. Three and Four Factor Solutions Using Principal Axes Factor Analysis - Orthogonal Solution</td>
<td>93</td>
</tr>
<tr>
<td>9. Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution</td>
<td>94</td>
</tr>
<tr>
<td>10. Simple Regression on DAP TT</td>
<td>98</td>
</tr>
<tr>
<td>11. Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, and Motor Scale</td>
<td>100</td>
</tr>
<tr>
<td>12. Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, and Motor Factor</td>
<td>101</td>
</tr>
<tr>
<td>13. Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, Gross Motor, and Draw-A-Design</td>
<td>102</td>
</tr>
<tr>
<td>14. Multiple and Stepwise Regression on DAP TT - BBCS, PASS, and CBCL</td>
<td>103</td>
</tr>
<tr>
<td>15. Multiple and Stepwise Regression on DAP TT - P, A, S, S</td>
<td>104</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>16</td>
<td>Comparisons of Means and SDs of Major Variables by Gender</td>
</tr>
<tr>
<td>17</td>
<td>Comparisons of Means and SDs of PASS Tasks by Gender</td>
</tr>
<tr>
<td>18</td>
<td>Comparisons of Means and SDs of the Components of the PASS Model by Gender</td>
</tr>
<tr>
<td>19</td>
<td>Comparisons of the Means and SDs of Motor Subscales and CBCL Subscales by Gender</td>
</tr>
<tr>
<td>20</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and McCarthy Motor Scale</td>
</tr>
<tr>
<td>21</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and McCarthy Motor Factor</td>
</tr>
<tr>
<td>22</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and Gross Motor Scale</td>
</tr>
<tr>
<td>23</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and Draw-A-Design</td>
</tr>
<tr>
<td>24</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and Bracken Basic Concepts</td>
</tr>
<tr>
<td>25</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and Child Behavior Checklist</td>
</tr>
<tr>
<td>26</td>
<td>Three and Four Factor Solutions Using Principal Axes Factor Analysis - Oblique Solution - Pass Tasks and Draw a Person</td>
</tr>
</tbody>
</table>
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Significant correlations ($p &lt; .01$) between the DAP and other variables</td>
<td>106</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

Background

A few standardized psychological tests might be described as classics since they are enduring, traditional, and have recognized value. The human figure drawing test meets these criteria for it has a long history, is used frequently, and is widely assumed to be a useful measure of intelligence and personality.

More than 60 years have passed since the publication of Goodenough's small volume, *Measurement of Intelligence by Drawings* (1926). This work contains a review of past interest in human figure drawings as well as a statistically supported system for evaluating a drawing of a man by children from ages 4 - 13. An expansion and revision of this method was published by Harris in 1963 as the *Goodenough-Harris Drawing Test*. Recently, a further refinement of the system, *Draw A Person: A Quantitative Scoring System* (Naglieri, 1988), has been nationally standardized and published. Among standardized tests purporting to measure intelligence, only the *Stanford-Binet* (Terman, 1916) with its 70 year history can claim a longer record of continuous use.
Although human figure drawings were first used systematically to estimate intelligence, their role as a measure of personality can also claim a long history. Goodenough (1926) was one of the first to suggest that drawings were likely sources of information about personality development as well as cognitive development. She proposed that "a method of scoring might be derived which would throw new light upon eccentricities of mental functioning during childhood" (p. 66). However, it was not until 1949 that the first catalogue of relationships between specific drawing characteristics and specific personality traits and emotional states was published (Machover, 1949). Machover (1949) and Hammer (1958) noted that their interest in the Draw a Person (DAP) task as a projective instrument was spurred by the fact that children could earn identical intelligence scores although their drawings differed greatly. The unique components of a drawing became the data used to infer personality characteristics and emotional states. Later, Hammer (1958), Levy (1958), Koppitz (1968), Handler (1967) and Saarni and Azara (1977) developed variations of the Machover system. All of these systems are in use today for individual personality assessment and research. The DAP test can claim a place with the Rorschach (1942) and TAT (Thematic Apperception Test, Murray, 1943) as a long-lasting technique in the measurement of personality.

Throughout its history, the DAP task has been one of the most frequently used psychological measures. Attributes which contributed to the popularity of the DAP include its ease and speed of administration, non-test-like nature, and usefulness when language barriers prevent adequate explanation of a more complex task. In addition, the same drawings could be used
to assess intelligence and personality since the materials, instructions, and time demands are similar. Details presented in drawings provide two levels of information, one which focuses on the unique components of the drawing, and one which allows comparison with aspects common to a particular age group.

The frequency of use of the DAP has been measured through numerous surveys of professional practices. Sundberg's (1961) review of surveys conducted between 1935 and 1961 found that both the Machover DAP personality assessment and the Goodenough intelligence system were among the top ten tests with average ranks of second and ninth respectively. Sundberg commented on the durability of the Goodenough Draw A Man and the Stanford-Binet as the only two tests which had placed in the top ten in every reported survey since 1935. Sundberg stated that this persistently high ranking was especially notable since test popularity was often ephemeral. In the decade from 1951 to 1961, as many as 76% of the top ten tests in one year did not reappear the next year.

A survey by Lubin, Wallis, and Paine (1971) which duplicated Sundberg's procedures indicated that the Goodenough Draw A Man task was no longer included among the tests used by at least 10% of the institutions surveyed. However, the Machover DAP task was ranked fifth. Scott (1981) noted that these surveys did not include practitioners in an educational setting. Therefore, the results reported by Sundberg (1961) and Lubin et al. (1971) are likely to be underestimates of the actual use and influence of the DAP through 1971.
Recent surveys indicate that the DAP is still widely used in clinical and educational settings. Of the school psychology practitioners responding to a National Association of School Psychologists (NASP) survey, 78% reported use of the DAP as a projective personality instrument. Of university training program representatives responding to the same survey, 75% reported their programs included supervised instruction in human figure drawing interpretation (Reschly, Genshaft, & Binder, 1986). Responses from 25% of over 100 inpatient settings surveyed by Sweeney, Clarkin, and Fitzgibbon (1987) indicated that the human figure drawing task was part of the standard assessment battery. The DAP ranked sixth in usage according to this survey. Patrick (1984) and Anderson (1985) reported that the Draw A Person task was one of seven evaluation measures used in a large Oklahoma school district to identify children at risk for school failure. It appears that the Draw A Person task is an important and frequently used component of psychological assessment in clinical and educational settings.

Validity of the DAP

Because of its long history and frequent use, it is important to understand what psychological constructs are tapped when a child is asked to produce a human figure drawing. It is particularly important since this task has been used to estimate both cognitive level and emotional status, two aspects of human functioning which would appear to be best measured by different tasks.

There is a large body of literature reporting results of studies of the reliability, validity, and utility of the various DAP assessment systems. This
literature shows greater agreement about the validity and usefulness of quantitative systems based on the concept originally developed by Goodenough than about the projective systems developed by Machover, Hammer, Handler, Levy, or Koppitz.

The validity of the DAP as a measure of personality and predictor of behavior has been heatedly debated from its inception to the present (Sims, Dana, and Bolton, 1983; Goldman and Velasco, 1980; Swensen, 1971, 1965). Kellogg (1979) attacks its construct validity by challenging the basic theoretical assumption that drawings project hidden thoughts and feelings.

To assume that some mental inadequacy - due to a brain defect, physical immaturity, or emotional instability - accounts for the distorted physiology of the HFD's in child art is an erroneous reaction of the adult mind that psychology and psychiatry perpetuate (p. 10).

Other criticisms center on the failure of controlled studies to establish the validity of any system of personality assessment using drawings (Anastasi, 1988).

In contrast to the projective systems, all versions of the DAP as a measure of intelligence or cognition have been shown to have important statistical properties which have been supported through experiment. Various types of validity have been established by

1. The linear relationship between age (approximately 5 to 12 years) and scoring criteria such as the amount of detail included, control of drawing strokes, and ability to represent proportion,

2. An average moderate correlation of the DAP with recognized tests of intelligence (Scott, 1981), and
3. The emergence of one primary factor, probably reflecting Spearman's "g", from factor analysis (Naglieri, 1988).

In addition, inter-rater, intra-rater, and inter-item reliability have all fallen in the high ranges as defined by Anastasi (1988).

Although the quantitative systems of evaluating a human figure drawing have an extensive, relatively consistent research base, they have also been subject to criticism. Numerous structural limitations of the Goodenough and Harris systems are summarized in Sattler (1982, 1988). These limitations have been addressed in the most recent modification by Naglieri (1988). It is now possible to replicate and expand many of the earlier studies of the DAP using a method which has addressed and controlled for the technical problems of the earlier Goodenough and Harris systems.

There are several validity questions surrounding the DAP which need to be examined. For many years, the claim that "the child draws what he knows" (Goodenough, 1926, p. 12) was generally accepted. Mortensen (1984) traces the passage of this phrase from Ricci in 1887 to Kerschensteiner and Levinstein, 1905, Luquet, 1913, Goodenough, 1926, and Piaget, 1951. Recently, the claim that the DAP is a measure of what a child knows has been strongly challenged (Goodnow, 1977, Kellogg, 1979, and Freeman, 1980). One of the strongest statements was made by Kellogg (1979) who called the use of a human figure drawing test as a means of estimating intelligence "pure commercial hogwash" (p. 16).

A further difficulty in understanding what the Draw a Person measures lies in the lack of agreement on the definition of the construct(s) under consideration. Earned scores have been termed measures of intelligence
(Goodenough, 1926, McWhinnie, 1971), cognitive skills and ability (Sattler, 1982), concept understanding (Harris, 1963, Reynolds, 1978), and conceptual maturity (Hargreaves, 1982). Examination of the relationship between the DAP, theories of cognition, and intellectual measures other than traditional intelligence tests is needed.

Data generated by studies based on multiple or vague definitions are also confounded by the often unexplained and generally unmeasured influence of developmental patterns. A characteristic of all DAP quantitative systems is their failure to discriminate among subjects above age 12. Harris attempted to explain this failure by placing the task in a Piagetian framework. However, Carlson's (1970) study showed no significant correlation between the DAP and several Piagetian tasks. At present researchers have not established a sound theoretical base which would account for the age-related discriminatory power of the test. Is it possible that the nature of the task is such that increased scoring is not possible?

Another poorly researched influence on the DAP is motor development. Although a small body of literature exists which shows a moderate correlation of the DAP with the visual-motor skills tapped by design copying tasks such as the Bender Visual Motor Gestalt Test (Bender, 1938) and McCarthy Draw A Design (McCarthy, 1972), other motor abilities have been ignored. Apparently, investigators have been satisfied to accept Goodenough's statement that "drawings made by young children ... are determined by concept development rather than by visual imagery or by manual skill" (1926, p. iii). The work of Williams (1983), Laszlo and Bairstow (1985), and Laszlo and Broderick (1985) suggests that kinesthetic awareness and overall motor
development have a strong influence on cognitive tasks at least through the age of eight.

In light of the continued use of the Draw A Person task, it is important to ask how accurately and efficiently does it serve any of its announced purposes? What factors contribute to the final product, a graphic representation of a person? Is its best use as a measure of cognition or personality, neither of these, or a mixture of both? What developmental criteria must be considered at different ages? It is impossible to answer these questions from the available research. Mortenson (1984) wrote that "the more detailed study of the essential problem, which factors influence the drawing of a person, has ... not been advanced much during recent years (p. 68)."

In order to meet the need to further examine the validity of the DAP, this study aims to examine the impact of four variables on human figure drawing scores obtained through a system related to intellectual ability. An intellectual scale is preferred because all of the three major systems (Goodenough, Goodenough-Harris, and Naglieri) have shown better reliability and validity than any of the personality scales. The areas which will be investigated are:

1. Gross and fine motor development
2. Basic concept understanding
3. Cognitive processing
4. Personality/Emotionality

The Naglieri (1988) *Draw A Person: A Quantitative Scoring System* is used to score all drawings because it is the most comprehensive and recent of the quantitative scoring methods. It has excellent construct validity and high inter-rater, intra-rater, and split-half reliability. A problem in motor
development, emotional stability, cognitive processing, or level of cognition should be revealed through a diminished score using this variation of Goodenough's method.

The four instruments described below are selected because they meet established standards (American Psychological Association, 1985) for test construction more consistently than other available instruments. These standards are that

1. The instrument was developed to reflect a clearly expressed theory based construct.
2. The validity and reliability of the items used in the instrument have been tested.
3. Each instrument has been used to examine the performance of a representative sample of children.
4. Age-based standards for performance were available.

The *McCarthy Motor Scale* is one of only three tests which were rated as a valid measure of motor development from a group of twelve reviewed by Laszlo and Bairstow (1985). The scale measures both fine and gross motor skills. It is anticipated that the child's overall level of motor production will be related to the child's ability to produce a human figure drawing. The kindergarten age group selected for this study is likely to show a relationship between motor ability and drawing score.

The teacher version of the *Child Behavior Checklist* (Achenbach & Edelbrock, 1986) is a highly regarded reliable and valid measure of behaviors (Barkley, 1987) which are associated with the personality traits and emotional conditions which have been studied in the DAP projective literature
(Machover, 1949, Koppitz, 1964, Daum, 1983, Sims et al., 1983). This literature shows that the most significant and consistent indicator of personality or emotional difficulties is omission of parts of the human figure such as eyes, body, or head (Koppitz, 1968, Swensen, in Goodstein and Lanyon, 1971, Goldman and Warren, 1976, Goldman and Velasco, 1980, Eno, Elliot, & Woehlke, 1981, Daum, 1983). Omission of parts also affects scores derived from any of the systems used to measure intelligence for these are based on inclusion of parts. For example, omission of a part of a human figure would result in a lower raw score of 3 to 7 points using the Naglier scoring system (1988). Obviously, omission of more than one part would result in further lowering of the raw score. Thus, based on one of the most consistent findings in the DAP projective literature, it is likely that a relationship would be found between scores derived from human figure drawings scored with the Naglieri system and scores derived from the Child Behavior Checklist, an instrument which is sensitive to personality and emotional problems such as depression, aggression, withdrawal, inattention, and self-destructiveness.

Bracken's Basic Concept Scale (1984) is a psychometric instrument which meets all the criteria established by the A.P.A. for test construction. It reflects a child's mastery of seven concept areas using auditory (language) and visual (pictures) modalities. It was designed to include all concepts important to school success. This task is especially valuable in light of Harris' redefinition of the DAP as a measure of conceptual maturity rather than "g", general intelligence. If a human figure drawing taps knowledge of concepts, a moderate to high degree of correlation would be anticipated between the DAP and Bracken's Basic Concept Scale scores.
The cognitive processing measures used in this study are experimental tasks based on the four cognitive processes described by Luria (1973) and operationalized by Das and Naglieri (1989). These processes are planning, attention, successive coding, and simultaneous coding. The tasks related to each process have been studied and found to be reliable and valid (Naglieri and Das, 1988). Because these tasks are theory-based, significant correlations with the DAP would help to establish a theoretical base for interpretation of DAP scores. It is hypothesized that the child’s ability to attend, code, or plan will be reflected in a DAP score.

Careful measurement of the contributions of motor skills, basic concept understanding, cognitive processing, and emotional status to the variation in DAP scores is important in light of the previously cited high level of use by practitioners in clinical and educational settings. This study focuses on the influence of these four factors on the human figure drawings of kindergarten children. Differences in the performance of boys and girls on all of the variables is also investigated.

The kindergarten age level has been chosen for several reasons. First, some of the four factors under investigation likely influence a child’s performance at this age and not at later ages. Second, DAP evaluations frequently form part of the data used to make decisions about a child’s entrance into school or promotion to the next grade. They are also used in test batteries designed to identify children at risk for school failure and thus eligible for special services or treatment. Any measure which is frequently employed in decisions which will have a life-long impact on a young child needs to be thoroughly understood and appropriately applied.
Data for this study was obtained through individual testing of approximately 100 kindergarten children. Data analysis included Pearson Product Moment correlations, factor analysis, analysis of variance and multiple regression to establish the significance of relationships between the Draw A Person test (Naglieri, 1988) and four variables. These variables are motor skills, as measured by the McCarthy Motor scale; concept understanding, as measured by the Bracken Basic Concepts Scale; planning, attention, simultaneous and successive processing, as measured by the tasks derived from the Das-Luria cognitive processing model; and personality or emotionality, as measured by the Teacher Version of the Child Behavior Checklist. It is hoped that this study will increase understanding of the underlying constructs which contribute to the child's drawing of a human figure.
CHAPTER II

REVIEW OF THE LITERATURE

Contemporary psychological uses of human figure drawings are derived from two powerful theories which have shaped Western culture, Darwin's theory of evolution and Freud's theory of personality development. The measurement of intelligence (Goodenough, 1926) or conceptual maturity (Harris, 1963) is based on the assumption that this important survival characteristic is distributed unevenly on a normal curve throughout a population and can be quantitatively estimated by analysis of the content and style of a drawing. The measurement of personality is based on the assumption that unconscious wishes, thoughts, and feelings influence behaviors and are expressed in the components of a human figure drawing. The validity of these two uses of human figure drawings continues to be debated (e.g. Freeman, 1980, Scott, 1981, Kellogg, 1973, 1979). Was Goodenough correct when she claimed that "the nature of drawings made by children in their early years is conditioned by their intellectual development" (1926, p. iii)? Was Machover correct when she claimed that "Prognostication of the course and treatment of a personality problem or mental disorder has frequently been accurately made solely on the drawings" (1949, p. 25)? Do intelligence and personality interact to affect cognitive scores and personality interpretation? Are there other important variables which influence human figure drawing
scores and interpretations? What does the Draw A Person task measure and how well does it measure what it measures? This review will focus on human figure drawing as a measure of intelligence or conceptual maturity. Theoretical interpretations of the task and the influence of motor ability, concept understanding, and personality will also be examined.

**Human Figure Drawings: From Darwin to Goodenough**

The long history, wide usage and presumed validity of the human figure drawing task as a measure of intelligence rests primarily upon its relationship to Darwin's theory of evolution. If one agrees with Woodward (1982) that "The idea for mental testing had sprung from Darwinian soil" (p. 12), then it can be shown that the human figure drawing test is more closely related than most mental tests to this Darwinian soil.

In the late nineteenth century, Darwin's biologically based theory of evolution was integrated into other philosophic and scientific areas (Richards, 1982). According to Haeckel (1891, as cited in Jaeger, 1982), "of all the branches of anthropology, not one is so affected and altered by the theory of descent as psychology." Borstleman (1963) described as "immediate heirs of Darwin" (p. 35) psychologists associated with the study of child development and psychological measurement such as Preyer, Hall, Gesell, Dewey, Binet, Terman, and Thorndike.

One of the most influential extensions of Darwinian theory was Haeckel's theory of recapitulation, a theory which was, for example, the basis for Hall's novel and influential studies of adolescence (Gottlieb, 1983, Hothersall, 1984). Haeckel postulated that human development recapitulated in an orderly manner all the adult stages of the species comprising its
evolutionary history (Gottlieb, 1983, Woodward, 1982, Kagan, 1979). Implicit in this theory was the expectation of differences between existing groups and an assumption that these groups could be arranged in an evolutionary hierarchy. This part of the theory, called the "culture-epoch" (Goodenough, 1950, p. 32) theory, used child development as the standard which determined a group's position within the hierarchy.

Interest in Haeckel's theory was intensified through European contact with "lower, primitive peoples" (Haeckel, 1891, cited in Jaeger, 1983, p. 434) as Western exploration and colonization of Africa and Asia expanded. "In the late 1800's, parallels between animals and children, primitive societies and the early history of humans were rampant" (Borstelman, 1983, p. 34). A new discipline, "Social Darwinism," (Brinton, Christopher, & Wolff, 1971, p. 861) emerged from studies based on Haeckel's "culture epoch" theory and claimed that the principles of natural selection validated the superiority of western societies. One of the tools used to establish this claim was the analysis and evaluation of human figure drawings.

The collection and study of drawings was especially attractive because, unlike many other methods of measurement, the experiment was short, required minimal equipment, and could be explained through rudimentary communication. In addition, samples of human figure drawings from vanished cultures existed which could be compared with those of modern cultures. Claparède, Luquet, Rouma, and Lamprecht found many similarities between prehistoric, primitive, and children's art which reinforced their idea that "nature always makes use of the same means to further the development of the individual as well as that of the species" (Eng, 1931, p. 213). Their
conclusions reinforced Social Darwinian assumptions since they judged superiority by attributes which appeared in western styles of drawing.

Lamprecht initiated a world-wide investigation of human figure drawings in order to test Haeckel's theory of recapitulation (Goodenough, 1926, Eng, 1931, Harris, 1963). One result of his study was a culture-classification system based on the quality of art. This system caused consternation among many Social Darwinists for it gave an unexpectedly high ranking to supposedly primitive cultures such as African Bushmen and North American Eskimos (Eng, 1931). Lamprecht's classification system was disputed by his colleague, Rouma, who believed the total history of a culture was the appropriate criterion for estimating social evolutionary level. Despite this argument, human figure drawings were considered an important tool for understanding relationships between and within groups by both Rouma and Lamprecht.

Another factor which facilitated the application of Darwinian ideas to cultures, and to the drawings produced by them, was the early 19th century development of probability theory by Gauss and Laplace (Goodenough, 1950). These mathematical theorems stated that the farther an event was from the mean the rarer its occurrence. Data plotted on a graph formed a bell-shaped curve. When these laws were applied to human physical characteristics such as height and weight by the Belgian Quetelet (Hothersall, 1984), a bell-shaped curve again emerged. Quetelet theorized that the average or "normal" characteristic was the ideal toward which humans were evolving (Goodenough, 1950). Galton, a scientist committed to Darwinian principles and the inheritance of intelligence, extended the concept of normal distribution to human mental characteristics. He, however, did not share Quetelet's opinion that average was ideal but adopted an elitist view that characteristics
at the upper end of the normal distribution were the ideal toward which humans were slowly evolving.

Galton believed that human evolution could and should be speeded by the manipulation of a group's genetic stock through institution of social programs which controlled marriage and progeneration (Hothersall, 1984). His ideas formed the basis for eugenics, a concept which would encourage the development and use of intelligence tests. Human figure drawings, originally collected in order to establish a cultural developmental hierarchy, easily adapted to both the mathematical and theoretical criteria for measurement of intelligence. For those who agreed with or were influenced by the ideas of the Social Darwinists, the validity of the human figure drawing as a measure of intelligence was obvious and would not be questioned for many years.

Information about human figure drawings was gained through two methods, cross-sectional and longitudinal studies. Cross-sectional studies were often monumental in scope, extending to thousands of subjects. These were epitomized by Kerschensteiner's collection in the early 1900's of over 100,000 drawings from the schoolchildren of Munich. Kerschensteiner intended to compare quality and/or content of drawings with mental ability, sex, and developmental stage (Mortensen, 1984, Harris, 1963, Eng, 1931, Goodenough, 1926). Other studies compared aspects of drawings to artistic talent, school achievement, and personality (Mortensen, 1984, Harris, 1963, Eng, 1931, Goodenough, 1926). These studies of the extent to which one variable, such as the content of a drawing, is related to another variable, such as school grades, were spurred by Galton and Pearson's development of the mathematical process of correlation (Hothersall, 1984).
Other investigators employed a longitudinal approach (Mortensen, 1984, Goodenough, 1950, 1926). Goodenough (1950, 1926) listed studies of children by Tiedemann, Darwin, Preyer, Stern, and Shinn. A contemporary of Goodenough, Eng (1931), recorded and evaluated the drawings of her niece, Margaret, from age 1 to age 8. She also compared Margaret's progress to that of individual children studied by Luquet, Major, Dix, Scupin, and Krotsch. In making these comparisons, Eng emphasized individual differences by citing the wide range of ages at which the various children achieved each drawing milestone. She also presaged modern conceptions of intelligence such as Sternberg's (1988) theory and Gardner's (1988) "multiple intelligences" by contrasting Margaret's precocity in drawing with her normal development in other areas. These cross-sectional and longitudinal studies fit into the hierarchical structure of Social Darwinism for they showed the existence of individual and group differences within clearly defined drawing stages.

Many researchers described the characteristics of drawing stages. Even the most primitive stage, scribbling, was subdivided into several sequentially ordered parts (Mortensen, 1984, Gardner, 1982, Kellogg, 1979, DiLeo, 1970, Grozinger, 1955, Harris, 1963, Eng, 1931) by early investigators such as Rouma (Mortensen, 1984) and Burt (Harris, 1963) and by recent investigators such as Kellogg (1979). According to Kellogg, there are 20 basic scribbles in the repertoire of the 2 year old which are then combined into 6 diagrams which form the components of all early drawings.

\[0 + \times || / \ldots\ldots\]

For example, the combination of the circle and cross results in a mandala, one of the archetypal symbols described by Jung (Mortenson, 1985, Harris, 1963). Kellogg believed that the mandala, drawn by all normal children of the
world between the ages of 3 and 4, was the forerunner of the classic child's
drawing of a person where the circle comprises both head and body and the
limbs stick out from the circle like rays.

Systematic descriptions of developmental changes in drawings were
also devised by Kerschensteiner, Rouma, Lowenfeld, and Piaget. Mortensen
(1984) summarized their efforts by grouping the changes into three stages:
1. scribbling - marks on paper with no recognizable representation
2. schematic - representational but stereotyped
3. naturalistic - increasingly realistic details and proportion
Many of these developmental criteria were incorporated by Goodenough into
her scoring system for figure drawings including Rouma's contention that a
profile drawing was one of the highest stages of drawing and
Kerschensteiner's opinion that depiction of three dimensions constituted the
most advanced drawing stage (Goodenough, 1926, Eng, 1931, & Mortensen,
1984).

Just as important as the development of criteria for scoring was devel-
opment of a rationale for the meaning of a score. Luquet claimed that children
draw what they know, not what they see. He based this assertion on the
observation that young children did not need a model in order to draw and
generally ignored one if it was present. Thus, a child's drawing must reflect
his mental concept of a person. The idea that a drawing reflects knowledge, a
product of intelligence, became the foundation for the human figure scoring
system of Goodenough (1926).

Reinforcement of the cognitive or intelligence element in drawing also
came from its developmental similarity to oral speech, a similarity which led
many theorists to call drawing a graphic language (Goodenough, 1926, Eng,
Eng specifically equated babbles with scribbles, single words with the circle, and formal expression of complex ideas with complete and complex drawings. In her scheme, drawing, language, and concepts all began as simple and stereotypical and became complex and intentional. The relationship of drawing and language was also noted in the order in which children applied language to drawings, first naming it after completion and later naming it before it was attempted (Gardner, 1982, Lowenfeld, 1954, Mortensen, 1984, Harris, 1963, Eng, 1931, Kellogg, 1979). Piaget (in Barrett, 1983) added an interim stage in which scribbles were labelled during drawing and modifications immediately attempted to make ensuing scribbles fit the label. The relationships between drawings, language, and concepts contributed to the human figure drawing task's status as a measure of cognition or intelligence.

At the time Goodenough published *Measurement of Intelligence by Drawings* (1926), demand for methods of classifying a population by level of intelligence was strong. This demand was promulgated by advocates of Eugenics, the purposeful application of Darwinian laws of selection to humans. Henry Goddard, a prominent proponent of eugenics in the United States, believed that the "chief determiner of human conduct is a unitary mental process which we call intelligence" (Goddard, 1922, p. 1.).

intelligence was genetically determined and little affected by environment. According to Goddard, each person's role in life needed to be equal to his intelligence in order to prevent "social inefficiency, individual unhappiness, misdemeanors and crime" (p. 37).

Goodenough was aware of and sympathetic to eugenic ideas. She described Goddard's imagination as "brilliant" (Goodenough, 1950, p. 54)
when he applied laws of genetic inheritance to human intelligence and praised the positive effect of his notorious work on the Kallikak's: "if ... the feeble-minded person is not merely a burden upon society but is likely to become an active menace as well, then it was high time that something be done about it" (p. 55). One alternative was classification of persons by intelligence and assignment to appropriate levels of treatment, training, and supervision. Goodenough minimized other results of eugenics which are vividly described by Hothersall (1984) who reported Goddard's membership on the "Eugenics Section of the American Breeders' Association" (p. 313), a group responsible for devising "practical methods of eliminating 'defective people' from the population of the United States" (p. 313). Proposals included the disenfranchisement of the unintelligent, sterilization of mental defectives and rejection of immigrants on the basis of how they looked. Other prominent psychologists such as Yerkes, Thorndike, and Terman, Goodenough's advisor, also ascribed to eugenic views. In 1926, the measurement of intelligence was justified by social necessity.

Need for quick and simple intelligence measures was also generated by the influx of immigrant children into the schools. In her 1950 text, *Mental Testing*, Goodenough described the population of a Dayton, Ohio, school in 1912-1913: "the primary grades in many cities, especially those having many children of foreign-born parents, became clogged with children whose ages ranged all the way from six to sixteen years" (p. 16). Examination of her data shows that in grade one students ranged in age from five to fourteen. In grade five, only 35% of the students were at the expected age of 10 and 12% were three or more years older than 10. Examination of the data presented in Goodenough's description of the schools she used in pre-standardization and
standardization of the Draw A Man test shows that school placement statistics were little changed ten years later.

Part of the population used for pre-standardization of the Draw A Man Test came from schools in Perth Amboy, New Jersey, which contained many "foreign or negro" (Goodenough, 1926, p. 36) students. It was Goodenough's conclusion that these schools were "furnishing far more than (the) normal quota of ... retarded children" (1926, p. 37). Retardation was defined as the failure to be promoted one or more times. Approximately 3,600 children from five New Jersey and New York schools later participated in standardization of the Draw A Man Test. Using Goodenough's definition of retardation, 27% of the students at age 7, 40% at age eight, 45% at age nine, and 57% at age ten were retarded. Because "The great amount of retardation led to the belief that norms based upon the total number of children without regard to grade location would be too low" (Goodenough, 1926, p. 39), Goodenough standardized her test on the basis of the sample's 2,306 subjects who had made normal progress in school. The influence of Social Darwinism is seen in Goodenough's acceptance of the idea that minority and immigrant children would be disproportionately represented among the retarded. It is curious that the human figure drawing test was later frequently described as a measure which was culture-fair (McWhinnie, 1971, Hilger, Klett and Watson, 1976, and Weiss, 1980, 1981).

**Construction of the Goodenough Draw A Man Test**

Early studies of drawing showed that the human figure was the most frequent subject of free drawings of children below age 10 (Ayer, 1916, Goodenough, 1926, Eng, 1931). Maitland (1895, in Ayer, 1916) found that
five to seven year old children most often chose to draw human figures followed by still lifes, plants, houses, and animals. By 11-13 years, the order had changed to still life, geometric design, plants, houses, animals, and humans. Although Goodenough (1926) originally hoped to develop a system applicable to any free drawing, she found this task impossible. Therefore, she selected a human figure as the subject of the drawing test because it was a subject with which all children were equally familiar. Following initial testing, she chose to use only a drawing of a man because the greater uniformity of male clothing increased scoring reliability.

The Draw A Man task contained 51 items which were scored as 1 (present) or 0 (not present). Three types of items earned points: the presence of a part or detail, such as the head or hair; correct depiction of proportion, such as eye width greater than height; and quality of drawing strokes, such as all lines firm and controlled. Each item had shown regular and rapid increase in its appearance in drawings as age increased. In addition, each item differentiated between children of the same age who were placed in different grades.

Points earned from a drawing were converted into a mental age equivalent. A ratio IQ was derived by dividing the child's mental age by the child's chronological age. Norms for children from ages three to 13 were determined but those at the extremes were obtained by extrapolation. Goodenough suggested that the test was most valid for children from age 5 to age 10. The norms were established in yearly intervals based on the midpoint of the year. Although Goodenough found significant differences between the performance of boys and girls, separate norms were not devised. Goodenough described the differences as "qualitative" (p. 57).
Reliability

Test-retest reliability (interval = 1 day) of .937 for a group of 194 first-grade children was obtained by Goodenough. This was considerably higher than McCarthy's (1944) figure of .68 (interval = 1 week) for a group of 386 normal eight to nine year olds or McCurdy's (1947) figure of .69 (interval = 3 months) for a group of 56 first-grade children. Other test-retest data ranged from .68 (interval = 6 weeks) to .91 (interval - 4 days) for groups of retarded males (Brill, 1937, Yepsen, 1929). Test-retest data was confounded by the varying intervals between testings and the use of norms with year intervals making them relatively insensitive to the child's age. A difference between scores could be an artifact of the norms table rather than a reflection of change in the child's performance. Goodenough (1926) obtained a split-half reliability of .77 for each age from five to ten giving a SEM of 5.4 points, equivalent to 16.2 months in mental age. This was considered comparable to the SEM on the Stanford-Binet of 14.8 months.

Inter- and intra-rater reliability generally fell in the high ranges for the Goodenough test. McCarthy (1944) obtained an inter-rater correlation of .90 and intra-rater reliability of .94. This consistency may be due to the extensive scoring directions and numerous practice samples.

Validity

Validity for Goodenough's test was determined using several criteria (Goodenough, 1926). First, raw scores increased regularly with age as would be expected in a task measuring what a child knows. Second, no child with a Goodenough IQ below 100 was promoted an extra grade and every
child with a Goodenough IQ below 70 was retained. Third, moderate correlations were found with other measures of intelligence such as the Army Alpha Test and Trabue Completion test. Finally, Goodenough reported a correlation of .741 between the DAM and the Stanford-Binet in a sample of 334 children. A low correlation between drawing scores and teacher estimates of intelligence led Goodenough to conclude that the DAM measured "an aspect of mentality likely to be disregarded by teachers in their estimates of pupil intelligence" (p. 50).

Goodenough's Draw A Man test quickly became a standard part of psychological evaluation batteries (Sundberg, 1961). As later studies (Harris, 1963) provided evidence that the task was likely sensitive to variables other than intelligence, questions about the validity of the drawing task were raised but its use continued.

Summary

The Goodenough test was published at a time when the views of Social Darwinists suggested that the laws of natural selection could and should be applied to the human race. One application was classifying a person on the basis of intelligence in order to determine his educational and vocational placement. A short, simple, inexpensive, non-language task such as Goodenough's Draw A Man Test was a welcome tool for measurement of intelligence. Although the Goodenough test is seldom used today, its basic design and purpose has strongly influenced the Harris (1963) and Naglieri (1988) expansions and modifications. The impact of Goodenough's theoretical position that children draw what they know rather than what they see was reinforced by a high correlation between the DAM and the Stanford-Binet.
Although later studies showed lower correlations between the DAM and individually administered intelligence tests, the initial impression that the DAP was primarily a measure of cognition persisted. The relative contribution of other factors to the score was neglected as was the search for a stronger theoretical explanation of the task.

**Harris' Revision of Goodenough's Scale**

In 1963, Harris published a revised and extended version of the Goodenough DAM task. Virtually all of the Goodenough items were incorporated into his scale and new items added to extend the discriminatory power of the test to age 15 years-11 months. In order to create an alternate form of the task, Harris also developed and standardized a scale for the drawing of a woman (DAW). In comparison to the original Goodenough DAM which contained 51 scoring items, the new DAM scale contained 73 items and the DAW scale, 71. Harris' criteria for inclusion of a scoring point were similar to Goodenough's.

Harris found significant differences between the performance of boys and girls. Therefore, separate norms (1 year intervals) for each sex were created (Harris, 1963, Harris, Roberts, and Pinder, 1970). Harris' directions also called for a drawing of the self but standardization data for this drawing was not developed. Instead it was suggested that the drawing be evaluated using the appropriate existing tables.

The standardization sample of 2,975 subjects was selected to reflect the 1950 census. Harris replaced the ratio IQ with a standard score with a mean of 100 and standard deviation of 15.
The Goodenough-Harris test (G-H DAP) was favorably reviewed by Pringle, Hunter, Wottmar, Kaplan, and Rose and less favorably by Dunn (in Buros, 1972). Dunn noted the sparse data provided for the revision and the reliance on statistics derived almost exclusively from studies of the Goodenough DAM, a less comprehensive measure with old norms and a different method of obtaining a score. This deficiency was soon corrected by more than 100 studies of the Goodenough-Harris test published between 1963 and 1977 (Scott, 1981).

Reliability

Inter-rater reliability (the consistency of scores obtained by two or more persons evaluating the same drawing) and intra-rater reliability (one person re-scoring the same drawing after a period of time) was found to be in the high range as defined by Anastasi (1988). For example, Dunn (1967a, 1967b, 1967c) reported inter-rater reliability of .88 and intra-rater reliability of .93 when two self-taught scorers evaluated 72 drawings of children in grades one to six. These reliability coefficients were similar to those reported by Quast and Ireton (1966), Yule, Lockyer, and Noone (1976), Yater, Barclay, and McGilligan (1969), Gayton, Bassett and Bishop (1970), Harris, Roberts, and Pinder (1970), Sinha (1970), Levy (1971), Phillips, Smith, and Broadhurst (1973), Gayton, Tavormina, Evans, and Schuh (1974), Evans, Ferguson, Davies, and Williams, (1975), Hilger et al. (1976), Lehman and Levy (1971), Pihi and Nimrod (1976), Naglieri and Maxwell (1981), Piersel and Santos (1982) and Gottling (1985).

Test-retest reliability (the consistency of scores earned by the same child over a period of time) was reported by Georgas and Papadopoulou
(1968), Levy (1971), McGilligan, Yater, and Heusing (1971). Reliability ranged from .45 to .87 for intervals from two weeks to seven months. The reliability of the Harris revision appeared to be similar to the Goodenough DAM.

Validity as a Test of Intelligence

The most comprehensive review of the validity of the G-H DAP as a measure of intelligence is that of Scott (1981). Of the 100+ studies reviewed, 40 used one or more individually administered intelligence tests as the criterion. The Stanford-Binet was investigated in 31 studies and the WISC-R in 30. Scott’s technique for summarizing results was to determine a mean score from studies grouped according to test and type of population. For example, a mean correlation between the DAM and the Stanford-Binet of .52 was derived from nine studies of normal children aged 5 - 11. Results from studies of heterogeneous (r = .51) and special populations (r = .49) were in the same range. Similar DAW correlations in a number of these studies indicated the DAW was a useful alternate form of the test. Correlations of .52 (normals), .50 (heterogeneous), and .49 (special) between the DAM and the WISC-R were almost identical to the Stanford-Binet. However, the two studies of the DAW and WISC-R showed high (.81) rather than moderate correlations.

The results of Scott’s survey, however, must be interpreted with caution. Grouping of studies masks the range of differences between obtained correlations. Further, most studies failed to compensate statistically for larger or smaller than normal standard deviations of scores (Dunn, 1967 b, c, Schaefer & Sternfield, 1971, Laosa, Swartz, and Holtzman, 1973). Dunn attributed high correlations between the Harris DAM and DAW and intelligence measures to the change from a ratio IQ to a standard score, but it is
likely that correlations were inflated due to the large standard deviations of the WISC, 22.6 in one study and 23 in the second. Two more recent studies (White, 1979, and Naglieri & Maxwell, 1981) obtained correlations between the DAM and the WISC-R of .72 and .70, slightly lower than Dunn's, when standard deviation correction procedures were employed. However, many of the studies included in Scott's review obtained much lower correlations than these in order to arrive at an average correlation of .50.

The value of the obtained correlations between the drawing task and intelligence tests brought very different conclusions from researchers. Yule et al. (1976) called DAP-WISC correlations of .33 to .41 "far too low to warrant the acceptance of the revised Goodenough-Harris Drawing Test as a test of general intelligence" (p. 111). Oakland and Dowling (1983), on the other hand, called a median DAP-WISC-R correlation of .55 indicative of strong support for its use as an intelligence measure.

Other methodological problems which confounded understanding of the Harris DAP included small samples, poorly defined groups, heterogeneous groups, use of varying forms of the intelligence measure, and differing standards for meaningful correlations. Curiously, many studies omitted the mean and standard deviation for the DAM or DAW while providing these figures for the WISC, WISC-R or Stanford-Binet (Dunn, 1967a, c, Lehman & Levy, 1971, Tramill, Edwards, & Tramill, 1980, Yule, Lockyer, & Noone, 1976, Laosa et al., 1973).

Investigators strongly suggested that the DAP was not useful in individual evaluations because it frequently yielded significantly lower standard scores than the WISC, WISC-R, or Stanford-Binet. They found discrepancies as large as 30 points and an average discrepancy of approximately 10 points.
There was some agreement that the DAP was more consistent with scores in the below average range on intelligence tests and less consistent with scores in the average and above average ranges (Scott, 1981, Reisman and Yamokoski, 1973, Lehman & Levy, 1971).

A final problem with the Goodenough-Harris DAP was the persistence of sex differences and the acceptance by reviewers and users of separate norms for boys and girls. These separate norms confound claims of presence or absence of sex differences in any study since male and female scores have not been determined in the same manner (For example, see Tramill and Edwards, 1980, Laosa, Swartz, and Holtzman, 1973, Stevenson, Parker, Wilkinson, Hegion, and Fish, 1976, Oakland and Dowling, 1983). Mortensen (1984) addressed the problem with Harris' scales:

"the great sex differences in the drawing test points against drawing as being purely cognitively determined. Outspoken sex differences in other intelligence tests usually lead to their exclusion from a test battery even in composite tests. When only a single test is used as here, such a big sex difference seems even more fatal. Not even the use of different norms for the two sexes justifies the use of it. Unless it is accepted that girls are generally more intelligent than boys - which it is not - it shows clearly a strong influence by factors other than purely intellectual ones. (p. 67).

Mortensen's concern is generally absent from the DAP literature where the unusualness of differing criteria is seldom voiced and the impact of separate but "equal" measures on test validity is seldom examined

Summary

There is general agreement that the Harris DAM and DAW are reliable tests which are highly correlated with their predecessor, the Goodenough DAM. On average, the drawing task correlates significantly but in the
moderate range with standardized tests of intelligence. The task appears to be more valid as a screening measure for groups than for individuals due to the significantly lower scores often obtained on the DAM than on the Stanford-Binet or WISC (R). An intelligence scale which requires separate norms for boys and girls is unprecedented and a serious challenge to the validity of a measure of intelligence or intellectual maturity.

**Draw A Person: A Quantitative Scoring System (DAP-Q)**

In 1988, Naglieri modified the two previous scoring systems for human figure drawings in an attempt to correct structural and statistical weaknesses (Sattler, 1982, Scott, 1981, Phillips, Smith, & Broadhurst, 1973). These weaknesses included norms which no longer reflected the United States population, different scoring criteria for drawings of the man and the woman, lack of criteria or norms for a drawing of the self, two sets of norms because of differences between the responses of boys and girls, relatively gross measurement because of yearly increments in the norms, outdated scoring items, and ambiguous or inadequate scoring criteria.

In the DAP-Q, three drawings are evaluated, man, woman, and self, using the same scoring criteria. A standard score is derived for each drawing as well as a Total standard score. The scores have a mean of 100 and standard deviation of 15. There are three general scoring elements, presence, detail, and proportion, for each of 14 areas. A fourth element, bonus, is earned if all previous elements are scored. Generalized scoring categories were created in an effort to minimize the effect of cultural changes in clothing or hair styles. Each drawing has a potential of 64 scoring points. Freeman (1980) was critical of the drawing task in that it assumed one drawing
revealed all the child knew. The DAP-Q offers a child three opportunities to
draw, a procedure which greatly increases the reliability of the Total Test
score. Although Harris (1963) offered two drawings, he considered then
alternate forms rather than two samples. When two drawings were obtained,
he recommended averaging the standard scores and did not create a norms
table based on two drawings.

The DAP-Q standardization sample of 2,662 subjects aged 5 to 17 was
selected to reflect the 1980 U.S. census. Schools rather than individuals were
used to determine socioeconomic status. Evaluation of the data indicated that
there were no significant differences between the raw scores of boys and
girls, although boys consistently scored lower than girls. The difference when
converted into standard score points was minimal, approximately one fifth of a
standard deviation. Raw score-age progressions indicated that 21 age divi-
sions were appropriate: quarter-year intervals for ages five to eight; half year
intervals for ages nine and ten. Because less discrimination was found in raw
score changes for older subjects, ages 11 to 17 were grouped together.

The manual reports results of numerous DAP-Q studies of groups
matched for age, sex, or race. In all cases, there were no significant differ-
ences between the scores earned by boys or girls, blacks or whites,
Hispanics or non-Hispanics (Naglieri, 1988).

Reliability

A median Total Test internal reliability of .86 was reported in the DAP-Q
manual. The range of reliability coefficients for ages 5 to 17 was .83 to .89.
These results are comparable to or better than those reported for either of the
previous measures. Inter-rater (.94) and intra-rater (.96) reliabilities (Gottling,
1985) suggested that the stability of the test scores was high. Test-retest reliability of .74 was similar to the reliabilities reported for the Goodenough and Goodenough-Harris versions of the test.

**Validity**

Initial studies indicated that the DAP-Q correlated in the high range with the Goodenough-Harris DAP (.75 to .87). Comparisons of the DAP-Q and tests of intelligence (Lillis, 1987, Wisniewski & Naglieri, 1988) were in the moderate range, .64 (Differential Abilities Scale) and .51 (WISC-R). These results are similar to those obtained using the two previous scoring systems.

**General Developmental Theory**

The three DAP scoring systems summarized in the previous section are based on a model of a developmental sequence in drawing which reflects levels of cognitive growth. Goodenough believed a child's general intelligence could be estimated by this model. Harris, however, moved away from general intelligence and suggested that the task was a reflection of conceptual maturity. Koppitz (1968) moved even further from the intelligence and cognitive aspects and called a set of thirty characteristics "derived from the Goodenough-Harris scoring system and from the writer's own experience" (p. 9) Developmental Items. According to Koppitz, a Developmental Item occurs only on relatively few HFDs of children of a younger age level and then increases in frequency of occurrence as the age of the children increases, until it gets to be a regular feature of many or most HFDs at a given age level. (p. 9)
Koppitz believed these items were related to age and maturation and not to artistic ability, school learning, instructions, or medium. These conclusions were similar to those reached by Goodenough or Harris. It appears that only the name applied to the scoring scale and the number and complexity of the items differentiates between the various scoring systems. Intelligence, conceptual maturity, and development are estimated so similarly that they appear to be interchangeable constructs.

Many of the developmental aspects of human figure drawing have been questioned. Hagen (in Freeman & Cox, Eds., 1985) used historical and cross-sectional methods to determine if the expected hierarchy of drawing geometry, elements of which were used by Rouma and Lamprecht to evaluate human figure drawings, actually existed. This hierarchy included three forms of drawing, orthogonal, affine and projective. In orthogonal drawing, all right angles remain right angles, all lines in parallel remain in parallel and size does not change regardless of the drawer's distance from the object. In affine production, right angles become acute or obtuse, but parallel lines remain parallel and size does not change with increasing distance. In projective drawing, right angles become acute and obtuse, parallel lines are replaced by lines which converge to a point, and size decreases as the drawer's distance from the object increases. Hagen's examination of representative art from ancient and modern societies showed that orthogonal geometry does not always emerge first. The earliest examples of Chinese, Japanese, Inca, Aztec, Mayan, and American Northwest Indian art use an affine system. Further, Hagen supported Lamprecht's early twentieth century findings that ice age artists (approximately 10,000 B.C.) and African Bushmen (approximately 1000 A.D.) used the most complex geometrical system of
projection. Only two other samples of projection as the dominant geometry in
drawing exist: Chinese landscapes (1000 A.D.) and Renaissance painting
(1500 A.D.). Hagen concludes that "I cannot find evidence that the art styles
of different cultures have developed systematically in an Orthogonal to Affine
to Projective progression. Even in Western art (including Egypt), one can
make this case only by choosing stylistic ancestors with great care and by
ignoring diversions such as Cubism" (p. 69).

Hagen also conducted a cross-sectional study of development in draw-
ing. She asked children aged 6, 8, and 10 and adults to draw a house. Less
than half of the adults used projective geometry when a model was present
and only 10% when no model was present. From 20% to 30% used the sup-
possedly primitive orthogonal style. When the instructions asked the drawers
to use perspective, the number of drawings using affine geometry increased
but the number of drawings using projective geometry remained minimal.
Hagen concludes

That drawing develops is a basic assumption underlying the use of
drawing tests to measure cognitive, perceptual, emotional and sexual
development. Is it possible to measure developmental level with a tool
that itself shows no development? Of course, representational draw-
ing, like representational art, involves the depiction of many aspects of
human experience other than that of spatial layout. It may be the case
that these aspects permit the valid use of drawing tests to measure
developmental level, but the issue invites some skepticism and consid-
erable empirical work. (p. 76-77)

Hagen's position is important for both the interpretation of human figure draw-
ing scores and the construction of the scoring criteria are based on develop-
mental assumptions and include points related to spatial layout.

One of the scoring areas of the Goodenough and Harris tests based on
developmental assumptions related to spatial layout is profile drawing. This
was included on the basis of Rouma's (in Goodenough, 1926) claim that it was one of the highest forms of drawing development. Goodenough was ambivalent about profile drawing: "Most of the literature on children's drawings tends to give the impression that the change to the profile position is a general rule which all children come to adopt in their drawings, but my own figures show that this is far from being the case" (p. 110). Harris (1963) eliminated profile as a scoring point for the DAW. However, he retained it for the DAM while at the same time suggesting that future revisions should remove it. Profile drawing was included in the pre-standardization form of the DAP-Q (Naglieri, 1988), but was removed when the item did not meet statistical standards.

Other scoring areas also reflect the assumption that development is defined by more and more realistic depiction by increasing use of perspective, the projective method of drawing. Harris' DAM/DAW scales gave points for using acute angles, foreshortening in feet, modeling of body contours such as chin, cheeks, breasts, and hips, and modeling of clothes to show pleats and drapes.

The idea that increasingly precise accuracy in drawing was related to higher levels of development and cognition was also challenged by perceptual psychologists such as Arnheim, Moustgard, and Gibson (in Mortenson, 1985). In their view, what a child saw was never based on photographic details but was primarily a registering of essential structure. As a result, children drew what they perceived which was likely to be different from what they knew and from what was optically present.

Perception, and the resulting drawing, could be altered by many variables. One of these variables was set or motivation (Barrett, Beaumont, & Jennett, 1985, McWhinnie, 1971). Lowenfeld (1954) observed that children
developed schemas by which they typically represented an object. However, if the child drew in response to a request, the schema might be adapted resulting in exaggeration, omission, or change in a part of the drawing. A second variable which influenced perception was cultural experience and learning. McWhinnie (1971) believed drawings were best explained by McFee's perceptual-delineation theory. This theory emphasized "the learned nature of perception and drawing" (p. 142) and considered a drawing "the result of the sum total of ... cultural and social experience" (p. 136).

McWhinnie cited the studies of Salome (1967) and Badri (1965 a, b) in support of McFee's theory. These studies showed that materials, experience, and learning influenced the quality and complexity of children's drawings. In a related study, Leviton and Kiraly (1974) found no significant differences in DAP scores earned by two groups of 4 year olds. However, when one group was exposed to an intensive enrichment program and another was not, significant differences appeared in post-treatment DAP scores. The authors concluded that DAP scores were sensitive to environmental experiences. These authors also questioned the assumption that drawings are primarily developmental in nature.

The DAP and Models of Intelligence

The meaning of the human figure drawing tests' correlation with other tests is dependent upon the constructs attributed to the related test. Early studies compared DAP scores to general intelligence scores derived from standardized tests of intelligence such as the Stanford-Binet (Terman, 1916). As the popularity of the concept of general intelligence declined, attempts were made to divide intelligence into more useful components. These
components might be the Verbal and Performance IQs derived from the WISC (R) (Wechsler, 1949, 1974), the 120 cells of Guilford's Structure of the Intellect (Anastasi, 1988) or the three factors found through factor analysis of the WISC-R (Kaufman, 1979).

Many studies of the DAP attempted to redefine it as a measure of non-verbal intelligence using the Verbal and Performance scales of the WISC (R) as the criteria. Harris (1963), in his review of the Goodenough DAM literature, believed that "From the evidence... the Draw-A-Man Test is not more allied with performance than with verbal abilities" (p. 99). However, the non-verbal format of the drawing task led other investigators to a different conclusion.

Dunn (1967c) found a correlation of .64, .59, and .62 between the DAM and the WISC Full Scale (FS), Verbal (V), and Performance (P) IQs respectively, and .32 and .17 between the DAM and the Verbal and Non-Verbal IQs of the California Test of Mental Maturity (CTTM). From this data, Dunn concluded: "Correlations with verbal abilities are consistently lower than correlations with skills such as Block Designs (.60) and spatial perception" (p. 301). This claim lacks the support of mathematically determined significant differences between the various correlations of the DAM and WISC scales and subtests. Further, the DAM - Information, Comprehension, Similarities, and Vocabulary correlations are slightly higher than the DAM - Picture Completion, Picture Arrangement, and Coding correlations while DAM - Vocabulary correlation is the same as the DAM - Object Assembly. In addition, Dunn ignored the data which showed that the DAM correlated significantly with the CTTM Verbal IQ but not with the Non-verbal IQ. Dunn's conclusion that the DAM best measures "the ability to develop and utilize
concrete functional-motoric concepts as contrasted with abstract-verbal concepts" (Dunn, 1967c, p. 301) is unsupported.

Gayton et al. (1974) found correlations of .48 and .57 between the DAP and the Verbal and Performance scales of the WISC. Significant differences between other correlations in the study were given. However, the significance of the difference between Verbal and Performance scale-DAP correlations was not reported leading to the assumption it was not significant. Gayton's conclusion, "Consistent with previous findings, the correlations between the figure-drawings and Performance IQ are higher than those between figure-drawings and Verbal IQ." (p. 370), while true, is misleading.

Laosa et al. (1973), claimed greater correlation between DAM-PIQ than DAM-VIQ in a study of sex differences. However, they used only two WISC subtests, Vocabulary and Block Design, and did not determine the significance of the difference between the correlations. Further, the inconsistent reporting of results for the two subtests makes it difficult to evaluate the relevance of the conclusions. Vocabulary scores are reported for both sexes at three ages for two drawings. Block Design scores are reported using a range and median. The correlations for boys are virtually identical at ages 8 and 14 and only show a dip in the Vocabulary correlation at age 11. The correlations for girls are more discrepant and suggest the likelihood of a significant difference.

The perpetuation of the non-verbal label is evident in Levinson and Block's (1977), study of the drawing performance of Orthodox Jewish children. They stated "We do not ... consider the Goodenough-Harris Test as a valid test of verbal intelligence. However, we do feel it is a good measure of
performance ability" (p. 155). Several flawed articles were cited to support this statement.

Similar repetitions of the high DAP - Performance IQ relationship are found in Mortenson (1984), Koppitz, (1968, 1984) and Oakland and Dowling (1983). Oakland and Dowling mention a "small but insignificant trend" (p. 527) in favor of DAP-PIQ over DAP-VIQ correlations among Anglos, blacks, and Mexican-Americans. Koppitz uses selected case studies to support her position that the DAP overestimates the ability of subjects with a high PIQ and low VIQ and underestimates the ability of those with the opposite profile. Mortensen (1984) perpetuates the non-verbal concept: "The fact that the performance part of the WISC corresponds better to the drawing test than the verbal part seems to be in opposition to Harris' idea of drawing as mainly reflecting concept formation" (p. 67). One article is cited in support of this sweeping statement.

The loose interpretation of data in the studies mentioned above is sufficient to challenge the contention that the drawing task is more related to Performance tasks than Verbal tasks. However, support can be found for the opposite conclusion, that the DAP is more closely related to measures of Verbal ability. Hartman (1972) found correlations between the DAP-VIQ of .37 and DAP-PIQ of .28 for a group of 50 normal children. The DAP-VIQ correlation was significant at p < .01 while the DAP-PIQ was significant at p < .05. Tramill, Edwards, and Tramill (1980) investigated DAP- WISC-R correlations for 100 children in academic difficulty. Contrary to expectations, overall variance for boys was equally divided among the verbal and performance subtests while the variance for girls was stronger for verbal subtests than performance. A recent study employing the new DAP-Q found a correlation of .54
with the WISC-R Performance scale and .42 with the Verbal scale (Naglieri, Prewett, & Bardos, in press). A t test for the difference between correlated correlations showed this difference was not significant. It appears that a differential relationship between a drawing score and verbal and nonverbal intelligences has not been established.

The studies which attempted to relate the DAP to nonverbal intelligence, although unsuccessful, did use a model of intelligence exemplified by validated tasks (the Verbal and Performance scales of the WISC-R). Other attempts to establish a theoretical model have been flawed by a lack of reliable and valid tasks. For example, when Harris (1963) called the drawing task a measure of conceptual maturity and used the title of his manual, *Children’s Drawings as Measures of Intellectual Maturity*, to modify the strong connection of the DAP with general intelligence, he relied on descriptive language without the support of experimental tasks. Freeman (1980) pointed out the poverty of Harris’ redefinition: "there are no independent criteria put out by Harris for identifying such increases [in intellectual maturity] or analyzing the underlying processes" (p. 53).

Harris also suggested that Piagetian theory might best explain drawing scores. He based this suggestion on the DAP’s ability to discriminate efficiently among children in the age groups corresponding to Piaget’s preoperational and concrete operational stages. While Harris again did not support his theorizing with experiment, others have conducted some research in this area.

Piaget and Inhelder (1969) called drawings "a copy of reality" (p. 631). From ages two to seven, drawing was a special case of imitation exemplified by rigid drawings because of the child’s passivity in the perceptual process.
At ages seven and eight, the child began to develop discrimination so that the drawing reflected the child's activity in choosing what was needed for his drawing purpose. Imitation became controlled wholly by intelligence.

Carlson (1970) investigated the relationship between the DAM, Raven's Progressive Matrices, and four Piagetian tasks of conservation of mass and weight. Using a sample of 221 Kindergarten - Grade three students, Carlson found that the DAM correlated significantly with the Raven’s test and with tests of conservation. However, when the variance of the DAM was held constant, the Matrices test was still significantly related to the Piagetian tasks. When the variance of the Raven's test was held constant, the DAM was not significantly related to conservation. Carlson concluded that there was only limited support for Harris' hypothesis that the DAM measured cognitive development similarly to Piagetian tasks.

Barrett (1983) reviewed studies of the 4 ontogenetic stages of drawing described by Piaget and Inhelder. These stages were: 1. Fortuitous realism, 2. Failed realism, 3. Intellectual realism, and 4. Visual realism. Barrett found that, according to the literature, these stages were not inexorable and that different stages were frequently present in the same drawing. For example, a drawing might have parts which were intellectually realistic (showing transparencies) and visually realistic (no transparencies and with correct perspective). Therefore, Barrett suggested that Piagetian ontogenetic stages applied only in the most general manner.

Madden (1986) investigated Piaget and Inhelder's theory that the child uses schemas (concepts) to interpret perceptual data and then generates an internal image which guides drawing. This study focused on Piaget and Inhelder's claim that children develop a "fore-image" two years before the
"mental image proper" and thus are more accurate in copying than in drawing. Children were asked to draw an array, copy it, or choose from predrawn samples. Piaget's claim that because of the "fore-image" a child would be more successful in copying than in drawing tasks was not confirmed. Madden's hypothesis that choice would result in even fewer errors was also not confirmed. The degree of success in both copying and drawing appeared to be related to the complexity of the task. If a task appeared complex and if conflicting cues were given, children used concrete operations to organize the information. Again, Piagetian concepts about the nature of the drawing task were only slightly supported.

Bremner (1985) found that children have innate geometric biases favoring right angles and symmetry. Human figure drawings with arms sticking out at right angles from both sides of the body are reflections of these biases rather than of knowledge or conceptual maturity. Other examples of these biases are houses with windows on either side of a door and chimneys placed at a 90 degree angle relative to a sloping roof. Bremner suggested that Piaget failed to account for this bias when conducting his water-in-the-tilted-glass experiments and Freeman (1980) called the experiment "worthless" (p. 348). Failure to account for the perpendicular bias distorted the "child's knowledge of the concepts involved in the task ... similar dangers exist when we try to interpret children's drawings" (Bremner, p. 311). This comment applies to the Goodenough, Harris, and Naglieri systems of scoring which base points on the presence or absence of a 90 degree angle between arms and the side of the body.
The DAP and Possible Links to the Das-Luria Cognitive Processing Theory

The unsatisfactory results of attempts to relate the DAP to a content-oriented theory of intelligence as exemplified by the Verbal-Performance scales of the WISC-R or to a cognitive developmental theory such as Piaget's has led some experimental psychologists to focus on the mental processes which must work together to produce a drawing. Their work, and the language used to describe the processes, supports the appropriateness of applying the Das-Luria model of cognitive processing to human figure drawings. Luria described "three principal functional units of the brain whose participation is necessary for any type of mental activity" (1973, p. 43). The first unit is responsible for cortical tone and maintenance of attention. The second receives and stores information in a successive or simultaneous format and the third controls regulation, evaluation, and modification of behavior. In the Das-Luria model this latter function is called planning. The four components of the theory have been examined through the development and testing of numerous tasks related to each component (Naglieri, 1988).

Methods of coding of information, similar to those described in the Das-Luria model, were proposed by Selfe (1985) as a factor in the different drawing behaviors of normal and autistic children. Some autistics, with limited experience, little language, minimal acquired knowledge, and lack of social responsiveness, do not experience the drawing difficulties with which non-autistic children struggle. They are, instead, able to create highly detailed, spatially correct drawings. Selfe cites Paivio's theory of "two distinct modes of internal representation in thinking processes" (p. 151) to explain these
differences. Some images are present simultaneously and recognized by their spatial arrangement while others are "specialized for sequential processing, for processing over time" (p. 151). Selfe suggests that autistic children draw more accurately and realistically because they are using a visual/spatial (simultaneous) processing mode without input or interference from the temporal/verbal (sequential) mode. Freeman (1980) has attributed the errors in normal children's human figure drawings to their attempts at integration of temporal and spatial processes. Freeman (1980) divided children into two groups according to the structure of their human figure drawings. The tadpole group made only one large body part from which limbs extended and the conventional group made a head and trunk. Freeman analyzed their responses to a systematically varied set of incomplete drawings and clay figures. Freeman found a typical serial order of construction or completion of a figure: head, (trunk), legs, arms. He proposed that this order reflected the well-known psychological principle that humans focus on end points in a series and pay less attention to its middle components.

It seems quite plausible that children are trying to make their temporal order of production map onto the spatial order of the page by anchoring on the top structural feature, but then are prone to end-anchor on the other terminal structural item, the legs. (p. 334)

Freeman also varied the circumstances under which children placed arms on drawings. He found that human figures composed of enclosed contours (circle, square, or rectangle) whose proportionate head-trunk size was systematically varied resulted in more frequent arm placement on the larger body part. Although this effect was most evident among the tadpole drawers, many of the conventional drawers either showed the same response or made
an initial pen movement toward the largest part before inhibiting their reaction if it was incorrect. In contrast, when circular heads of varied sizes were paired with a collapsed contour trunk (a line or filled in stick), both groups of children were more accurate in their arm placement. According to Freeman, this was likely due to the salience of vertical cues as well as the strong contrast between the circle and the line. Freeman called the children's responses to body-proportion effect and end-anchoring examples of "rational strategies for dealing with spatial planning" (p. 338). The generation of strategies is one of the principal functions of the planning factor in the Das-Luria cognitive processing model. Although Freeman described spatial (simultaneous) and temporal (successive) processes in the execution of a human figure drawing, he believed that "any topic in drawing [should] be considered as a planning problem" (p. 302). His work is consonant with the Das-Luria construct of the interdependence of the functional units of the brain.

Freeman emphasized the importance of theory and testing of it through carefully controlled experiments. In the past, he suggested, vague theories such as the child draws what he knows have been accepted and repeated. But, Freeman writes, "one cannot simply issue knowledge a blank check by asserting that it is dominant in the drawing-process - for what interests us is the mechanism whereby it can exert its effect. The literature is searched in vain for that" (p. 353).

Crook (1985) has shown that the interaction of knowledge and planning often results in drawings which adults perceive as immature or bizarre. For example, the child who knows many attributes of the object he is drawing may try to portray all of them. Thus he may fail to use the occlusion (overlap) which would constitute a realistic depiction. This would explain
transparencies in human figure drawings and the teacup drawings first studied by Cox (Freeman, 1980). In the latter, children draw the handle even though the cup is placed in such a position that the handle cannot be seen and the instructions ask the child to draw exactly what is seen. Because the child's knowledge of the object to be drawn exceeds his knowledge of or experience with the conventions of drawing, he does not respond to the task in the same way as an adult. Quantitative scoring systems which award points based on the use of occlusion may reflect the ability to plan an appropriate drawing strategy rather than knowledge of concepts related to the human figure.

The relationship between planning and drawing has been investigated by Laszlo and Bairstow (1985) and Laszlo and Broderick (1985). Laszlo and Bairstow developed a series of tests for kinaesthetic sensitivity. Of the children diagnosed as clumsy by this measure, 74% were dyskinesiesthetic, functionally affected in their writing and drawing performance by their poor kinaesthetic processing ability. When these children participated in training designed to increase their awareness of their motor plans, significant improvement in their writing and drawing skills were noted.

Laszlo and Broderick compared the copying and tracing of three simple designs by 491 children from ages 5 to 12. They found that five and six year old children who could not copy accurately were able to describe how their drawings should look. Thus, their drawing failures could not be attributed to a knowledge deficit. In addition, the design to be copied was available throughout the drawing task so that difficulties in copying could not be attributed to memory. Laszlo and Broderick proposed that inefficient motor planning was an important variable in the children's success. When the motor
planning component was diminished by asking the children to trace the
designs, speed and accuracy were enhanced.

The copying of 5 to 6 year olds, who were unable to form angles accu-
rately or to close the figures, was significantly different from that of 6 to 7 year
olds. By age 7, closure was no longer a problem and angles improved.
Laszlo and Broderick offered several reasons for these differences. First, the
younger children had inadequate information stored in memory due to inexpe-
rience and insufficient kinesthetic coding. As a result, the motor program
used in the copying task was inappropriate and inefficient. Kinesthesia, the
"sense of movement and position" (p. 357), provides information about "static
position, movement extent and direction, velocity and force" (p. 360). It is the
"primary channel for detecting any discrepancy between the movement and
the goal set for the task" (p. 360). When the motor planning demands were
reduced by asking the children to trace the design, they were able to make
the appropriate movements to form angles and close the figure although some
errors were still present. The authors concluded that motor planning, which is
dependent on adequate kinesthetic information, was the source of difficulty for
young children. This inadequacy also affected the ability of the children to
detect errors and make corrections in the motor program. The authors note the
minimal importance of visual information in a copying or drawing task as the
pencil mark cannot be seen until the motor plan is implemented. Further,
many children made viewing their work-in-progress difficult by the way in
which they gripped the pencil or arranged their hands and head.

Laszlo and Broderick concluded that "Most clumsy children find writing
and drawing difficult, some even impossible" (p. 369). Young children of
kindergarten age and older children with kinesthetic processing problems were not able to

put to paper the internal representations they build of their world. Consequently, accepting a child's drawings as a faithful representation of his mental images would lead us to underestimate his mental processes .... The astounding thing is that the child can draw at all, not that he draws inaccurately. (p. 370)

The importance of planning in the acquisition of motor skills is further reason for investigating the relationship between motor ability and human figure drawings.

Motor Development

The most frequently studied motoric aspect of human figure drawing has been the visual-motor component. Success in drawing has been compared to success in paper and pencil design copying tests such as the Bender Visual-Motor Gestalt Test or McCarthy Draw A Design subtest. Hartman (1972) and Oakland and Dowling (1983) obtained correlations of .57 (n = 50, p < .01) and .34 (n = 188, p < .01) respectively for the Harris DAP and the Bender-Gestalt. Reynolds (1978) found a correlation of .48 (group administration) and .52 (individual administration) between the McCarthy Draw A Child and Draw A Design subtests for a group of 83 K-2nd grade children.

Another study of a paper and pencil task with cognitive demands different from design copying was conducted by Stanley and Watson (1980). They compared the success of dyslexic (n = 10) and non-dyslexic (n = 10) boys in two motor production tasks, writing an essay and drawing a person. No group differences emerged in the drawing task which was scored for time
on task and Harris points, but group differences did emerge in the writing task in the number of spelling errors and time on task.

In other studies, the motoric component of the human figure drawing task has been estimated by varying the materials used to construct the human figure. Brittain and Chien (1980) gave preschool children materials such as paper and pencil, a puzzle, and clay to make or assemble a man. When no significant differences in the children's degree of success was found, the authors concluded that the critical task demand was cognitive, not motoric. In contrast, Golomb (1973, 1974) found that preschool children were unable to create a clay figure resembling a human but could draw a primitive tadpole figure with pencil and paper. Given puzzle pieces which made a human being, they were able to construct a sophisticated figure. The author concluded that the critical difference in success was the type of motor skill and motor planning needed for success. Golomb's findings supported Olson (1970) who felt that perceptual skills were always ahead of productive skills. As a result a freehand drawing task such as the DAP was more difficult than assembling a puzzle but less difficult than molding clay.

Studies which compare the DAP to tests specifically designed to measure motor ability or development are rare, perhaps because of Goodenough's (1926) dismissal of the influence of visual imagery and manual skill on drawings and strong belief in their cognitive nature. When Harris (1963) found a significant correlation \( r = .43 \) between the DAP and the motor component of the SRA Primary Abilities Test in a study of 164 kindergarten children, he discounted the importance of this finding by citing the unreliable nature of the motor measure.
It is not uncommon to find motor ability measured with a task not validated for the purpose. An example is Koppitz' (1968) study of the influence of artistic ability on human figure drawings. Koppitz believed that children with artistic ability were likely to have "good visual-motor perception and good fine motor coordination (p. 26)", attributes she claimed were measured by the Performance scale of the WISC. Mortensen (1984) criticized several aspects of Koppitz' methods:

It is probable that a good visual-motor perception and a good motoric coordination are necessary conditions for being a good drawer, but they may not be sufficient, and for example the motoric abilities measured by the WISC are not necessarily identical with those needed for drawing. (p. 34)

Koppitz' conclusion that artistic ability defined as visual-motor perception and good fine motor coordination did not influence human figure drawing scores was questionable.

Some studies contain information about the DAP and motor development but do little to explain the findings because the motor aspect is not a primary focus of the study. For example, Pihl and Nimrod (1976) used a bipolar teacher rating scale of the achievement of 44 Grade five children in academic and specialty areas to examine the validity of the DAP. Significant correlations were found between Harris IQ and teacher ratings of overall achievement, reading, and written expression but not for arithmetic, art, or physical education. The authors concluded that these results supported use of the DAP as an intelligence measure as well as Goodenough's claim that it was not influenced by artistic ability. The non-significant correlations between arithmetic and physical education were ignored.
Studies by Salome (1967) and Badri (1965 a,b) which challenged developmental assumptions about the motor abilities needed for drawing were reviewed by McWhinnie (1971). Two axioms were disputed by Salome (1967); one, that children ages 3-7 need large pencils and crayons for drawing because of immature fine motor skills and two, that small children are not motorically ready to put fine lines and minute details in their drawings. He found that kindergarten children could successfully learn to use a variety of drawing materials and that the material used to create the drawing had a strong impact on its qualities. According to Salome, the motor skills needed for drawing were learned rather than developmental. Salome’s position was supported by previous studies by Badri (1965a, b) who compared the human figure drawings of children who differed primarily in their previous experience with paper and pencil activities. The inexperienced children made drawings which were judged inferior to those of the experienced children. Motor learning, not motor development, was the important variable.

The paucity of studies specifically designed to measure the relationship between the DAP and motor development suggests that any statement pertaining to the influence or lack of influence of motor development on the drawing task is a reflection of opinion rather than a conclusion based on the results of well-designed studies.

**Personality**

The use of human figure drawings as a measure of personality traits and emotional states was a byproduct of its use as an intelligence measure. Goodenough (1926) had noticed subtle qualitative differences between the drawings of children with problems and normal children and suggested that
another scoring system might be developed which would "throw light on functional mental disorders" (p. 82). However, she was aware that her conclusions were based on clinical judgment and "subject to a very wide margin of error" (p. 66).

Pre- and Post-World War II enthusiasm for study of the dynamics of psycho analytic theory and self-image theory (Hammer, in Rabin, 1981) led to the development of techniques of drawing analysis which are still in use today. One of the first proponents of projective analysis of human figure drawings was Karen Machover. In 1949, she published a monograph, *Personality Projection in the Drawing of the Human Figure*, which outlined the principles of drawing interpretation. Machover described the origin of her system. "In the course of administering Goodenough's Drawing-of-a-Man test for the usual IQ purposes, it was discovered that careful study of the individual drawings often yielded rich clinical material not related to the intellectual level of the subject. Children securing the same mental age rating would frequently do strikingly different and individualized drawings" (Machover, 1949, p. 15).

Machover's system required analysis of body parts, clothing, structural and formal features and conflict indicators. Structure and form included theme, action, order, balance, size, placement, perspective, and line quality. Conflict included omissions, erasures, and shading. Traits found in abnormal subjects were used to confirm abnormal traits in "so-called normals" (p. 24) whose drawings frequently "revealed neurotic conflicts ... and perhaps even full-blown schizophrenia" (p. 24). Machover did not investigate the frequency of the trait in a large, representative group.

Another early advocate of systematic evaluation of drawings, Emanuel Hammer, also credited Goodenough with the first suggestion that human figure
drawings by children were "tapping personality factors in addition to intellectual capabilities" (Hammer, 1958, p. 19). Hammer's system had two categories, projection and expression. Projection was shown through facial expression and body parts; expression was the drawing process used. There was much overlap between Machover's and Hammer's systems.

Levy (in Hammer, 1958) expressed concern that methods such as Machover's could lead to simplistic, literal, pseudo-Freudian interpretation of drawings. In order to emphasize the high level of training and experience required to analyze a drawing according to psychodynamic principles, Levy outlined these prerequisites for the drawing interpreter: courses in the pluralism of William James, understanding of the linguistic symbols of James Joyce, knowledge of Freud's system of dream interpretation, acquaintance with the symbolism of Stekel and Jung, and "a personal psychoanalysis" (p. 85). Levy's attitude was a contrast to the facile enthusiasm of Machover who exclaimed that drawings conveyed the concerns and feelings of children "with striking literalness" (p. 35).

Koppitz (1968) devised a scoring scale for human figure drawings which consisted of 30 developmental items and 30 emotional indicators. Many of the developmental items were similar to those found in the Goodenough and Goodenough-Harris scoring systems, and many of the emotional indicators similar to those of Machover. Emotional indicators were divided into three categories, quality signs, special features, and omissions. The number of indicators in a drawing had diagnostic significance. Unlike previous systems, this method was normed on a population of normal schoolchildren and a reliability study was conducted. Koppitz rejected Machover's body-image psychoanalytical approach in favor of Sullivan's
Interpersonal Relationship Theory. Drawings reflected the "inner child of the moment" (p. 4). Handler (Sims, Dana, and Bolton, 1983), who devised a scoring system to measure anxiety, also rejected the psychodynamic theory and adopted the position that drawings measure state rather than trait. Nevertheless, his 21 anxiety indicators were similar to items found in Machover, Hammer, Levy, and Koppitz.

Reviews of the many attempts to establish the validity of projective interpretation of human figure drawings were published by Swensen (in Goodstein and Lanyon, 1971), Roback (1968), and Hammer (in Rabin, 1981). Hammer's conclusion reflects the past and the present state of projective interpretation: "The research studies in the field of projective drawings are, by and large, so contradictory that the writer finds himself taking a deep breath as he settles down to try to make sense of the mosaic" (p. 151).

These reviews cite flaws such as uncontrolled variables which could influence drawing performance, lack of longitudinal studies, lack of control groups, and failure to establish the reliability of the personality measures and the validity of the constructs.

Many validity studies have been based on the degree of correlation between human figure drawings and scores earned on rating scales such as the Maudsley Personality Inventory, the Manifest Anxiety Scale, the Test Anxiety Scale for Children, the General Anxiety Scale for Children, the IPAT Anxiety Scale, the Rorschach Concept Test (Sims et al. 1983), Coopersmith Self-Esteem Inventory (Prytula and Thompson, 1973), and Rosenberg Self-Esteem Scale (Delatte and Hendrickson, 1982). Sims et al. found insignificant correlations between human figure drawings and the self-report measures of anxiety. Similarly, Prytula and Thompson found no significant
differences in self drawings by children low or high in self-esteem as measured by the Coopersmith scale. Delatte and Hendrickson found limited support for a few Machover items related to self-esteem for boys but no relationship for girls. These studies indicate that the constructs of anxiety and self-esteem as defined by analysis of drawings differed from the constructs measured by the rating scales. A problem with many of these studies may be related to Coopersmith's claim that self-concept is too immature in pre-adolescent or early adolescent children to be measured adequately (Delatte and Hendrickson, 1982).

The insignificant correlations between measures purporting to examine the same construct was further confused by failure to control for the effect of drawing quality and/or intelligence (Sims et al., 1983). For example, two signs of anxiety, shading and erasure, occurred most often in the drawings which earned the highest scores according to the Goodenough or Goodenough-Harris systems or were rated as the best drawings.

The validity of other items from Machover's catalogue of pathological indicators has also been poorly supported in studies. Shading, shifting to the left, hands in pockets or behind the back, large head size, and symbolic items such as guns, clubs, knives, teeth, and prominent shoulders and chin were considered indicative of juvenile delinquency by Machover, Hammer, and Koppitz. However, Montague and Prytula (1975) found that only transparencies distinguished between a delinquent group and a non-delinquent group. Daum (1983) classified delinquents according to their offense and formed three groups; withdrawn (truants and runaways), aggressive (two or more hostile crimes), controls (other offenses) and non-delinquents. Only 2 of 12
signs of aggression were significant while 3 signs related to within.

Another Machover sign, drawing a figure close to the edge of the paper, was said to indicate anxiety and insecurity. When Holmes and Stephens (1984) asked 76 college students to complete three tests, Memory for Designs, Bender Visual Motor Gestalt Test, and the DAP, 28 students edged on at least one of the tasks. The authors concluded that edging occurred too frequently to be considered indicative of pathology and its inconsistent use by the same person across the three tests lacked the stability of a personality trait.

One of the most frequent findings in the projective literature is that omission of features or details is related to pathology (Daum, 1983; Goldman and Warren, 1976; Goldman and Velasco, 1980; Swensen, 1971; Koppitz, 1968; and Eno, Elliott, and Woehlke, 1981). Since omissions would impact quantitative scores, it would seem likely that correlations between IQ and emotional or personality indicators would have been obtained routinely. This, however, has not been the case.

One study (Pohl and Nimrod, 1976) found a significant correlation (r = .27) between Goodenough-Harris IQ and Koppitz Emotional Indicators. Another evaluated drawings according to nine variables for groups which were divided according to intelligence and mental health (Maloney and Glasser, 1982). Two of nine selected variables failed to distinguish between any of the groups; seven distinguished between the retarded and psychiatric nonpsychotics, and six between the retarded and normals. None, however, separated the mentally retarded and the psychiatric psychotic, all of whom had IQs scores above 90 from a previous evaluation. Omissions, distortions,
and simplifications of head and body parts were among the variables studied. Similarly, Koppitz (1968) found that omissions were the only Emotional Items which distinguished between kindergarten students rated as high or low in achievement according to teacher rating and standardized testing.

Many of the variables which have shown the greatest reliability and validity as emotional or personality indicators are similar to items included in the quantitative scoring systems designed to measure the construct labeled intelligence by Goodenough, conceptual maturity by Harris, and development by Koppitz. In Maloney and Glasser's study, omissions, distortions, and simplifications would all be reflected in quantitative scoring systems. Koppitz' list of emotional indicators contains numerous items which are virtual duplicates of items on her developmental list, a list which was "derived from the Goodenough-Harris scoring system and from the writer's own experience" (1968, p. 9). Fourteen of the emotional items or almost 50% would affect a score obtained from the DAP-Q of Naglieri. The extensive overlap of items considered diagnostic of psychopathology with quantitative cognitive systems suggests that an interaction between intelligence and emotional/personality exists. This is an area which has been overlooked in studies of the Goodenough and Goodenough-Harris system and should be studied for the newest version of the DAP.

The DAP and Kindergarten Children

The confused picture of the DAP's relationship to the four variables under consideration in this study also extends to the few studies which include drawings of kindergarten age children and one or more of these variables. In these studies, the DAP was not the primary variable under
investigation and kindergarteners were only one of the groups studied. Nevertheless, some information can be inferred about the meaning of DAP scores for this age group. Serwer, Shapiro, and Shapiro (1972) found no relationship between the DAP and the Metropolitan Achievement Test in a study of 46 kindergarten children exhibiting learning problems. Parts of the Metropolitan Achievement Test tap understanding of concepts similar to those in the Bracken Basic Concepts Scale. However, other intelligence, readiness, and perceptual-motor tasks fared poorly also. The best predictor of achievement in this study was a teacher rating scale. In contrast, Stevenson et al. (1976) found teacher ratings inferior to test results in predicting school achievement. Their battery of tests included measures of cognitive processes, psychometric tasks such as the DAP, PPVT, WRAT and Bender and a five point scale which rated children in 13 areas. In this study, the DAP was a modest predictor of later reading ability but not arithmetic. In addition, the educational level of the father interacted significantly with the boy's scores on the DAP.

Dudek, Goldberg, Lester, and Harris (1969) tested 103 kindergarten students using an individual and group intelligence measure, the Goodenough-Harris DAP, Rutgers Drawing Test and the Lincoln Oseretzsky Motor Scale. A child psychiatrist evaluated the child after a home interview. These results were then compared with California Achievement Test results at the end of first and second grades. The DAM and DAW were modest but significant predictors of achievement. Surprisingly, one of the best three predictors was the Motor Scale. Personality deficits interacted significantly with the DAP scores. This study suggested that motor skills and personality were likely related to the DAP.
This review has shown that past studies of relationships between conceptual knowledge, cognitive processing, motor development, personality (as exemplified through behavior) and the DAP have led to confusing and frequently contradictory conclusions. Further study using the newest version of the drawing task is needed in order to establish guidelines for the valid interpretation of children's human figure drawing scores.
CHAPTER III
METHOD

Subjects

The sample consists of 103 students, 51 female and 52 male, who attended regular kindergarten classes located in two Central Ohio school districts. Minorities comprise 13.6% of the sample with 10.7% black and 2.9% Asian. Although individual SES information is not available, the diverseness of the sample can be inferred from public information. One school is located in a suburb where elementary students earn scores above national norms on the Metropolitan Achievement Test, slightly more than 40% of the high school graduates go on to college or post-high school training and federal funding accounts for less than 0.1% of the school budget. The other school population consists of a large number of disadvantaged students as evidenced by government funding for a free breakfast program and a full-time aide for each kindergarten teacher. Although the total sample represents all levels of parent educational and employment levels, it is skewed toward the lower end of the SES scale.

Table 1 shows there are no significant differences between males and females in terms of age ($F = 1.37, p > .10$) or MAT-EF ($F = .817, p > .10$) scores. The average MAT score of 103 for the Total Sample suggests that the
sample is of average intellectual ability. This score should be considered somewhat inflated because low raw scores on the MAT-EF earn relatively high standard scores in the younger age ranges.

Table 1
Description of Sample

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>52</td>
<td>51</td>
<td>103</td>
</tr>
<tr>
<td>X Age (mos.)</td>
<td>75.19</td>
<td>74.04</td>
<td>74.62</td>
</tr>
<tr>
<td>S.D. Age</td>
<td>5.30</td>
<td>4.67</td>
<td>5.01</td>
</tr>
<tr>
<td>MAT-EF Mean</td>
<td>104.71</td>
<td>102.49</td>
<td>103.61</td>
</tr>
<tr>
<td>MAT-EF S.D.</td>
<td>12.50</td>
<td>12.44</td>
<td>12.46</td>
</tr>
</tbody>
</table>

Procedures

Each student was individually administered all experimental tasks by trained examiners. All testing was conducted in the students' school by advanced graduate students in the school psychology program. One testing session of approximately 90 minutes was required for each subject. The teachers of the students completed a behavior rating scale for each student.

The experimental battery consisted of three human figure drawings, man, woman, and self, scored according to Draw A Person: A Quantitative
Scoring System (Naglieri, 1988), the Bracken Basic Concept Scale (Bracken, 1984), the motor scale from the McCarthy Scale of Children's Abilities (McCarthy, 1972), nine PASS cognitive processing tasks (Das and Naglieri, 1989), and the Teacher Version of the Child Behavior Checklist (Achenbach and Edelbrock, 1986). The tests were administered in counter-balanced order. Four complete sets of the 24 possible test sequences were assigned to students in alphabetical order. Seven additional sequences were needed to reach the total of 103. Students were selected for testing in random order.

All subtests were administered in standard order as described in the manuals. For the experimental PASS tasks, the order was: Visual Search, Planned Connections, Figure Memory, Matrix Analogies, Successive Word Recall, Sentence Repetition, Sentence Questions, Expressive Attention-Words (1, 2, and 3), and Number Finding (1 and 2). The Child Behavior Checklist was filled out by the teacher of each subject at the teacher's convenience with no order specified.

Experimental Tasks

Draw A Person

For each drawing, the subject was provided with an 8 1/2" × 11" sheet of plain white paper and a number two lead pencil with an eraser. The child was first asked to make the very best picture of a man, then of a woman, and finally of themselves. There was a five minute time limit per drawing. Scoring is based on fourteen categories (for example, fingers) which yield points for
presence, detail(s), and proportion. A one point bonus is earned if all previous points in the category are scored. There are 64 scoring points per drawing. Standard scores with a mean of 100 and standard deviation of 15 are obtained for each drawing and for the total score of all 3 drawings. One set of scoring criteria and standardized scores is used for the three drawings. A separate norms table is used to obtain a standard score from the total raw score of all three drawings. This total test score is the most reliable and useful score and is used for all statistical analyses.

The manual for Draw A Person: A Quantitative Scoring System reports standardization data as well as estimates of the task’s validity and reliability. Two studies obtain correlations between the Naglieri scoring system and the 1962 Goodenough-Harris which range from .75 to .87 (Naglieri, 1988). These results suggest the tests are measuring similar construct(s). Therefore, it is likely that users of the test would make inferences about scores which would correspond to those used for Goodenough-Harris scores. Construct validity of the DAP is also established through examination of the increase in scoring points earned as the subjects’ age increases. There is a significant correlation between mean raw score points and age from ages 5 to 11. In addition, factor analysis yields one factor, a result which provides further support for the use of the test as a nonverbal estimate of general ability (Naglieri, 1988). Criterion validity is examined in studies showing the relationship of the DAP to the Differential Ability Scales (DAS) (Lillis, 1987) and to the Matrix Analogies Test-Short Form (MAT-SF) and the Multilevel Academic Survey Test (MAST) (Naglieri, 1988). Significant correlations are found with all these measures. Item analysis using Cronbach’s alpha (r = .86), test-retest (r = .73), inter-rater
(r = .95), and intra-rater (r = .97) comparisons establish the excellent overall reliability of the scoring system.

**Bracken Basic Concept Scale**

Each subject was individually administered the *Bracken Basic Concept Scale*. Each item of the test consists of a set of four pictures, one of which corresponds to a word spoken by the examiner. The subject was asked to point to the picture which matched the word.

Standard scores are obtained for each of seven categories; School Readiness Composite, Direction/Position, Social/Emotional, Size, Texture/Material, Quantity, and Time/Sequence as well as for the Total Test. Category standard scores have a mean of 10 and a standard deviation of 3. The Total test standard score has a mean of 100 and standard deviation of 15. The Total test standard score is the most reliable score and was used in all statistical analyses.

The *Bracken Basic Concept Scale* is designed to measure the subject's knowledge of basic concepts in 11 areas such as color, direction, size, relationships, or shapes. Bracken adopts Kagan's position that concepts are the "fundamental units of intelligence" (Bracken, 1984, p. 6) and constitute a child's "most rudimentary functional vocabulary" (p. 6). A child must understand the meaning which his surrounding culture places on a word. The concepts which are utilized in the test appear in 13 preschool and primary measures of cognition and achievement and in popular curriculum materials. The test is standardized in accordance with the 1980 census data. The total test score has an internal consistency at age 5 of .94 and at age 6 of .96. Three
validity studies with Form A of the *Boethm Test of Basic Concepts* (1971) result in correlations of .88, .86, and .78. Similarly, three studies with the *Peabody Picture Vocabulary Test-Revised* (Dunn & Dunn, 1981) result in correlations of .74, .88, and .84. In addition, the BBCS is able to discriminate between a group of deaf children, a population known to be delayed in concept formation, and a group of normal children. The deaf children earn scores 2 standard deviations lower than the normals. The BBCS appears to be a psychometrically sound instrument which measures many of the basic concepts related to a child's ability to construct a drawing of a human figure.

**Motor Scale, McCarthy Scales of Children's Abilities**

Each student was individually tested on the five subtests of the McCarthy Motor Scale. The motor scale is comprised of two sets of tasks, one designed to measure gross motor skills and the other, fine motor skills (McCarthy, 1972, Kaufman & Kaufman, 1977). Points are earned on each subtest either through pass/fail (1, 0) or through finer evaluation of the quality of the performance of the task on a 0 to 3 scale. For example, a child's score when asked to bounce a small ball using the open palm is determined by ranges of successful bounces (3 - 5 bounces = 3). In contrast, a child either successfully imitates an action of the examiner (1) or does not (0). The MSCA manual (pp. 36-39) reports correlations between the General Cognitive Index (intelligence estimate) and the motor scale at ten ages. In the population which will be examined in this study the uncorrected correlations are .67 (5 1/2), .68 (6 1/2), and .37 (7 1/2) respectively. Uncorrected correlations include Draw-A-Design and Draw-A-Child, subtests which are also used on
the Perceptual-Performance Scale. The motor scale is included by McCarthy as part of the MSCA because of her belief that motor behaviors and cognition are closely related among younger children. The McCarthy Motor Scale is an easily administered, reasonably reliable test which is attractive to children.

Five scores derived from the McCarthy Motor scale are used in this study. The first score is the Motor Scale T score (Mean = 50, SD = 10) provided in the MSCA manual. This score is derived from the total raw score of the five subtests included in the Motor Scale. In order to remove the effect of the Draw-A-Child subtest, a second score, the Motor Factor T score is created. This is done by converting raw scores for each motor subtest to a T score using data provided by Naglieri and Harrison (1979). They derive the T scores using raw score data from the standardization of the MSCA. T scores for four (Draw-A-Child was excluded) of the five motor subtests are averaged to obtain a Motor Factor T score. The third and fourth scores are based on McCarthy's (1972) division of the scale into Fine Motor and Gross Motor tasks. She considers three subtests, Leg Coordination, Arm Coordination, and Imitative Action, measures of gross motor ability. T scores of these tasks are averaged to produce a Gross Motor score. McCarthy considers two subtests, Draw-A-Design and Draw-A-Child, measures of fine motor ability. T scores of these tasks are averaged to produce a Fine Motor score. In order to remove the influence of Draw-A-Child from the Fine Motor score, the Draw-A-Design T score is also analyzed separately. In summary, the five scores based on McCarthy Motor scale subtests are the McCarthy Motor Scale score (five subtests), McCarthy Motor Factor score (four subtests), Gross Motor score (three subtests), Fine Motor score (two subtests), and Draw-A-Design.
Cognitive Assessment System: Experimental Version

Planning Tasks

Visual Search. In this timed task, the individual was asked to find and point to a picture, number, or letter exactly the same as the picture, number, or letter located in a box in the center of the search field. A separate search was placed in the top and bottom halves of each 8-1/2 x 11 inch page. Timing began immediately upon presentation of the search and continued until both searches were completed. A sample item was used for training at the beginning of the task and each time (3) the difficulty of the task increased. The total score is the sum of the times for the sixteen search pages. This task is based on work by Teuber, Battersby, and Bender (1949). Later studies by Ashman and Das (1980), Naglieri and Das (1988), and Naglieri, Prewett, and Bardos (in press) find the task loads on a planning factor.

Planned Connections. This task required the subject to connect a series of numbers in proper sequence. Two items, one demonstration and one practice, were completed by the subject to assure understanding of the task. Six items were presented to each subject. The score is the time needed to complete all six items. This task is based on the Trail Making task which is part of the Army Individual Test of General Ability (1944) and later used in studies by Reitan (1955) and Spreen and Gaddes (1969). Recent studies by Ashman and Das (1980) and Naglieri and Das (1988) and Naglieri, Prewett, and Bardos (in press) find this task loads on a factor defined as planning.
Simultaneous Processing

Figure Memory In this task, a geometric figure was shown to the subject for five seconds. After the stimulus was removed, the subject was immediately presented with a complex drawing which contained the previously displayed figure. The subject was asked to trace the original design without making any additions or omissions. This task is used by Das, Kirby, and Jarman (1979), Naglieri and Das (1987), and Naglieri, Prewett, and Bardos (in press) as a measure of simultaneous processing. Again demonstration and sample items preceded administration of the scored items.

Matrix Analogies-Expanded Form The subject was required to choose, by pointing to one of six options, the design which best completed a visual analogy presented in a 4 or 9 box matrix. Four groups of matrices were presented and the score is based on the number of analogies correctly completed within a specified time (a maximum of 12 minutes per group). This test was developed by Naglieri (1985) and is shown in studies by Naglieri and Das (1987), Naglieri, Prewett, and Bardos (in press) to load on a factor defined as simultaneous processing.

Successive Processing

Word Recall The subject was asked to repeat in correct order a series of orally presented single syllable words. The simplest series contained 2 words and the longest series contained 9 words. The score is the total number of items correctly repeated by the child. Studies by Das et al. (1979), Naglieri and Das (1987), Naglieri, Prewett, and Bardos (in press) find that this task loads on a factor defined as successive processing. Again, training items were administered.
**Sentence Repetition and Questions.** In part 1, the subject was asked to repeat sentences in which expected nouns and verbs were replaced by color words. For example, a subject might be asked to repeat the sentence "The red is yellow." The score is the number of sentences correctly repeated. In part 2, the subject was required to listen to a sentence in which color words replaced expected nouns and verbs and then answer a question about the sentence. For example, a subject who was asked "Who is greening?" in response to the statement, "The yellow was greening the blue." would reply, "The yellow." The score is the number of questions answered correctly. Studies by Naglieri and Das (1988) and Naglieri, Prewett, and Bardos (in press) find this task loads on the factor defined as successive processing. Sample items were administered and corrections made on the first item if the subject failed it. No further help was given.

**Attention**

**Expressive Attention- words.** Following administration of training items to insure that the subject knew both the name and size (large or small) of a series of pictured animals, the subject was presented with the first of 3 tasks. In the first task, he was asked to tell the size of each of a series of animals pictured in rows. Each picture is similar in size regardless of the relative size of the animals. Next, the subject was asked to tell the size of the animals when the pictures corresponded to their relative size. In other words, a small animal's picture is smaller than that of a large animal. Finally, the subject was asked to tell the size of each animal when the pictures were the opposite size of the animal's actual size. In other words, a bear appears in a small picture while a rabbit appears in a large picture. A few pictures matching the animal's
relative size are interspersed among the discrepant pictures. This procedure is used so that the subject could not use a planning process instead of an attention process. The subject was asked to perform each task as quickly as possible. For each of the three tasks a score is recorded for both time and number correct. This version of the task was developed by Das and Naglieri (1989) based on the Stroop test (Golden, 1978). In a study by Naglieri, Prewett, and Bardos (in press) the task is found to load on a factor defined as attention.

**Number Finding.** In this task, the subject was presented with a page which contained three key numbers at the top. Below this set of key numbers were 18 rows of numbers. The subject located and underlined all numbers which matched any member of the stimulus group as quickly as possible. Since no subject was able to complete either of the tasks in the allotted time, the number correct and the number of errors form the basis for the score.

**Teacher Version of the Child Behavior Checklist**

The *Teacher Version of the Child Behavior Checklist* (1983) consists of a series of questions which can be answered by a teacher, aide, or other school employee familiar with the behavior of the child. Although an Adaptive Behavior scale is also available as part of the checklist, this study uses only the Behavior Problems scale. The Behavior Problems scale consists of 118 items with a three-step response, not true (0), sometimes true (1) and often true (2). Most of the items duplicate those found on the parent version of the checklist (Achenbach and Edelbrock, 1983). Substitutions were made for nine items which were deemed inappropriate for teacher response and slight
alterations were made in four others. The original set of items for the parent checklist was derived from review of past literature, responses of child behavior specialists, and interviews with parents. The average completion time for teachers is estimated by the authors at ten minutes.

Each checklist yields eight scale scores which combine to form an Internalizing, Externalizing, and Total Score. All raw scores are converted to normalized T scores with a mean of 50 and standard deviation of 10. Significantly high scores are indicative of problems.

The items on the checklist are grouped into nine separate scales (syndromes) derived from factor analysis of the Teacher Report Forms (TRFs) of 1700 children referred to school or mental health specialists for behavioral or social-emotional problems. However, the checklist is standardized on a group of 1100 normal children comprised of 50 boys and 50 girls from grades 1 - 10.

There are separate profiles for boys and girls in two age groups, 6 - 11 and 12 - 16. Of the nine scales, seven are the same for boys and girls. However, the boy's checklist includes an Obsessive-Compulsive scale while the girl's scale includes a Depressed scale. Some items appear on more than one scale. The authors compare this duplication to the appearance of a fever in numerous diseases.

Test-retest reliability is .84 for an interval of 15 days and .90 for an interval of 7 days. Stability over a two month period is .74 and over a four month period, .68. The authors suggest that the checklist is extremely useful for charting of changes in student behavior. Comparisons of the CBCL-TRF and other popular rating scales are of the same magnitude as correlations
between major tests of intelligence. For example, correlations of .89 and .85, respectively, are obtained between Conner’s Revised Teacher Rating Scale and the Externalizing factor of the CBCL-TRF and the Total score of the CBCL-TRF. According to the manual, all demographic effects on ratings are small but students of lower SES earn higher scores than students in middle and above SES groupings. The authors state that the checklist could be used for research purposes with children younger than or in grades different from those in the standardization sample. The children in this study attained a mean age of 6 years 3 months and were in the last six weeks of their kindergarten year. It is felt that the subjects in the sample are sufficiently close in age and in school experience to those aged six and in first grade who are included in the CBCL standardization group to warrant the use of the scale’s standardized T scores.

**Research Questions and Data Analysis**

All data is analyzed by a Macintosh computer using *Statview 512+* (1986, Brainpower, Inc.). Means and standard deviations are derived for all tests. Raw scores are converted to standard scores with a mean of 100 and standard deviation of 15 for all PASS tasks. A significance level of p < .01 level is used in all analyses of variance, multiple regression, and correlation in order to maintain an appropriate error rate for a repeated measures design.
The following hypotheses are examined.

1. There will be a significant Pearson Product Moment correlation between the a) McCarthy Motor Scale T Score and DAP Total Test Score, b) McCarthy Gross Motor T Score and DAP Total Test Score, c) McCarthy Fine Motor T Score and DAP Total Test Score, d) McCarthy Motor Factor T score (All subtests except Draw A Child) and DAP Total Test Score, and e) McCarthy Draw-A-Design T Score and DAP Total Test Score.

2. There will be a significant Pearson Product Moment correlation between the a) Child Behavior Checklist-Teacher Report Form (CBCL-TRF) Externalizing T score and DAP Total Test Score, b) CBCL-TRF Internalizing T score and DAP Total Test Score, and c) CBCL-TRF Total Behavior Problems T score and DAP Total Test Score.

3. There will be a significant Pearson Product Moment correlation between the a) Bracken Basic Concepts Scale Total Score and the b) DAP Total Test Score.

4. The PASS tasks will conform to the Das-Luria theoretical model as shown by significant loadings on predicted factors when principal axes factor analyses are employed. Three and four factor solutions will be explored.

5. There will be a significant Pearson Product Moment correlation between the Planning Scale Score and DAP Total Test Score. The Planning Scale score is derived from the average standard scores
of Visual Search and Planned Connections tasks. These tasks represent the Planning component of the PASS model.

6. There will be a significant Pearson Product Moment correlation between the Attention Scale score and DAP Total Test Score. The Attention Scale score is derived from the average standard scores of Expressive Attention and Number Finding tasks. These tasks represent the Attention component of the PASS model.

7. There will be a significant Pearson Product Moment correlation between the Simultaneous Scale Score and DAP Total Test Score. The Simultaneous Scale score is derived from the average standard scores of the Figure Memory and Matrix Analogies tasks. These tasks represent the Simultaneous Processing component of the PASS model.

8. There will be a significant Pearson Product Moment correlation between the Successive Scale Score and DAP Total Test Score. The Successive Scale is derived from the Word Recall and Sentence Repetition and Questions tasks. These tasks represent the Successive Processing component of the PASS model.

9. The group of variables, motor skills as represented by the McCarthy Motor Scale T score, behavior problems as represented by the Child Behavior Checklist Total T score, basic concept attainment as represented by the Bracken Basic Concepts Total score, and cognitive processes as represented by the PASS Total score, will account for a significant amount of variation in DAP TT scores when multiple regression analysis is used. Stepwise multiple regression
analysis will be used to determine which of the variables in this
group are significant predictors of DAP TT scores.

10. The group of variables, motor skills as represented by the
McCarthy Motor Factor $T$ score, behavior problems as represented
by the Child Behavior Checklist Total $T$ score, basic concept
attainment as represented by the Bracken Basic Concepts Total
score, and cognitive processes as represented by the PASS Total
score, will account for a significant amount of variation in DAP TT
scores when multiple regression analysis is used. Stepwise multiple
regression analysis will be used to determine which of the variables
in this group are significant predictors of DAP TT scores.

11. The group of variables, motor skills as represented by the
McCarthy Draw-A-Design $T$ score and McCarthy Gross Motor $T$
score, behavior problems as represented by the Child Behavior
Checklist Total $T$ score, basic concept attainment as represented
by the Bracken Basic Concepts Total score, and cognitive pro-
cesses as represented by the PASS Total score, will account for a
significant amount of variation in DAP TT scores when multiple
regression analysis is used. Stepwise multiple regression analysis
will be used to determine which of the variables in this group are
significant predictors of DAP TT scores.

12. Scores of the components of the PASS model, the Planning stan-
dard score, Attention standard score, Simultaneous Processing
standard score, and Successive Processing standard score, will
account for a significant amount of variation in DAP TT scores when
multiple regression analysis is used. Stepwise multiple regression analysis will be used to determine which of the variables in this group are significant predictors of DAP TT scores.

13. The variables, behavior problems as represented by the Child Behavior Checklist Total T score, basic concept attainment as represented by the Bracken Basic Concepts Total score, and cognitive processes represented by the PASS Total score, will account for a significant amount of variation in DAP TT scores when multiple regression analysis is used. Stepwise multiple regression analysis will be used to determine which of the variables in this group are significant predictors of DAP TT scores.

14. There will be significant differences between scores earned on the a) McCarthy Motor Scale, b) McCarthy Motor Factor, c) CBCL-TRF Total Behavior Problems, d) Bracken Basic Concepts Scale Total, and e) PASS Total when subjects are grouped according to sex. A mixed design (Kennedy, 1978) or split-plot (Keppel, 1982) analysis of variance will be used to determine if significant differences are present. A one factor analysis of variance will be used to determine which differences are significant.

15. There will be significant differences between scores earned on the a) Attention, b) Successive, c) Simultaneous, and d) Planning Scales when subjects are grouped according to sex. A mixed design (Kennedy, 1978) or split-plot (Keppel, 1982) analysis of variance will be used to determine if significant differences are
present. A one factor analysis of variance will be used to determine which differences are significant.

16. There will be significant differences between scores earned on the a) McCarthy Gross Motor, b) McCarthy Fine Motor, c) McCarthy Fine Motor Factor, d) CBCL Internalizing, and e) CBCL Externalizing scores when subjects are grouped according to sex. A mixed design (Kennedy, 1978) or split-plot (Keppel, 1982) analysis of variance will be used to determine if significant differences are present. A one factor analysis of variance will be used to determine which differences are significant.
CHAPTER IV
RESULTS

This chapter begins with a summary of statistical transformations of raw score and standardized score data. Results of four types of data analysis are then presented. The data analyses are Pearson Product Moment correlations and Pearson Product Moment correlations with age partialled out of all variables, principal axes factor analysis of the eight tasks representing the Das–Luria cognitive processing model, multiple and stepwise multiple regression analysis of six combinations of variables on the DAP TT, and repeated measures ANOVAs of the influence of gender on subsets of the variables.

PASS Tasks

The means and standard deviations for the raw scores of the PASS tasks are shown in Table 2. Examination of these raw scores shows that there is sufficient variability of item difficulty so that neither ceiling nor floor effects influence the scores of most of the tasks. For example, Figure Memory has a mean score of 6.39 and a standard deviation of 2.13. This means that three standard deviations are possible in either direction. Three tasks, however, Matrix Analogies, Sentence Repetition and Sentence Questions show a
Table 2
Raw Score Means and Standard Deviations of PASS Variables

<table>
<thead>
<tr>
<th>PASS TASKS</th>
<th>Males (52)</th>
<th>Females (51)</th>
<th>Total (103)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual</td>
<td>Mean 285.15</td>
<td>266.80</td>
<td>276.07</td>
</tr>
<tr>
<td>Search</td>
<td>S.D. 122.24</td>
<td>120.58</td>
<td>121.18</td>
</tr>
<tr>
<td>Planned Connection</td>
<td>Mean 184.67</td>
<td>186.08</td>
<td>185.37</td>
</tr>
<tr>
<td>Figure</td>
<td>S.D. 61.07</td>
<td>69.38</td>
<td>65.00</td>
</tr>
<tr>
<td>Memory</td>
<td>Mean 6.65</td>
<td>6.12</td>
<td>6.39</td>
</tr>
<tr>
<td>Matrix</td>
<td>S.D. 2.35</td>
<td>1.86</td>
<td>2.13</td>
</tr>
<tr>
<td>Analogies</td>
<td>Mean 12.25</td>
<td>9.62</td>
<td>10.95</td>
</tr>
<tr>
<td>Word</td>
<td>S.D. 7.64</td>
<td>6.47</td>
<td>7.17</td>
</tr>
<tr>
<td>Recall</td>
<td>Mean 6.94</td>
<td>6.90</td>
<td>6.92</td>
</tr>
<tr>
<td>Repetition</td>
<td>S.D. 2.51</td>
<td>2.35</td>
<td>2.42</td>
</tr>
<tr>
<td>Sentence</td>
<td>Mean 2.35</td>
<td>2.37</td>
<td>2.36</td>
</tr>
<tr>
<td>Sentence</td>
<td>S.D. 1.75</td>
<td>1.71</td>
<td>1.72</td>
</tr>
<tr>
<td>Question</td>
<td>Mean 2.23</td>
<td>1.80</td>
<td>2.02</td>
</tr>
<tr>
<td>Question</td>
<td>S.D. 2.24</td>
<td>1.63</td>
<td>1.96</td>
</tr>
<tr>
<td>Expressive Time</td>
<td>Mean 70.52</td>
<td>75.16</td>
<td>72.81</td>
</tr>
<tr>
<td>Attention Time</td>
<td>S.D. 20.85</td>
<td>25.42</td>
<td>23.22</td>
</tr>
<tr>
<td>Expressive Error</td>
<td>Mean 3.37</td>
<td>2.55</td>
<td>2.96</td>
</tr>
<tr>
<td>Attention Error</td>
<td>S.D. 6.62</td>
<td>5.19</td>
<td>5.94</td>
</tr>
<tr>
<td>Number Finding 1</td>
<td>Mean 19.27</td>
<td>19.47</td>
<td>19.37</td>
</tr>
<tr>
<td>(No. Corr - Errors)</td>
<td>S.D. 7.22</td>
<td>6.04</td>
<td>6.63</td>
</tr>
<tr>
<td>Number Finding 2</td>
<td>Mean 15.37</td>
<td>16.76</td>
<td>16.06</td>
</tr>
<tr>
<td>(No. Corr - Errors)</td>
<td>S.D. 6.31</td>
<td>6.21</td>
<td>6.27</td>
</tr>
</tbody>
</table>
minor floor effect in that the maximum the scores could vary below the mean is about 1.5 standard deviations. No tasks show a ceiling effect.

For statistical analyses, the raw scores of each PASS task are transformed into standard scores with a mean of 100 and standard deviation of 15. Although nationally standardized scores are available for the Matrix Analogies task, the sample's means and standard deviations are used to compute a standard score.

Several transformations of the data are made to obtain the eight task scores which comprise the PASS model. First, the standard scores for Sentence Repetition and Sentence Questions are averaged to form one score, Sentence Repetition & Questions. Second, the Expressive Attention score is obtained by standardizing and averaging the time score and error score for each subject. The highest scores are earned by subjects who complete the task swiftly but make few or no errors and the lowest scores by subjects who complete the task slowly and make many errors. Third, a raw score for each page of the Number Finding task is obtained by subtracting the number of errors from the number correct. These adjusted raw scores are then standardized. This procedure distinguishes between students who maintain attention consistently and those who do not. The two Number Finding standard scores are averaged to obtain one Number Finding score.

As a result of these transformations and combinations each component of the PASS model is comprised of two task scores with a mean of 100 and standard deviation of 15. Visual Search and Planned Connections represent the Planning process, Expressive-Attention and Number Finding the Attention process, Figure Memory and Matrix Analogies the Simultaneous process,
Table 3
Means and SDs of the Four Components of the PASS Model

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planning</td>
<td>100.00</td>
<td>13.66</td>
</tr>
<tr>
<td>Attention</td>
<td>99.99</td>
<td>11.54</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>99.99</td>
<td>12.61</td>
</tr>
<tr>
<td>Successive</td>
<td>100.00</td>
<td>12.75</td>
</tr>
<tr>
<td>PASS Total</td>
<td>99.99</td>
<td>9.29</td>
</tr>
</tbody>
</table>

and Word Recall and Sentence Repetition & Questions the Successive process. Table 3 shows the means and standard deviations which result when the two scores comprising each PASS component are averaged to form a standard score for Planning, Attention, Simultaneous, and Successive Processing. A PASS total score is derived by averaging the Planning, Attention, Simultaneous, and Successive Processing scores.

**Standard Scores**

Means and standard deviations of scores for tasks which have been nationally standardized are shown in Table 4. Four scores in addition to the overall McCarthy Motor Scale T score are obtained using T scores or combinations of T scores of Motor Scale subtests. The subtest T scores were
Table 4

Means and SDs for All Variables Which Have Been Nationally Standardized

<table>
<thead>
<tr>
<th>Tests</th>
<th>Mean</th>
<th>SD</th>
<th>Tests</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAP Total</td>
<td>92.43</td>
<td>15.35</td>
<td>McC Fine</td>
<td>51.64</td>
<td>8.26</td>
</tr>
<tr>
<td>Man</td>
<td>93.57</td>
<td>14.20</td>
<td>McC Mot Fact</td>
<td>46.57</td>
<td>7.71</td>
</tr>
<tr>
<td>Woman</td>
<td>93.83</td>
<td>13.99</td>
<td>McC. Mot.</td>
<td>46.56</td>
<td>9.86</td>
</tr>
<tr>
<td>Bracken</td>
<td>94.41</td>
<td>16.40</td>
<td>Achen. Ext</td>
<td>56.12</td>
<td>10.31</td>
</tr>
<tr>
<td>McC Gross</td>
<td>44.96</td>
<td>8.53</td>
<td>Achen. Tot</td>
<td>54.58</td>
<td>9.46</td>
</tr>
<tr>
<td>McC DAD</td>
<td>51.43</td>
<td>10.88</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: DAP and Bracken Scores are Standard Scores with a Mean of 100 and SD of 15. McCarthy and Achenbach Scores are T Scores with a mean of 50 and SD of 10.

derived by Naglieri and Harrison (1979) from the raw scores of the standardization sample of the McCarthy Scales of Children’s Abilities (MCSA)

McCarthy (1972) describes Leg Coordination, Arm Coordination, and Imitative Action as gross motor tasks. Therefore, the average of the T scores for these three subtests comprises the Gross Motor score. McCarthy
considers Draw-A-Child and Draw-A-Design fine motor tasks. $T$ scores derived from the raw scores of these two subtests are used to obtain a Fine Motor score.

In order to remove the influence of the Draw-A-Child subtest from the correlations between the McCarthy Motor scores and the DAP TT scores, two new motor scores are created. These are the Fine Motor Factor (Draw-A-Design) and the Motor Scale Factor (Leg Coordination, Arm Coordination, Imitative Action and Draw-A-Design). Raw scores for these subtests are changed into $T$ scores using the tables created by Naglieri and Harrison (1979). The $T$ scores for the four subtests comprising the Motor Factor are averaged to obtain an overall Motor Factor score.

Table 4 shows that the subjects in this sample earn a mean score somewhat lower than the standard mean score of 100 on the DAP TT (-0.5 SD approximately) and BBCS (-0.30 SD approximately). However, standard deviations for these tests, 15.35 and 16.40 respectively, are close to the expected standard deviation of 15. The mean score for the McCarthy Motor Scale is also somewhat lower (-0.34 SD approximately) than the standard mean score of 50. The standard deviation (9.86) for this score is similar to the expected standard deviation of 10. CBCL Externalizing and CBCL Total scores are somewhat higher (+0.6 and +0.46 SDs respectively) than the expected mean score of 50. Standard deviations, however, are close to the expected standard deviation of 10. The mean CBCL Internalizing score of 50.40 is close to the expected mean score of 50, but the standard deviation (8.12) is somewhat lower than expected (10). Higher scores on the CBCL scales indicate more problematic behaviors are noted.
Correlations of DAP TT with McCarthy Motor, BBCS, and CBCL Variables

Table 5 shows the Pearson Product Moment correlations between the DAP TT and 16 variables. Bentler and Lettieri (1976, in Bentler, Lettieri, & Austin) utilize the following formula to determine the minimum sample size for correlational and multiple regression analyses

\[ n = 60 + 10 \sqrt{x} \]

where \( x \) is the number of variables in the correlation matrix. When 16, the number of variables in the correlation matrix shown in Table 5, is substituted for \( x \), a sample size of 100 is obtained as the minimum sample size. The number of subjects, 103, who participated in this study meets this criteria. Because a number of variables show a significant correlation with chronological age, partial correlations with age held constant are determined. These are shown in Table 6. The partial coefficients are used to determine the significance of the correlations between variables.

Correlations between the DAP TT and five combinations of motor subtest scores are obtained. The McCarthy Motor Scale is comprised of five subtests: Leg Coordination, Arm Coordination, Imitative Action, Draw-A-Design and Draw-A-Child. A standardized \( T \) score was available from the MSCA manual for this scale. The McCarthy Motor Factor Score is the average \( T \) score of these subtests with Draw-A-Child removed. The subtest \( T \) scores of the Motor Scale were also used to create a Gross Motor \( T \) Score (Leg Coordination, Arm Coordination, and Imitative Action) and Fine Motor \( T \)
Scor... of the Draw-A-Child subtest from the Fine Motor score, The Draw-A-Design subtest T score was used as a separate Fine Motor measure.

The McCarthy Motor scale score (Leg Coordination, Arm Coordination, Imitative Action, Draw-A-Design, Draw-A-Child) correlated significantly \( (p < .01) \) with the DAP TT score as did the McCarthy Motor Factor scale score (Leg Coordination, Arm Coordination, Imitative Action, Draw-A-Design). The McCarthy Gross Motor scale score (Leg Coordination, Arm Coordination, Imitative Action) and the McCarthy Fine Motor Scale score (Draw-A-Design, Draw-A-Child) also correlate significantly \( (p < .01) \) with the DAP TT score. Finally, Draw-A-Design correlates significantly \( (p < .01) \) with the DAP TT score. These results support Hypotheses 1a, 1b, 1c, 1d, and 1e which state that each of five combinations of McCarthy Motor Scale tasks will show a significant correlation with the DAP TT.

The correlation between the DAP TT and the Child Behavior Checklist Internalizing Scale is non-significant \( (p > .05) \). The CBCL Externalizing Scale and CBCL Total scores correlates with the DAP TT at the \( p < .05 \) level. The experiment-wise error rate for the study is \( p < .01 \). This result suggests weak support for Hypotheses 2a and 2c and no support for Hypothesis 2b.

The Bracken Basic Concept Scale correlates significantly \( (p < .01) \) with the DAP TT. This supports Hypothesis 3. The significance of the correlations between the PASS components, Planning, Attention, Simultaneous Processing and Successive Processing will be discussed following presentation of the results of factor analyses of the PASS tasks.
Table 5 Continued

<table>
<thead>
<tr>
<th></th>
<th>Atten</th>
<th>Simul</th>
<th>Succ.</th>
<th>PASS</th>
<th>CBCL-I</th>
<th>CBCL-E</th>
<th>CBCL-T</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simul.</td>
<td>.412 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success.</td>
<td>.211*</td>
<td>.267 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASS</td>
<td>.769 **</td>
<td>.729 **</td>
<td>.606 **</td>
<td>.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL-Int</td>
<td>-.235*</td>
<td>-.147</td>
<td>-.095</td>
<td>-.249*</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL-Ext</td>
<td>-.391 **</td>
<td>-.035</td>
<td>-.124</td>
<td>-.309 **</td>
<td>.135</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL-Tot</td>
<td>-.382 **</td>
<td>-.014</td>
<td>-.151</td>
<td>-.304 **</td>
<td>.396 **</td>
<td>.894 **</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.034</td>
<td>.155</td>
<td>-.005</td>
<td>.001</td>
<td>.129</td>
<td>.062</td>
<td>.102</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: * = p < .05  ** = p < .01

Table 5
Intercorrelations of Variables

<table>
<thead>
<tr>
<th></th>
<th>DAP</th>
<th>Mot Tot</th>
<th>Mot Fact</th>
<th>Gross M</th>
<th>Fine Mo</th>
<th>Dr-A-D</th>
<th>Brack</th>
<th>Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAP-TT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mot Total</td>
<td>.561 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mot Fact</td>
<td>.480 **</td>
<td>.820 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Mot</td>
<td>.360 **</td>
<td>.698 **</td>
<td>.942 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine Mo</td>
<td>.750 **</td>
<td>.732 **</td>
<td>.618 **</td>
<td>.362 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr-A-Des</td>
<td>.513 **</td>
<td>.681 **</td>
<td>.617 **</td>
<td>.319 **</td>
<td>.898 **</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bracken</td>
<td>.424 **</td>
<td>.509 **</td>
<td>.521 **</td>
<td>.447 **</td>
<td>.468 **</td>
<td>.425 **</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>.341 **</td>
<td>.386 **</td>
<td>.471 **</td>
<td>.411 **</td>
<td>.417 **</td>
<td>.367 **</td>
<td>.487 **</td>
<td>1</td>
</tr>
<tr>
<td>Attention</td>
<td>.212*</td>
<td>.418 **</td>
<td>.414 **</td>
<td>.348 **</td>
<td>.404 **</td>
<td>.354 **</td>
<td>.492 **</td>
<td>.669 **</td>
</tr>
<tr>
<td>Simul.</td>
<td>.272 **</td>
<td>.270 **</td>
<td>.272 **</td>
<td>.148</td>
<td>.411 **</td>
<td>.423 **</td>
<td>.605 **</td>
<td>.462 **</td>
</tr>
<tr>
<td>Success.</td>
<td>.217*</td>
<td>.209*</td>
<td>.237*</td>
<td>.212*</td>
<td>.151</td>
<td>.174</td>
<td>.439 **</td>
<td>.290*</td>
</tr>
<tr>
<td>PASS</td>
<td>.358 **</td>
<td>.435 **</td>
<td>.476 **</td>
<td>.382 **</td>
<td>.470 **</td>
<td>.448 **</td>
<td>.688 **</td>
<td>.832 **</td>
</tr>
<tr>
<td>CBCL-In</td>
<td>-.092</td>
<td>-.163</td>
<td>-.190</td>
<td>-.190*</td>
<td>-.084</td>
<td>-.091</td>
<td>-.177</td>
<td>-.257 **</td>
</tr>
<tr>
<td>CBCL-Ext</td>
<td>-.22*</td>
<td>-.304 **</td>
<td>-.313 **</td>
<td>-.273 **</td>
<td>-.252 **</td>
<td>-.245*</td>
<td>-.150</td>
<td>-.362 **</td>
</tr>
<tr>
<td>CBCL-Tot</td>
<td>-.219*</td>
<td>-.290 **</td>
<td>-.316 **</td>
<td>-.297 **</td>
<td>-.216*</td>
<td>-.197*</td>
<td>-.140</td>
<td>-.350 **</td>
</tr>
<tr>
<td>Age</td>
<td>-.279**</td>
<td>-.227*</td>
<td>-.376**</td>
<td>-.345**</td>
<td>-.255*</td>
<td>-.253**</td>
<td>-.327**</td>
<td>-.107</td>
</tr>
</tbody>
</table>

Note: * = p < .05 ** = p < .01
Table 6

Intercorrelations of Variables with Age Partialled Out

<table>
<thead>
<tr>
<th></th>
<th>DAP</th>
<th>Mot Tot</th>
<th>Mot Fact</th>
<th>Gross M</th>
<th>Fine Mo</th>
<th>Dr-A-D</th>
<th>Brack</th>
<th>Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAP-TT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mot Total</td>
<td>.53 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mot Fact</td>
<td>.42 **</td>
<td>.81 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gross Mot</td>
<td>.29 **</td>
<td>.68 **</td>
<td>.93 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fine Mo</td>
<td>.73 **</td>
<td>.72 **</td>
<td>.58 **</td>
<td>.30 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dr-A-Des</td>
<td>.48 **</td>
<td>.66 **</td>
<td>.58 **</td>
<td>.26 **</td>
<td>.89 **</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bracken</td>
<td>.37 **</td>
<td>.47 **</td>
<td>.45 **</td>
<td>.38 **</td>
<td>.42 **</td>
<td>.37 **</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>.33 **</td>
<td>.37 **</td>
<td>.47 **</td>
<td>.40 **</td>
<td>.41 **</td>
<td>.35 **</td>
<td>.48 **</td>
<td>1</td>
</tr>
<tr>
<td>Attention</td>
<td>.21 *</td>
<td>.42 **</td>
<td>.43 **</td>
<td>.36 **</td>
<td>.41 **</td>
<td>.36 **</td>
<td>.51 **</td>
<td>.67 **</td>
</tr>
<tr>
<td>Simul.</td>
<td>.33 **</td>
<td>.32 **</td>
<td>.36 **</td>
<td>.22 *</td>
<td>.47 **</td>
<td>.48 **</td>
<td>.70 **</td>
<td>.49 **</td>
</tr>
<tr>
<td>Success</td>
<td>.22 *</td>
<td>.21 *</td>
<td>.25 *</td>
<td>.22 *</td>
<td>.15</td>
<td>.18</td>
<td>.46 **</td>
<td>.29 **</td>
</tr>
<tr>
<td>PASS</td>
<td>.37 **</td>
<td>.45 **</td>
<td>.51 **</td>
<td>.41 **</td>
<td>.49 **</td>
<td>.46 **</td>
<td>.73 **</td>
<td>.84 **</td>
</tr>
<tr>
<td>CBCL-Int</td>
<td>-.06</td>
<td>-.14</td>
<td>-.15</td>
<td>-.16</td>
<td>-.05</td>
<td>-.06</td>
<td>-.14</td>
<td>-.25 *</td>
</tr>
<tr>
<td>CBCL-ExI</td>
<td>-.21 *</td>
<td>-.30 **</td>
<td>-.31 **</td>
<td>-.27 **</td>
<td>-.24 *</td>
<td>-.24</td>
<td>-.14</td>
<td>-.36 **</td>
</tr>
<tr>
<td>CBCL-Tot</td>
<td>-.20 *</td>
<td>-.28 **</td>
<td>-.30 **</td>
<td>-.28 **</td>
<td>-.20 *</td>
<td>-.18</td>
<td>-.11</td>
<td>-.34 **</td>
</tr>
</tbody>
</table>

Note:  * = p < .05  ** = p < .01
Table 6 Continued

<table>
<thead>
<tr>
<th></th>
<th>Attent.</th>
<th>Simul.</th>
<th>Succ.</th>
<th>PASS</th>
<th>CBCL-I</th>
<th>CBCL-E</th>
<th>CBCL-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simul.</td>
<td>.42**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td>.21*</td>
<td>.27 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASS</td>
<td>.77 **</td>
<td>.74 **</td>
<td>.61 **</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL-Int</td>
<td>-.23*</td>
<td>-.147</td>
<td>-.10</td>
<td>-.25*</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBCL-Ext</td>
<td>-.39 **</td>
<td>-.05</td>
<td>-.12</td>
<td>-.31 **</td>
<td>.13</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>CBCL-Tot</td>
<td>-.38 **</td>
<td>-.03</td>
<td>-.15</td>
<td>-.31 **</td>
<td>.39 **</td>
<td>.89 **</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: * = p < .05  ** = p < .01


The PASS Model

The intercorrelations of the 8 tasks based on the four factor PASS model are obtained and are presented in Table 7. None of the PASS tasks show a significant correlation with age. Therefore, no adjustment of the correlations is necessary. The correlations are further investigated through factor analysis in order to examine the consistency of the PASS model with the assignment of tasks to scales. These tasks have previously been shown

Table 7
Intercorrelations of PASS Tasks

<table>
<thead>
<tr>
<th></th>
<th>VS</th>
<th>PC</th>
<th>FM</th>
<th>L-MAT</th>
<th>WR</th>
<th>SR&amp;Q</th>
<th>ExA</th>
<th>NF</th>
</tr>
</thead>
<tbody>
<tr>
<td>VS</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>.659**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FM</td>
<td>.383**</td>
<td>.422**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAT</td>
<td>.290**</td>
<td>.319**</td>
<td>.413**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WR</td>
<td>.242*</td>
<td>.295**</td>
<td>.162</td>
<td>.120</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR&amp;Q</td>
<td>.152</td>
<td>.209*</td>
<td>.163</td>
<td>.318**</td>
<td>.446**</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ExA</td>
<td>.515**</td>
<td>.432**</td>
<td>.258**</td>
<td>.170</td>
<td>.129</td>
<td>.149</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>NF</td>
<td>.569**</td>
<td>.592**</td>
<td>.417**</td>
<td>.338**</td>
<td>.156</td>
<td>.187</td>
<td>.513**</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: Visual Search - VS, Planned Connections - PC, Figure Memory - FM, Matrix Analogies Test - MAT, Word Recall - WR, Sentence Repetition & Questions - SR&Q, Expressive Attention - ExA, Number Finding - NF

Note: * = p < .05  ** = p < .01
Factor Analysis of PASS Tasks

The PASS model is characterized by the independence and interdependence of the components of the model. Principal axes factor analyses which use $R^2$ in the diagonal are performed on the eight PASS subtests. The factor analysis criterion of at least a ten to one ratio between subjects and number of variables (Bentler & Lettieri, 1976, p.7, in Bentler, Lettieri, & Austin, Eds.) is exceeded. Although Bentler (p. 152, in Bentler, Lettieri, & Austin, Eds.) considers five variables per factor as desirable for having confidence in derived factors, he also states that when fewer variables are present, several factors can still be interpreted with confidence. Three and four factor solutions were explored.

The four factor solution yielded eigenvalues of 3.37, 1.27, .966 and .676 and accounted for 78.5% of the variance. The three factor solution accounted for 70% of the variance. Factor loadings of the eight tasks are presented in Tables 8 and 9. The doubled critical value of $r$ for a sample of 103 at the .01 level of significance (.504) is used to estimate the strength of the loading of a task on a factor. This procedure has been previously used by Nalgieri, Prewett, and Bardos (in press) and is derived from the work of Stevens (1986, p. 343-344) and Cliff and Hamburger (1967). Loadings greater than .35 (a traditional cutoff for inclusion on a factor) but less than .504 are also noted.

In the three factor orthogonal solution shown in Table 8, Planning and Attention tasks load strongly on the first factor while Figure Memory has a
moderate loading, below the criterion recommended by Stevens (1986) and Cliff and Hamburger (1967) but above the traditional .35 criterion. SUCCESSIVE

Table 8
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Orthogonal Solution

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th>4 Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/ATT</td>
<td>Succ</td>
<td>Simul</td>
<td>Plan</td>
</tr>
<tr>
<td>VS</td>
<td>.791 **</td>
<td>.104</td>
<td>.114</td>
<td>.709</td>
</tr>
<tr>
<td>PC</td>
<td>.748 **</td>
<td>.183</td>
<td>.168</td>
<td>.791</td>
</tr>
<tr>
<td>FM</td>
<td>.432 *</td>
<td>.094</td>
<td>.361 *</td>
<td>.399</td>
</tr>
<tr>
<td>MAT</td>
<td>.207</td>
<td>.137</td>
<td>.852 **</td>
<td>.157</td>
</tr>
<tr>
<td>WR</td>
<td>.189</td>
<td>.831</td>
<td>.031</td>
<td>.275</td>
</tr>
<tr>
<td>SR&amp;Q</td>
<td>.121</td>
<td>.514</td>
<td>.250</td>
<td>.020</td>
</tr>
<tr>
<td>ExA</td>
<td>.619 **</td>
<td>.037</td>
<td>.050</td>
<td>.408</td>
</tr>
<tr>
<td>NF</td>
<td>.734 **</td>
<td>.041</td>
<td>.223</td>
<td>.573</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2r (p < .01)  * factor loading ≥1.5 r (p < .01)
Table 9

Three and Four Factor Solutions Using Principal Axes Factor Analysis

Oblique Solution

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th>4 Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/ATT</td>
<td>Succ</td>
<td>Simul</td>
<td>Plan</td>
</tr>
<tr>
<td>VS</td>
<td>.793 **</td>
<td>.055</td>
<td>.009</td>
<td>.621 **</td>
</tr>
<tr>
<td>PC</td>
<td>.722 **</td>
<td>.133</td>
<td>.051</td>
<td>.803 **</td>
</tr>
<tr>
<td>FM</td>
<td>.359 *</td>
<td>.038</td>
<td>.314</td>
<td>.339</td>
</tr>
<tr>
<td>MAT</td>
<td>.000</td>
<td>.045</td>
<td>.876 **</td>
<td>.007</td>
</tr>
<tr>
<td>WR</td>
<td>.058</td>
<td>.859 **</td>
<td>.122</td>
<td>.249</td>
</tr>
<tr>
<td>SR&amp;Q</td>
<td>.021</td>
<td>.503 **</td>
<td>.214</td>
<td>.233</td>
</tr>
<tr>
<td>ExA</td>
<td>.637 **</td>
<td>.000</td>
<td>.045</td>
<td>.052</td>
</tr>
<tr>
<td>NF</td>
<td>.719 **</td>
<td>.019</td>
<td>.123</td>
<td>.387 *</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2r (p < .01)  * factor loading ≡ 1.5 r (p < .01)

Tasks form a strong second factor. One Simultaneous task loads strongly and one moderately on a third factor. In the four factor orthogonal solution, Planning tasks load together as a first factor with a secondary loading of Figure Memory and Attention tasks. Successive tasks again form a strong second factor. Simultaneous tasks comprise the third factor with a strong
loading of Matrices and moderate loading of Figure Memory. Attention tasks define the fourth factor.

The oblique solutions take into account the intercorrelations of the variables. Because the Das-Luria theory assumes some interdependence among the factors comprising the model, the oblique procedure should be the most accurate reflection of the factor structure of the tasks.

Examination of the three factor oblique solution shows that the first and second factors are similar to those found in the orthogonal solution. However, the third factor is primarily represented by Matrices. Figure Memory does not reach the .35 level required for inclusion on the factor although it is the only other task close to a critical value. The correlations between the Planning/Attention and Successive factors ($r = .27$), Planning/Attention and Simultaneous factors ($r = .38$) and Simultaneous and Successive factors ($r = .24$) reflect the expected modest interdependence of these three factors.

In the four factor oblique solution, the first factor is composed of the Planning tasks with a secondary loading of one Attention task, Number Finding. The Successive and Simultaneous factors each are represented by the tasks assigned to these factors. One Attention task loads strongly and one moderately on a fourth factor. The intercorrelation between the factors defined as Planning and Attention is strongest ($r = .64$), between Planning and Simultaneous factors moderate ($r = .40$) and lowest between Planning and Successive ($r = .30$), Simultaneous and Successive ($r = .22$), Simultaneous and Attention ($r = .27$) and Successive and Attention ($r = .24$). The strength of the correlations between the four factors conforms to the theoretical model, reflects the physiological relationships between the functional units of the
brain (Luria, 1973) and is similar to results of previous studies (Naglieri & Das, 1987, 1988, Naglieri, Das, Stevens, & Ledbetter, 1989). Despite some minor discrepancies, the four factor oblique principal axes analysis closely resembles the PASS model and provides support for Hypothesis 4.

**Correlations of DAP TT with Planning, Attention, Simultaneous and Successive Processing Scores**

The correlations of the DAP TT and the four components of the PASS model are presented in Table 5. These correlations with age held constant are presented in Table 6. Although Planning, Attention, Simultaneous and Successive processing do not correlate significantly with age, correlations from Table 6 are used in order to be consistent with previous discussions of DAP TT correlations.

The Planning Scale and Simultaneous Scale scores correlate significantly ($p < .01$) with the DAP TT. These results support Hypotheses 5 and 7. The Attention Scale and Simultaneous Scale scores correlate at the $p < .05$ level with the DAP TT. The experiment-wise error rate for the study is $p < .01$. This result suggests weak support for Hypotheses 6 and 8. In order to further examine the relationship between the PASS model and the DAP TT, a $t$ test for correlated correlations is applied to the four PASS-DAP TT correlations. There is no significant difference between the DAP TT / Planning and DAP TT / Attention correlations ($t = 1.687, p > .10$), between the DAP TT / Planning and DAP TT / Simultaneous Processing correlations ($t = .714, p > .10$) and the DAP TT / Planning and DAP TT / Successive
Processing correlations \((t = 1.116, p > .10)\). There also is no significant difference between the DAP TT / Attention and DAP TT / Simultaneous Processing correlations \((t = .579, p > .10)\), between the DAP TT / Attention and DAP TT / Successive Processing correlations \((t = -.041, p > .10)\) and between the DAP TT / Simultaneous Processing and DAP TT / Successive Processing correlations \((t = .477, p > .10)\). These results suggest that all four components of the PASS model contribute to the success of the human figure drawing.

**Prediction of DAP TT Scores using Regression**

Three types of regression analyses are conducted. These are simple regression, multiple regression, and stepwise multiple regression. Results of simple regression analyses are shown in Table 10.

The paper and pencil tasks which comprise the Fine Motor score account for the greatest amount of DAP TT variance while the least amount of variance is accounted for by the CBCL scales. Multiple regression analyses are conducted in order to determine which subsets of variables are the most effective predictors of DAP TT. Because of the previously discussed finding that age correlated significantly with several of the variables, age is included as a variable in each analysis. This procedure prevents inflation of the influence of a variable significantly correlated with age. Results of the multiple regression analyses are shown in the first row (All Variables) of Tables 11, 12, 13, 14, and 15. The adjusted \(R^2\) can be used to compare subsets of unequal numbers of variables (Snedecor and Cochran, 1980). According to this criteria, the subset of variables which is comprised of age,
McCarthy Motor Scale, BBCS, PASS, and CBCL accounts for the greatest amount of variance (33%). The subset of variables comprised of age and the four components of the PASS model, Planning, Attention, Simultaneous, and Successive processing, accounts for the least amount of variance (18.4%).

Table 10
Simple Regression on DAP TT

<table>
<thead>
<tr>
<th>Variable</th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adj.$R^2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>McC Mot Scale</td>
<td>.561</td>
<td>.315</td>
<td>.308</td>
<td>46.396</td>
</tr>
<tr>
<td>McC Mot Factor</td>
<td>.48</td>
<td>.231</td>
<td>.223</td>
<td>30.256</td>
</tr>
<tr>
<td>Gross Motor</td>
<td>.36</td>
<td>.13</td>
<td>.121</td>
<td>15.068</td>
</tr>
<tr>
<td>Fine Motor</td>
<td>.75</td>
<td>.563</td>
<td>.559</td>
<td>130.175</td>
</tr>
<tr>
<td>Draw-A-Design</td>
<td>.513</td>
<td>.263</td>
<td>.256</td>
<td>36.085</td>
</tr>
<tr>
<td>BBCS</td>
<td>.424</td>
<td>.18</td>
<td>.172</td>
<td>22.199</td>
</tr>
<tr>
<td>PASS</td>
<td>.358</td>
<td>.128</td>
<td>.12</td>
<td>14.405</td>
</tr>
<tr>
<td>Planning</td>
<td>.341</td>
<td>.116</td>
<td>.108</td>
<td>13.286</td>
</tr>
<tr>
<td>Attention</td>
<td>.212</td>
<td>.045</td>
<td>.036</td>
<td>4.759</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>.272</td>
<td>.074</td>
<td>.065</td>
<td>8.085</td>
</tr>
<tr>
<td>Successive</td>
<td>.217</td>
<td>.047</td>
<td>.037</td>
<td>4.968</td>
</tr>
<tr>
<td>CBCL Internal</td>
<td>.092</td>
<td>.008</td>
<td>-.001</td>
<td>.856</td>
</tr>
<tr>
<td>CBCL External</td>
<td>.22</td>
<td>.048</td>
<td>.039</td>
<td>5.128</td>
</tr>
<tr>
<td>CBCL Total</td>
<td>.219</td>
<td>.048</td>
<td>.039</td>
<td>5.085</td>
</tr>
</tbody>
</table>
However, all five subsets of variables used in multiple regression analyses account for significant amounts of DAP TT variance ($p < .001$) and support Hypotheses 9 - 13. In order to determine which of the variables in a subset make significant contributions to DAP TT variance, stepwise multiple regression analyses are conducted.

The computer program Statview 512+ (Feldman & Gagnon, 1986) conducts stepwise multiple regression using step-forward analysis. This method, discussed in Snedecor and Cochran (1980), selects the variable with the highest simple regression $F$ value as the first entry into the regression equation. Other variables are then examined for their contribution to further reduction in the residual sums of squares. The strength of this reduction is shown by an $F$ value. This process continues until all variables making a significant reduction have been included. Central to this process is the choice of a guideline for acceptable reduction in the residual sums of squares. Snedecor and Cochran state that "the 5% significance level has been used but this is not a problem in testing significance" (p.360). Draper & Smith (1981) recommend the 10% significance level. They believe too much information is lost when more stringent levels are set. Therefore, in this study, an $F$ value of 2 is chosen for entry into the equation. This value is close to the 10% significance level for degrees of freedom 4/100.

The Statview 512+ program allows variables to be forced into the stepwise multiple regression equation. Age is forced into each regression equation first so that subsequent reductions in the residual sums of squares are not influenced by age.
Results of the first analysis, which includes chronological age, *Bracken Basic Concept Scale* (BBCS), PASS Total, the *Child Behavior Checklist* (CBCL) Total, and the McCarthy Motor Scale scores, are shown in Table 11. These measures together account for 33% of the variance in the DAP TT scores ($F = 11.052, p = .01$). Stepwise multiple regression shows that the McCarthy Motor Scale ($F = 23.79, p < .01$) and the BBCS ($F = 3.26, p < .05$) make significant contributions to the reduction in residual mean square error. The DAP TT variance accounted for by the addition of the PASS and CBCL variables is not significant.

Table 11

Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, and Motor Scale

<table>
<thead>
<tr>
<th></th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adj $R^2$</th>
<th>$F$ Enter</th>
<th>df</th>
<th>$F$ remove</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Var.</td>
<td>.60</td>
<td>.36</td>
<td>.33</td>
<td>5/97</td>
<td></td>
<td></td>
<td>11.05***</td>
</tr>
<tr>
<td>Chron Age#</td>
<td>.28</td>
<td>.08</td>
<td>.07</td>
<td>1/101</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor Scale</td>
<td>.58</td>
<td>.34</td>
<td>.33</td>
<td>39.51***</td>
<td>2/100</td>
<td></td>
<td>23.79***</td>
</tr>
<tr>
<td>BBCS</td>
<td>.60</td>
<td>.36</td>
<td>.34</td>
<td>2.44*</td>
<td>3/99</td>
<td></td>
<td>3.26**</td>
</tr>
</tbody>
</table>

Note: # variable forced into equation before others

Note: * $p < .10$, ** $p < .05$, *** $p < .01$
Results of the second analysis, which includes the variables chronological age, McCarthy Motor Factor, *Bracken Basic Concept Scale* (BBCS), PASS Total, and *Child Behavior Checklist* (CBCL) Total scores, are shown in Table 12. Together, these measures account for 25% of the variance in the DAP TT scores ($F = 7.79, p = .01$).

Table 12

Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, and Motor Factor

<table>
<thead>
<tr>
<th></th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adj $R^2$</th>
<th>$F$ Enter</th>
<th>df</th>
<th>$F$ remove</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Var.</td>
<td>.54</td>
<td>.29</td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
<td>7.79***</td>
</tr>
<tr>
<td>Chron Age#</td>
<td>.28</td>
<td>.08</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor Fact</td>
<td>.49</td>
<td>.24</td>
<td>.23</td>
<td>21.62***</td>
<td>2/100</td>
<td>10.34***</td>
<td></td>
</tr>
<tr>
<td>BBCS</td>
<td>.53</td>
<td>.28</td>
<td>.26</td>
<td>4.91***</td>
<td>3/99</td>
<td>14.91***</td>
<td></td>
</tr>
</tbody>
</table>

Note: # variable forced into equation before others

Note: * $p < .10$, ** $p < .05$, *** $p < .01$

In the stepwise analysis, the McCarthy Motor Factor (Draw-A-Child excluded) is the best predictor of DAP TT scores ($F = 10.34, p < .01$) but the BBCS also makes a significant contribution ($F = 14.91, p < .01$) The PASS and CBCL variables do not make a significant contribution. Removal of
Draw-A-Child reduces the amount of variance accounted for by the McCarthy tasks and increases the amount of variance accounted for by the BBCS. Results of the the third analysis, which includes chronological age, McCarthy Gross Motor, Draw-A Design, BBCS, PASS Total, and CBCL Total scores, are shown in Table 13. These measures account for 30.3% of the variance in the DAP TT scores ($F = 8.406, p = .01$). Draw-A-Design is the most efficient predictor of DAP TT scores ($F = 18.06, p < .01$) and the BBCS adds significantly ($F = 5.62, p < .01$) to the prediction equation. The Gross Motor, PASS, and CBCL scores do not make a significant addition.

Table 13

Multiple and Stepwise Regression on DAP TT - BBCS, PASS, CBCL, Gross Motor, and Draw-A-Design

<table>
<thead>
<tr>
<th></th>
<th>$R$</th>
<th>$R^2$</th>
<th>Adj $R^2$</th>
<th>$F$ Enter</th>
<th>df</th>
<th>$F$ remove</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Var.</td>
<td>.59</td>
<td>.34</td>
<td>.30</td>
<td>6/96</td>
<td></td>
<td></td>
<td>8.41***</td>
</tr>
<tr>
<td>Chron Age#</td>
<td>.28</td>
<td>.08</td>
<td>.07</td>
<td></td>
<td>1/101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draw-Des.</td>
<td>.54</td>
<td>.29</td>
<td>.27</td>
<td>29.32***</td>
<td>2/100</td>
<td></td>
<td>18.06***</td>
</tr>
<tr>
<td>BBCS</td>
<td>.57</td>
<td>.33</td>
<td>.31</td>
<td>5.62***</td>
<td>3/99</td>
<td></td>
<td>5.62***</td>
</tr>
</tbody>
</table>

Note: # variable forced into equation before others
Note: * $p < .10$, ** $p < .05$, *** $p < .01$
The results of the first three analyses which include subsets of McCarthy Motor Scale tasks show that the paper and pencil tasks, Draw-A-Child and Draw-A-Design, are the best predictors of DAP TT scores. The measure of basic concepts, BBCS, accounts for increasing amounts of error variance as the Motor Scale subsets are changed. When the Motor Scale is in the set, the BBCS adds 1.5% of additional variance which is significant at the 0.10 level. When the Motor Factor is in the set, the BBCS adds 2.9% additional variance which is significant at the 0.01 level. When Draw-A-Design and Gross Motor scores are in the set, the BBCS adds 3.2% additional variance which is significant at the 0.01 level.

Results of the fourth analysis, which includes chronological age, BBCS, PASS Total, and CBCL Total scores, are shown in Table 14.

Table 14
Multiple and Stepwise Regression on DAP TT - BBCS, PASS, and CBCL

<table>
<thead>
<tr>
<th></th>
<th>R</th>
<th>R²</th>
<th>Adj R²</th>
<th>F Enter</th>
<th>df</th>
<th>F remove</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Var.</td>
<td>.49</td>
<td>.24</td>
<td>.21</td>
<td></td>
<td>4/98</td>
<td></td>
<td>7.59***</td>
</tr>
<tr>
<td>Chron Age#</td>
<td>.28</td>
<td>.08</td>
<td>.07</td>
<td></td>
<td>1/101</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PASS</td>
<td>.45</td>
<td>.21</td>
<td>.19</td>
<td>16.19***</td>
<td>2/100</td>
<td></td>
<td>2.81*</td>
</tr>
<tr>
<td>BBCS</td>
<td>.47</td>
<td>.22</td>
<td>.20</td>
<td>2.28*</td>
<td>3/99</td>
<td></td>
<td>2.28*</td>
</tr>
</tbody>
</table>

Note: # variable forced into equation before others
Note: * p < .10, ** p < .05, *** p < .01
These measures account for 20.5% of the variance in DAP TT scores \((F = 7.585, p < .01)\). The PASS score enters the equation first. The size of the \(F\) value (16.185) indicates it significantly reduces the error variance. However, when the BBCS is added to the regression equation, the amount of variance attributed to PASS scores is greatly reduced due to the significant intercorrelations of the PASS and BBCS scores \((r = .73)\). Before the BBCS is accepted into the equation, PASS has a partial \(F\) of 16.19 \((p < .01)\). Following the inclusion of the BBCS, PASS has a partial \(F\) of 2.81 \((p < .10)\).

Results of the fifth analysis, shown in Table 15, includes chronological age and the four component scores of the PASS model, Planning, Attention,

<table>
<thead>
<tr>
<th></th>
<th>(R)</th>
<th>(R^2)</th>
<th>Adj (R^2)</th>
<th>(F) Enter</th>
<th>df</th>
<th>(F) remove</th>
<th>(F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Var.</td>
<td>.47</td>
<td>.22</td>
<td>.18</td>
<td>5/97</td>
<td></td>
<td></td>
<td>5.61***</td>
</tr>
<tr>
<td>Chron Age#</td>
<td>.28</td>
<td>.08</td>
<td>.07</td>
<td>1/101</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>.42</td>
<td>.18</td>
<td>.16</td>
<td>12.44***</td>
<td>2/100</td>
<td>4.60**</td>
<td></td>
</tr>
<tr>
<td>Simult.</td>
<td>.46</td>
<td>.21</td>
<td>.19</td>
<td>4.09**</td>
<td>3/99</td>
<td>4.09***</td>
<td></td>
</tr>
</tbody>
</table>

Note: \# variable forced into equation before others

Note: * \(p < .10\), ** \(p < .05\), *** \(p < .01\)
Simultaneous, and Successive Processing. These measures account for 18.4% of the variance in DAP TT scores ($F = 5.61, p < .01$). Simultaneous processing and Planning tasks make significant contributions to the regression equation. Successive processing and Attention do not. The overlap in variance accounted for by Simultaneous processing and Planning tasks is shown by the reduction of the Simultaneous partial $F$ ($F = 12.43, p < .01$ to $F = 4.60, p < .05$) when Planning is added to the regression equation. The moderate correlation ($r = .49$) between these components of the Das-Luria model is reflected in the regression analysis.

In summary, the regression analyses show that the McCarthy Motor scale sets with the exception of the Gross Motor scale are the best predictors of DAP TT scores. The BBCS consistently makes a significant although small contribution to the reduction in DAP TT error variance in the presence of Motor Scale tasks. When the Motor tasks are removed from consideration, the PASS components Total score is the best predictor. However, the BBCS is strong enough to enter the equation after the PASS Total and reduces the effect of the PASS Total. This reduction occurs when two tasks are correlated. A similar effect is found when the four components of the PASS model are compared. The Simultaneous and Planning components are the best predictors of DAP TT scores. Although Simultaneous processing enters the regression equation first, its partial $F$ value is effectively reduced when Planning is introduced. This reflects the significant correlation between Simultaneous processing and Planning. The stepwise regression analyses show that many of the variables used in this study are measuring related aspects of the DAP TT despite variable names which suggest they are
measuring distinct aspects. The complex multidimensional nature of the DAP appears in Figure 1.

![Diagram showing significant correlations](image)

**Figure 1** Significant correlations ($p < .01$) between the DAP and other variables
Gender Differences

One of the difficulties with previous scoring systems for the human figure drawing is the appearance of significant differences between scores earned by males and females. This difference is noted by Goodenough (1926) and Harris (1963). Harris finds the differences so great that he devises separate scoring criteria for males and females. Therefore, any study of the DAP needs to examine gender differences.

Results of the investigation of differences between scores earned on the DAP TT, McCarthy Motor Factor and Motor Scale, BBCS, PASS scales, and the CBCL when the sample is divided by gender are shown in Table 16.

Table 16
Comparisons of Means and SDs of Major Variables by Gender

<table>
<thead>
<tr>
<th></th>
<th>Males (52)</th>
<th></th>
<th>Females (51)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
<td>S.D.</td>
<td>Difference</td>
</tr>
<tr>
<td>DAP TT</td>
<td>89.69</td>
<td>15.90</td>
<td>95.22</td>
<td>14.39</td>
<td>5.53</td>
</tr>
<tr>
<td>Bracken</td>
<td>95.50</td>
<td>16.91</td>
<td>93.29</td>
<td>15.96</td>
<td>2.21</td>
</tr>
<tr>
<td>McC Mot. Fact</td>
<td>46.56</td>
<td>7.37</td>
<td>46.58</td>
<td>8.11</td>
<td>.02</td>
</tr>
<tr>
<td>McC. Mot.</td>
<td>46.79</td>
<td>10.05</td>
<td>46.33</td>
<td>9.77</td>
<td>.46</td>
</tr>
<tr>
<td>CBCL. Tot</td>
<td>56.62</td>
<td>9.77</td>
<td>52.51</td>
<td>8.74</td>
<td>4.11</td>
</tr>
<tr>
<td>PASS Tot</td>
<td>100.5</td>
<td>10.31</td>
<td>99.49</td>
<td>8.19</td>
<td>1.01</td>
</tr>
</tbody>
</table>
Age is not a factor in these analyses as Table 1 (Chapter 3) shows there is no significant difference between males and females in age. Results of the ANOVA indicate there are no significant differences between scores earned by males and females on the DAP or the other five measures \( (F = .07, p > .10) \).

Means and standard deviations of scores earned on the eight tasks which comprise the four scales of the PASS model are shown in Table 17. The results of the ANOVA show no significant differences between scores earned by males and females on these tasks \( (F = .305, p > .10) \).

Means and standard deviations of scores earned by males and females on the four component scales of the PASS model are shown in Table 18. The results of the ANOVA reveal no significant differences between the scores earned by males and females \( (F = .305, p > .10) \).

Means and standard deviations of scores earned by males and females on the motor and behavior subscales are shown in Table 19. The results of the ANOVA show there are no significant differences between males and females on these measures \( (F = 4.428, p > .03) \).

The scores earned by both males and females on all of the variables used in this study are similar. The results do not support hypotheses 14a, b, c, d, and e, 15a, b, c, and d, and 16a, b, c, d, and e which anticipate gender differences on all of the variables.
Table 17
Comparisons of Means and SDs of PASS Tasks by Gender

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means</td>
<td>S.D.</td>
<td>Means</td>
<td>S.D.</td>
<td></td>
<td>Diff/Means</td>
</tr>
<tr>
<td>Planning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual Search</td>
<td>98.88</td>
<td>15.13</td>
<td>101.15</td>
<td>14.93</td>
<td>2.26</td>
<td></td>
</tr>
<tr>
<td>Planned Connections</td>
<td>100.16</td>
<td>14.09</td>
<td>99.84</td>
<td>16.01</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expressive Attention</td>
<td>100.23</td>
<td>12.99</td>
<td>99.76</td>
<td>11.85</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Number Finding</td>
<td>99.06</td>
<td>14.84</td>
<td>100.96</td>
<td>13.47</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>Simultaneous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Figure Memory</td>
<td>101.86</td>
<td>16.55</td>
<td>98.08</td>
<td>13.11</td>
<td>3.78</td>
<td></td>
</tr>
<tr>
<td>Loc MAT-EF</td>
<td>102.72</td>
<td>15.99</td>
<td>97.23</td>
<td>13.53</td>
<td>5.49</td>
<td></td>
</tr>
<tr>
<td>Successive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Recall</td>
<td>100.14</td>
<td>15.55</td>
<td>99.89</td>
<td>14.58</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>Sentence Repetition &amp; Questions</td>
<td>100.97</td>
<td>16.43</td>
<td>98.99</td>
<td>13.46</td>
<td>1.98</td>
<td></td>
</tr>
</tbody>
</table>
Table 18
Comparisons of Means and SDs of the Components of the PASS Model by Gender

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>Diff/Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means</td>
<td>S.D.</td>
<td>Means</td>
</tr>
<tr>
<td>Planning</td>
<td>99.52</td>
<td>13.23</td>
<td>100.49</td>
</tr>
<tr>
<td>Attention</td>
<td>99.64</td>
<td>11.96</td>
<td>100.36</td>
</tr>
<tr>
<td>Simultaneous</td>
<td>102.29</td>
<td>13.98</td>
<td>97.66</td>
</tr>
<tr>
<td>Successive</td>
<td>100.56</td>
<td>14.22</td>
<td>99.44</td>
</tr>
</tbody>
</table>

Table 19
Comparisons of the Means and SDs of Motor Subscales and CBCL Subscales by Gender

<table>
<thead>
<tr>
<th></th>
<th>Males Mean</th>
<th>Males SD</th>
<th>Females Mean</th>
<th>Females SD</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>McC Gross</td>
<td>44.44</td>
<td>8.43</td>
<td>45.48</td>
<td>8.69</td>
<td>1.05</td>
</tr>
<tr>
<td>McDAD</td>
<td>51.60</td>
<td>8.39</td>
<td>51.68</td>
<td>8.21</td>
<td>0.08</td>
</tr>
<tr>
<td>CBCL Int</td>
<td>52.33</td>
<td>9.23</td>
<td>48.43</td>
<td>6.31</td>
<td>3.90</td>
</tr>
<tr>
<td>CBCL Ext</td>
<td>57.40</td>
<td>10.33</td>
<td>54.80</td>
<td>10.22</td>
<td>2.60</td>
</tr>
</tbody>
</table>
Additional Analyses

In order to further understand the theoretical basis for the DAP TT, additional factor analyses are conducted with each of the variables used in this study. Results are shown in Tables 20 through 26. The McCarthy Motor Scale, Motor Factor, and Gross Motor scale load on a factor with planning and attention tasks. Bracken Basic Concept Scale and Draw-A-Design load on a simultaneous factor while the CBCL loads on the factor with planning and attention tasks. This shows that the numerous items of the CBCL related to attention correspond with performance on attention and planning tasks.
Table 20
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and McCarthy Motor Scale

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th>4 Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PL/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Att/PI</td>
</tr>
<tr>
<td>McC Motor</td>
<td>.434 *</td>
<td>.115</td>
<td>.011</td>
<td>.487 *</td>
</tr>
<tr>
<td>Vis Search</td>
<td>.783 **</td>
<td>.024</td>
<td>.001</td>
<td>.586 **</td>
</tr>
<tr>
<td>Pain Conn</td>
<td>.722 **</td>
<td>.104</td>
<td>.046</td>
<td>.422 *</td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.351 *</td>
<td>.000</td>
<td>.341</td>
<td>.094</td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.040</td>
<td>.021</td>
<td>.868 **</td>
<td>.226</td>
</tr>
<tr>
<td>Word Rec</td>
<td>.101</td>
<td>.612 **</td>
<td>.112</td>
<td>.000</td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.096</td>
<td>.747 **</td>
<td>.124</td>
<td>.011</td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.658 **</td>
<td>.013</td>
<td>.071</td>
<td>.872 **</td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.718 **</td>
<td>.029</td>
<td>.125</td>
<td>.647 **</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 r (p < .01) * factor loading ≈ 1.5 r (p < .01).
Table 21
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and McCarthy Motor Factor

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th>4 Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/Att</td>
<td>Sim</td>
</tr>
<tr>
<td>Mc Mot Fac</td>
<td>.510 *</td>
<td>.052</td>
</tr>
<tr>
<td>Vis Search</td>
<td>.790 **</td>
<td>.013</td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.726 **</td>
<td>.051</td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.388 *</td>
<td>.302</td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.001</td>
<td>.929 **</td>
</tr>
<tr>
<td>Word Rec</td>
<td>.078</td>
<td>.143</td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.044</td>
<td>.119</td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.637 **</td>
<td>.035</td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.710 **</td>
<td>.142</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 $r$ ($p < .01$) * factor loading $\geq 1.5$ $r$ ($p < .01$).
Table 22
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and Gross Motor Scale

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th></th>
<th>4 Factors</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pl/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Plan</td>
<td>Succ</td>
<td>Sim</td>
<td>Sim</td>
<td>Att</td>
</tr>
<tr>
<td>Gross Motor</td>
<td>.478*</td>
<td>.175</td>
<td>.151</td>
<td>.442*</td>
<td>.147</td>
<td>.131</td>
<td>.132</td>
<td></td>
</tr>
<tr>
<td>Vis Search</td>
<td>.785**</td>
<td>.054</td>
<td>.050</td>
<td>.725**</td>
<td>.021</td>
<td>.095</td>
<td>.147</td>
<td></td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.693**</td>
<td>.129</td>
<td>.115</td>
<td>.734**</td>
<td>.089</td>
<td>.164</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.348*</td>
<td>.021</td>
<td>.346</td>
<td>.334</td>
<td>.016</td>
<td>.397*</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.023</td>
<td>.002</td>
<td>.905**</td>
<td>.035</td>
<td>.082</td>
<td>.814**</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Word Rec</td>
<td>.097</td>
<td>.799**</td>
<td>.056</td>
<td>.223</td>
<td>.602**</td>
<td>.075</td>
<td>.093</td>
<td></td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.011</td>
<td>.551**</td>
<td>.157</td>
<td>-.114</td>
<td>.742**</td>
<td>.184</td>
<td>.104</td>
<td></td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.629**</td>
<td>.008</td>
<td>.000</td>
<td>.339</td>
<td>.002</td>
<td>.008</td>
<td>.605**</td>
<td></td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.673**</td>
<td>.030</td>
<td>.206</td>
<td>.523**</td>
<td>.023</td>
<td>.254</td>
<td>.256</td>
<td></td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 $r$ ($p < .01$)  * factor loading $\geq 1.5$ $r$ ($p < .01$).
Table 23
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and Draw-A-Design

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th></th>
<th>4 Factors</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pl/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Plan</td>
<td></td>
</tr>
<tr>
<td>Draw-A-Des</td>
<td>.085</td>
<td>.019</td>
<td>.565**</td>
<td>.072</td>
<td>.048</td>
<td>.585**</td>
<td>.087</td>
<td></td>
</tr>
<tr>
<td>Vis Search</td>
<td>.825**</td>
<td>.004</td>
<td>.006</td>
<td>.524**</td>
<td>.019</td>
<td>.108</td>
<td>.387*</td>
<td></td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.653**</td>
<td>.071</td>
<td>.182</td>
<td>.346*</td>
<td>.004</td>
<td>.291</td>
<td>.432*</td>
<td></td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.183</td>
<td>.011</td>
<td>.539**</td>
<td>.003</td>
<td>.043</td>
<td>.651**</td>
<td>.095</td>
<td></td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.030</td>
<td>.125</td>
<td>.575**</td>
<td>.034</td>
<td>.126</td>
<td>.568**</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td>Word Rec</td>
<td>.063</td>
<td>.716**</td>
<td>.128</td>
<td>.027</td>
<td>.370*</td>
<td>.078</td>
<td>.420*</td>
<td></td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.230</td>
<td>.682**</td>
<td>.281</td>
<td>.002</td>
<td>1.01**</td>
<td>.127</td>
<td>.003</td>
<td></td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.621**</td>
<td>.035</td>
<td>.015</td>
<td>.772**</td>
<td>.080</td>
<td>.030</td>
<td>.044</td>
<td></td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.626**</td>
<td>.067</td>
<td>.261</td>
<td>.440*</td>
<td>.045</td>
<td>.416*</td>
<td>.000</td>
<td></td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 r (p < .01)  * factor loading ≥ 1.5 r (p < .01).
Table 24
Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and Bracken Basic Concepts

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th>4 Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Plan</td>
</tr>
<tr>
<td>Bracken</td>
<td>.225</td>
<td>.275</td>
<td>.532 **</td>
<td>.062</td>
</tr>
<tr>
<td>Vis Search</td>
<td>.802 **</td>
<td>.034</td>
<td>.008</td>
<td>.664 **</td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.726 **</td>
<td>.109</td>
<td>.032</td>
<td>.807 **</td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.342</td>
<td>.001</td>
<td>.339</td>
<td>.363 *</td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.031</td>
<td>.021</td>
<td>.854 **</td>
<td>.018</td>
</tr>
<tr>
<td>Word Rec</td>
<td>.076</td>
<td>.650 **</td>
<td>.086</td>
<td>.145</td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.144</td>
<td>.699 **</td>
<td>.226</td>
<td>.205</td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.631 **</td>
<td>.015</td>
<td>.016</td>
<td>.252</td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.703 **</td>
<td>.029</td>
<td>.142</td>
<td>.506 **</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 r (p < .01) * factor loading ≡ 1.5 r (p < .01).
Table 25

Three and Four Factor Solutions Using Principal Axes Factor Analysis

Oblique Solution - Pass Tasks and Child Behavior Checklist

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th>4 Factors</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>PI/Att</td>
</tr>
<tr>
<td>CBCL</td>
<td>.652**</td>
<td>.097</td>
<td>.329</td>
<td>.636**</td>
</tr>
<tr>
<td>Vis Search</td>
<td>.624**</td>
<td>.000</td>
<td>.259</td>
<td>.579**</td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.610**</td>
<td>.094</td>
<td>.253</td>
<td>.546**</td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.052</td>
<td>.002</td>
<td>.658**</td>
<td>.078</td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.073</td>
<td>.151</td>
<td>.594**</td>
<td>.000</td>
</tr>
<tr>
<td>Word Rec</td>
<td>.100</td>
<td>.668**</td>
<td>.069</td>
<td>.009</td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.169</td>
<td>.676**</td>
<td>.217</td>
<td>.000</td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.610**</td>
<td>.025</td>
<td>.077</td>
<td>.682**</td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.592**</td>
<td>.041</td>
<td>.320</td>
<td>.634**</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 $r$ ($p < .01$) * factor loading $\geq 1.5~r$ ($p < .01$).
Table 26

Three and Four Factor Solutions Using Principal Axes Factor Analysis
Oblique Solution - Pass Tasks and Draw a Person

<table>
<thead>
<tr>
<th></th>
<th>3 Factors</th>
<th></th>
<th></th>
<th>4 Factors</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PI/Att</td>
<td>Succ</td>
<td>Sim</td>
<td>Att/Pl</td>
<td>Succ</td>
<td>Sim</td>
<td>Plan</td>
</tr>
<tr>
<td>DAP</td>
<td>.168</td>
<td>.168</td>
<td>.179</td>
<td>.001</td>
<td>.138</td>
<td>.173</td>
<td>.236</td>
</tr>
<tr>
<td>Vis Search</td>
<td>.815 **</td>
<td>.016</td>
<td>.025</td>
<td>.564 **</td>
<td>.019</td>
<td>.028</td>
<td>.265</td>
</tr>
<tr>
<td>Plan Conn</td>
<td>.737 **</td>
<td>.109</td>
<td>.035</td>
<td>.198</td>
<td>.023</td>
<td>.018</td>
<td>.768 **</td>
</tr>
<tr>
<td>Fig Memory</td>
<td>.345</td>
<td>.002</td>
<td>.318</td>
<td>.219</td>
<td>.014</td>
<td>.348</td>
<td>.133</td>
</tr>
<tr>
<td>Matrix Anal.</td>
<td>.077</td>
<td>.034</td>
<td>.913 **</td>
<td>.055</td>
<td>.002</td>
<td>.926 **</td>
<td>.043</td>
</tr>
<tr>
<td>Word Rec</td>
<td>.062</td>
<td>.609 **</td>
<td>.097</td>
<td>.003</td>
<td>.438 *</td>
<td>.076</td>
<td>.217</td>
</tr>
<tr>
<td>Sent R &amp; Q</td>
<td>.189</td>
<td>.780 **</td>
<td>.168</td>
<td>.016</td>
<td>1.02 **</td>
<td>.081</td>
<td>.138</td>
</tr>
<tr>
<td>Exp Attent</td>
<td>.643 **</td>
<td>.000</td>
<td>.067</td>
<td>.818 **</td>
<td>.045</td>
<td>.067</td>
<td>.156</td>
</tr>
<tr>
<td>Numb. Find</td>
<td>.714 **</td>
<td>.037</td>
<td>.104</td>
<td>.682 **</td>
<td>.002</td>
<td>.140</td>
<td>.021</td>
</tr>
</tbody>
</table>

Note: ** factor loading > 2 \( r (p < .01) \) * factor loading \( \approx 1.5 \) \( r (p < .01) \).
CHAPTER V
DISCUSSION

The purpose of this study was to investigate four variables which reflect the way in which professionals interpret the human figure drawing task. These variables represent cognitive processes related to intelligence, concept understanding, personality, and motor functioning.

Review of the literature shows that no firm theoretical base has ever been established for understanding the relationship between the human figure drawing task and cognition. Past studies which focus on the DAP and standardized tests of intelligence (Scott, 1981) do little to define the construct(s) of intelligence measured by these tests. Therefore, a primary aim of this study was to discover if there is a relationship between the Draw A Person and constructs included in a theoretical model representing a current concept of intelligence (Naglieri, 1989). This concept of intelligence is based on Luria's model of cognitive processing.

A second aim of this study was to discover if there is a relationship between the DAP and the child's conceptual understanding (Harris, 1963) using a test of basic concepts. Additionally, the relationship between the DAP and personality using a behavioral problems checklist was tested. Finally the relationship between the DAP and a measure of fine and gross motor functioning was examined. In general the results demonstrate that the
DAP is a measure that is complex and multidimensional but mainly influenced by simultaneous and planning processes.

**The DAP and PASS Processing**

The correlations between the DAP and the PASS variables show the human figure drawing task is significantly related to two components of this model of cognitive functioning, simultaneous and planning processes.

Simultaneous processing allows the child to perceive the human figure as a whole as well as the interrelationships between all of the parts. The significant relationship between the DAP and the simultaneous processing component of the PASS model shows that the human figure drawing task involves this cognitive process.

The significant correlation between the DAP and the McCarthy Draw-A-Design task further supports the hypothesis that simultaneous processes are utilized in the human figure drawing task. Luria considered figure copying an example of uncomplicated simultaneous processing (Luria, 1966) and previous studies have shown that figure copying tasks load on a simultaneous factor (Das, 1984, Ashman & Das, 1980, Das, Kirby, & Jarman, 1979, Das, Cummins, Kirby, & Jarman, 1979, Das, Kirby & Jarman, 1975, Das & Molloy, 1975). In this study, factor analysis of Draw-A-Design and the eight PASS tasks shows that Draw-A-Design loads on a simultaneous factor.

The correlation between the DAP and the Planning component of the PASS model shows the human figure drawing task is significantly related to this control process as well as simultaneous processing. Planning requires that the child be able to set a drawing goal and make an initial plan for
reaching it. Then, the child must evaluate the success of the drawing.

Finally, the individual must be able to modify the plan if necessary in order to reach the original goal or be able to set and plan for a new goal. The significant correlation between Planning and the DAP supports the research of Freeman (1980), Freeman and Cox (1985), Crook (1985), and Laszlo and Bairstow (1985), all of whom suggest that planning is strongly involved in the drawing process.

Further support for the importance of planning in the DAP can be found by examining the respective contributions of planning and simultaneous processing to DAP and Draw-A-Design variances. Simultaneous processing makes a significant contribution to Draw-A-Design variance but planning does not make an additional contribution. This is in contrast to the DAP where the contributions of simultaneous processing and planning are approximately equal. Because the child can refer to a model throughout the drawing activity, success in the Draw-A-Design task is more influenced by simultaneous processing than planning.

The difference between the DAP and the Gesell Incomplete Man (Ilg, Ames, Haines, & Gillespie, 1978), tasks which have been called "analogous" by Kaufman & Kaufman (1977, p.143), also illustrates how simultaneous and planning processes may be utilized in drawing. The Incomplete Man presents the subject with an already drawn figure which has important parts missing. The child must have a conception of the whole human figure and the interrelationship of its parts in order to complete the figure. However, the child does not have to struggle with decisions about initial size of the figure, placement of the figure on the page, or proportions between various parts. Similar to the
design copying task, a model is available for many of the parts which need to be added.

The Incomplete Man and Draw-A-Design diminish planning demands but require simultaneous processing. When drawing the human figure, however, an initial plan is particularly important for there are no guidelines other than the edges of the paper which suggest where to begin or what size the figure should be. These differences between the DAP, Draw-A-Design, and the Incomplete Man make it clear that the cognitive demands of the DAP include both planning and simultaneous processing and that the demands of the task influence the process or processes used.

The DAP and Conceptual Understanding

The correlation between the DAP and the Bracken Basic Concept Scale (BBCS) shows the human figure drawing task is significantly related to this measure of conceptual understanding. The concepts presented in the BBCS are representative of those used in preschool and primary assessment instruments and curriculum materials (Bracken, 1984). In the BBCS manual, Bracken distinguishes between intelligence and understanding of basic concepts. He states that "while a child low in intelligence cannot be taught directly to be more intelligent, a child low in conceptual knowledge can and should be taught the basic concepts he or she does not understand" (p. 9). The finding of a significant relationship between the BBCS and DAP can be viewed as support for Harris' (1963) contention that the human figure drawing task should be regarded as a measure of conceptual maturity. It also suggests
that the child's exposure to environmental and educational experiences influences success on this task.

Defining the DAP as a measure of conceptual maturity does not provide information about the cognitive processes which must be applied for efficient acquisition of the concepts needed for success in the DAP and the BBCS. This study shows that the BBCS is significantly related to all four components of the PASS model but the most significant relationship is with simultaneous processing. The correlation of the BBCS with simultaneous processing is significantly greater than that of the BBCS with attention, planning, or successive processing as determined by tests of correlated correlations. Additionally, factor analysis of the eight PASS tasks and the BBCS results in a strong loading of the BBCS on the simultaneous factor. Thus, the BBCS can be considered a measure of simultaneous processing for this sample.

The significant relationship of the DAP to the BBCS lends further support to the importance of simultaneous processing in the human figure drawing task. Regression analysis of the DAP-BBCS relationship also provides support for the contribution of planning to the DAP. When the four PASS components and the BBCS are compared for significant contributions to DAP variance, the BBCS makes the greatest contribution and absorbs the contributions of three of the four PASS components. Planning, however, adds a small but significant contribution.

There is evidence from regression analysis to support the view that the DAP is related to basic concepts even when all of the PASS components are partialled out. It should be noted, however, that the additional increase in variance accounted for is minimal. Nevertheless, other factors which may be
contributing to mastery of basic concepts must be considered. Bracken's contention that concepts can be taught supports the likelihood that the child's cultural background and school experience may be variables which are reflected in the DAP-BBCS relationship. When Goodenough (1926) found differences between DAP scores earned by children from urban areas such as New Jersey and California and rural areas of the south, she interpreted these as showing "a true difference in the average mental capacity of the several groups" (p.40). However, modern researchers suggest that cultural background (Koppitz & Casullo, 1983) and educational experience are important variables related to the acquisition of concepts. McWhinnie's (1971) review of DAP research cited numerous studies which support this view. One is the study by Delta (1967) which found that children enrolled in Head Start performed at a significantly higher level on the DAP than counterparts who did not have similar educational experiences. Other evidence for the influence of culture and education is provided by Hilger et al. (1976) and Weiss (1981). Hilger found significant differences in drawings made by two racially and culturally different groups in Japan and attributed these differences to the minority status of one of the groups. Weiss attributed significant differences between the drawings of Peruvian children to the degree of exposure to formal education. It is possible that cultural background and educational experience are elements which are part of the relationship between the BBCS and the DAP. Some support for this is found in the significantly different BBCS scores earned by the two school groups. Only the PASS Total, which correlates highly with the BBCS, shows a similar difference when subjects are grouped by school.
The DAP and Motor Functioning

The correlation between the DAP and the McCarthy Motor Scale shows there is a significant relationship between this measure of motor functioning and the human figure drawing task. When the Draw-A-Child subtest is removed and the remaining subtests grouped to form the McCarthy Motor Factor, a significant relationship still appears. When the Motor Scale subtests are separated into Fine and Gross motor categories, significant correlations with the DAP again emerge. A significant correlation is also found when Draw-A-Design is examined singly. Thus, in this study, the DAP is related to both gross and fine motor functioning. This finding suggests that Harris (1963) should have investigated further the significant correlation which he found between the Goodenough-Harris DAP and the motor component of the SRA Primary Abilities Test in a study of 164 kindergarten children.

The high correlation between the DAP and the Fine Motor scale (Draw-A-Child and Draw-A-Design) reflects the influence of the Draw-A-Child subtest. Previous studies show a high correlation between this subtest and the more comprehensive Goodenough-Harris scoring system (Naglieri & Maxwell, 1981) as well as a high correlation between the Goodenough-Harris and Naglieri scoring systems (Naglieri, 1988, Gottling, 1985). The correlation between the DAP and the second fine motor task, Draw-A-Design, is similar to that found in previous studies by Hartman (1972) and Oakland and Dowling (1983) who used the Goodenough-Harris DAP and the Bender Visual-Motor Gestalt Test. The interpretation of figure copying tasks as simultaneous processing has been discussed in a preceding section.
The significant correlation between the DAP and the Gross Motor Factor is a new finding which must be considered when human figure drawings are interpreted for this age group. Kaufman & Kaufman (1977) write that the McCarthy "gross motor tests are noncognitive, whereas the fine motor tests are cognitive" (p. 143). However, the significant correlation between the Gross Motor Factor and the DAP, BBCS, PASS total, Planning, and Attention challenge this assertion.

An interpretation of the McCarthy Gross Motor Factor's correlation with the DAP and the BBCS is suggested by Bracken (personal communication, January 25, 1990). He believes the construct shared by these 3 measures is the child's conceptual understanding. According to Bracken (1985), understanding of 38 basic concepts is necessary for a child to follow the instructions used in the *McCarthy Scales of Children's Abilities* (1972). All of these concepts are incorporated into the BBCS (Bracken 1984). However, inspection of the BBCS concepts which correspond to the Gross Motor task directions indicates they are mastered by over 90% of children the same age as those in this study (Bracken, 1985). Additionally, demonstration routinely accompanies the verbal direction and more than one trial is often mandatory if the child does not earn the maximum score. Therefore, the Gross Motor tasks with their emphasis on demonstration and practice appear less dependent on the understanding of verbal concepts than other sections of the MSCA.

Another explanation for the significant relationship between the BBCS and the Gross Motor tasks is that both are influenced by two components of the PASS model, planning and attention. It is unlikely that simultaneous processing, which the BBCS primarily represents, accounts for the relationship
since the Gross Motor tasks have an insignificant correlation with this processing component. However, both the BBCS and the Gross Motor tasks correlate significantly with planning and attention.

Further understanding of the cognitive nature of the Gross Motor scale is gained through factor analysis. This analysis places the Gross Motor scale on the factor formed by PASS planning and attention tasks. The planning tasks load strongly and the attention tasks moderately. Thus, the Gross Motor scale can be described as a measure of planning and attention for this sample.

The establishment of Draw-A-Design and the BBCS as simultaneous tasks and the Gross Motor tasks as a planning and attention task affirms an important aspect of the PASS model; namely that the processes needed to solve tasks are independent of modality, type of stimuli, and method of presentation (Naglieri et al., in press). Figure Memory and Draw-A-Design are visual-spatial tasks which require a motor response and Matrices is a visual-spatial task which requires either a verbal or pointing response. However, the BBCS, which loads on the same factor, is a verbal-visual task for which either a verbal or simple motor response is appropriate. The two tasks used to identify the planning component of the model, Visual Search and Planned Connections, and one of the tasks used to identify the attention component, Number Finding, are paper and pencil tasks which are performed while the child is seated. In contrast, the gross motor tasks are whole body responses which take place while the child is standing. The significant correlations of the DAP with these measures indicate the DAP is a complex task related to three important measures, all of which fit within the PASS processing model.
These relationships affirm the DAP’s validity as a task influenced by the
cognitive processes of simultaneous coding and planning.

The DAP TT and the Teacher Version, Child Behavior Checklist

The hypothesis that the DAP scoring system developed by Naglieri
would correlate with a behavior checklist was weakly supported in this study.
These results are similar to the work of Prytula and Thompson (1972) and
Delatte and Hendrickson (1982) who found minimal relationships between the
Goodenough-Harris DAP and measures of anxiety, self-esteem, and self-con-
cept in pre-adolescent children.

Although this study suggests that the human figure drawing task is of
limited usefulness as an indicator of emotional problems and social malad-
justment in kindergarten age children, a conclusion that it has no possibility of
usefulness is premature. Some of the weakness in the DAP-CBCL relation-
ships may be due to the statistical problems created by the nature of a behav-
ior problems checklist. Scores from such a checklist form a highly skewed
distribution as most children earn scores in the normal range of behavior. This
is true in this study and likely contributes to the minimal correlations with the
DAP. Also contributing to the skewed distribution is the grouping of raw score
data for a number of CBCL T scores, a procedure which further condenses
the data. However, when raw score data was analyzed, significant correla-
tions between the DAP and the CBCL were still not attained when age was
partialled out.

It is possible that other criteria for scoring a human figure drawing such
as those presented by McNeish (1989) might show a significant correlation
with an instrument such as the CBCL. McNeish (1989) created and provided some validity support for a 38 item scale which differentiates between the drawings of elementary age children with clinical problems and normal children. Some items appear in the DAP projective literature and many emerged from analysis of drawings from the sample used to standardize the *Draw A Person: A Quantitative Scoring System* (Naglieri, 1988). Although the relationship of the DAP to personality appears minimal in this study, it is possible that additional criteria for scoring the DAP, better methods of grouping data from a checklist, and other forms of statistical analysis would support the appropriateness of the DAP as a screening measure for personality problems in young children.

**The Influence of Gender**

No differences emerge between boys and girls on any of the measures used in this study. This result does not support previous research which finds differences between boys and girls in their motoric, cognitive, verbal, and behavioral development. For example, when Kaufman and Kaufman (1977) investigate the McCarthy Motor Scale, they find significant differences in favor of boys on the Arm Coordination subtest and in favor of girls on the Leg Coordination and Draw-A-Child subtests. This is attributed to boys "greater experience in catching and throwing" (p. 100) and to girls "more rapid maturation" (p. 100). The subjects in this study do not differ on any of these subtests. The similarity in the motor functioning of boys and girls supports the work of Laszlo and Bairstow (1985) who use a process approach to measure motor skills. They find that "Motor programming processes, as measured by novel
tasks, develop at an equal rate and to an equal degree in boys and girls, indicating that there is no innate sex differences in motor abilities" (p. 162). The lack of sex differences in the PASS processes which the McCarthy motor tasks represent lends further support to Laszlo and Bairstow's conclusion.

Sex differences which are frequently noted in two other areas do not appear in this study. One area is in scores earned on cognitive tests of language, spatial perception, or mathematical proficiency (Gardner, 1982) and the second is the predominance of boys diagnosed as suffering from Attention Deficit Disorder with or without Hyperactivity (Gardner, 1982, Howell, Huessey, & Hassuk, 1985). In this study, boys and girls do not differ in their performance on a verbal measure, the BBCS, spatial measures such as Figure Memory, Draw-A-Design or Matrices, or the mathematical concepts included in subscales of the BBCS. Boys and girls also do not differ on two measures of attention, one based on performance and one based on teacher ratings. This latter finding is of interest since it raises the question of when Attention Deficit Disorder is diagnosed and on what basis. It is particularly interesting since the age of the children participating in this study matches the DSM III-R and Barkley (1987, 1978) criteria for age of onset of ADD-H.

As noted in the literature review, boys and girls performed differently on the DAP using the criteria developed by Goodenough and expanded by Harris. The newest scoring system (Naglieri, 1988) uses general categories within which typical items chosen by boys and girls earn equal credit. No significant difference between boys and girls appeared during standardization (Naglieri, 1988) and the results of this study support this finding. The failure of sex differences to emerge on any of the measures used in this study means
that the relationships which have been established between them and the DAP are not confounded by gender differences.

**Age Effects**

One of the reasons for studying the DAP performance of kindergarten children was its frequent use in assessments which determine entrance into school, promotion to the next grade, and eligibility for special services. Both schools participating in this study used a pre-kindergarten evaluation for the purpose of estimating the child's readiness for school. It is assumed that children who are ready will do well and children who are not ready will do well if they wait another year to enter school (Ilg et al., 1978). The ages of 15% of the children in this study indicate that they were either held back a year before entering kindergarten or were repeating a year of kindergarten. A significant negative correlation of age with tasks related to schooling indicates that these older children in general are scoring poorly. The benefits of delayed entrance into school are not clear.

**Future Research**

The relationships between the DAP and the variables used in this study need to be investigated for other age groups in order to understand if developmental factors influence the DAP. The literature suggests that planning, particularly, is a cognitive process which is physiologically and behaviorally developmental.

The relationship between the DAP and children's educational, cultural and economic backgrounds needs to be established for this age group and
others. It is possible the negative correlation between age and many of the variables used in this study is related to these factors.

One of the criticisms of the DAP is that some children who earn high scores on the WISC-R or Stanford-Binet earn much lower scores on the DAP (Scott, 1981, Reisman & Yamokoski, 1973, Lehman & Levy, 1971). Naglieri (1989) suggests that the WISC-R and Stanford-Binet are primarily reflections of simultaneous and successive coding processes. Since the DAP also reflects planning, it is possible that inefficient application of this important process is related to low DAP and high intelligence test scores. This is an area which needs further study.

The arguments presented in the literature to support personality evaluation through drawings are powerful but have continued to elude experimental support. It is possible that McNeish’s criteria, described previously, may be applicable to the drawings of kindergarten age children as well as older children. It would be of interest to discover if McNeish’s 38 items or a combination of these with the Naglieri system would show a significant correlation with the CBCL.

Perhaps the most important area for future research is increased understanding of the relationship between the DAP and the cognitive components of the PASS model for this age group. Further studies of the PASS model should include use of more than two tasks per process and a wider mix of modality and content within these tasks. A unique characteristic of the PASS model is its capacity to cut across traditional methods of measuring cognitive abilities through content (verbal or non-verbal) and modality (auditory, visual, motoric). In this study, visual presentations with simple
(pointing) or complex (drawing) motor responses predominated. Other PASS tasks are available with different combinations of the visual, auditory, or kinesthetic modalities, verbal and non-verbal content, and materials other than paper and pencil (Naglieri et al., 1989) and should be tested for their relationship to the PASS model and the DAP.

A second area related to the PASS model which needs further investigation is the dual loading of one of the attention tasks on the first and fourth factors, defined as planning and attention respectively. Although a close relationship between these two processes is anticipated given Luria’s description of the physiological and behavioral connections between them (1973, 1980), other tasks might be more efficient measures of attention for this age group. It is important to identify tasks which differentiate between these two processes so that the interpretation of the DAP as a task related to planning is clear.

Another area for investigation is the relationship of the DAP to complex concepts. The strong relationship between the DAP and the simple one word BBCS concepts appears to be related to simultaneous processing. A question for future research should be whether understanding of more complex concepts is also significantly related to the DAP and to simultaneous processing. Examples of tasks which measure complex concepts are the Conceptual Grouping subtest of the McCarthy Scales of Children's Abilities (1972) and the Tokens task (Naglieri & Das, 1988). These subtests require the ability to understand verbal directions and to combine concepts such as size, shape, color, number, and position. It is important to establish how these
tasks fit within the PASS model for the kindergarten age group and to establish their relationship to the DAP.

Conclusion

This study addresses the issue of the validity of the human figure drawing task. It was labelled by Goodenough (1926) as an intelligence measure, by Harris (1963) as a conceptual measure, by Dunn (1967) and Gayton et al. (1974) as a measure of non-verbal ability and by Machover (1949) as a personality measure. These labels persist in the literature and in psychological practice despite insufficient or conflicting support.

This study shows that tests which purport to measure several of the variables listed above are significantly related to the DAP for the population which participated in this study. These variables include cognitive processes included in an expanded definition of intelligence (the PASS tasks), concept understanding (BBCS), and non-verbal ability (MAT-EF and Draw-A-Design). In addition, a variable virtually ignored in the literature, motor functioning, is also significantly related to the DAP.

However, the most important finding in this study is that all of the variables correlate significantly with one or more of the PASS components. This means that the previously vague definitions such as intelligence, verbal, non-verbal, or visual-motor ability, concept understanding or conceptual maturity can all be interpreted within the Das-Luria theory of cognitive processing. In this study, the DAP is influenced by simultaneous and planning processes, the BBCS and Draw-A-Design by simultaneous processing, the Gross Motor scale and the CBCL by planning and attention. If future studies support and
expand these findings, it will be possible to interpret DAP scores as well as scores of other frequently used tests within a coherent theoretical framework.
REFERENCES


Army Individual Test of General Ability (1944).


APPENDIX A
PARENT NOTIFICATION LETTER

April 12, 1989

Dear Parents,

Your school has agreed to participate in a study I will be conducting during the next few weeks. My work will be supervised by Jack A. Naglieri, Ph.D., a member of the faculty of The Ohio State University. Your child will be asked to complete tasks commonly used in kindergarten screenings or individual evaluations. For example, your child will be asked to draw pictures of a man, a woman, and self, pick a picture from a group of four which best corresponds to a word, repeat a series of numbers or words, or attempt simple motor tasks such as standing on one foot, skipping, and catching and throwing a beanbag. These activities will require approximately one hour of your child's participation. The purpose of the study is to understand the relationships between these tasks for children of kindergarten age. I believe your child will find the tasks interesting and enjoyable.

Your child will not be identified by name and only general individual information such as age and sex will be collected. The test sessions will be scheduled so that your child will not miss any special school events.

I appreciate the cooperation of the Reynoldsburg school district, your principal, Ms. Dawson, and your child's kindergarten teacher, Mrs. Maurer, in supporting this important study. I look forward to meeting with your child in the next few weeks.

Sincerely,

Suzanne H. Gottling
Ph.D. Candidate,
School Psychology,
The Ohio State University

Jack A. Naglieri, Ph.D.
Advisor