1997

Predictions of Urban, Ethnically Diverse, Early School-Age Children's Academic Achievement Using a Neuropsychological Screening Battery

Stephen R. Clingerman
Loyola University Chicago

LOYOLA UNIVERSITY OF CHICAGO

PREDICTIONS OF URBAN, ETHNICALLY DIVERSE, EARLY SCHOOL-AGE CHILDREN'S ACADEMIC ACHIEVEMENT USING A NEUROPSYCHOLOGICAL SCREENING BATTERY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

BY

STEPHEN R. CLINGERMAN

CHICAGO, ILLINOIS

JANUARY, 1997

ACKNOWLEDGMENTS

The author is very grateful and deeply indebted to Karen Wills, Ph.D., for chairing this dissertation committee and for her encouragement, guidance, and support. This writer expresses his appreciation to Joseph Durlak, Ph.D., Alan DeWolfe, Ph.D., and Martha Ellen Wynne, Ph.D., for their continuing contributions to my research and clinical knowledge through their comments and suggestions as committee readers. Thanks are extended to J. Clifford Kaspar, Ph.D., for providing inspiration for this project, and for permission to reprint the directions and stimuli from the Cards Test. The author appreciates the assistance of Sister Joyce, Principal of St. Jerome's Grammar School, and Ms. Rosealee Keane and Ms. Nancy Kelly, Principals of St. Ignatius Grammar School, for allowing the data collection at their institutions. Thanks are expressed to the participating children, parents, and teachers for their cooperation, as well as to the undergraduate psychology student examiners and their Loyola University instructors. Finally, this writer thanks his wife, Cindy Antonsen, for her support.

To the memory of Mary Beery, my grandmother

TABLE OF CONTENTS

Page

ACKNOWLEGEMENTS iii

LIST OF TABLES vii

ABSTRACT ix

Chapter

I. INTRODUCTION 1

Aims of this Study 1

Significance of Early Screening 2

Contributions of Neuropsychology to Predictive Screening 3

II. REVIEW OF RELATED LITERATURE 7

Theories of Learning Abilities 8

LD Subtyping and Test Selection 10

Pediatric Neuroanatomical Evidence of Alexia and Acalculia 12

Alexia and Dyslexia 14

Acalculia and Dyscalculia 20

Use of Neuropsychological Tests in Predicting Young Children's Academic Achievement 22

General Conclusions of the Literature Review 33

III. METHOD 34

Participants 34

Procedure 37

General 37

Examiners 37

Measures 38

Development of Predictive Screening Battery and Data Reduction 38

Development of Criterion Variables and Data Reduction 51

Development of Partitioning of Achievement Distribution 53

Main Research Questions and Analyses 53

Discriminant Function Analyses to Test Relationship of Predictor to Criterion Variables 53

MANOVA/t-Tests to Assess Cultural Fairness of Measures 55

Exploratory Research Questions and Analyses 57

Discriminant Function Analyses to Assess Cultural Fairness 57

Discriminant Function Analyses to Assess Age Differences 57

IV. RESULTS 60

Main Research Questions and Analyses 60

Discriminant Function Analyses to Test Relationship of Predictor to Criterion Variables 60

MANOVAS/t-Tests to Assess Cultural Fairness of Measures and Effects of Handedness 80

Exploratory Research Questions and Analyses 84

Discriminant Function Analyses to Assess Cultural Fairness 84

Discriminant Function Analyses to Assess Age Differences 110

V. DISCUSSION 133

Six Aims of the Study 133

Issues in Predictive Accuracy of Screening Batteries 134

Issues in Test Selection and Strength of Predictors 139

Exploratory Issues in Prediction 146

Predicting Different Achievement Criteria 157

Limitations of the Study and Further Suggestions for Future Research 159

Conclusion 164

Appendix

1. LETTER OF PERMISSION 167

2. CARDS TEST 169

REFERENCES 172

VITA 189

LIST OF TABLES

Table Page

1. The Relationship of Presumed Brain Substrates of Alexia and Acalculia with Initial Screening Battery Measures 18

2. Early Neuropsychological Screening Batteries' Measures 25

3. Early Screening Batteries' Tests Classified as a Function of Five Neuropsychological Areas 27

4. Townes et al. (1980)-Teeter (1985) Battery as a Function of Five Neuropsychological Areas 31

5. Ethnic Background, Gender, and Mean Age of Participating Children by Grade 35

6. The Project's Final Battery Measures, Means, and Standard Deviations as a Function of Six Neuropsychological Areas 40

7. Correlations of Sub-Test Scores within Three Composite Neuropsychological Domains 49

8. Correlations Among the Final Battery's Six Neuropsychological Domains 52

9. Three-group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence) 61

10. Three-group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence) 66

11. Two-group Classifications of Achievement Predictions and Predictors Selected in Analyses (Including Intelligence) 71

12. Two-group Classifications of Achievement Predictions and Predictors Selected in Analyses (Excluding Intelligence) 75

13. The t-Tests Between Ethnic Minority and Ethnic Majority Children's Scores on Criterion Variables 81

14. Ethnic Group Means and Standard Deviations on Predictors and ANOVAS between Scores of Ethnic Minority and Ethnic Majority Groups 83

15. Three-group Classifications of Achievement Predictions and Strengths of Predictors for Two Ethnic Groups (Including Intelligence) 86

16. Three-group Classifications of Achievement Predictions and Strengths of Predictors for Two Ethnic Groups (Excluding Intelligence) 94

17. Two-group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence) for Two Ethnic Groups 97

18. Two-group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence) for Two Ethnic Groups 104

19. Three-group Classifications of Achievement Predictions and Strengths of Predictors for Two Age Groups (Including Intelligence) 112

20. Three-group Classifications of Achievement Predictions and Strengths of Predictors for Two Age Groups (Excluding Intelligence) 117

21. Two-group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence) for Two Age Groups 122

22. Two-group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence) for Two Age Groups 128

ABSTRACT

This study developed a relatively culture-fair neuropsychological screening battery to predict Low Average, Average, and High Average academic achievement in an ethnically diverse, urban, young school-age population. Children in grades kindergarten through third participated (N = 64). Of the total, 41% were Ethnic Minority children and 59% were Ethnic Majority children, while 49% were Younger children (kindergarten and first grade) and 51% were Older children (second and third grade). The predictive interval was 6 months.

The predictor measures sampled six neuropsychological domains: Intelligence, Language, Attention-Memory, Visual-Spatial, Somesthetic, and Fine Motor. The criterion achievement measures were based upon the three WRAT-R subtests (Jastak & Wilkinson, 1984): Combined Academic (average of the three subtest standard scores), Language-Related (average of the Word Recognition and Spelling subtest standard scores), and Arithmetic (Arithmetic subtest standard score). Low Average (standard score < 90), Average (standard score = 90 - 110), and High Average (standard score > 110) achievement was predicted for each criterion.

Stepwise discriminant function analyses (with and without the Intelligence predictor) were used to predict the three levels of achievement (overall analyses) and examine ethnic-group and age-group comparisons. Two-group analyses, Low Average versus "Average-Plus" (combined Average and High Average scores) also were performed. Student's t-test, MANOVA, and ANOVA analyses were performed to test ethnic group differences.

The predictive accuracy rates (6-months) for the three achievement groupings were

significant for each of the criterion variables (44-64%) and accuracy using two achievement groupings was higher (63-84%). The battery's accuracy generally is comparable to that of other studies. There were ethnic-group differences and age-group differences in predictive accuracy and in the strength of predictors.

The more complex multifactorially determined predictors (Intelligence and Attention-Memory) were the most sensitive predictors, but the more "basic" Somesthetic and Fine Motor tasks also proved useful. Language and Visual-Spatial predictors also represented differentiated, discrete areas of functioning that correlated with Intelligence and contributed independent variance to prediction.

CHAPTER I

INTRODUCTION

This study attempted to develop a relatively culture-fair neuropsychological screening battery to predict low average, average, and high average academic achievement in an ethnically diverse, urban, young school-age population. The first part of this introduction will focus on the aims of the study. The second section of the introduction will examine the significance of early identification of children likely to achieve above or below the average range. The third part of the introduction will discuss the importance of neuropsychology to prediction and screening in the early school years.

Aims of the Study

This study sought to meet several needs in the predictive screening of young children's academic achievement. First, improvement in short-term accuracy prediction rates are needed for all screening batteries (Satz & Fletcher, 1988), especially in assessing average and mildly impaired children. Second, predictive work identifying children with academic talents or strengths is needed, as most learning research has been devoted to identifying academic problems (Jansky, 1978). The present investigation attempted to meet both needs by studying children in regular classrooms and by seeking to predict high average, average, and low average academic achievement.

Third, more work with predictive screening using young school-age children is needed (Deysach, 1986), as most neuropsychologists investigating academic abilities study older children (i.e., 8 years of age and older). The children in this study (ages 5-8 years old) were fairly young compared to many studies.

Fourth, "culture-fair" screening measures are needed for urban, low income, and ethnically diverse populations (Morrison & Hinshaw, 1988). This study attempted to generate a relatively culture-fair screening battery by developing a stronger rationale guiding test selection (compared to that used in prior studies) and by including ethnic minority children as participants. More than two-fifths of the present study's sample was comprised of ethnic minority children.

Fifth, better use of neuropsychological tests in predictive batteries is needed (Hinshaw, Carte, & Morrison, 1986), as most screening batteries have ignored many domains (e.g., fine motor and memory) that recent research has shown underlie academic abilities (Deysach, 1986). This project attempted to assess fine motor, attention/memory, and language domains in more comprehensive ways than previous studies.

Sixth, prediction and screening studies of academic achievement have focused almost exclusively on reading, while it has recently been noted that prediction of spelling and mathematics abilities also is needed (Teeter, 1985). The present study predicted overall academic, language-related, and mathematics achievement.

Significance of Early Screening

The importance of early identification of variations in children's academic achievement is based upon two assumptions. First, it is assumed that educational programs and other "environmental interventions" may change and shape developing central brain functions in children (Shapiro, Palmer, Wachtel, & Capute, 1984). Second, many investigators think that central nervous system plasticity decreases with age (Deysach, 1986; Shapiro et al., 1984). Therefore, early identification may provide greater opportunities to provide individualized curricula and instruction that promote better neural development in all children during the time of greatest plasticity, which presumably would maximize cognitive and academic gains.

These hypotheses have implications for children across all levels of achievement. Early identification of children with learning disabilities (LD) is needed because LD children often have had years of school failure by the time resources typically are mobilized to help (Cruikshank, 1968; Green, Lyles, & Eissenfeldt, 1980; LaTorre, 1985). Perhaps because of these delays and school failure, learning disabled children have a higher incidence of emotional and psychiatric disturbances than other children (Kroll, 1984). Follow-up studies of "early identified" learning disabled children demonstrate that they have improved long-term outcome compared to their counterparts identified "late" (Muehl & Forell, 1973; Shenk, Fitzsimmons, Bullard, & Satz, 1980). Early identification and remediation might reduce emotional problems among these children as well as provide cognitive interventions during the time of greatest plasticity.

Screening batteries also are useful in identifying average students and gifted children (Jansky, 1978). Prediction and identification of average and gifted students receive less attention in the literature than assessment of LD students, but children performing in the average or above average academic ranges are important for all prediction studies, as they represent the complement of the learning disabled population. More accurate prediction of the non-impaired groups would lower false positive rates for the LD children and lower false negative rates for other children. Also, identification of gifted students may allow them placement in more challenging academic milieus using more demanding curricula.

Contributions of Neuropsychology to Predictive Screening

Neuropsychology, the study of brain-behavior relations, provides a way of thinking about human functioning utilizing cognitive domains and known or hypothesized behavioral-neuroanatomical associations (Luria, 1966). Areas of ability including intelligence, language, attention, memory, motor, and sensory-perception comprise distinct cognitive domains (Lezak, 1983). Neuropsychology also involves certain brain-behavior principles that allow one to infer the integrity of neural structures. For example, the knowledge that the body is largely "cross-wired" for motor and sensory functions often is useful in assessing lateralized damage (Reitan & Wolfson, 1985).

The application of neuropsychological theory to children and academic achievement has led directly to improved prediction of academic achievement, as well as identification, diagnosis, and remediation of learning disabled children's deficits (Gaddes, 1980; Obrzut & Hynd, 1991; Rourke, 1989, 1991; Silver & Hagin, 1990). Neuropsychological screening batteries predicting academic skills have achieved short term (1-3 years) accuracy rates of 70-75% and false positive rates of 20-30% in the identification of moderately and severely impaired readers (Jansky, 1978; Satz & Fletcher, 1988; Silver & Hagin, 1990). Adequate short-term classification of average versus mildly impaired children has not yet been achieved (Jansky, 1978; Satz & Fletcher, 1988; Silver & Hagin, 1990). Ultimately, in order for screening batteries to prove useful to educators and psychologists, predictive accuracy rates must improve and false positive rates must decline (Silver & Hagin, 1990).

Improving predictive accuracy using neuropsychological screening batteries may be difficult. Gaddes (1981) and Spreen (1978a) note that the predictive screening batteries' short-term accuracy (70-75%) may represent a "ceiling or optimum" and suggest that the rest of the variance may be attributable to such factors as motivation, the quality of teachers, and family views on education. Silver (1978) and Jansky (1978) point out that false positive rates (20-30%) may be lowered by lowering the "vulnerability" cutoff, although at the cost of increasing false negatives.

Recent neuropsychological predictive screening protocols have begun to broaden their sampling of attention/memory, motor, and sensation-perception abilities (Teeter, 1985; Townes, Turpin, Martin, & Goldstein, 1980), while achieving similar rates of success compared to other batteries. These studies suggest that more work on test selection in these areas may prove fruitful in improving prediction and reducing false positive rates. Also, the neuropsychological predictive screening literature has been marked by a lack of theory in test selection (Satz & Fletcher, 1988) that may have limited predictive success. The present study sought to incorporate theory (as will be discussed later) in selecting tests.

Another primary need in early screening is development of culture-fair test batteries suitable for low income and minority populations (Morrison & Hinshaw, 1988). Neuropsychological measures that draw upon the cognitive underpinnings of academic abilities better than traditional educational measures (Silver, 1978), may yield a more equitable battery for use with ethnic minority and low socioeconomic children (Morrison & Hinshaw, 1988). This study sought to select measures that would give rise to a culture-fair battery.

Several competing hypotheses have been generated about how the predictive power of the various neuropsychological domains may change as development proceeds. Satz and associates (Satz, Taylor, Friel, & Fletcher, 1978) suggest that perceptual-sensory-motor markers of academic problems are more useful in younger children and that language and conceptual ability markers are more useful in older children. In contrast, Silver and Hagin (1990) indicate that perceptual, sensory, and motor processes are important as "red flags" of academic problems in older as well as younger children. Jansky (1978) assumes that language-based abilities may be equally predictive of academic abilities in younger children as in older children. The present study sought to confirm Satz's contention that optimal neuropsychological predictors of academic achievement vary with age.

Many reviews have discussed thoroughly the importance and relevance of neuropsychological tests to learning disability classification and diagnosis (e.g., Denckla, 1979; Hynd & Cohen, 1983; Mattis, 1978; Obrzut & Hynd, 1991; Rourke, 1989, 1991), but few investigators have noted the importance of neuropsychological tests in assessing average and above average achieving children (Jansky, 1978). In past studies, average and above average children have been combined in distinguishing them from impaired children. However, because scores on most neuropsychological tests are normally distributed, there is good reason to think that use of these measures could predict average and above average achievement, as well as below average scores. The present study sought to predict performance across the full range of academic achievement.

CHAPTER II

REVIEW OF RELATED LITERATURE

The review of the literature will proceed as follows: Two areas of literature related to learning ability will be reviewed. First, the theories of learning ability will be noted. Second, the literature on subtyping of learning disabled children will be surveyed. The brief subtyping review will focus on identifying test domains that measure component neuropsychological functions underlying academic achievement. Assuming that abilities are normally distributed, the study of LD subtypes may reveal the relevant functional domains that require assessment in order to predict the full range of academic achievement.

Next, two areas related to neurological and neuropsychological bases of academic achievement will be reviewed. First, the literature regarding areas of the brain that may be critically involved in reading and mathematics, based on children with focal brain damage, will be reviewed. The neuropsychological literature provides the best available evidence with which to identify brain functions that need to be assessed (and evidence on how to select tests to assess those functions) in order to predict academic achievement in all children. Assuming academic skills are normally distributed, tests that predict the lowest part of the distribution also may predict the middle and upper levels. Lastly, the methodology, test selection, and predictive results of previous neuropsychological screening batteries for young children will be reviewed.

Theories of Learning Abilities

Children in regular classrooms often have been used in early prediction studies, but the focus of concern and theories about academic achievement has been learning disabled children, not average and above average achieving children (Denckla, 1979; Pirozzolo & Campanella, 1981). Theories developed to account for LD may be extended, however, to account for all ranges of academic functioning. Theories of learning were once classified according to "nature-nurture" or "heredity-environment" etiology but the dichotomy is now mostly a heuristic device, as it is now known that the extreme positions are not independent, mutually exclusive constructs.

Adherents of the nature-heredity position presumed that genetic causes, illnesses, and injuries led to "bad wiring" and intrinsic learning problems (Critchley, 1966; Hermann, 1959; Silver & Hagin, 1964). The strongest evidence of the genetic role in learning disabilities are the results of twin studies and the consistent findings of boys' higher prevalence rates of developmental disabilities (DeFries & Gillis, 1991; Harris, 1986; Lubs et al., 1991; Smith, 1986; Smith, Pennington, Fain, & Ing, 1989). Studies of brain damage and neurological disorders in children provide additional evidence that genetic and innate factors are associated with long-term academic disabilities (Denckla, 1979; Pirozollo & Campanella, 1981).

Analogously, it was presumed (but not usually stated) that genetic causes and good health led to "good wiring" and average or above average functioning. Evidence of the genetic role in average and above average achieving children remains to be explored. The role of neural integrity in the functioning of average and above average children has not been investigated because of the invasive nature of the necessary procedures (Hiscock & Kinsbourne, 1987).

Alternately, the importance of social-environmental factors including socioeconomic status (SES), family size, and position in the sibship hierarchy has been noted (Rutter & Yule, 1975; Satz & Friel, 1978). Physical deprivation and emotional stress are cited as causes of below average achievement, and conversely, enriched environments and emotionally healthy homes are thought to give rise to average to above average abilities.

The midpoint of the nature-nurture dimension of learning theories include developmental and maturational theories in which "windows" of growth occur and in which delays may or may not be remediated (Bender, 1957; Gesell & Thompson, 1938; Satz et al., 1978). Academic disabilities are viewed as the result of lags or delays in development, and academic success is presumably the result of adequate or flourishing maturational processes. Remediation and promotion of growth occurring within developmental time-windows may be especially effective, and if the opportunities are missed, remediation may be made more difficult and growth may be stunted (Critchley, 1966; Spreen, 1978a).

Complicating the conclusions from studies seeking to explore or confirm the importance of either end of the nature-nurture continuum are several confounds of SES and brain functioning making the dichotomy apparent and artificial. For example, in low SES families, the mothers are likely to have had poor health care, poorer paying jobs, and family history of LD. Low SES children are susceptible to increased prenatal and perinatal risk factors which in turn are associated with learning disabilities. Also, it is now known that severe environmental stress and deprivation affect physiology (e.g., promote immunosuppression) and can even damage areas of the brain such as the hippocampus (Uno, Tarara, Else, Suleman, & Sapolsky, 1989). Likewise, environmental stimulation and nurturance enhance neural development and functioning such that at the average and high average SES levels, health care for mothers and children is better and certain environmental stressors are reduced (Deysach, 1986).

In recognition of the breakdown of the nature-nurture dichotomy, the current learning theories are multicausal and suggest that genetic influences, neural functioning, developmental/maturational factors, and the child's environment and health all contribute to learning abilities (Bannatyne, 1971; Deysach, 1986; Rabinovitch, 1968). Multicausal theories are most consistent with the accumulating evidence of the interaction of genetic factors, brain function, and sociocultural influences in children.

The increasing evidence of the confluence between genetic and other innate influences and the environmental and developmental factors provides a foundation upon which to suggest that neuropsychological prediction approaches with children may be particularly useful.

LD Subtyping and Test Selection

Prior to the late 1970s, most investigators maintained that a single-function deficit (that is, a deficit in one specific cognitive process) underlies the learning problems of children (Rourke, 1985). Intelligence testing led to one of the first subtypes identified: a group of children low in IQ and relatively low in all academic areas (Rutter, 1978). Using tests based on single subtypes, investigators have focused alternatively on visual-spatial impairments (Bender, 1957; Gesell & Thompson, 1938; Hermann, 1959; Orton, 1937), auditory-visual (cross modality) integration deficits (Ayers, 1975; Birch, 1962; Strauss & Lehtinen, 1947), or auditory-language deficits (Downing, 1973; Vellutino, 1978, 1991; Vellutino & Scanlon, 1985). Investigators identifying language-impaired subtypes have focused variously on naming (Denckla, 1979), or phonological (Tallal, Stark, & Mellits, 1985; Vellutino, 1978), semantic, or syntactic processing (Vellutino, 1991).

The validity of the intellectual subtypes (Morris, Blashfield, & Satz, 1986; Rutter, 1978), visual subtypes (Rourke, 1989; Silver & Hagin, 1990; Spreen, 1978a) and auditory-language subtypes (Benton, 1978; Denckla, 1979; Vellutino, 1979) is generally accepted, but the auditory-visual integration/cross-modality assimilation theories have been severely criticized by several groups (e.g., Blank, Weider, & Bridger, 1968; Rudnick, Sterritt, & Flax, 1967; Senf & Freundl, 1971) and do not appear valid (Hynd & Cohen, 1983; Vellutino, 1978).

Neuropsychologists and educators currently have largely rejected the single-function approach (Benton, 1978) and have suggested that learning disabled children's problems probably need to be conceptualized in terms of the various processes that underlie each academic activity (Boder, 1973; Doehring, 1978; Mattis, French, & Rapin, 1975; Rourke, 1978). Among the most prominent of the multiple subtyping schemata were those including tests of visual and auditory functioning (Johnson & Mykleburst, 1967; Kinsbourne & Warrington, 1963; Pirozollo, 1981). A second common subtyping schema is an extension of the two-subtype model (visual-spatial and auditory-linguistic), adding a third, combined-deficit group (Boder, 1973; Ingram, Mason, & Blackburn, 1970; Mattis, 1978; Satz & Morris, 1981). Combined-deficit children are those youngsters who have problems performing visual-spatial and auditory-linguistic tests.

Multiple subtyping models have added motor, sensory, and sequential reasoning subtypes to the previous paradigms. Test findings of abnormal motor functioning such as expressive speech problems and hand-writing impairments as well as perceptual deficits (auditory, visual, and tactile) have led investigators to suggest a "perceptual-motor" subtype (Ayers, 1975; Denckla, 1979; Frostig & Maslow, 1973; Joschko & Rourke, 1985; Mattis et al., 1975; Rourke, 1989). Many investigators have identified sequential processing deficits in groups of children (Bakker 1983; Denckla, 1979; Mattis, 1978; Rourke & Strang, 1983), although a few criticisms have been made of the validity of the subtype (Blank et al., 1968; Vellutino, 1978). More recently, Tallal (Tallal, Townsend, Curtiss, & Wulfeck, 1991) and Spreen (Spreen & Haaf, 1986) have demonstrated the importance of nonverbal sequential processing in language.

Two other subtypes that have been identified are attentional-hyperactivity problems and "no deficit" groups. Use of attention test results has identified large groups of active distractible, and impulsive children with attention, hyperactivity, and academic problems (Denckla, 1979; Kerasotes & Walker, 1983; Whalen, Henker, & Hinshaw, 1985). Several studies have identified groups of children with academic impairments lacking other known neuropsychological deficits (Morris et al., 1986; Lyon, Stewart, & Freedman, 1982; Silver & Hagin, 1990). Silver and Hagin (1990) suggest that these youngsters may have emotional problems or chaotic home environments.

In conclusion, a review of the literature indicates that it is well established that valid "learning problem" subtypes in children include intellectual, auditory-language, perceptual-motor (language-motor and graphomotor), and attention-deficit hyperactivity clusters (Denckla, 1979). Despite criticisms, the visual-spatial and sequencing/temporal order subtypes also appear valid (Benton, 1978; Tallal et al., 1985). These clusters are thought to reflect dissociable neuropsychological (or cognitive processing) domains. Tests tapping these five processing domains are potentially important in predicting all ranges of achievement and, as will be seen below, guided the selection of measures in the present battery.

Pediatric Neuroanatomical Evidence of Alexia and Acalculia

A second way to approach selection of measures in predictive screening batteries is to identify the regions of the brain implicated in learning ability and select tests known to be associated with those neural sites. Evidence of the neuroanatomical substrates of academic achievement will be subdivided into disorders of reading (alexias) and mathematics (acalculias). Assuming that neurological functions are normally distributed, a review of brain dysfunctions or lesions associated with low functioning also may provide the best available information regarding the critically functional brain sites of reading and mathematics abilities for all children. Therefore, a brief review of lesions that affect children's reading and mathematics abilities follows.

The relationship between test performance and brain functioning is based upon a chain of assumptions (Caramazza & Berndt, 1978). It is assumed that if a behavior (such as reading) is disrupted by lesioning a specific brain area, that area is critically involved in performance of the behavior by the intact brain. A complex behavior is made up of simpler component functions, and each component function is associated with certain cortical and/or subcortical areas. It is further assumed that tests which reflect such simple functions associated with particular areas will be sensitive predictors of later performance of the relevant complex behavior (e.g., reading, mathematics). A large literature on neuropsychological tests specifically associated with certain areas of the brain has emerged (Lezak, 1983; Heilman & Valenstein, 1985; Kolb & Wishaw, 1980).

The study and interpretation of pediatric acquired reading or mathematics disorders (alexias and acalculias) is complicated by several factors. First, research is very sparse and classification of the disorders difficult. Children provide fewer autopsies, physicians are less likely to use invasive procedures with children, and children rarely have vascular infarctions (the most common cause of acquired academic disorders in adults) (Hiscock & Kinsbourne, 1987; Taylor & Fletcher, 1983). Second, the investigation of pediatric alexia and acalculia is further complicated by a number of developmental issues (Taylor & Fletcher, 1983). The age of the child at onset of dysfunction or damage and his or her pre-existing knowledge and skills in reading and mathematics vary across samples. Outcome is related to the remaining integrity of associated abilities including language, attention, and memory, and individuals' capacities to recover or develop academic abilities.

Alexia and Dyslexia

Alexia may be defined as a central brain impairment which gives rise to an inability to read or understand written or printed language or the symbolic significance of words (Reitan & Wolfson, 1985). As will be seen, available research on anatomical correlates of pediatric alexia supports multisite causes (Duane, 1991; Hynd & Semrud-Clikeman, 1989) but emphasizes general left hemisphere and left temporal or bi-temporal dysfunction (Denckla, 1979).

A case study of a child (Drake, 1968) with a reading disorder implicated structural problems in the areas of the left angular gyrus, splenium of the corpus callosum, and cerebellum. The role of the various cortical and subcortical dysfunctions was unclear in Drake's (1968) patient because of the large number of areas damaged.

One neuroimaging study indicated that of learning disabled children (75% of whom had reading disabilities) presenting with subtle neurological lateralizing signs, as few as 20% (5 of 32 children) had definitively abnormal computed tomography (CT) scans (Denckla, LeMay, & Chapman, 1985). Among the five abnormal scans, four had larger than normal lateral ventricles and the fifth child had a slight midline shift to the right. Of the four children with enlarged lateral ventricles, two children had bilateral enlargements, one had an enlarged left ventricle and the fourth had an enlarged right ventricle.

Several group studies of comparative morphology indicate that dyslexic children's right temporoparietal or parietal-occipital areas are equal or larger in size to corresponding areas in the left side, in contrast to the usual pattern (left hemisphere larger than right) exhibited by normal children (Hier, LeMay, Rosenberger & Perlo, 1978; Hynd, Semrud-Clikeman, Lorys, Novey, & Eliopulos, 1990; Pirozollo & Campanella, 1981; Rumsey et al., 1986). More specifically, Haslam, Dalby, Johns, and Rademaker (1981) found that significantly more dyslexic children had symmetric temporal areas, in contrast to the expected pattern (left temporal larger than right temporal) exhibited by most normal children. These data support theories of dysfunction and agenesis in the left hemisphere (Hynd et al., 1990), and possible compensation by right-sided growth. Alternatively, inadequate pruning of the right hemisphere also would be consistent with these results. (Pruning refers to the normal developmental process by which certain synaptic connections in the brain are eliminated.)

Recent postmortem evidence from cytoarchitectonic studies of child dyslexics also reveals multiple sites of cerebral malformation or dysfunction. A review of autopsy and surgical microscopic studies revealed structural cell abnormalities in bilateral frontal, left temporal, and thalamic regions (Hynd & Semrud-Clikeman, 1989).

Right-hemisphere processing dysfunction in children with reading comprehension problems has been emphasized by Rourke (1989). Rourke's theory is supported by data showing a pattern of neuropsychological performance in some LD children similar to that shown by children and adults who have documented right hemisphere damage. Rourke (1989) suggests that right hemisphere damaged children have circumscribed learning deficits including relatively poor reading comprehension and mathematics abilities, in the presence of relatively strong word decoding abilities. He also notes a cluster of learning disabled children who have emotional difficulties (e.g., depression), as well as tactile, visual-spatial, attention, memory, and complex psychomotor deficits, all problems ascribed to right-hemisphere functioning (Rourke, 1989).

Welsh and Pennington's group has noted the importance of executive functioning, based upon the development of children's frontal (Welsh & Pennington, 1988) and prefrontal brain systems (Welsh, Pennington, & Groisser, 1991), in performing various tasks including school work. Pennington's group also uses the logic that tests which are selectively impaired subsequent to acquired frontal damage in children and adults may be used to assess integrity of frontal functions in all persons. Executive functions consist of a wide range of abilities including planning and strategizing, inhibiting impulses, maintaining cognitive sets, developing mental representations, using working memory, and employing self-monitoring (Welsh & Pennington, 1988; Welsh et al., 1991).

Studies of dyslexic and other learning disabled children's neurophysiological correlates suggest involvement of the corpus callosum, frontal motor areas, and the left hemisphere temporo-parieto-occipital junction (Hynd et al., 1990). A carefully controlled study comparing dyslexic boys' brain electrical activity to normals found that the dyslexic group exhibited lower electrical activity in the frontal motor strip and the temporo-parieto-occipital junction (Duffy, Denckla, Bartels, & Santini, 1980). The Duffy et al. (1980) study has been criticized because of the small number of dyslexic subjects ($n = 8$) and over-representation of left-handed children ($n = 4$) (Taylor & Fletcher, 1983).

The importance placed on the neurophysiological studies, comparative morphology research, and cytoarchitectonic work must be tempered. Several studies have failed to find significant differences between learning disabled and normal populations (Haslam et al., 1981; Obrzut, 1989), and criticisms of this research include methodological problems, theoretical lapses, and validity concerns (Hiscock & Kinsbourne, 1987; Obrzut, 1989; Taylor & Fletcher, 1983).

Association of subcortical areas with reading or behavioral problems has been hypothesized. Investigators have attributed attention deficits in dyslexic and hyperactive children to dysfunctions in the ascending reticular activation system (RAS) (Dykman, Wallis, Suzuki, Ackerman, & Peters, 1970, 1971) and RAS-limbic system connections (Denckla, 1979;

Hynd & Cohen, 1983; Silver, 1971). The limbic-RAS theories have been criticized as not specifically accounting for reading disorders' greater prevalence compared to mathematics disorders (Spreen, 1978a). One investigative team has suggested that cerebellar-vestibular dysfunction is central to the academic problems of these children (Frank & Levinson, 1973). Given that cerebellar dysfunction markers are absent in many dyslexic children (Spreen, 1989), and not all motor problems in dyslexic children are cerebellar in origin (Hynd & Cohen, 1983), most investigators dismiss this theory. Further study is needed to clarify the role of RAS, limbic and cerebellar structures in reading.

In conclusion, considerable documentation suggests that reading is a global, multisite brain activity, and tests comprising a predictive screening battery should probably be sensitive to the various discrete cortical areas of the brain, not merely global functioning, in order to adequately predict reading. The largest proportion of research suggests that the frontal, central, and posterior portions of the left hemisphere (in right-handed patients) are critically involved in reading. Evidence of the crucial role of the right hemisphere in reading comprehension (although not in reading "decoding") has emerged in well conducted recent studies. Interestingly, as neuroimaging, neurophysiological, and cytoarchitectonic techniques have improved, results of research have not narrowed the sites thought to be associated with reading. In marked contrast, more cortical areas associated with pediatric alexia and dyslexia have been identified. The possibility of critically functional roles of limbic, RAS, and other subcortical areas in reading has been raised recently, but confirmatory evidence is presently lacking.

Table 1 notes the various cortical areas thought to be critically associated with reading and this project's initial tests that were selected because of their sensitivity to damage or dysfunction at these sites. The left hemisphere is associated with verbal abilities and with

Table 1

The Relationship of Presumed Brain Substrates of Alexia and Acalculia with Initial Screening Battery Measures

Brain Substrates	Initial Screening Battery Measures
Left Hemisphere	Aphasia Screening Test (Reitan, 1974)
	WISC-R Information (Wechsler, 1974)
	WISC-R Digit Span and Arithmetic (Wechsler, 1974)
	Finger Tapping Test (Reitan, 1974)[a]
	Speed of Motor Performance Test (Schulman et al., 1969)[a]
	Finger Localizing Test (Reitan, 1974)[a]
	Fingertip Symbol Perception Test (Reitan, 1974)[a]
Bi-Frontal Functioning	WISC-R Arithmetic/Digit Span/Coding (Wechsler, 1974)
	Cards Test (Schulman et al., 1965)
	Finger Tapping Test (Reitan, 1974)
	Speed of Motor Performance Test (Schulman et al., 1969)
Right Hemisphere	Visual Motor Gestalt Test (Bender, 1946)
	WISC-R Coding (Wechsler, 1974)
	Cards Test (Schulman et al., 1965)
	Finger Tapping Test (Reitan, 1974)[b]
	Speed of Motor Performance Test (Schulman et al., 1969)[b]
	Finger Localizing Test (Reitan, 1974)[b]
	Fingertip Symbol Perception Test (Reitan, 1974)[b]

a. Right Hand Scores

b. Left Hand Scores

right-handed movement and sensation. Generalized left hemisphere functioning has been assessed with language tests such as the Aphasia Screening Test (Reitan, 1974) and verbal intelligence tests such as the Information, Digit Span, and Arithmetic WISC-R (Wechsler, 1974) subtests (Lezak, 1983; Heilman & Valenstein, 1985). Left parietal abilities may be assessed with right hand Finger Localizing and Fingertip Symbol Perception (Reitan, 1974) measures (Lezak, 1983; Heilman & Valenstein, 1985; Spreen, 1978a).

The frontal lobes are associated with motor control, working memory, attention, organization, and planning. Measures of fine motor speed such as the Finger Tapping Test (Reitan, 1974) and Speed of Motor Performance Tests (Schulman, Buist, Kaspar, Child, & Fackler, 1969) may be used to sample and assess bi-frontal functioning (Lezak, 1983; Heilman & Valenstein, 1985; Spreen, 1978a). Fine motor functioning on the right and left hands most consistently lateralizes to the contralateral hemisphere, allowing separate assessment of left and right hemispheres (Lezak, 1983; Heilman & Valenstein, 1985; Spreen, 1978a). Verbal attention measures such as the Arithmetic and Digit Span WISC-R (Wechsler, 1974) subtests and nonverbal attention measures such as the Coding WISC-R (Wechsler, 1974) subtest and the Cards Test (Schulman, Kaspar, & Throne, 1965) assess left and right frontal functioning, respectively, and when considered together, bi-frontal functioning (Lezak, 1983; Heilman & Valenstein, 1985; Spreen, 1978a).

The right hemisphere commonly is associated with visual-spatial pattern perception and construction, as well as left-handed movement and sensation. Right-hemisphere abilities (and specifically right parietal) are often assessed using such visual-spatial measures as the Visual Motor Gestalt Test (Bender, 1946) and the left hand measures of the Finger Localizing Test and the Fingertip Symbol Perception (Reitan, 1974) tests (Lezak, 1983; Heilman & Valenstein, 1985; Spreen, 1978a). Attempts were made to represent frontal, as well as left and right hemisphere areas particularly well (see Table 1).

Acalculia and Dyscalculia

Acalculia may be defined as a central brain impairment in the ability to understand the symbolic function of numbers and the nature of arithmetical processes (Reitan & Wolfson, 1985). Henschen first used the term "acalculia" and he described two types: alexic (number recognition impairments) and anarithmetia (calculation dysfunctions) (Benson & Denckla, 1969; Cohn, 1961; Warrington, 1982). These disorders are associated with left hemisphere damage in adults (Levin & Spears, 1985). A third type of arithmetic disorder, spatial acalculia, was first identified by Hecaen, Angelergues, and Houiller (as cited in Levin & Spiers, 1985). Patients with spatial acalculia may invert numbers, may have deficits in use of columns and rows, and may have impairments in carrying operations (Benson & Weir, 1972; Dahmen, Hartje, Bussing, & Sturm, 1982). This spatial disorder is associated with right hemisphere damage in adults (Levin & Spears, 1985).

The pediatric literature in the area of acalculia is smaller than the children's alexia literature. Case studies, usually of two to four children, reveal a variety of deficits in all 3 of the acalculia types: impairments in reading numbers and signs, deficits in mathematical visual-spatial abilities, and impairments in performing calculations (Levin & Spiers, 1985). Posterior right hemisphere dysfunction is often considered the cause of children's arithmetic problems in visual-spatial domains, especially in early school years (Gaddes, 1981; Johnson & Mykleburst, 1971). Semrud-Clikeman and Hynd (1990) acknowledge the importance of temporo-parieto-occipital left-hemisphere functioning in mathematics abilities, but note several evoked potential studies that implicate posterior right hemisphere dysfunction in children's arithmetic abilities. However, substantive criticisms of methodological and theoretical problems of evoked potential research have been made

(Hiscock & Kinsbourne, 1987; Obrzut, 1989; Taylor & Fletcher, 1983).

One prominent pediatric syndrome, Developmental Gerstmann's Syndrome, includes acalculia among other features. The hypothesized syndrome is thought to be a disorder of body schema integrity that gives rise to acalculia, agraphia (impaired writing ability), finger agnosia (inability to localize finger touch), and left-right disorientation (Benson & Weir, 1972; Kinsbourne & Warrington, 1963). Certain investigators have denied the syndrome's existence (Critchley, 1966; Levin and Spiers, 1985) or questioned its definitional characteristics (Benson & Weir, 1972; Benton, 1977; Levin & Spiers, 1985; Roeltgen, Sevush, & Heilman, 1983). Investigators using careful methodology have continued to identify children with Developmental Gerstmann's Syndrome (Rourke & Strang, 1978; Spellacy & Peter, 1978) and it is currently suggested that the syndrome is valid (Semrud-Clikeman & Hynd, 1990). Investigators attribute the origin of the syndrome to left-parietal dysfunction (Levin & Spiers, 1985), although others have posited right-hemisphere deficits (Rourke & Strang, 1978) or bilateral-parietal dysfunction (Weinberg & McLean, 1986). Benton (1979) has integrated theories of damage in the left or right parietal lobes by proposing that deficient verbal skills may be associated with "left-hemisphere Gerstmann syndromes" and visual-spatial problems may be linked to "right-hemisphere Gerstmann syndromes."

In conclusion, children with right hemisphere and left hemisphere damage may manifest different subtypes of acalculia. Specifically, left hemisphere damage (in right-handed persons) more commonly gives rise to alexic disorders in reading numbers and arithmetical signs. Evidence suggests that either right or left parietal damage, or both, may be critically involved in spatial impairments in mathematics.

These conclusions may be used as a basis upon which to select measures to predict mathematics achievement (see Table 1). These data suggest that mathematics, like reading, is a multisite brain activity. The sensorimotor measures in the battery (i.e., Finger tapping, Speed of Motor, Finger Localizing, and Fingertip Symbol Perception) are presumed to be associated with bilateral posterior-frontal and bilateral anterior-parietal functioning. The Wechsler Factor III is thought to be linked with bilateral frontal functioning, and the Cards Test is presumed to be associated with right-hemisphere frontal functioning. The Aphasia Screening Test is presumed to measure critical functioning in the left hemisphere and the Bender Gestalt is thought to be linked to right hemisphere functioning.

Use of Neuropsychological Tests in Predicting Young Children's Academic Achievement

Beginning with Monroe's work in the 1930s, the psychoeducational academic-readiness literature documents the use of educational tests and measures in the prediction of reading, spelling, and arithmetic (de Hirsch, Jansky, & Langford, 1966). The neuropsychological way of thinking provides a non-traditional perspective by emphasizing the theoretical understanding of the cognitive and neurophysiological underpinnings of academic abilities (Silver, 1978). These theoretical differences translate into practical differences in that neuropsychologists may more fully assess sensory-motor, attention-memory, and language domains (Deysach, 1986). Neuropsychological predictive screening studies of young children's academic achievement began in the 1960s as coherent theory and practice began to be prove useful to school psychologists (Gaddes, 1983). Predictive efforts using these batteries have increased during the past three decades (Satz & Fletcher, 1988).

Three groups of investigators have been influential using neuropsychological screening batteries to predict later academic achievement (impaired and normal reading groups) in kindergarten children: Jansky and de Hirsch's group (1972; de Hirsch et al., 1966), Satz's group (Satz & Friel, 1974, 1978; Satz et al., 1978), and Spreen's group (1978b, 1989). All of these batteries have demonstrated fairly good short-term classification of moderate to severely learning disabled children. Most studies included boys and girls (de Hirsch et al., 1966; Jansky & de Hirsch, 1972; Spreen, 1978b, 1989), but only Jansky and de Hirsch worked with low SES, ethnically diverse children. Most of Satz's samples were comprised entirely of Middle-Class, Caucasian male children, and Spreen's sample was largely Caucasian. Some studies have excluded children with below average IQ or emotional problems (de Hirsch et al., 1966; Spreen, 1978b, 1989), while others included such children (Jansky & de Hirsch, 1972; Satz & Friel, 1974, 1978; Satz et al., 1978).

The most frequently used approach to develop predictive screening batteries is to select a number of measures presumed or demonstrated to predict reading, give them to a large number of children, and use correlation-based analyses to select a small number of screening tests that best predict reading (e.g., de Hirsch et al., 1966; Jansky & de Hirsch, 1972; Satz & Friel, 1974, 1978).

One disadvantage in basing test selection on correlational techniques is that varying batteries usually emerge from the analyses of different samples. For example, while Satz's group has developed the best predictive protocol (Gaddes, 1981; Teeter, 1985), of the 22 variables examined in the beginning phase of the project, 5 variables (Satz et al., 1978), 8 variables (Satz & Friel, 1978) and two (different) protocols of 4 variables (Satz & Friel, 1974; Fletcher & Satz, 1984) have been selected in various stages of the project based upon different samples. Another disadvantage to empirically-based test selection is that the approach is atheoretical. Lack of theory may hinder achieving an understanding of the mechanisms underlying academic achievement, and ultimately may limit predictive results.

A second approach to test selection is to consciously apply theory. For example, Hinshaw et al. (1986), like Spreen (1978b), continue the more formal use of theory in selecting measures from three neuropsychological areas: verbal/language, spatial/perceptual, and motor coordination, in designing their study. While use of the theoretical approach avoids the problem of shifting battery composition among samples, the empirical approach may have greater predictive power for a particular sample.

In all predictive screening batteries, criterion reading measures are given concurrently or at a later date, and rank order correlations (de Hirsch et al., 1966) or discriminant analyses (Fletcher & Satz, 1982; Jansky & de Hirsch, 1972; Satz & Friel, 1974, 1978; Spreen, 1978b, 1989) are used to assess accuracy of prediction of poor readers. Jansky and de Hirsch (1972) assessed their sample's reading ability 2.5 years later, Satz examined his groups after 2, 4 and 7 years (Fletcher & Satz, 1982; Satz & Friel, 1974, 1978), and Spreen (1989) retested his participants after 6 years. Reading criterion measures have included the Gates Advanced Primary and Gates-MacGinitie Paragraph Reading Test (Gates & MacGinitie, 1965) (used by Jansky & de Hirsch, 1972), a 10-point teacher rating scale (Satz & Friel, 1974), and IOTA and Classroom Reading Measures (Satz & Friel, 1978). "Problem" or "failing" readers are variously defined as 1/2 to 2 standard deviations below the mean, or 1-2 years behind in grade level on criterion reading measures (de Hirsch et al., 1966; Jansky & de Hirsch, 1972; Satz & Friel, 1974, 1978; Spreen, 1978b).

The tests comprising the four predictive batteries (de Hirsch et al., 1966; Jansky & de Hirsch, 1972; Satz & Fletcher, 1982; Spreen, 1978b) are listed in Table 2. All four batteries seem to have fairly similar predictive power: correctly identifying 60-80% of the kindergarten children who will later become problem readers, but incorrectly identifying 20-40% of children as likely to have future reading difficulty. All investigators used discriminant analysis to maximize the identification of youngsters with high risk of reading problems while minimizing "false positives" (i.e., children mistakenly identified as having learning problems) and "false

Table 2

Early Neuropsychological Screening Batteries' Measures

Investigator	Screening Battery Measures
de Hirsch et al., (1966)	Visual Motor Gestalt Test (Bender, 1946)
	Word Matching (Gates & MacGinitie, 1965)
	Pencil Use (de Hirsch et al., 1966)
	Word Reversals & Story Words (de Hirsch et al., 1966)
	Categories (de Hirsch et al., 1966)
	Auditory Discrimination Test (Wepman, 1958)
	Word Recognition I & II (de Hirsch et al., 1966)
	Word Reproduction (de Hirsch et al., 1966)
Jansky & de Hirsch (1972)	Visual Motor Gestalt Test (Bender, 1946)
	Word Matching (Gates & MacGinitie, 1965)
	Picture & Letter Naming (de Hirsch et al., 1966)
	Sentence Memory (Terman & Merrill, 1937)
Satz & Fletcher (1982)	VMI[a] (Beery & Buktenika, 1967)
	Peabody Picture Vocabulary Test (Dunn, 1965)
	Alphabet Recitation (Satz & Fletcher, 1982)
	Recognition-Discrimiation Test (Small, 1969)
Spreen (1978)	Benton Visual Retention Test-Revised (1963)
	Peabody Picture Vocabulary Test (Dunn, 1965)
	Coloured Matrices (Raven, 1965)
	Teacher Rating (Spreen, 1978)

[a] Developmental Test of Visual-Motor Integration

negatives" (i.e., children mistakenly identified as not having learning problems). Specifically, the predictive batteries correctly identified 91% (de Hirsch et al., 1966) and 77% (Jansky & de Hirsch, 1972) of failing readers over 2.5 years, 60-80% of severely disabled readers (Fletcher & Satz, 1982; Gaddes, 1981) over 2-7 years, and 63-86% of impaired readers over 6 years (Spreen, 1989). The higher true positive percentages in these studies (e.g., 80%) are associated with assessment over 2 years and the lower percentages (e.g., 60%) with 6-7 year follow-up studies. Short term false positives rates among the studies consistently ranged from 19-30% (de Hirsch et al., 1966; Fletcher & Satz, 1982; Gaddes, 1981; Jansky & de Hirsch, 1972; Spreen, 1989).

While these germinal efforts by Jansky and de Hirsch, Satz, and Spreen are important, three long-standing issues in the area of early neuropsychological screening for academic problems have not been fully addressed by these studies: narrow sampling of neuropsychological abilities, inadequate measures for use with ethnically diverse and lower socioeconomic populations, and inadequate prediction of mildly impaired or average children (including high false positive rates).

First, perhaps the most significant criticism of the early screening batteries is that their tests are quite limited in their sampling of neuropsychological abilities, as illustrated in Table 3. All of these investigators have classified their variables in similar functional systems and their nosologies formed the basis of the table. Also, Lezak (1983) as well as Spreen and Strauss (1991) classify the tests in these functional categories. The visual modality is dominant in all of the batteries and accounts for 5 of 10 (de Hirsch et al., 1966), 4 of 5 (Jansky & de Hirsch, 1972), 3 of 5 (Fletcher & Satz, 1982), and 3 of 3 (Spreen, 1978b, 1989) tests used in the predictive batteries. (Spreen's fourth measure is a teacher rating scale, not a test per se.) De Hirsch et al. (1966), Jansky and de Hirsch (1972) and Satz and Fletcher's

Table 3

Early Screening Batteries' Tests Classified as a Function of Five Neuropsychological Areas

Investigative Team	Neuropsychological Area		
	Intelligence	Language	Attention-Memory
de Hirsch et al. (1966)	[Visual Motor Gestalt][a]	Word Reversals & Story Words	[Auditory Discrimination][a]
	[Categories]	Word Matching	[Word Recognition I & II][a]
		Categories & Word Reproduction	[Word Reproduction][a]
		Auditory Discrimination	
		Word Recognition I & II	
Jansky & de Hirsch (1972)	[Visual Motor Gestalt][a]	Picture & Letter Naming	Sentence Memory
		Word Matching	[Word Matching]
Satz & Fletcher (1982)	[PPVT][a,b]	PPVT[b]	[Visual Motor Integration][a]
	[Visual Motor Integration][a]	Alphabet Recitation	[Recognition-Discrimination][a]
Spreen (1978)	Raven's Coloured Matrices	PPVT[b]	[Revised Visual Retention][a]
	[PPVT][a,b]		[Raven's Coloured Matrices][a]

(continued)

Table 3 (continued)

Investigative Team	Neuropsychological Area	
	Sensation-Perception	Fine Motor
de Hirsch et al. (1966)	Visual Motor Gestalt	[Visual Motor Gestalt][a]
	Pencil Use	[Pencil Use][a]
	[Word Reversals][a] & [Word Matching][a]	[Word Reproduction][a]
	[Auditory Discrimination][a]	
Jansky & de Hirsch (1972)	Visual Motor Gestalt	[Visual Motor Gestalt][a]
	[Word Matching][a]	
Satz & Fletcher (1982)	Visual Motor Integration	[Visual Motor Integration][a]
	Recognition-Discrimination	
Spreen (1978)	Revised Visual Retention	[Revised Visual Retention][a]
	[Raven's Coloured Matrices][a]	

a. Tests in brackets "[ ]" represent secondary use in assessment of an area.

b. "PPVT" refers to the Peabody Picture Vocabulary Test (Dunn, 1965).

(1982) batteries rely very little on auditory input, each protocol using only one measure. Other domains of neuropsychological function also receive little attention. De Hirsch et al. (1966) sampled fine motor ability using only two measures, and in the other three batteries, fine motor ability is sampled using only one test. All of the batteries' fine motor tests involve a strong visual component. The lack of adequate sampling of abilities may be related to the prediction "ceilings."

Second, cultural fairness continues to be a concern, because while de Hirsch and Jansky's batteries were designed and used with urban low income populations having large minority representations, Satz and Spreen's batteries were not. Jansky and de Hirsch (1972) rely upon a separate conversion score tables for Blacks and Whites (for all tests) in constructing their Screening Index, which represents an effort to maintain cultural fairness. However, separate conversion scores for Blacks and Whites raise questions of cultural bias in the predictive measures used. The initial Satz study used a sample of White Middle-Class boys, as did all of their follow-up studies with one exception, a cross-validation study including girls and a small sample of African American children (Satz & Friel, 1978). Spreen's Battery (1978b) was developed using Canadian, mostly White, Middle-Class children.

Satz's and Spreen's batteries also may be criticized as lacking cultural fairness in use with low-income and minority children because of their extensive dependence on the Peabody Picture Vocabulary Test (PPVT: Dunn, 1965) and revised version (PPVT-R: Dunn & Dunn, 1981). The PPVT and PPVT-R have been described as being "especially dangerous" in assessing the skills and abilities of ethnic-minority children (Sattler, 1988). The PPVT significantly underestimates cognitive functioning of Hispanic (Laosa, 1984) and Native American children (Naglieri & Yazzie, 1983). With respect to African-American children, de Hirsch et al. (1966) administered the PPVT to urban ethnically diverse, low-income New York city children and found that the test did not correlate significantly with reading or writing (although it correlated with spelling). The issue of cultural fairness has not been adequately addressed by any of these investigators.

Third, no screening protocol has yet been successful in short term classification of average or mildly impaired children (Jansky, 1978; Silver & Hagin, 1990). These populations have received less attention in the literature (Janksy, 1978). This problem is perhaps most apparent in the consistent short-term false positive rates of 20-30% that are considered unacceptably high for practical classroom use (Silver, 1978). While, it is encouraging that Fletcher and Satz (1984) have demonstrated that their battery predicts achievement more accurately than teacher ratings, mistakenly identifying 1 in 3 or even 1 in 5 children as problem learners negates the batteries' practical value.

In response to these concerns, Townes et al. (1980) developed a neuropsychological screening protocol based upon 10 subtests from the best standard battery for this age group: the Reitan-Indiana Younger Children's Battery (Reitan, 1974). This screening battery is reviewed below, and Table 4 lists the tests that comprise the battery as a function of the neuropsychological area. Townes et al. (1980) comment in general regarding the neuropsychological domain tapped by the tests, and Joshko and Rourke (1985), Selz (1981), and Reitan (1974) explicitly classify the tests in these categories. Of particular note is the broadened sampling of motor, sensory, and reasoning domains among this battery, compared to the older batteries. Teeter (1985) essentially used the same battery (excluding the Grip Strength Test) in her prediction study. Townes et al. (1980) studied Caucasian, Middle-Class kindergarten and second grade children and compared the battery's predictive ability to that of four WISC-R subtests. Teeter (1985) tested Middle-Class students in kindergarten and followed them up one year later (first grade), using the Townes et al. (1980) screening battery

Table 4

Townes et al. (1980)-Teeter (1985) Battery as a Function of Five Neuropsychological Areas

Neuropsychological Area	Townes et al. (1980)-Teeter[a] (1985) Measures
Intelligence	Color Form (Reitan, 1974)
	Progressive Figures (Reitan, 1974)
	Matching Pictures (Reitan, 1974)
Language	Reitan-Indiana Aphasia Screening (Reitan, 1974)
Attention-Memory	Target Test (Reitan, 1974)
Sensory-Perception	Finger Localizing/Fingertip Writing (Reitan, 1974)
	Reitan-Indiana Aphasia Copy Errors (Reitan, 1974)
	Matching Figures (Reitan, 1974)
	Matching Vs (Reitan, 1974)
	Star & Concentric Square (Reitan, 1974)
Fine Motor	Finger Tapping (Reitan, 1974)
	Test of Grip Strength (Reitan, 1974)

a. The Teeter (1985) Battery is the same as the Townes et al. (1980) Battery except the latter includes the Test of Grip Strength and the entire McCarthy Scales of Children's Abilities to predict academic achievement.

The three largest areas of agreement among the Townes et al. (1980) and prior protocols lay in their use of visual-spatial tasks requiring pencil skills, visual-spatial matching tasks, and language tasks. However, the tests in Townes' battery measure attention or immediate memory and diverse language abilities, in contrast to the tests used in the older batteries. Also, there is a focus on more basic motor ("non-pencil"), sensory-perceptual (tactile), and conceptual abilities (changing cognitive sets) in the Townes battery, in comparison to prior protocols.

Townes et al. (1980) and Teeter (1985) used discriminant function analysis to classify children's concurrent academic outcome into "high" and "low" achievement (above and below the 50th percentile, respectively). Townes et al. (1980) battery was as effective as the full WISC-R in concurrently classifying 70-75% of the kindergartners and second graders on most Stanford Reading Achievement measures. Teeter's concurrent high and low classification rates of kindergartners using the Townes and the full McCarthy Batteries were equivalent and excellent (93-96%). Teeter's one-year follow-up rates for the first graders, were equivalent for both batteries (76-80%).

Teeter (1985) also classified children's performance, by grade, into three groups based on achievement score: above the 80th percentile, between the 60th and 80th percentile, and below the 60th percentile. For kindergartners, the concurrent prediction rates for the Reitan screening battery and the full McCarthy Scales were 61% and 71%, respectively. The one-year follow-up analyses (first grade data), revealed equivalent predictive accuracy rates for the Reitan and the McCarthy batteries (57-61%) which were lower than the concurrent rates.

Accuracy rates of predicting achievement using the Townes et al. (1980) and Teeter (1985) batteries are equivalent to, but no better than, the power of the older batteries. It remains for batteries to improve on short-term accuracy and cultural-fairness making such batteries practical and useful to educators and psychologists (Silver and Hagin, 1990).

General Conclusions of the Literature Review

Review of the literature points the way in several specific directions towards a theoretically and empirically-based selection of measures for a neuropsychological screening battery suitable for use with young urban children. First, the literature indicates that individual differences in learning and achievement are multicausally determined. Because both biological and social factors contribute to academic achievement, ideally both biological and social predictors should be used. But, given the desire to find culture-fair predictors, as well as limited time and sample size, the present study will focus on neuropsychological measures thought to tap the biological functions and thought to minimize variance due to social differences.

Second, the literature indicates that there are multiple LD subtypes and that complex academic skills (such as reading and arithmetic computation) are subserved by multiple brain areas and/or systems. Therefore, it makes sense to select measures that tap a broad array of discrete neuropsychological domains.

Third, the literature suggests that previous neuropsychological batteries predicted fairly well, but at rates too low for practical classroom use. Also, few were theory-directed, few predicted above-average performance, few were used to predict Arithmetic as well as Reading achievement, and few have explicitly examined the cultural fairness of the predictor or criterion measures. The present study seeks to expand and improve upon previous batteries by addressing these limitations.

Chapter III

METHOD

Participants

Offers to participate in the project were extended to three private-parochial elementary schools and one public school located in a large, Midwestern city. The four schools were selected because they were in close proximity to each other (within a two-square mile area) and because each school served students from a lower middle-class ethnically diverse population. Two of the parochial schools accepted the offer and provided the children for the project. The children comprised a nonreferred population drawn from regular education classrooms. The introductory materials stated that parents might withdraw their child from the project at any time without penalty (other than the presumed loss of benefits accrued from their child's participation). No parent withdrew his or her child from the project.

Chi-square tests performed on the data from the two schools revealed that the two samples did not differ with respect to gender (X^2 = .01, $\underline{df}$ = 1, $\underline{p}$ = $\underline{ns}$), handedness (X^2 = .96, $\underline{df}$ = 1, $\underline{p}$ = $\underline{ns}$), or ethnicity (X^2 = 5.70, $\underline{df}$ = 3, $\underline{p}$ = $\underline{ns}$). However, in the analysis of ethnic groups, 63% of the cells had an expected frequency of less than 5 (a violation of the assumption of 20% or fewer cells), so the ethnic minority children were combined into a single group, and the chi-square test rerun between the two schools (no assumptions being violated). The proportion of ethnic minority children in combined ethnic samples from the two schools were not significantly different (X^2 = 1.88, $\underline{df}$ = 1, $\underline{p}$ = $\underline{ns}$). Table 5 provides the demographic characteristics, including participants' ages, ethnic backgrounds, and gender, by

Table 5

Ethnic Background, Gender, and Mean Age of Participating Children by Grade

Grade	Children	Gender		Ethnic Background				Age	
		Male	Female	European American	African American	Hispanic American	Asian American	(In Months)	
	n	n	n	n	n	n	n	M	SD
K	21	13	8	13	5	2	1	67.0	4.5
1	10	6	4	5	2	0	3	77.9	4.7
2	13	6	7	9	2	1	1	89.3	4.3
3	20	11	9	11	3	4	2	103.9	3.4
Totals	64	36	28	38	12	7	7	84.8	15.8

grade. One student in School A (5% of sample) was one grade year behind, given his or her age, as were two students in School B (5% of sample). The proportion of left-handers in the total sample (18%) is equivalent to estimates of the population average (14%) (Lezak, 1983).

In comparing the ethnic representation of School B participants to the total population of School B, European Americans are somewhat over-represented (54% of participants versus 39% of the school), Asian Americans (12% versus 17%) and Hispanic Americans (18% versus 26%) slightly under-represented, and African Americans (16% versus 19%) are proportionately represented. With respect to socioeconomic status, forty-eight percent of School B's students during the 1991-1992 school year (the first year these statistics were available) were considered lower middle socioeconomic status in that they received supplemental funding for school lunches. The ethnic and socioeconomic data for the school-at-large were unavailable for School A. Throughout the rest of this document, the combined African American, Hispanic American, and Asian American children in this study's sample will be called the "Ethnic Minority" group, and the combined European American children will be called the "Ethnic Majority" group.

A t-test between the mean ages of the students of the two schools revealed that children in School B (M = 87.6 months, SD = 16.7) were significantly older than students in School A (M = 79.0 months, SD = 12.1) (t = -2.12, df = 62, p = .038). The two schools' samples did not differ with respect to the percentage of students from the four grades: Kindergarten through third grade (X^2 = 7.38, df = 3, p = ns). However, 25% of the cells had an expected frequency of less than 5, so a Younger children's group (combined Kindergarten and first grade) and an Older children's group (combined second and third grade) were formed; the test was rerun and again was nonsignificant (X^2 = 2.27, df = 1, p = ns). In comparing the number of Older and Younger Ethnic Minority and Ethnic Majority children, there was not a significant age-difference between the two ethnic groups ($X^2 = 0.04$, $\underline{df} = 1$, $\underline{p} = \underline{ns}$).

Procedure

General

In the Fall of 1988, teachers in grades kindergarten through third sent introductory letters and parental consent forms home with all children, and all children returning the signed consents were tested. The letter and consent to the parents emphasized the investigator's interest in learning problems, and described the project benefits as including feedback on academic abilities in the Spring, and if appropriate, referral for further testing and assessment.

The investigator was on-site during all data collection. The neuropsychological predictive battery data collection (approved by the Institutional Human Subjects Review Board of Loyola University) took place in the Fall of 1988, between October 24 and November 17, 1988. In the Spring of 1989, between April 22 and May 25, the academic achievement tests were administered. (The measures are described in detail below.) The length of time between the Fall and Spring sessions ranged between 162 and 203 days ($\underline{M} = 177.9$; $\underline{SD} = 6.5$ days). The children were debriefed following the academic achievement testing.

Parental feedback forms, notifying parents of the level of performance of their child (i.e., average, above average, superior) were mailed to all parents whose children scored in at least the average range ($\underline{n} = 44$). The investigator personally met with parents of children performing below average (standard score < 90) on any of the academic achievement tests ($\underline{n} = 20$), and made rèferrals for psychoeducational evaluations if indicated.

Examiners

Nine undergraduate psychology students, enrolled in psychology research courses, were trained to administer the neuropsychological test battery. Each student examiner received extra credit or partial course credit for their work. The trainees signed a contract agreeing to spend 4-6 hours per week for 6-8 weeks learning to administer the measures and testing the school children. The student examiners were not informed about the research questions until the completion of testing, at which time they were debriefed.

Trainees completed two 2-hour testing introduction sessions, focusing on the test materials and test administration. Copies of specific test administration instructions were provided, as well as four handouts emphasizing general issues in assessment, potential difficulties using the tests, and proper decorum and behavior in the schools. Examiners then completed two 2-hour administration training sessions. Trainees practiced administering the battery to a child (relative or friend of the family). Lastly, the investigator observed and critiqued the examiners administering the test battery to each other.

The primary investigator collected all of the criterion academic achievement data in the Spring of 1989, but had not yet scored the predictive neuropsychological test data from the Fall of 1988 and was thus not informed about the children's neuropsychological test performances.

Measures

Development of Predictive Screening Battery and Data Reduction

A subset of tests from the neuropsychological battery routinely used by the Charles I. Doyle, S. J. Center staff was used as the predictor set in the present study. The center's Neuropsychological Research Group selected this subset of tests for brief screening (i.e., tests that could be administered in 20-30 minutes), as the complete battery took 10 hours to administer.

The first objective guiding screening test selection was to economically and reliably sample five theoretically important areas of neuropsychological functioning: Intelligence, Attention-Memory, Language, Perception-Sensation, and Fine Motor ability. Based upon preliminary analyses, the tests chosen to represent the Perception-Sensation area were further divided into two areas: Somesthetic and Visual-Spatial. Thus, six neuropsychological domains comprised the final battery. The second objective guiding test selection was cultural fairness. In considering cultural fairness, it was thought that neuropsychological measures that tap the underpinnings of academic abilities better than traditional measures (Silver, 1978) might yield a more culture-fair battery. For example, Morrison and Hinshaw (1988) demonstrated that while socioeconomic status is related to IQ and academic achievement, neuropsychological measures correlated with academic achievement, but not with IQ or socioeconomic status.

Table 6 lists the tests comprising each of six neuropsychological domains in the screening battery and the sample and available standardization means and standard deviations. Each of the original measures selected for this screening battery has proven highly reliable and has been standardized on large numbers of children (Brown, Rourke, & Cicchetti, 1989; Kaufman, 1979; Koppitz, 1970; Reitan, 1974, 1987; Rourke & Strang, 1983; Schulman, Buist, Kaspar, Child, & Fackler, 1969; Spreen & Haaf, 1986). Each test is available from its respective publisher with the exception of the Speed of Motor Performance Test (Schulman et al., 1969) and the Cards Test (Schulman et al., 1965). Shulman et al. 1969 provides a detailed description of the Speed of Motor Performance Test and the complete directions for administration. See Appendices 1 and 2 for the letter of permission to reprint the Cards Test and the description of the test. All four WISC-R subtests were retained to maintain the consistency of the battery across ages, but scores were pro-rated for the very youngest children. Investigators have used pro-rating of similar measures in their studies of young children (e.g., Townes et al., 1980).

Table 6

The Project's Final Battery Measures, Means, and Standard Deviations as a Function of Six Neuropsychological Areas

Neuropsychological Area	Measures	Sample: M (SD)	Standardization: M (SD)
Intelligence	WISC-R Information subtest (Wechsler, 1974)	10.7 (3.9)	10.0 (3.0)
Language	Reitan-Indiana Aphasia Screening (Reitan, 1974)	46.4[a] (16.4)	50.0 (10.0)
Attention-Memory	Composite (sum) of the following:	32.1 (6.8)	30.0 (--)
	WISC-R Digit Span (Wechsler, 1974)	9.9 (3.3)	10.0 (3.0)
	WISC-R Arithmetic (Wechsler, 1974)	10.4 (2.8)	10.0 (3.0)
	WISC-R Coding (Wechsler, 1974)	11.9 (2.8)	10.0 (3.0)
Fine Motor	Composite (sum) of the following:	67.2 (15.4)	70.0 (--)
	Finger Tapping (Reitan, 1974)	22.7 (7.2)	20.0 (--)
	Speed of Motor Test (Schulman et al., 1969)	44.4 (11.2)	50.0 (--)

(continued)

Table 6 (continued)

Neuropsychological Area	Measures	Sample: M (SD)	Standardization: M (SD)
Somesthetic	Composite (sum) of the following:	227.8 (26.7)	200.0 (--)
	Finger Localizing Test (Reitan, 1974)	107.8[a] (19.2)	100.0 (--)
	Fingertip Symbol Recognition Test (Reitan, 1974)	120.0[a] (12.6)	100.0 (--)
Visual-Spatial	Visual Motor Gestalt Test (Bender, 1946)	47.9[b] (8.9)	50.0[b] (10.0)

Note. Dashes indicate the standard deviation was not available for combined scores (e.g., only available for left and right hand scores, considered separately).

a. Data not log-transformed to allow comparison with standardization mean.

b. Scores based upon Koppitz's (1970) developmental scoring system.

Standard scores were computed for as many variables as possible in the battery, and natural logarithmic transformation of scores was performed for non-normally distributed variables. The frequency distribution of each variable was obtained and means, medians, and kurtosis were examined to ensure that each variable was normally distributed. Each variable determined to be non-normally distributed was log-transformed, with a constant (i.e., 1) added whenever zero could occur as a raw score, to avoid such scores being treated as missing data.

Intelligence. The Information subtest of the WISC-R (Wechsler, 1974) was selected as the measure of Intelligence and ability to acquire new learning. Standard administration and scoring of the subtest was used and the possible raw scores range was 0 - 30 points. Raw scores were transformed into age-scaled standard scores (M = 10, SD = 3; Wechsler, 1974). The Information subtest was chosen because it is brief, simple, and has the second highest correlation, following Vocabulary (r = .72), with Full Scale WISC-R IQ in the standardization sample (Sattler, 1992).

Language. The Aphasia Screening Test was selected as the Language measure as the test is well established as measuring language-related abilities (Joschko & Rourke, 1985; Selz, 1981). The screening test is comprised of 22 language-related items including naming and copying figures, naming numbers and letters, performing simple arithmetic, demonstrating awareness of body parts, and exhibiting knowledge of left and right. The standard administration of the Reitan-Indiana Aphasia Test was used and Reitan's (1987) "weighted error" system was used to score this test. The raw score range using Reitan's weighted error score system was 0 - 60 error points. Reitan (1987) also provides means and standard deviations for the 5-8 year olds weighted Aphasia Error Totals such that a T-score (M = 50, SD = 10) was generated for each child's score. The T-scores were then log-transformed, because the distribution of scores was non-normal. Townes et al. (1980) and Teeter (1985) have demonstrated that the Aphasia Screening Test is among the most sensitive in the Reitan-Indiana Battery for detecting learning problems in kindergarten and first grade children (ages 5 to 6 years old), although no data exist on the measure's predictive validity for second and third grade children (ages 7-8 years old).

Attention-Memory. The Freedom From Distractibility (Kaufman, 1975), Factor III (Wechsler, 1974) subtests (Arithmetic, Digit Span, and Coding) of the WISC-R were included to assess Attention-Memory. Standard administration of each WISC-R subtest was used. The possible raw score range for Arithmetic was 0 - 18 points, the possible range for Digit Span was 0 - 28 points, and the possible raw score range for Coding was 0 - 50 points. The raw score for each subtest was transformed (Wechsler, 1974) to age-scaled standard scores (M = 10, SD = 3). The three subtests' standard scores were summed to derive the score used in the analyses. Kaufman (1979) has discussed the factor's use as a measure of attention-concentration, and Factor III scores have been depressed (compared to the standardization sample) in studies of Attention-Deficit Hyperactivity disordered (ADHD) children (Sutter, Bishop, & Battin, 1987). These three WISC-R subtests also have been considered a sequential reasoning measure (Bannatyne, 1971). Other investigators have discussed the factor as an immediate memory measure (Cohen, 1957; McFie, 1961). The factor has proven useful in identifying children with developmental disorders and learning disabilities (Kaspar et al., 1992) as well as brain-damaged patients with attention problems and epilepsy (Dennerll, 1964; Tarter, 1972).

Gutkin and Reynolds (1981) analyzed the WISC-R standardization sample and demonstrated that these three subtests yielded the least difference between African Americans and European Americans. These data suggest that these three subtests are the most culture-fair WISC-R measures, and thus also were included in the battery for this reason.

The Cards Test (Schulman et al., 1965) was the second measure of Attention-Memory. The Cards Test is a vigilance task which requires a child to observe a series of 200 flip cards (one per second) and attempt to pick out 20 illustrations of a baby from 180 pictures of rabbits. The standard administration was used and the Cards Test was scored with respect to attention errors (i.e., not responding when a picture of a baby was shown) and impulsivity errors (i.e., saying "baby" when a rabbit was shown). The possible raw score range of attention errors was 0 - 20 errors, while the possible range for impulsivity errors was 0 - 180 errors. Both sets of data were log-transformed as the distributions were non-normal. The Cards Test has proven useful as a measure in three separate studies differentiating children with attentional problems associated with brain damage or soft neurological signs from normal children (Kaspar & Koshaba, 1974). As discussed below, however, this test was dropped from the final analyses because it did not correlate with other measures of Attention-Memory.

Sensation-Perception: Visual-Spatial. The Bender Visual Motor Gestalt Test (Bender, 1946) and the Reitan-Klove Finger Localizing Test and the Symbol Recognition Test (Reitan, 1974) were selected as the Sensation-Perception tests, because all three tests involve shape perception and because intact parietal functioning is well recognized as critical to adequate performance on all three measures (Lezak, 1983). The Bender was chosen because of its long history of use in screening batteries (e.g., de Hirsch et al., 1966; Jansky & de Hirsch, 1972), and its structure, which requires organizational as well as visual-perceptual and visual-motor skills. Koppitz's (1970) Bender developmental scoring system of distortions, rotations, integration problems, and perseverations was used and gives rise to possible raw error scores of 0 -30 errors. Raw error scores were transformed into age-scaled Developmental Error T-scores (M = 50, SD = 10) using Koppitz's norms (1970). As discussed below, the Bender Gestalt did not correlate with the Reitan-Klove Finger Localizing and Fingertip Symbol

Perception tests for this sample. The Bender score was considered as a separate predictor from the scores of finger and fingertip sensation in the final battery, and was retained as a predictor because visual-organizational and visual-spatial skills are very important in academic performance (Benton, 1978; Kevale, 1982; Spreen, 1978a).

Sensation-Perception: Somesthetic. The Reitan-Klove Finger Localizing Test and the Reitan-Klove Fingertip Symbol Recognition Test (Reitan, 1974) were selected as representative Sensation-Perception measures because they are among the most well recognized measures of sensory ability (Joschko & Rourke, 1985; Spreen & Strauss, 1991). The tests have repeatedly demonstrated the ability to predict academic achievement (Teeter, 1985; Townes et al., 1980). In the Finger Localizing Test, the examiner uses a screen to block the child's vision and randomly touches the fingers of one hand at the site where the base of the nail meets the rest of the finger; following a 2-second delay, the child points to the touched finger with the opposite hand (Reitan, 1974). Four trials with each finger are conducted in the test and the possible raw error score ranges from 0 - 20 errors per hand (0 - 40 errors for both hands). In the Fingertip Symbol Recognition Test, the examiner uses a screen to block the child's vision and a stylus to trace an `X' or an `O' on the bottom of the fingertip (Reitan, 1974). The symbols are randomly assigned and four trials per finger are conducted. The standard administration was used for both tests and the score for each measure is the total number of errors made with each hand. The possible error score ranges from 0 - 20 errors per each hand (0 - 40 errors for both hands). Knights and Norwood's (1980) age-scaled norms were used to generate T-scores for the Finger Localizing and Fingertip Symbol Perception subtests for each hand (Reitan, 1974). The T-score distribution of the Finger Localizing and Fingertip Symbol Perception subtests was non-normal, so the scores for both hands for these two variables were log-transformed and summed to derive the predictor.

It has been suggested that finger recognition ability is not developmentally stable until the age of 10 years (Ellison, 1983), and Fletcher, Taylor, Morris, & Satz (1982) have raised questions about the reliability and validity of Small's (Finger) Discrimination test (1969) such that it was dropped from Satz's predictive battery (Fletcher et al., 1982). However, the Reitan-Klove tests appear much more reliable and developmentally appropriate tests than Small's (1969) measures with respect to number of trials (20 trials per hand versus 5 trials, respectively) and method of identifying fingers (pointing with the other hand versus numbering fingers or pointing to analogous finger of model's hand, respectively). Reitan (1969) specifically rejected as developmentally inappropriate the type of tasks that Small (1969) used with younger children.

Fine Motor. The Reitan-Indiana Finger Tapping Test (Reitan, 1974) and the Speed of Motor Performance Test (SMPT) (Schulman et al., 1969) were two face-valid measures selected to assess Fine Motor performance. These motor-speed measures relied less on visual-spatial abilities and pencil skills than the measures used in other screening batteries. The Reitan Finger Tapping test is among the most widely used measures of fine motor speed and dexterity (Lezak, 1983; Selz, 1981). The standard administration of the tapping test was used and the score is the mean number of index finger taps (5 trials per hand) within a 10-second interval for each hand. Possible raw scores range from 0 - 51 taps for each hand. Age-scaled standard scores ($\underline{M}$ = 10, $\underline{SD}$ = 3) were generated for the score of each hand on the Finger Tapping Test (Reitan, 1974) using Klonoff and Low's (1974) norms.

The Speed of Motor Performance Test (SMPT, Schulman et al., 1969) was originally comprised of 4 subtests: Tapping, Pegs, Picks, and Beads, of which the latter three are used in this study. [The electric Reitan tapper was used instead of the SMPT Blood Counter tapper because the Reitan tapper has proven more reliable in use with children 5-8 years old than tappers such as the SMPT (Reitan, 1969)]. The standard administration of the tests was used. The Pegs and Picks tests are measures of dominant and nondominant hand speed and dexterity, and the raw score for both tests is the time it takes a child to complete the task (one score for each hand). In the Pegs Subtest, a child puts 6 square pegs in 6 square holes, first using only his dominant then his nondominant hand. In the Picks subtest, the children put 15 toothpicks in a small styrofoam ball. Possible raw scores (times) ranged from 0 - 18 seconds (or longer) per hand for Pegs, and 0 - 95 seconds (or longer) per hand for Picks. The Beads subtest is a measure of bilateral motor coordination and generates a single score: the time it takes to use both hands to string 10 beads on a 16 inch shoelace. Possible raw scores (times) ranged from 0 - 113 seconds (or longer) for the Beads subtest. The five SMPT (2 each for Pegs and Picks, and 1 for Beads) raw scores were transformed into age-scaled standard scores (M = 10, SD = 3) and summed (Schulman et al., 1969). The Speed of Motor Performance test has proven a useful measure in separate studies differentiating brain-damaged children from normals and in distinguishing borderline intellectual functioning children from children of normal intelligence (Kaspar & Sokolec, 1980).

Intercorrelations supporting data reduction. Single measures assessed the Intelligence and Language areas, but multiple measures were used for the Attention-Memory, Sensation-Perception, and Fine Motor areas. A matrix of Pearson product-moment correlation coefficients was generated using standard scores (or natural log-transformed scores of non-normally distributed variables) for all measures. The intercorrelations among multiple scores comprising the Attention-Memory, Fine Motor, and Sensation-Perception areas were examined within each area in order to assess construct validity. Measures were retained only if (a) each measure of conceptual area comprised of multiple measures correlated significantly [with at least intermediate levels of correlation (i.e., $r > .30$)] with at least one-third of the other measures within its conceptual area; and (b) the measures did not correlate highly (i.e., $r > .60$) with variables measuring other constructs. Overall scores for Attention-Memory, Fine Motor, and Sensation-Perception areas then were derived by using the sum of the standard scores (or log-transformed scores) of all measures in each domain.

The area of Intelligence was the one exception to measures' inclusion despite intercorrelations with other domains. Because of the global multi-faceted nature of "intelligence," measures from different conceptual areas were permitted to correlate significantly with Intelligence.

Table 7 lists the subtest intercorrelations for each composite area: Attention-Memory, Fine-Motor, and Sensation-Perception. (The Intelligence and Language areas were each comprised of one measure.) The correlational analyses strongly supported combining the measures in the Fine Motor area, but only partially supported combining the initially selected set of measures in the Attention-Memory and Sensation-Perception areas. The two measures derived from the Cards Test were dropped from further analyses because they did not correlate significantly with the other Attention-Memory measures. While the Bender developmental score was expected to be associated with the other Sensation-Perception measures, the correlations with those other measures were low. These analyses indicated that the Somesthetic variables (Finger Localizing and Fingertip Symbol Recognition Tests) were distinct from the Visual-Spatial (Bender Gestalt) variable. The importance of the Visual-Spatial domain is substantiated in the LD subtyping literature (Denckla, 1979), as well as the predictive screening investigations (Benton, 1978), and so the Bender was retained in the analyses as representing a separate sixth neuropsychological area: the Visual-Spatial domain.

The majority of sample means and standard deviations in the tests and six domains in the final battery were very close to the known standardization means, supporting the

Table 7

Correlations of Sub-Test Scores within Three Composite Neuropsychological Domains

Attention-Memory

Subtest	Digit-Span	Arithmetic	Coding	Cards Test: A[a]
Arithmetic	**.40****			
Coding	**.33***	**.41****		
Cards Test: A[a]	.00	.08	.13	
Cards Test: B[b]	.11	.11	-.08	-.10

Fine Motor

Subtest	Finger Tapping: D[c]	Finger Tapping: ND[d]	Pegs: D[c]	Pegs: ND[d]	Picks: D[c]	Picks: ND[d]
Finger Tapping: ND[d]	**.79****					
Pegs: D[c]	**.37***	.28				
Pegs: ND[d]	.27	**.37***	**.58****			
Picks: D[c]	.16	.14	**.50****	**.31***		
Picks: ND[d]	.26	**.32***	**.46****	**.40****	**.66****	
Beads	**.31***	.23	**.35***	**.66****	**.51****	**.57****

(continued)

Sensation-Perception

Subtest	Bender-Gestalt	Finger Localization: D[c]	Finger Localization: ND[d]	Fingertip Symbol: D[c]
Finger Localization: D[c]	-.10			
Finger Localization: ND[d]	-.08	**.55****		
Fingertip Symbol: D[c]	.09	.24	.16	
Fingertip Symbol: ND[d]	.28	.24	**.31***	**.56****

a. "A" signifies impulsivity errors.

b. "B" signifies attentional errors.

c. "D" signifies dominant hand.

d. "ND" signifies nondominant hand.

*$p < .05$. **$p < .01$.

contention that the children are drawn from a normal population (see Table 6).

Table 8 lists the intercorrelations among the final six neuropsychological areas. Intelligence correlated significantly with Language, Visual-Spatial, and Attention-Memory areas, while the Fine Motor and Somesthetic areas were related. The importance of Intelligence and its diversity of expression is suggested by its correlation with what might be thought of as the "higher cognitive domains" in the battery. The relationship of Fine Motor and Somesthetic areas may be a function of the neural proximity of the motor strip (posterior frontal) to the secondary and tertiary sensory areas (anterior parietal) hypothesized to be associated with the motor and sensory measures (Lezak, 1983).

Development of Criterion Variables and Data Reduction

The criterion measures were derived from the three WRAT-R (Jastak & Wilkinson, 1984) subtests: Word Identification, Spelling, and Arithmetic. Age-based standard scores (M = 100, SD = 15) provided by Jastak and Wilkinson (1984) were used to derive three criterion variables. First, a "Combined Academic" variable was calculated by averaging the standard scores of Word Identification, Spelling, and Arithmetic subtests. Second, a "Language-Related" variable was obtained by averaging standard scores of the Spelling and Word Identification subtests. Third, the Arithmetic standard score was used as an "Arithmetic" criterion variable.

Intercorrelations supporting data reduction. The frequency, mean, median, and kurtosis was generated and examined for each of the original WRAT-R scores and for the three criterion variables to ensure that each variable was normally distributed. All WRAT-R scores and criterion variables were normally distributed. Intercorrelations were run on the Word Identification, Spelling, and Arithmetic WRAT-R subtests, and fully support the proposed criterion variables. That is, all three academic scores were significantly correlated (r > .46),

Table 8

Correlations Among the Final Battery's Six Neuropsychological Domains

Domain	Intelligence	Language	Visual-Spatial[a]	Attention-Memory[a]	Fine Motor[a]
Language	**.46****				
Visual-Spatial	**.30***	.14			
Attention-Memory[a]	**.41****	.23	.23		
Fine Motor[a]	.12	.24	.00	.17	
Somesthetic[a]	.08	.08	-.01	.05	**.31***

[a] Composite domain comprised of several subtests.

$^{*}p < .05$. $^{**}p < .01$.

supporting the Combined Academic variable. The Word Identification and Spelling were more strongly correlated (r = .84) than was either subtest with Arithmetic (r = .46 and r = .50, respectively), supporting the Language-Related and Arithmetic variables.

Development of Partitioning of Achievement Distribution

The subjects' achievement distribution was partitioned into three levels, Low Average, Average, and High Average, for each of the three criterion variables: Combined Academic, Language-Related, and Arithmetic measures. Data were originally partitioned by using both absolute performance thresholds (based on test standard scores), and terciles (based on sample distributions). The results of the two distributions were almost identical, and therefore, it was decided to partition scores based on absolute performance. The "Low Average" group's mean Combined Academic standard score (M = 88, n = 18) was derived from children whose scores were below 95, the "Average" group's mean standard score (M = 99, n = 19) was based upon children whose scores lay between 95 and 105, and the "High Average" group's mean standard score (M = 114, n = 27) represented children whose scores were above 105. Univariate analyses of variance (ANOVA) confirmed that these three levels of achievement groupings were significantly different from each other for each of the three criterion variables: Combined Academic (F = 130.1, df = 2, 61, p = .001), Language-Related (F = 142.7, df = 2, 61, p = .001), and Arithmetic (F = 124.1, df = 2, 61, p < .001).

Main Research Questions and Analyses

Discriminant Function Analyses to Test Relationship of Predictor to Criterion Variables

The first three research questions were related to the ability of neuropsychological tests to adequately distinguish High Average, Average, and Low Average academic achievement groups as defined by the three criterion variables. The battery had not been used in prior studies, and determining the relative strengths of the unique predictors was of strongest interest, so stepwise discriminant function analysis was the statistical method chosen for the classification analyses. Three stepwise discriminant function analyses were performed, one for each of the three criterion variables: Combined Academic, Language-Related, and Arithmetic variables. Each of the six neuropsychological predictors were allowed opportunity for selection into the stepwise analysis during each of the three discriminant function procedures.

A second set of Wilks' stepwise discriminant function analyses were performed for each of the three criterion variables, excluding Intelligence as a predictor, for two reasons. First, the Intelligence predictor correlates with the Language, Attention-Memory, and Visual-Spatial predictors such that colinearity may obfuscate the value of the other predictors (when Intelligence is included). Second, because of the social and political problems generated in using "intelligence tests" with Ethnic Minority children (see for example Sattler, 1992, pp. 566 - 572), the ability of the other domains to predict achievement was of interest.

The Wilks' stepwise discriminant procedure was used to assess the importance and strength of the predictors' ability to account for unique variance among the criterion groups. At each step, the single predictor that best minimized the value of Wilks' Lambda was selected (and the Wilks' Lambda statistic and probability level of the function reported) and then the other predictors were re-evaluated to assess unique predictive variance. The Wilks' Lambda statistic generated determines the importance of the predictor in the discriminant function, and the importance of the predictor <u>decreases</u> as the value of the function approaches the limit of 1.0. The process continued, as long as unselected variables had a tolerance ($1 - R^2_i$) of greater than 0.001 (i.e., the predictor does not closely approximate a linear combination of other variables) and as long as the predictor had an $\underline{F}$-value of over 1.0 when re-evaluated following the latest step. (The F-value is the ratio of the within-groups sum of squares to the between-groups sum of squares for the discriminant function grouping variable.)

(Question 1.) It was expected that the first function derived from the selected predictors of the screening battery would classify children's Combined Academic scores for the three groups of achievement at a higher rate (approximately 85%) than previous batteries (and misclassify fewer of their scores). Obtaining this rate of accuracy would represent a more clinically useful rate than previous three-group classification screening rates of 61% for regular classroom students (Teeter, 1985) and two-group success rates (60-75%) for identifying problem readers (Benton, 1978; Jansky, 1978; Satz & Fletcher, 1988; Silver & Hagin, 1990).

(Question 2.) It was hypothesized that each of the first functions generated from the selected predictors of the battery would be equally effective at classifying the Language-Related and the Arithmetic groups. It was expected that each of the first functions derived from the selected predictors would correctly classify approximately 85% of the three groups (Low Average, Average, and High Average) using Language-Related and Arithmetic scores.

(Question 3.) It was hypothesized that the Intelligence measure would not be selected in the stepwise analyses of any of the three criterion measures. This expectation was based upon the idea that the Intelligence measure would be a poor predictor for Ethnic Minority children (two-fifths of the sample) and kindergarten and first grade children (half the sample). In contrast, it was hypothesized that the Language, Attention-Memory, Fine Motor, Somesthetic, and Visual-Spatial variables would be selected and prove to be equally good predictors of all three criterion achievement measures.

MANOVAS/t-Tests to Assess Cultural Fairness of Measures

The next research question investigated the relative cultural fairness of the six neuropsychological areas. It was presumed that obtained test differences between Ethnic

Minority and Ethnic Majority children did not reflect SES or educational opportunity, because the demographic data between the two ethnic groups were similar (all of the children lived in the same neighborhood and attended the same schools). Therefore, for the present purposes, any differences between groups of Ethnic Minority and Ethnic Majority children were attributed to the "cultural unfairness" of the tests.

First, power analyses were performed on the two ethnic groups' six predictor variables to assess the probability of failing to reject a null hypothesis (when the null would be false). Given sufficient power, the analyses to test these research questions were then contingent on the results of the assessment of dependence/independence of the six areas. A multiple analysis of variance (MANOVA) was performed to assess cultural fairness (lack of significant differences) comparing scores of the Ethnic Majority children (n = 38) to the combined group of Ethnic Minority children (n = 26) across the six neuropsychological areas, because significant correlations among the areas demonstrated their dependence. (If the conceptual domains had been found to be independent, individual t-tests would have been conducted on each of the five constructs, instead.) Sample effect-size calculations were performed on significant differences between ethnic groups to quantify the clinical significance of sample size group differences.

(Question 4.) It was expected that the "Intelligence" measure would be found to be culturally biased, in that scores on the Intelligence test would be significantly higher for Ethnic Majority children than for the Ethnic Minority children.

(Question 5.) It was hypothesized that the Attention-Memory, Language, Fine Motor, Visual-Spatial, and Somesthetic areas would be found to be culturally-fair predictors. That is, it was expected that scores for these five predictors would not differ significantly between Ethnic Minority and Ethnic Majority children.

Exploratory Research Questions and Analyses

Discriminant Function Analyses to Assess Cultural Fairness

Two exploratory research questions were related to the cultural fairness of the six neuropsychological measures in predicting achievement for the separate groups of Ethnic Minority and Ethnic Majority children. Three Wilks' stepwise discriminant function analyses, one for each of the criterion variables (Combined Academic, Language-Related, and Arithmetic), analogous to the discriminant function analyses for the full sample, were conducted separately for each of the two ethnic groups. In all three analyses for each ethnic group, measures from all six predictors were allowed opportunity for selection in the stepwise function predicting the classification of the three achievement groups (Low Average, Average, and High Average). A complete second set of Wilks' stepwise discriminant function analyses was conducted, excluding Intelligence, to determine how well the remaining predictors performed.

(Question 6.) It was expected that the Intelligence measure would be selected in the Ethnic Majority children's achievement analyses, but would not be selected in the Ethnic Minority children's achievement analyses, for all three criterion variables because of the cultural bias of the Intelligence predictor.

(Question 7.) It was hypothesized that the Language, Fine Motor, Attention-Memory, Visual-Spatial, and Somesthetic areas would be selected and predict all three achievement criteria equally well, for the Ethnic Majority and Ethnic Minority groups.

Discriminant Function Analyses to Assess Age Differences

Two additional exploratory questions were related to predicting achievement for children of different age levels. Controversy exists regarding whether neuropsychological predictors of younger and older children's achievement scores change with maturation. There are at least three competing views. The first view is that tests of perceptual, sensory, and motor processes are more important predictors of reading in younger (than older) children, while tests of conceptual and language processes are better predictors of reading in older (than younger) children (Satz et al., 1978). A second perspective suggests that tests of visual-perceptual processes are predictive of learning problems throughout development (Silver & Hagin, 1990). A third view is that tests of language processes are predictive of reading problems throughout development (Jansky, 1978).

To assess such age-related differences in predictors of achievement, the sample was divided roughly in half by age. The "Younger" children's group was comprised of kindergartners and first graders ($\underline{n}$ = 31) and the "Older" children's group was comprised of second and third graders ($\underline{n}$ = 34). For each age group three Wilks' stepwise discriminant function analyses were conducted, predicting each of the three academic criterion variables. Scores from each of the six neuropsychological predictors for each of the two age groups were allowed opportunity for selection in the stepwise analyses for each of the three criterion variables. As with the previous research questions, a complete set of stepwise discriminant function analyses were performed excluding Intelligence to assess how well the battery's remaining five domains predicted achievement.

(Question 8.) It was hypothesized that the Aphasia Screening Test would be selected in the Older children's achievement analysis but not the Younger children's achievement analysis for the Language-Related and Combined Academic criterion variables, consistent with Satz et al. (1978). It was expected that the Aphasia Screening Test would not significantly predict Arithmetic better for either Younger children or Older children.

(Question 9.) It was expected that the Fine Motor, Somesthetic, Attention-Memory, and Visual-Spatial predictors would be selected and predict achievement in the Younger but not the Older children, for all three achievement criteria, consistent with Satz et al. (1978).

CHAPTER IV

RESULTS

Main Research Questions and Analyses

Discriminant Function Analyses to Test Relationship of Predictor to Criterion Variables

Overall three-group analyses. The main research questions involved assessing the extent to which scores grouped into High Average, Average, and Low Average achievement groupings on each criterion variable may be discriminated from each another using the six predictor variables: Intelligence, Language, Attention-Memory, Visual-Spatial, Somesthetic, and Fine Motor. There were three criterion variables (Combined Academic, Language-Related, and Arithmetic abilities), and thus three separate discriminant function analyses were performed, one for each criterion variable. Stepwise selection discriminant function analysis (Wilks' method) was used in all classification analyses to ascertain the unique relative strength of the six neuropsychological measures to predict academic achievement. Two complete sets of discriminant function analyses were performed, one set including and the other excluding Intelligence as a predictor, in order to examine the ability of the remaining variables to predict achievement.

Three-group overall analyses: Including Intelligence. The discriminant function group classifications and relative strengths of predictors (including Intelligence) that met inclusion criteria, for the Combined Academic, Language-Related and Arithmetic variables are shown in Table 9.

Table 9

Three-Group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence)

Combined Academic Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	28	17 (60.7%)	7 (25.0%)	4 (14.3%)
Average	18	7 (38.9%)	8 (44.4%)	3 (16.7%)
Low Average	18	1 (5.6%)	2 (11.1%)	15 (83.3%)

Note. Total cases correctly classified = 62.5%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.72	.0001
2.	Somesthetic	.64	.0001
3.	Attention-Memory	.58	.0001

(continued)

Table 9 (continued)

Arithmetic Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	27	19 (70.4%)	5 (18.5%)	3 (11.1%)
Average	20	6 (30.0%)	8 (40.0%)	6 (30.0%)
Low Average	17	3 (17.6%)	1 (5.9%)	13 (76.5%)

Note. Total cases correctly classified = 62.5%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.78	.0006
2.	Attention-Memory	.74	.001
3.	Fine-Motor	.70	.002

(continued)

Table 9 (continued)

Language-Related Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	21	14 (66.7%)	5 (23.8%)	2 (9.5%)
Average	18	5 (27.8%)	8 (44.4%)	5 (27.8%)
Low Average	25	6 (24.0%)	5 (20.0%)	14 (56.0%)

Note. Total cases correctly classified = 56.3%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.81	.002
2.	Attention-Memory	.72	.0006

The Combined Academic stepwise discriminant function (including Intelligence) was highly significant ($X^2 = 32.6$, df = 6, p = .0001), and produced a correct classification of 63% of children's scores into achievement groups. The true-positive classification of the Low Average group was 83%, but only 44% and 61%, respectively, for the Average and High Average groups. (The misclassification rates for all three-group analyses will be discussed at the end of each set of three-group analyses.) The Box's M statistic was significant (Box's M = 27.1, df =12, 13023, p = .01), raising concerns regarding multivariate normality and homogeneity of covariance. The strongest predictor selected in the stepwise function was Intelligence, followed by Somesthetic, and Attention-Memory measures, in that order.

The Arithmetic stepwise discriminant function (including Intelligence) was significant ($X^2 = 21.2$, df = 6, p = .002), and produced a correct classification rate of 63%. The Low Average and High Average groups's true-positive rates were 77% and 70%, respectively, and the Average group's success rate was lower (40%). The Box's M statistic was nonsignificant for this analysis (Box's M = 6.4, df = 12, 13340, p = ns), suggesting that the assumptions for discriminant analysis have been met using this conservative test. The predictors that met selection criteria were Intelligence, Attention-Memory, and Fine Motor, in descending order of strength.

The discriminant function for Language-Related achievement (including Intelligence) was significant ($X^2 = 19.6$, df = 4, p = .0006) and produced an overall success rate of 55%. True-positive prediction rates of Language-Related achievement for the Low Average group was 56%, for the Average group was 44%, and for the High Average group was 67%. The Box's M statistic was nonsignificant (Box's M = 8.8, df = 6, 61073, p = ns). Only two predictors met selection criteria: Intelligence was the strongest predictor followed by Attention-Memory.

Three-group overall analyses: Excluding Intelligence. The stepwise discriminant function group classifications and relative strength of selected predictors (excluding Intelligence) for the Combined Academic, Language-Related and Arithmetic criteria are presented in Table 10.

The Combined Academic stepwise discriminant function (excluding Intelligence) was highly significant (X^2 = 25.1, df = 6, p = .0003), and produced a total classification success rate of 64% which was equivalent to the rate including Intelligence (63%). The true-positive classification rates were 72%, 50%, and 68%, respectively, for the Low Average, Average, and High Average groups. The Box's M statistic was highly significant (Box's M = 49.0, df = 12, 12023, p = .0001). With Intelligence excluded, Attention-Memory became the strongest predictor, followed by Somesthetic, and Language, in that order.

The Arithmetic stepwise discriminant function (excluding Intelligence) was significant (X^2 = 17.1, df = 8, p = .03), but the classification success rate produced (52%) was lower than the rate that included the Intelligence predictor (63%). The High Average and Low Average groups's true-positive rates were 63% and 59%, respectively, whereas the Average group's rate was poor, even slightly lower than chance (30%). The Box's M statistic was highly significant for the three-group analysis (Box's M = 57.2, df = 20, 10175, p = .0001). Again, Attention-Memory was selected as the strongest predictor (excluding Intelligence), followed by Fine Motor, Language, and Somesthetic measures, in that order.

The Language-Related discriminant function (excluding Intelligence) was highly significant (X^2 = 11.4, df = 2, p = .003), but again the classification success rate (44%) was lower than the rate of the function that included Intelligence (55%). True-positive prediction rates of Language-Related achievement for the Low Average group was 56%, and for the High Average group was 43%, but was poor for the Average group 28% (slightly below the level of

Table 10

Three-Group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence)

Combined Academic Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	28	19 (67.9%)	7 (25.0%)	2 (7.1%)
Average	18	7 (38.9%)	9 (50.0%)	2 (11.1%)
Low Average	18	2 (11.1%)	3 (16.7%)	13 (72.2%)

Note. Total cases correctly classified = 64.1%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.79	.0008
2.	Somesthetic	.70	.0003
3.	Language	.66	.0003

(continued)

Table 10 (continued)

Arithmetic Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	27	17 (63.0%)	7 (25.9%)	3 (11.1%)
Average	20	7 (35.0%)	6 (30.0%)	7 (35.0%)
Low Average	17	5 (29.4%)	2 (11.8%)	10 (58.8%)

Note. Total cases correctly classified = 51.6%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.85	.008
2.	Fine Motor	.82	.02
3.	Language	.78	.02
4.	Somesthetic	.75	.03

(continued)

Table 10 (continued)

Language-Related Achievement

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	21	9 (42.9%)	4 (19.0%)	8 (38.1%)
Average	18	8 (44.4%)	5 (27.8%)	5 (27.8%)
Low Average	25	5 (20.0%)	6 (24.0%)	14 (56.0%)

Note. Total cases correctly classified = 43.8%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.83	.003

chance). The Box's M statistic was nonsignificant (Box's M = 2.8, df = 2, 8008, p = ns). Attention-Memory was the single predictor that met inclusion criteria.

Misclassifications. Given the three groupings of the criterion variable, by chance alone, one would normally expect a 33% misclassification rate in each of the two other groupings for a total misclassification rate of 66%. The misclassification rate in these analyses was generally lowest for the Low Average group (in 4 of 6 analyses), at a middle range for the High Average group, and always was highest for the Average group. The worst misclassification rates were for the Average group's Arithmetic and Language-Related achievement in which Intelligence was excluded. The relative difficulty in predicting Average scores probably occurred as an artifact of restricted range (see Discussion). There was no consistent pattern of the Average group being misclassified as either High Average or Low Average, across analyses. Overall, misclassifications of the Low Average and High Average group "extremes" were predominantly in the Average group (in 8 of 12 comparisons). The misclassification rate for Language-Related achievement was higher than for the Combined Academic or Arithmetic criteria. The "miss" rates with and without Intelligence as a predictor were nearly the same for the Combined Academic, but the exclusion of Intelligence in the Arithmetic and Language-Related analyses increased misclassifications by 11% for each of these two criteria.

Two-group overall analyses. Two-group stepwise discriminant function analyses, Low Average versus "Average-Plus" (combined Average and High Average scores), were performed in addition to the three-group analyses. The two-group analyses were added for two reasons. First, most previous research only has sought to identify "below average" learners, so contrasting the Low Average group with all other scores provides a similar grouping in order to compare predictions of this battery with those of other studies. Second, given the significant Box's M for three of six initial three-group analyses, the two-group analyses were performed in an attempt to satisfy the conservative test of multivariate normality and homogeneity of covariance in all analyses by increasing the sample-size (n) per cell. The Low Average group was defined as including all standard scores on the criterion measures that were below 95, while the Average-Plus group included all standard scores on the criterion measure at or above 95. Each set of discriminant function analyses again were performed twice (including and excluding Intelligence as a predictor).

Two-group overall analyses: Including Intelligence. Table 11 reveals the two-group classifications resulting from the stepwise discriminant function analyses and the relative strength of predictors (including Intelligence) that met the selection criteria for all three criterion variables.

The two-group Combined Achievement stepwise discriminant function (including Intelligence) was significant ($X^2 = 27.7$, df = 4, p = .0001), and produced a correct classification rate of 83%. The true-positive rate for the Low Average group was 89% and for the Average-Plus group was 80%. The Box's M statistic was again significant (Box's M = 31.6, df = 10, 4951, p = .002). The strongest predictor was Intelligence, followed by Somesthetic, Attention-Memory, and Fine Motor, in that order.

The two-group Arithmetic stepwise discriminant function (including Intelligence) was significant ($X^2 = 14.8$, df = 3, p = .002), and accurately classified 72% of the scores. The discriminant function accurately classified 82% of the Low Average group and 68% of the Average-Plus group. The Box's M statistic was nonsignificant (Box's M = 5.8, df = 6, 5633, p = ns). Intelligence, Attention-Memory, and Fine Motor were the strongest three predictors, in descending order.

The two-group Language-Related criterion stepwise discriminant function (including

Table 11

Two-Group Classifications of Achievement Predictions and Predictors Selected in Analyses (Including Intelligence)

Combined Academic Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	46	37 (80.4%)	9 (19.6%)
Low Average	18	2 (11.1%)	16 (88.9%)

Note. Total cases correctly classified = 82.8%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.77	.0001
2.	Somesthetic	.70	.0001
3.	Attention-Memory	.64	.0001
4.	Fine Motor	.63	.0001

(continued)

Table 11 (continued)

Arithmetic Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	47	32 (68.1%)	15 (31.9%)
Low Average	17	3 (17.6%)	14 (82.4%)

Note. Total cases correctly classified = 71.9%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.84	.0009
2.	Attention-Memory	.80	.001
3.	Fine Motor	.78	.002

(continued)

Table 11 (continued)

Language-Related Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	39	31 (79.5%)	8 (20.5%)
Low Average	25	8 (32.0%)	17 (68.0%)

Note. Total cases correctly classified = 75.0%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.83	.0007
2.	Intelligence	.79	.001

Intelligence) was significant (X^2 = 13.6, df = 2, p = .001), and produced a correct classification rate of 75%. The true-positive rates for the Low Average group was 68% and for the Average-Plus group was 80%. The Box's M statistic was nonsignificant (Box's M = 1.6, df = 3, 115800, p = ns). Attention-Memory was the strongest predictor followed by Intelligence.

Two-group overall analyses: Excluding Intelligence. The two-group classifications derived from the stepwise discriminant function analyses and the relative strength of predictors (excluding Intelligence) that met the selection criteria for all three criterion variables are shown in Table 12.

The two-group Combined Achievement criterion stepwise discriminant function (excluding Intelligence) was significant (X^2 = 21.5, df = 4, p = .0002), and produced a true-positive classification rate (83%) equal to that of the analysis including Intelligence. The true-positive rates for the Low Average and Average-Plus groups were equal (83%). The Box's M statistic was significant (Box's M = 51.9, df = 10, 4951, p = .0001). Attention-Memory, Somesthetic, Language, and Fine Motor were the strongest unique predictors, in descending order.

The two-group Arithmetic criterion stepwise discriminant function, excluding the Intelligence predictor, was significant (X^2 = 11.2, df = 4, p = .02) and produced a success rate (73%) essentially equal to that of the analysis including Intelligence (72%). The discriminant function correctly classified 71% of Low Average and 75% of Average-Plus groups's scores. The Box's M statistic was significant (Box's M = 32.2, df = 10, 4247, p = .001). Attention-Memory again was the strongest predictor, followed by Language, Fine Motor, and Somesthetic, in that order.

The two-group Language-Related stepwise discriminant function, excluding Intelligence as a predictor, was significant (X^2 = 11.4, df = 1, p = .0007), and produced a

Table 12

Two-Group Classifications of Achievement Predictions and Predictors Selected in Analyses (Excluding Intelligence)

Combined Academic Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	46	38 (82.6%)	8 (17.4%)
Low Average	18	3 (16.7%)	15 (83.3%)

Note. Total cases correctly classified = 82.8%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.82	.0005
2.	Somesthetic	.74	.0001
3.	Language	.72	.0002
4.	Fine Motor	.70	.0003

(continued)

Table 12 (continued)

Arithmetic Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	47	35 (74.5%)	12 (25.5%)
Low Average	17	5 (29.4%)	12 (70.6%)

Note. Total cases correctly classified = 73.4%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.89	.007
2.	Language	.87	.01
3.	Fine Motor	.85	.02
4.	Somesthetic	.83	.02

(continued)

Table 12 (continued)

Language-Related Achievement

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	39	26 (66.7%)	13 (33.3%)
Low Average	25	11 (44.0%)	14 (56.0%)

Note. Total cases correctly classified = 62.5%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.83	.0007

classification success rate of 63% that was lower than the rate of the analysis including the Intelligence predictor (75%). The true-positive rate for the Low Average group was 56% and for the Average-Plus group was 67%. The Box's M statistic was nonsignificant (Box's M = 1.6, df = 1, 10041, p = ns). The single predictor that met selection criteria was Attention-Memory.

Three- and two-group overall analyses: Summary. While use of the best predictors did not fulfill expectations by producing an improvement upon the three-group predictive accuracy of prior batteries (Research Question 1), the Combined Academic two-group classification rates (including and excluding Intelligence) approached the expected rate and demonstrated an improved rate of prediction compared to other batteries. (As expected mathematically, the two-group classification success rates were higher than the three-group rates.) The classification results in all three- and two-group analyses were significant and all demonstrated predictability equivalent to results from prior batteries for the Combined Academic (with or without the Intelligence predictor) and for the Arithmetic criterion (with Intelligence). Also, two-group success rates equivalent to those of previous batteries emerged for Arithmetic (with or without Intelligence) and Language-Related (with Intelligence).

Exclusion of Intelligence as a predictor did not lower the overall Combined Academic classification hit rate (64%), but it lowered each of the other two overall achievement rates by about 11%. For the two-group analyses, exclusion of Intelligence as a predictor did not lower Combined Academic or Arithmetic two-group success rates, but it did lower the Language-Related rate by 8%.

As stated in Research Question 2, the battery's measures were expected to predict both areas of achievement equally well, however, Language-Related achievement was more difficult to predict using this battery than Arithmetic achievement, irrespective of inclusion or exclusion of Intelligence as a predictor. (For the youngest children, the "Language-Related" criterion measures, copying symbols and identifying letters, probably are not adequate measures of language, so conclusions regarding the battery's ability to predict kindergartners' Language-Related achievement should be considered with caution.)

Three of six Box's M analyses were significant for both the three-group and two-group sets of analyses. [Those were: Combined Academic analyses (with and without Intelligence) and Arithmetic analysis (without Intelligence).] While the significant Box's M results raise questions regarding the assumptions of multivariate normality and homogeneity of covariance for the analyses, the Language-Related and Arithmetic (including Intelligence) analyses meet the assumptions using the conservative test.

Regarding Research Question 3, the expectation that Attention-Memory would predict achievement well was supported, but other variables failed to emerge consistently as good predictors, in part because of the predictive power of Intelligence, the strongest predictor for five of six analyses in which it was included (second only to Attention-Memory for the two-group Language-Related analysis). Attention-Memory emerged as a significant predictor (with or without Intelligence) of Combined Academic, Arithmetic, and Language-Related achievement. Attention-Memory was the strongest single predictor in all six analyses in which Intelligence was excluded. Somesthetic predicted Combined Academic (with or without Intelligence), and Arithmetic (only if Intelligence was excluded). Fine Motor predicted Arithmetic (with or without Intelligence). Language only emerged as a predictor of Arithmetic and Combined Academic (when Intelligence was excluded). Visual-Spatial was not a significant unique predictor in any of the analyses.

The pattern of results with respect to relative strengths of various predictors was identical for the three-group and two-group analyses, with one exception: Fine Motor emerged as one of the selected predictors of two-group Combined Academic achievement (with or without Intelligence) as well as of Arithmetic (with or without Intelligence). For the three-group analyses, Fine Motor was significantly predictive only of the Arithmetic criterion.

MANOVA/t-Tests to Assess the Cultural Fairness of Measures and Effects of Handedness

To answer the research questions regarding the cultural fairness of criterion and predictor measures, multivariate analysis of variance in MANOVA and Student's t-tests were used to compare the scores of Ethnic Majority and Ethnic Minority children on each measure. The t-tests were used to compare the groups on the criterion variables, because the Combined Academic criterion is a linear combination of the other two variables. A MANOVA was used to compare the two ethnic groups on the predictor variables because the intercorrelations between several of the predictors were significant. Follow-up univariate tests of analysis of variance in ANOVA also were performed.

Power analyses were conducted prior to comparing ethnic group means on the criterion variables in order to ascertain the likelihood of finding significant results, assuming such differences exist. The power to detect real differences between Ethnic Minority and Ethnic Majority groups for the criterion variables of Combined Academic, Arithmetic, and Language-Related achievement, was .73, .59, and .85, respectively. Given these data that were interpreted as demonstrating fair power overall (Cohen & Cohen, 1975), it was decided to perform t-tests to determine if there were significant differences between the two ethnic groups on the achievement measures.

Table 13 lists the results of t-tests performed between groups of Ethnic Majority and Ethnic Minority children's scores on the criterion variables. No significant differences between ethnic groups emerged in the analyses of either Combined Academic, Language-Related, or

Table 13

The t-Tests Between Ethnic Minority and Ethnic Majority Children's Scores on Criterion Variables

Variable	Ethnic Group	n	M (SD)	t-Value	df	p
Combined Academic	Ethnic Minority	26	100.1 (13.4)	-1.13	62	ns
	Ethnic Majority	38	103.6 (10.8)			
Language-Related	Ethnic Minority	26	99.9 (15.0)	-0.79	62	ns
	Ethnic Majority	38	102.6 (12.9)			
Arithmetic	Ethnic Minority	26	100.6 (14.3)	-1.44	62	ns
	Ethnic Majority	38	105.4 (12.1)			

Arithmetic variables.

Power analyses were also performed on the eight predictor variables to ascertain their ability to detect significant differences (if differences exist) with respect to the separate groups of Ethnic Minority and Ethnic Majority children. The power for Intelligence was .98, for Language was .63, for Attention-Memory was .93, for Visual-Spatial was .58, for Fine Motor was .94, and for Somesthetic was .99. It was decided to perform the MANOVA and follow-up ANOVA analyses between the ethnic groups on the predictor variables, because the power ranged from fair to good for the predictor variables (Cohen & Cohen, 1975).

A MANOVA was used to test for differences between the Ethnic Minority and Ethnic Majority children's scores (because of the intercorrelations among pairs of predictors) and the resulting F-ratio was significant ($F = 2.28$, $df = 6, 57$, $p = .05$). Table 14 lists the means and standard deviations of the six predictor variables by ethnic group and the results of univariate analysis of variance (ANOVA) follow-up tests on each of the six predictors. The ANOVAS indicated that the Ethnic Majority children's scores were significantly higher than the scores of the Ethnic Minority children on two predictors: Intelligence and Language.

In order to quantify the clinical significance of sample-size differences between ethnic groups on the Intelligence and Language variables, effect sample-size analyses were performed, establishing the proportion in standard deviation units represented by the difference between group means. The Intelligence and Language effect sizes were .77 and .63, respectively, which may be considered fairly large (Cohen & Cohen, 1975). However, the level of performance clinical descriptors (Wechsler, 1974) assigned to the Intelligence measure's ethnic group means were "Average" and "High Average" for the Ethnic Minority and Ethnic Majority groups, respectively. The level of performance clinical descriptors for the two ethnic groups on the Language measure were "Low Average" and "Average." The

Table 14

Ethnic Group Means and Standard Deviations on Predictors and ANOVAS Between Scores of Ethnic Minority and Ethnic Majority Groups

Predictor	Ethnic Group	n	M (SD)	F	df	Significance of F
Intelligence	Ethnic Minority	26	9.0 (3.7)	9.18	1,62	**.004**
	Ethnic Majority	38	11.9 (3.7)			
Language	Ethnic Minority	26	40.7 (20.9)	6.59	1,62	**.01**
	Ethnic Majority	38	50.2 (11.1)			
Attention-Memory	Ethnic Minority	26	31.6 (8.6)	0.23	1,62	ns
	Ethnic Majority	38	32.4 (5.3)			
Fine Motor	Ethnic Minority	26	66.3 (16.6)	0.13	1,62	ns
	Ethnic Majority	38	67.7 (14.6)			
Somesthetic	Ethnic Minority	26	16.2 (0.61)	0.59	1,62	ns
	Ethnic Majority	38	16.1 (0.69)			
Visual-Spatial	Ethnic Minority	26	45.8 (9.8)	2.34	1.62	ns
	Ethnic Majority	38	49.3 (8.0)			

distinction between these sets of descriptors for the ethnic group means are of small or marginal clinical significance.

Culture fairness: Summary. As expected in the hypothesis of the fourth Research Question, Intelligence scores were higher for Ethnic Majority children than for Ethnic Minority children. Unexpectedly, it was found that the Language measure scores also were higher for Ethnic Majority children than for Ethnic Minority children (Research Question 5). There were no differences between ethnic groups with respect to Attention-Memory, Visual-Spatial, Somesthetic, and Fine Motor predictors (Research Question 5).

Handedness. To ensure that hand dominance was not a significant factor to consider in the analyses, t-tests (one-tailed) were performed on the three criterion measures comparing group scores of right- and left-handed children. The results demonstrated that left-handed children did not have more difficulty in school than their right-handed peers with respect to Combined Academic ($t = 1.17$, $df = 62$, $p = ns$), Language-Related ($t = .98$, $df = 62$, $p = ns$), or Arithmetic ability ($t = 1.13$, $df = 62$, $p = ns$).

Exploratory Research Questions and Analyses

Discriminant Function Analyses to Assess Cultural Fairness

Three-group analyses within ethnic groups. The first set of exploratory research questions involved assessing whether neuropsychological measures could predict achievement for Ethnic Minority as well as Ethnic Majority children (analyzing these two ethnic groups's scores separately). Just as for the full sample, three stepwise (Wilks' method) discriminant function analyses (one for each of the three criterion variables) differentiating the three levels of academic achievement were conducted for each of the two ethnic groups. Each set of stepwise discriminant function analyses again were performed twice (including and excluding Intelligence as a predictor).

<u>Three-group analyses within ethnic groups: Including Intelligence</u>. The stepwise discriminant function group classifications for the criterion variables, by ethnic group, and the relative strengths of predictors that met selection criteria (including Intelligence) are presented in Table 15.

The stepwise discriminant function (including Intelligence) of Ethnic Minority children's Combined Academic scores was significant (X^2 = 18.3, df = 4, p = .001), and produced a 69% classification success rate. The true-positive classification rates for Low Average, Average, and High Average Ethnic Minority children's groups were 88%, 57%, and 64%, respectively. The Box's M statistic was nonsignificant for the analysis (Box's M = 3.9, df = 6, 6188, p = ns). Intelligence was the strongest predictor followed by Attention-Memory.

The Ethnic Majority children's Combined Academic stepwise discriminant function (including Intelligence) was significant (X^2 = 24.7, df = 8, p = .002), and produced an overall true-positive rate of 71%. The classification success rate was 70% for the Low Average group, 77% for the High Average group, and 64% for the Average group. The Box's M statistic was highly significant for the three-group analysis (Box's M = 62.3, df = 20, 3008, p = .0002). Intelligence was the strongest predictor followed by Somesthetic, Attention-Memory, and Language, in that order.

The Ethnic Minority children's Arithmetic criterion stepwise discriminant function (including Intelligence) was significant (X^2 = 18.1, df = 8, p = .02), and produced a classification hit rate of 69%. The true-positive rate was 67% for the Low Average group, 67% for the Average group, and 75% for the High Average group. The Box's M statistic was significant (Box's M = 59.9, df = 20, 1840, p = .002). Attention-Memory, Somesthetic, Visual-Spatial, and Language were the strongest predictors (in descending order). Intelligence was included as a predictor, but (atypically) was not selected in the analysis, hence the

Table 15

Three-Group Classifications of Achievement Predictions and Strengths of Predictors for Two Ethnic Groups (Including Intelligence)

Combined Academic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	7 (63.6%)	2 (18.2%)	2 (18.2%)
Average	7	2 (28.6%)	4 (57.1%)	1 (14.3%)
Low Average	8	0 (0.0%)	1 (12.5%)	7 (87.5%)

Note. Total cases correctly classified = 69.2%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.59	.002
2.	Attention-Memory	.44	.001

(continued)

Table 15 (continued)

Combined Academic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	17	13 (76.5%)	3 (17.6%)	1 (5.9%)
Average	11	2 (18.2%)	7 (63.6%)	2 (18.2%)
Low Average	10	1 (10.0%)	2 (20.0%)	7 (70.0%)

Note. Total cases correctly classified = 71.1%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.73	.004
2.	Somesthetic	.62	.002
3.	Attention-Memory	.51	.001
4.	Language	.48	.002

(continued)

Table 15 (continued)

Arithmetic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	8	6 (75.0%)	2 (25.0%)	0 (0.0%)
Average	9	1 (11.1%)	6 (66.7%)	2 (22.2%)
Low Average	9	1 (11.1%)	2 (22.2%)	6 (66.7%)

Note. Total cases correctly classified = 69.2%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.77	.05
2.	Somesthetic	.57	.01
3.	Visual-Spatial	.49	.02
4.	Language	.43	.02

(continued)

Table 15 (continued)

Arithmetic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	19	13 (68.4%)	4 (21.1%)	2 (10.5%)
Average	11	3 (27.3%)	5 (45.5%)	3 (27.3%)
Low Average	8	1 (12.5%)	0 (0.0%)	7 (87.5%)

Note. Total cases correctly classified = 65.8%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.78	.01
2.	Somesthetic	.72	.02
3.	Attention-Memory	.67	.04
4.	Fine Motor	.63	.05

(continued)

Table 15 (continued)

Language-Related Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	8	4 (50.0%)	3 (37.5%)	1 (12.5%)
Average	8	4 (50.0%)	2 (25.0%)	2 (25.0%)
Low Average	10	1 (10.0%)	2 (20.0%)	7 (70.0%)

Note. Total cases correctly classified = 50.0%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.61	.003

(continued)

Table 15 (continued)

Language-Related Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership: High Average	Average	Low Average
High Average	13	10 (76.9%)	1 (7.7%)	2 (15.4%)
Average	10	1 (10.0%)	6 (60.0%)	3 (30.0%)
Low Average	15	3 (20.0%)	5 (33.3%)	7 (46.7%)

Note. Total cases correctly classified = 60.5%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.76	.009
2.	Attention-Memory	.68	.01

classification and prediction analyses including and excluding Intelligence were exactly the same.

The Ethnic Majority's Arithmetic stepwise discriminant function (including Intelligence) was significant (X^2 = 15.3, df = 8, p = .05), and produced a classification success rate of 66%. The true-positive hit rates for Low Average, Average, and High Average groups were 88%, 46%, and 68%, respectively. The Box's M statistic was nonsignificant (Box's M = 32.6, df = 20, 1893, p = ns). Four predictors were selected in the stepwise analysis: Intelligence, Somesthetic, Attention-Memory, and Fine Motor, in descending order of strength.

The Ethnic Minority Language-Related stepwise discriminant function (including Intelligence) was significant (X^2 = 11.5, df = 2, p = .003), and produced a classification hit rate of 50%. The true-positive rate was 70% for the Low Average group, a poor 25% for the Average group (below the chance level of 33%), and 50% for the High Average group. The Box's M statistic was nonsignificant (Box's M = 1.8, df = 2, 1153, p = ns). Attention-Memory was the sole predictor that met selection criteria. Intelligence was allowed but failed to be selected in this analysis, hence the classification and prediction analyses including and excluding Intelligence were exactly the same.

The Ethnic Majority's Language-Related criterion stepwise discriminant function (including Intelligence) was significant (X^2 = 13.1, df = 4, p = .01), and produced a classification success rate of 61%. The true-positive hit rates for Low Average, Average, and High Average groups were 47%, 60%, and 77%, respectively. The Box's M statistic was nonsignificant for the analysis (Box's M = 5.4, df = 6, 15520, p = ns). Intelligence was selected as the strongest predictor, followed by Attention-Memory.

Three-group analyses within ethnic groups: Excluding Intelligence. The stepwise discriminant function group classifications (excluding Intelligence) for the criterion variables and the relative strengths of selected predictors, for each of two ethnic groups, are presented in Table 16. (Tables for Ethnic Majority children's Language-Related and Arithmetic analyses, in which Intelligence was excluded as a predictor, were omitted because the stepwise discriminant functions were nonsignificant.)

The Combined Academic criterion stepwise discriminant function for Ethnic Minority children's scores (excluding Intelligence) was significant ($X^2 = 16.0$, $\underline{df} = 60$, $\underline{p} = .01$). The total classification success rate of this analysis (73%) was slightly higher than the rate including Intelligence (69%). The true-positive classification rates for Low Average, Average, and High Average Ethnic Minority groups were 75%, 71%, and 73%, respectively. The Box's $\underline{M}$ statistic was nonsignificant (Box's $\underline{M} = 14.2$, $\underline{df} = 12, 1842$, $\underline{p} = \underline{ns}$). The selected predictors (excluding Intelligence) were Attention-Memory, Language, and Visual-Spatial, in descending order of strength.

The Ethnic Majority children's Combined Academic discriminant function (excluding Intelligence) was significant ($X^2 = 17.6$, $\underline{df} = 6$, $\underline{p} = .007$), and produced a classification rate (71%) that was identical to the rate of the analysis including Intelligence. The classification rate for true positives was 70% for the Low Average group, 82% for the High Average group, and 55% for the Average group. The Box's $\underline{M}$ statistic was highly significant (Box's $\underline{M} = 52.3$, $\underline{df} = 12, 3911$, $\underline{p} = .0001$). Attention-Memory was the strongest predictor (excluding Intelligence), followed by Somesthetic, and Language, in that order of selection.

The Arithmetic criterion discriminant function (excluding Intelligence) for Ethnic Minority children was identical to the analysis that had included Intelligence, since Intelligence was not selected as a significant predictor (see Table 15). The Ethnic Majority's Arithmetic criterion stepwise discriminant function was nonsignificant for the analysis excluding the Intelligence predictor ($X^2 = 7.2$, $\underline{df} = 4$, $\underline{p} = \underline{ns}$).

Table 16

Three-Group Classifications of Achievement Predictions and Strengths of Predictors for Two Ethnic Groups (Excluding Intelligence)

Combined Academic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	8 (72.7%)	2 (18.2%)	1 (9.1%)
Average	7	1 (14.3%)	5 (71.4%)	1 (14.3%)
Low Average	8	1 (12.5%)	1 (12.5%)	6 (75.0%)

Note. Total cases correctly classified = 73.1%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.61	.003
2.	Language	.54	.008
3.	Visual-Spatial	.48	.01

(continued)

Table 16 (continued)

Combined Academic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	17	14 (82.4%)	1 (5.9%)	2 (11.8%)
Average	11	4 (36.4%)	6 (54.5%)	1 (9.1%)
Low Average	10	1 (10.0%)	2 (20.0%)	7 (70.0%)

Note. Total cases correctly classified = 71.1%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.83	.04
2.	Somesthetic	.68	.009
3.	Language	.60	.007

(continued)

The Ethnic Minority Language-Related discriminant function (excluding Intelligence) was identical to the function including Intelligence, since in the latter analysis Intelligence was not selected (see Table 15). The Ethnic Majority's Language-Related criterion stepwise discriminant function was nonsignificant for the analysis excluding the Intelligence predictor ($X^2 = 2.4$, df = 2, p = ns).

Misclassifications. The misclassification rates in these ethnic group analyses again were generally lowest for the Low Average and High Average groups, and highest for the Average group (in 19 of 20 comparisons). As discussed below, the relatively high Average score miss rate is probably an artifact of restricted range. Overall, misclassifications of the Low Average and High Average groups generally lay in the Average group (in 9 of 11 comparisons). In the four of six possible comparisons, the misclassifications, including and excluding Intelligence as a predictor, were equivalent. The overall misclassification rates were generally equivalent comparing the two ethnic groups, with the exception of Language-Related achievement of Ethnic Majority children being predicted better than that of Ethnic Minority children. In two of three Ethnic Minority analyses, exclusion of Intelligence increased Low Average misclassification rates (11 - 13 points), and decreased Average and High Average misclassification rates (9 - 13 points). No other consistent pattern of misclassifications emerged from inspection of the data.

Two-group analyses. Two-group stepwise discriminant function analyses were conducted for each of the two ethnic groups, including and excluding Intelligence as a predictor.

Two-group analyses within ethnic groups: Including Intelligence. Table 17 shows the classification tables of two-group (Low Average versus Average-Plus) achievement and relative strengths of predictors (including Intelligence) that met selection criteria for all three

Table 17

Two-Group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence) for Two Ethnic Groups

Combined Academic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	18	14 (77.8%)	4 (22.2%)
Low Average	8	1 (12.5%)	7 (87.5%)

Note. Total cases correctly classified = 80.8%

Step	Predictor	Wilks' Lambda	p of Function
1.	Intelligence	.60	.0005
2.	Attention-Memory	.58	.002

(continued)

Table 17 (continued)

Combined Academic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	28	25 (89.3%)	12 (25.5%)
Low Average	10	1 (10.0%)	9 (90.0%)

Note. Total cases correctly classified = 89.5%

Step	Predictor	Wilks' Lambda	p of Function
1.	Attention-Memory	.84	.01
2.	Somesthetic	.70	.002
3.	Intelligence	.63	.001

(continued)

Table 17 (continued)

Arithmetic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	17	12 (70.6%)	5 (29.4%)
Low Average	9	3 (33.3%)	6 (66.7%)

Note. Total cases correctly classified = 69.2%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.83	.04
2.	Somesthetic	.78	.05

(continued)

Table 17 (continued)

Arithmetic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	30	25 (83.3%)	5 (16.7%)
Low Average	8	1 (12.5%)	7 (87.5%)

Note. Total cases correctly classified = 84.2%

Step	Predictor	Wilks' Lambda	p of Function
1.	Intelligence	.87	.03
2.	Somesthetic	.80	.02
3.	Attention-Memory	.75	.02

(continued)

Table 17 (continued)

Language-Related Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	16	13 (81.3%)	3 (18.8%)
Low Average	10	2 (20.0%)	8 (80.0%)

Note. Total cases correctly classified = 80.8%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.64	.001
2.	Intelligence	.59	.002
3.	Fine Motor	.55	.004

criterion measures, considering Ethnic Minority and Ethnic Majority groups separately. (The table for Ethnic Majority children's Language-Related analysis was omitted, because the discriminant function was nonsignificant.)

The Ethnic Minority children's two-group Combined Academic discriminant function (including Intelligence) was significant ($X^2 = 12.7$, df = 2, p = .002), and produced an 81% success rate. The Low Average group's true-positive rate was 88%, and the Average-Plus group's rate was 78%. The Box's M test was nonsignificant (Box's M = 2.5, df = 3, 3573, p = ns). Intelligence was selected as the strongest predictor, followed by Attention-Memory.

The Ethnic Majority children's two-group discriminant function for Combined Academic scores (including Intelligence) was significant ($X^2 = 15.7$, df = 3, p = .001), and produced a 90% true-positive hit rate. The classification success rates were essentially equal for the two groups: 89% and 90% for the Average-Plus and Low Average groups, respectively. The Box's M test was significant (Box's M = 17.5, df = 6, 1749, p = .02). Attention-Memory was the strongest predictor, followed by Somesthetic and Intelligence, in that order.

The Ethnic Minority children's two-group Arithmetic criterion discriminant function (including Intelligence) was significant ($X^2 = 5.8$, df = 2, p = .05), and produced a classification success rate of 69%. The true-positive rate for the Low Average group was 67%, and the rate for the Average-Plus group was 71%. The Box's M test was nonsignificant (Box's M = 6.8, df = 3, 6491, p = ns). Intelligence was selected as the strongest predictor, followed by Somesthetic.

The Ethnic Majority children's two-group Arithmetic stepwise discriminant function (including Intelligence) was significant ($X^2 = 9.8$, df = 3, p = .02), and produced a classification success rate of 84%. The hit rates for the Low Average and Average-Plus groups were 88% and 83%, respectively. The Box's M test was nonsignificant (Box's M =

12.7, df = 6, 946, p = ns). Intelligence was the strongest predictor, followed by Somesthetic and Attention-Memory, in that order.

The Ethnic Minority two-group Language-Related stepwise criterion discriminant function (including Intelligence) was significant (X^2 = 13.3, df = 3, p = .004), and produced a hit rate of 81%. Classification success rates for the Low Average and Average-Plus Ethnic Minority groups were equivalent (80% and 81%, respectively). The Box's M test was nonsignificant (Box's M = 12.3, df = 6, 2389, p = ns). Attention-Memory was the strongest predictor, followed by Intelligence and Fine Motor, in that order.

The Ethnic Majority two-group Language-Related discriminant function (including Intelligence) was non-significant (X^2 = 2.6, df = 1, p = ns).

Two-group analyses within ethnic groups: Excluding Intelligence. A second complete set of discriminant function analyses was performed, excluding the Intelligence predictor. The classification tables of two-group (Low Average versus Average-Plus) achievement and relative strengths of selected predictors (excluding Intelligence) for all three criterion measures, considering Ethnic Minority and Ethnic Majority groups separately, are presented in Table 18. (Tables for Ethnic Minority Arithmetic and Ethnic Majority Language-Related achievement were omitted as both of the discriminant functions of theses analyses were nonsignificant.)

The Ethnic Minority children's Combined Academic stepwise discriminant function (excluding Intelligence) was significant (X^2 = 10.1, df = 3, p = .02), and produced a predictive success rate (85%) that was slightly higher than the rate that included Intelligence (81%). The true-positive rate for the Low Average group was 75%, and the rate for the Average-Plus group was higher (89%). The Box's M test was nonsignificant (Box's M = 8.3, df = 6, 1170, p = ns). Attention-Memory was selected as the strongest predictor, followed by Language and Visual-Spatial, in that order.

Table 18

Two-Group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence) for Two Ethnic Groups

Combined Academic Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	18	16 (88.9%)	2 (11.1%)
Low Average	8	2 (25.0%)	6 (75.0%)

Note. Total cases correctly classified = 84.6%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.79	.02
2.	Language	.71	.02
3.	Visual-Spatial	.64	.02

(continued)

Table 18 (continued)

Combined Academic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	28	23 (82.1%)	5 (17.9%)
Low Average	10	3 (30.0%)	7 (70.0%)

Note. Total cases correctly classified = 79.0%

Step	Predictor	Wilks' Lambda	p of Function
1.	Attention-Memory	.84	.01
2.	Somesthetic	.70	.002

(continued)

Table 18 (continued)

Arithmetic Achievement: Ethnic Majority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	30	23 (76.7%)	7 (23.3%)
Low Average	8	3 (37.5%)	5 (62.5%)

Note. Total cases correctly classified = 73.7%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.90	.06
2.	Somesthetic	.82	.03

(continued)

Table 18 (continued)

Language-Related Achievement: Ethnic Minority Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	16	12 (75.0%)	4 (25.0%)
Low Average	10	2 (20.0%)	8 (80.0%)

Note. Total cases correctly classified = 76.9%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.64	.001
2.	Fine Motor	.60	.003

The Ethnic Majority children's two-group discriminant function for Combined Academic achievement (excluding Intelligence) was significant (X^2 = 12.6, df = 2, p = .002), and produced a classification success hit rate of 79%. The rates for the Low Average and Average-Plus groups were 70% and 82%, respectively. The Box's M test was significant (Box's M = 17.1, df = 3, 4659, p = .001). Attention-Memory was the strongest predictor followed by Somesthetic.

The Ethnic Minority children's two-group Arithmetic criterion discriminant function (excluding Intelligence) was nonsignificant (X^2 = 5.8, df = 3, p = ns).

The Ethnic Majority children's Arithmetic stepwise discriminant function (excluding Intelligence) was significant (X^2 = 7.1, df = 2, p = .03), and produced a hit rate of 74%. The classification success rate for the Low Average group was 63% and the rate for the Average-Plus group was 77%. The Box's M test was significant (Box's M = 11.1, df = 3, 2256, p = .02). Attention-Memory was selected as the strongest predictor, followed by Somesthetic.

The Ethnic Minority children's Language-Related stepwise discriminant function (excluding Intelligence) was significant (X^2 = 11.7, df = 2, p = .003), and produced a true-positive classification rate of 77%. The hit rate for the Low Average group was 80% and the rate for the Average-Plus group was 75%. The Box's M test was significant (Box's M = 9.3, df = 3, 13418, p = .04). Attention-Memory was the strongest predictor, followed by Fine Motor.

The Ethnic Majority Language-Related discriminant function, excluding the Intelligence predictor, was nonsignificant (X^2 = 2.9, df = 52 p = ns).

Three-and two-group analyses within ethnic groups: Summary. Three-group classification rates (with and without Intelligence) for the two ethnic groups were equivalent for Combined Academic (69-73%) and Arithmetic (66-69%) criteria, and these rates were equivalent to previous studies' rates. Three-group Language-Related achievement was higher for Ethnic Majority (61%) than Ethnic Minority children (50%). Two-group Combined Academic classification success rates for the selected battery measures (including or excluding Intelligence) produced correct classification rates (69-89%) for Ethnic Minority and Ethnic Majority groups at least as good and sometimes better than overall two-group rates derived from discriminant functions based on the full sample.

Three of eight Box's M analyses were significant in the three-group analyses and four of nine Box's M were significant in the two-group analyses. The Combined Academic (with and without Intelligence) for Ethnic Majority children and the Ethnic Minority Arithmetic analyses (without Intelligence) yielded significant Box's Ms statistics for three- and two-group analyses. The two-group Ethnic Minority Language-Related analysis (excluding Intelligence) also produced a significant Box's M.

The strongest predictor for Combined Academic for both ethnic groups was Intelligence, contradicting the hypothesis that Intelligence would not be selected in predicting Ethnic Minority achievement (Research Question 6). However, for Ethnic Minority children's Arithmetic and Language-Related three-group analyses, Attention-Memory was the strongest predictor and Intelligence, as hypothesized, failed to meet selection criteria. In contrast, Intelligence was the strongest predictor for Ethnic Majority children in all three-group analyses (in which it was included). Intelligence was such a strong unique predictor for Ethnic Majority children's three-group analyses that when it was excluded from the analyses, the Arithmetic and Language-Related discriminant functions were nonsignificant.

Intelligence emerged as a strong predictor of two-group Combined Academic achievement for both Ethnic Minority and Ethnic Majority children (Research Question 6). However, the pattern of scores suggests that Intelligence predicts both Low Average and

Average-Plus performance well for Ethnic Majority children and predicts Low Average performance well for Ethnic Minority children, but tends to underestimate Average-Plus scores for Ethnic Minority children. Thus, excluding Intelligence lowered correct classification rates for Ethnic Majority children by about 10%, and also lowered correct identification of the Low Average scores among Ethnic Minority children. However, identification of Average-Plus achievement among Ethnic Minority children was improved by about 10% by excluding Intelligence from the set of predictors.

Per Research Question 7, it was expected that all variables except Intelligence would predict equally well, but there were several differences between the ethnic groups in selected predictors. The pattern of results for the Attention-Memory and Language variables as predictors of achievement within ethnic groups was approximately the same as for the pooled full sample. In fact, as for the full sample, Attention-Memory predicted strongly (with or without Intelligence) for both Ethnic groups in all analyses. Language predicted three-group Combined Academic achievement for the Ethnic Majority children (with or without Intelligence) and three- and two-group Combined Academic for Ethnic Minority children (but only without the Intelligence predictor). Somesthetic predicted Combined Academic (with or without Intelligence) and Arithmetic (with and without Intelligence) for Ethnic Majority children, but predicted only Arithmetic (with or without Intelligence) for Ethnic Minority children. Visual-Spatial was an important predictor of Combined Academic and three-group Arithmetic for Ethnic Minority children, but was not selected in any analyses for Ethnic Majority children. Fine Motor only was selected as a three-group predictor of Arithmetic (with Intelligence included) for Ethnic Majority children.

Discriminant Function Analyses to Assess Age Differences

Three-group analyses within age groups. The second group of exploratory research questions assessed the ability of the battery's neuropsychological measures to predict the three levels of achievement for both Older and Younger age groups using stepwise discriminant function analysis.

Three-group analyses within age groups: Including Intelligence. The classification tables and relative strengths of unique predictors (including Intelligence) that met selection criteria are presented in Table 19.

The Older children's Combined Academic criterion discriminant function (including Intelligence) was significant ($X^2 = 27.8$, df = 6, p = .0001), and produced a classification success rate of 67%. The true-positive hit rate was 75% for the Low Average group, 73% for the High Average group, and only 50% for the Average group. The Box's M statistic was significant (Box's M = 32.9, df = 12, 4116, p = .006). Attention-Memory was the strongest selected predictor, followed by Somesthetic and Intelligence, in that order.

The Younger children's Combined Academic stepwise discriminant function (including Intelligence) was significant ($X^2 = 17.5$, df = 6, p = .008), and produced an overall classification success rate of 74%. The hit rates for the Low Average, Average, and High Average groups were 83%, 50%, and 82%, respectively. The Box's M statistic was highly significant for the three-group analysis (Box's M = 70.5, df = 12, 1125, p = .0001). Language was selected as the strongest predictor, followed by Somesthetic and Fine Motor, in that order. (Intelligence and Attention-Memory were included, but failed to meet the stepwise selection criteria and were not selected in the analysis. Hence the classification and prediction analyses including and excluding Intelligence were exactly the same.)

The Older children's Arithmetic criterion stepwise discriminant function (including Intelligence) was significant ($X^2 = 18.4$, df = 6, p = .005), and produced a classification success rate of 55%. The Low Average hit rate was 64%, the Average rate was 36% (near

Table 19

Three-Group Classifications of Achievement Predictions and Strengths of Predictors for Two Age Groups (Including Intelligence)

Combined Academic Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	8 (72.7%)	2 (18.2%)	1 (9.1%)
Average	10	3 (30.0%)	4 (50.0%)	2 (20.0%)
Low Average	12	0 (0.0%)	3 (25.0%)	9 (75.0%)

Note. Total cases correctly classified = 66.7%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.63	.0009
2.	Somesthetic	.48	.0003
3.	Intelligence	.38	.0001

(continued)

Table 19 (continued)

Combined Academic Achievement: Younger Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	17	14 (82.4%)	2 (11.8%)	1 (5.9%)
Average	8	3 (37.5%)	4 (50.0%)	2 (12.5%)
Low Average	6	1 (16.7%)	0 (0.0%)	5 (83.3%)

Note. Total cases correctly classified = 74.2%

Step	Predictor	Wilks' Lambda	p
1.	Language	.68	.005
2.	Somesthetic	.57	.004
3.	Fine Motor	.52	.008

(continued)

Table 19 (continued)

Arithmetic Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	7 (63.6%)	4 (36.4%)	0 (0.0%)
Average	11	3 (27.3%)	4 (36.4%)	4 (36.4%)
Low Average	11	1 (9.1%)	3 (27.3%)	7 (63.6%)

Note. Total cases correctly classified = 54.5%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.73	.009
2.	Attention-Memory	.61	.006
3.	Somesthetic	.53	.005

(continued)

Table 19 (continued)

Language-Related Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	10	9 (90.0%)	0 (0.0%)	1 (10.0%)
Average	8	1 (12.5%)	6 (75.0%)	1 (12.5%)
Low Average	15	2 (13.3%)	3 (20.0%)	10 (66.7%)

Note. Total cases correctly classified = 75.8%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.74	.01
2.	Attention-Memory	.60	.005
3.	Somesthetic	.54	.007

chance levels), and the High Average classification rate was 64%. The Box's M statistic was nonsignificant for the three-group analysis (Box's M = 19.3, df = 12, 4362, p = ns). Intelligence, Attention-Memory, and Somesthetic, in descending order of strength, were the three predictors strong enough to meet selection criteria.

The Younger children's Arithmetic criterion discriminant function (including Intelligence) was nonsignificant (X^2 = 11.7, df = 8, p = ns).

The Older Children's Language-Related discriminant function analysis (including Intelligence) was significant (X^2 = 17.8, df = 6, p = .007), and produced a classification success rate of 76%. The Low Average group's hit rate was 67%, the Average group's rate was 75%, and the High Average group's rate was 90%. The Box's M statistic was nonsignificant for the three-group analysis (Box's M = 23.8, df = 12, 2503, p = .07). Intelligence was the strongest predictor, followed by Attention-Memory and Somesthetic, in that order.

The Younger children's Language-Related criterion discriminant function (including Intelligence) was nonsignificant (X^2 = 7.7, df = 4, p = ns).

Three-group analyses within age groups: Excluding Intelligence. A second set of analyses was performed (excluding Intelligence) and Table 20 presents the classification tables and relative strengths of selected predictors.

The Older children's Combined Academic criterion discriminant function (excluding Intelligence) was significant (X^2 = 21.5, df = 4, p = .0003), and produced an overall hit rate of 73% that was slightly higher than the rate for the analysis including Intelligence (67%). The classification success rate was 75% for the Low Average group, 80% for the Average group, and 64% for the High Average group. The Box's M statistic was significant (Box's M = 23.3, df = 6, 19093, p = .002). Attention-Memory was the strongest predictor, followed by

Table 20

Three-Group Classifications of Achievement Predictions and Strengths of Predictors for Two Age Groups (Excluding Intelligence)

Combined Academic Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	7 (63.6%)	3 (27.3%)	1 (9.1%)
Average	10	1 (10.0%)	8 (80.0%)	2 (10.0%)
Low Average	12	1 (8.3%)	2 (16.7%)	9 (75.0%)

Note. Total cases correctly classified = 72.7%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.63	.0009
2.	Somesthetic	.48	.0003

(continued)

Table 20 (continued)

Arithmetic Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	11	5 (45.5%)	5 (45.5%)	1 (9.1%)
Average	11	3 (27.3%)	5 (45.5%)	3 (27.3%)
Low Average	11	1 (9.1%)	4 (36.4%)	6 (54.5%)

Note. Total cases correctly classified = 48.5%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.76	.02
2.	Somesthetic	.68	.02

(continued)

Table 20 (continued)

Language-Related Achievement: Older Children

Actual Group	n	Predicted Group Membership		
		High Average	Average	Low Average
High Average	10	6 (60.0%)	3 (30.0%)	1 (10.0%)
Average	8	2 (25.0%)	4 (50.0%)	1 (25.0%)
Low Average	15	0 (0.0%)	5 (33.3%)	10 (66.7%)

Note. Total cases correctly classified = 60.6%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.77	.02
2.	Somesthetic	.71	.04

Somesthetic.

The Combined Academic criterion discriminant function (excluding Intelligence) for Younger children was identical to the analysis that had included Intelligence, since Intelligence was not selected as a significant predictor (see Table 19).

The Older children's Arithmetic discriminant function (excluding Intelligence) was significant (X^2 = 11.6, df = 4, p = .02), and produced a classification success rate of 48%. The Low Average, Average, and High Average classification success rates were 54%, 46%, and 46%, respectively. The Box's M statistic was highly significant for the three-group analysis (Box's M = 14.6, df = 6, 22431, p = .04). Attention-Memory was the strongest predictor followed by Somesthetic.

The Younger children's Arithmetic criterion discriminant function (excluding Intelligence) was nonsignificant (X^2 = 11.7, df = 8, p = ns).

The Older Children's Language-Related discriminant function analysis (excluding Intelligence) was significant (X^2 = 10.2, df = 4, p = .04), with a classification success rate of 61%. The Low Average group's hit rate was 67%, the Average group's rate was 50%, and the High Average group's rate was 60%. The Box's M statistic was significant (Box's M = 17.3, df = 6, 7112, p = .02). Attention-Memory was selected as the strongest predictor followed by Somesthetic.

The Younger children's Language-Related criterion discriminant function, excluding the Intelligence predictor, was nonsignificant (X^2 = 7.7, df = 4, p = ns).

Misclassifications. The misclassification rates in these analyses again were generally highest for the Average group (in 6 of 8 analyses). The misclassifications of the Low Average and High Average groups were predominantly in the Average grouping (in 12 of 15 group comparisons). The overall misclassification rates for Older and Younger children on

Combined Academic achievement (the only available criterion comparison) were approximately equal (including or excluding Intelligence). Compared to Older children's rates including the Intelligence predictor, the exclusion of Intelligence resulted in a slightly lower Combined Academic miss rate, but it slightly increased misclassifications in Arithmetic groupings and substantially increased misclassifications (15%) in Language-Related achievement.

Two-group analyses within age groups. Two-group stepwise discriminant function analyses (including and excluding Intelligence) were conducted for each of the two age groups.

Two-group analyses within age groups: Including Intelligence. Table 21 reveals the classification tables and the relative strengths of unique predictors that met selection criteria in the significant two-group discriminant functions (Low Average versus Average-Plus) for the three criterion variables, considering Older children and Younger children's groups separately. Three discriminant function analyses, one for each of three criterion variables, were performed. (Tables for Younger children's Arithmetic and Language-Related achievement were omitted because the discriminant functions were nonsignificant.)

The Older Children's two-group Combined Academic criterion stepwise discriminant function (including Intelligence) was highly significant ($X^2 = 23.2$, $df = 3$, $p = .0001$), and produced a true-positive classification rate of 85%. The hit rate was 83% for the Low Average group and 86% for the Average-Plus group. The Box's M statistic was significant (Box's $M = 14.4$, $df = 6, 3402$, $p = .05$). Attention-Memory was the strongest predictor, followed by Intelligence and Somesthetic, in that order.

The Younger children's two-group Combined Academic stepwise discriminant function (including Intelligence) was significant ($X^2 = 15.6$, $df = 4$, $p = .004$), and produced a classification success rate of 87%. The true-positive classification rate for the Low Average

Table 21

Two-Group Classifications of Achievement Predictions and Strengths of Predictors (Including Intelligence) for Two Age Groups

Combined Academic Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	21	18 (85.7%)	3 (14.3%)
Low Average	12	2 (16.7%)	10 (83.3%)

Note. Total cases correctly classified = 84.9%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.63	.0002
2.	Intelligence	.52	.0001
3.	Somesthetic	.46	.0001

(continued)

Table 21 (continued)

Combined Academic Achievement: Younger Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	25	22 (88.0%)	3 (12.0%)
Low Average	6	1 (16.7%)	5 (83.3%)

Note. Total cases correctly classified = 87.1%

Step	Predictor	Wilks' Lambda	p
1.	Language	.77	.007
2.	Somesthetic	.66	.003
3.	Fine Motor	.60	.003
4.	Visual-Spatial	.56	.004

(continued)

Table 21 (continued)

Arithmetic Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	22	17 (77.3%)	5 (22.7%)
Low Average	11	2 (18.2%)	9 (81.8%)

Note. Total cases correctly classified = 78.8%

Step	Predictor	Wilks' Lambda	p
1.	Intelligence	.76	.004
2.	Attention-Memory	.67	.003

(continued)

Table 21 (continued)

Language-Related Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	18	15 (83.3%)	3 (16.7%)
Low Average	15	4 (26.7%)	11 (73.3%)

Note. Total cases correctly classified = 78.8%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.77	.005
2.	Intelligence	.74	.01

group was 83% and the rate for the Average-Plus group was 88%. The Box's M statistic was significant (Box's M = 42.4, df = 10, 359, p = .002). The selected predictors were Language, Somesthetic, Fine Motor, and Visual-Spatial, in descending order of strength. (Intelligence was included as a predictor, but was not selected in the analysis, so the classification and prediction analyses including and excluding Intelligence were exactly the same.)

The Older children's Arithmetic stepwise discriminant function, allowing all six predictors entry, was significant (X^2 = 12.0, df = 2, p = .003), and produced a correct classification rate of 79%. The true-positive classification rate for Low Average scores was 82% and the rate for the Average-Plus group was 77%. The Box's M statistic was nonsignificant for the analysis (Box's M = 5.3, df = 3, 9207, p = ns). Intelligence was selected as the strongest predictor, followed by Attention-Memory.

The Younger children's two-group Arithmetic stepwise discriminant function (including Intelligence) was nonsignificant (X^2 = 5.9, df = 3, p = ns).

The Older children's Language-Related stepwise discriminant function (including Intelligence) was significant (X^2 = 9.2, df = 2, p = .01), and produced a correct classification rate of 79%. The true-positive classification rate for the Average-Plus group was 83% and the rate for the Low Average group was 73%. The Box's M statistic was nonsignificant for the analysis (Box's M = 6.1, df = 3, 899337, p = ns). Attention-Memory was selected as the strongest predictor, followed by Intelligence.

The Younger children's two-group Language-Related stepwise discriminant function, allowing all six predictors entry, was nonsignificant, but approached significance (X^2 = 5.1, df = 2, p = .08).

Two-group analyses within age groups: Excluding Intelligence. A second complete set of stepwise discriminant function analyses (excluding Intelligence) also was performed. Table

22 presents the classification tables and relative strengths of selected predictors. (Tables for Younger children's Arithmetic and Language-Related achievement were omitted because the discriminant functions were nonsignificant.)

The Older children's Combined Academic discriminant function (excluding Intelligence) was significant ($X^2 = 18.0$, df = 2, p = .0001), and produced a classification success rate of 85%. The true-positive Average-Plus group's rate was 91%, and the rate for the Low Average group was 75%. The Box's M statistic was significant (Box's M = 10.8, df = 3, 15383, p = .02). Attention-Memory was the strongest predictor, followed by Somesthetic.

The Combined Academic criterion discriminant function (excluding Intelligence) for Younger children was identical to the analysis that had included Intelligence, since Intelligence was not selected as a significant predictor (see Table 20).

The Older children's Arithmetic stepwise discriminant function (excluding Intelligence) was significant ($X^2 = 8.3$, df = 1, p = .004), and produced a success hit rate of 79%. The true-positive classification rate for the Average-Plus group was 82% and the rate for the Low Average group was 73%. The Box's M statistic was nonsignificant, but approached significance (Box's M = 3.2, df = 1, 2028, p = .08). Attention-Memory was the sole predictor that met selection criteria.

The Younger children's two-group Arithmetic stepwise discriminant function (including Intelligence) was nonsignificant ($X^2 = 5.9$, df = 3, p = ns).

The Older children's Language-Related discriminant function (excluding Intelligence) was significant ($X^2 = 7.9$, df = 1, p = .005), and produced a success hit rate of 70%. The true-positive classification rate for the Average-Plus group was 72%, and the rate for the Low Average group was 67%. The Box's M statistic was significant (Box's M = 5.2, df = 1, 2812, p = .03). Attention-Memory was the sole predictor that met selection criteria in this analysis.

Table 22

Two-Group Classifications of Achievement Predictions and Strengths of Predictors (Excluding Intelligence) for Two Age Groups

Combined Academic Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	21	19 (90.5%)	2 (9.5%)
Low Average	12	3 (25.0%)	9 (75.0%)

Note. Total cases correctly classified = 84.9%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.63	.0002
2.	Somesthetic	.55	.0001

(continued)

Table 22 (continued)

Arithmetic Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	22	18 (81.8%)	4 (18.2%)
Low Average	11	3 (27.3%)	8 (72.7%)

Note. Total cases correctly classified = 78.8%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.76	.004

(continued)

Table 22 (continued)

Language-Related Achievement: Older Children

Actual Group	n	Predicted Group Membership	
		Average-Plus	Low Average
Average-Plus	18	13 (72.2%)	5 (27.8%)
Low Average	15	5 (33.3%)	10 (66.7%)

Note. Total cases correctly classified = 69.7%

Step	Predictor	Wilks' Lambda	p
1.	Attention-Memory	.77	.005

The Younger children's two-group Language-Related stepwise discriminant function (excluding Intelligence) was nonsignificant, but approached significance ($X^2 = 5.1$, $\underline{df} = 2$, $\underline{p} = .08$).

Three- and two-group analyses within age groups: Summary. Three-group classification rates (with and without Intelligence) for the two age groups were generally equivalent for Combined Academic (67-74%), but rates for Arithmetic and Language-Related could not be compared because the discriminant functions were nonsignificant for Younger children. Older children's three-group Language-Related (76%) rate (including Intelligence) was equivalent to previous studies' rates, but Arithmetic (48-55%) was lower (with and without Intelligence) than other rates. Two-group Combined Academic classification success rates for the selected battery measures (including or excluding Intelligence) produced correct classification rates (70-87%) for Older and Younger groups, generally equivalent to the two-group ethnic-group analyses and higher than several of the full-sample two-group analyses.

Five of seven Box's M analyses were significant in the three-group analyses and four of seven Box's M were significant in the two-group analyses. The Box's M statistics for Combined Academic (with and without Intelligence) for both age groups were significant for three- and two-group analyses. The two-group Arithmetic (excluding Intelligence) and the three- and two-group Language-Related analyses (both excluding Intelligence) also produced significant Box's Ms. The age group analyses considered as a group were the poorest among the various analyses in this report meeting the conservative test of assumptions, and none of the Younger children's analyses met the criterion.

For Older children, results are much the same as for the full sample (i.e., Intelligence, Attention-Memory, and Somesthetic were strong predictors) except that Language (in analyses without the Intelligence predictor) and Fine Motor (with or without Intelligence) dropped out as significant predictors. In contrast, Language was the strongest predictor (three- and two-group achievement) of Younger children's Combined Academic achievement, contradicting expectations (Research Question 8). Intelligence did not emerge as a significant predictor for Younger children.

As expected (Research Question 9), the Somesthetic and Fine Motor domains were the other important unique predictors for Younger children's Combined Academic (three- and two-group analyses). Contrary to expectations, however, the Somesthetic domain also emerged as an important predictor for Older children's Combined Academic (three- and two-group analyses), Arithmetic (three-group), and Language-Related (three-group) analyses. Contrary to hypotheses (Research Question 9), the Attention-Memory domain was important for all Older children's analyses, but none of the Younger children's analyses.

Chapter V.

DISCUSSION

Six Aims of the Study

This study demonstrated that a theoretically-based, brief neuropsychological screening battery can predict academic achievement in a young, urban, ethnically diverse population. Six findings were especially important and related to the six aims of the study. First, accuracy rates (44-64%) for predicting three achievement groupings (Average, Low Average, and High Average) were equivalent to (but did not significantly exceed) predictive accuracy rates of other batteries using three groups (61%). Predictive accuracy (83%) for the two achievement groupings (Average-Plus versus Low Average) met expected levels for predicting Combined Academic achievement, but were lower than expected for Arithmetic and Language-Related. Second, the battery classified scores in Low Average and High Average groups better than those of the Average group. Third, the best overall unique predictors for the full sample were consistently Intelligence and Attention-Memory, followed by Somesthetic, Fine Motor, and Language measures. Fourth, the exclusion of Intelligence as a predictor lowered Arithmetic (except in two-group analyses) and Language-Related classification rates, but did not affect overall Combined Academic success rates. Fifth, the predictive utility of the battery (including and excluding Intelligence) varied with ethnicity and age, as did the relative strengths of the predictors. Sixth, of the three achievement criteria, the battery predicted Combined Academic and Arithmetic achievement better than Language-Related achievement.

Issues in Predictive Accuracy of Screening Batteries

Predicting three-group achievement. The first aim was to improve the short term accuracy of prediction compared to that of previous neuropsychological batteries. For the Combined Academic three-group analyses, the overall true-positive identification rate (63%) for the full sample only approached and did not rise above other batteries' rates of 70-75% (de Hirsch et al., 1966; Fletcher & Satz, 1982; Gaddes, 1981; Jansky & de Hirsch, 1972; Spreen, 1989; Townes et al., 1980). As expected the Arithmetic classification success rate (63%) was equal to the Combined Academic rate, but contrary to expectations, the Language-Related true-positive three-group rate (56%) was lower than the other achievement rates. Nevertheless, these data show that the battery can predict Arithmetic performance, a criterion rarely employed in previous studies. (Misclassification rates are discussed in a separate section below.)

Results from the three-group analyses of this study, however, are not comparable to other batteries' predictions, as it is mathematically easier to predict two rather than three groups. The other screening batteries would be expected to have higher predictive accuracy rates because [except for Teeter (1985) discussed below] they typically predicted two rather than three groups: a "low" or "impaired" group versus an all-other-children's group.

Probably the best comparisons of this battery's effectiveness at predicting the three achievement criteria using three-group analyses are to be made with Teeter's (1985) work. Teeter is the only previous investigator to have used a neuropsychological predictive screening battery to compare three groups of children (above the 80th percentile, 60-80th percentile, or below the 60th percentile on achievement) concurrently as well as one year later. The overall three-group true-positive rates obtained for Combined Academic and Arithmetic (63%) in this study (over a 6-month interval) compare well to her battery's success rate of 61% for both concurrent and one-year follow-up predictions.

Predictions across the full range of achievement. A second aim of the study was to predict achievement for Average and High Average as well as Low Average achievers. The results demonstrated differential rates of prediction in that the battery consistently predicted Low Average scores best (64 - 90%), followed by High Average children's scores. The middle grouping (Average children's scores), was more difficult to predict that the extreme groupings, as is the case in all such correlational analyses.

Another explanation of the difficulty in predicting Average scores is their restricted range. The more restricted the range of scores, the smaller the correlation will be (because the predictors are correlated with the achievement variables). That is, the range of standard scores (SS) for the three criterion variables may account for the difficulty in predicting average children; the Average range (95 > SS > 105) was more restricted (10 points) than the score ranges of the Low Average (76 > SS < 95) (18 points) and High Average (105 > SS > 133) (27 points) groups. Future research studies could address the issue of restricted range by using larger samples and by ensuring that Low Average (e.g., 70 > SS < 90) and High Average groups (e.g., 110 > SS < 130) cover standard score *ranges* (e.g., 20 points) similar to the Average group (90 > SS < 110). The difficulty in predicting Average achievement also underscores the need for more work exploring the abilities that differentiate Average from High and Low Average children's scores.

Misclassifications. Determining the relative importance of accurate classification for each group's scores and examining the costs of misclassification for each of the groups with respect to "real world" significance are vital issues in predictive research. Most predictive screening studies have made use of two groups (e.g., a reading failure group versus an average group) (Benton, 1978; Satz & Fletcher, 1988). Likewise, two groups of misclassifications are produced: false positives (e.g., children whose scores are mistakenly predicted as failing) and false negatives (children whose scores are mistakenly predicted as average). Attention in the literature has been devoted almost equally to the false positives and false negatives, but for very different reasons. If a screening battery is used as the basis of educational placement decisions, the minimization of false negatives is necessary thereby ensuring that the greatest number of children who need additional help would receive such help. Otherwise, regular classroom children who need special help may be overlooked and fail academically. The false positive rates need to be minimized so that children avoid the possible social and emotional costs of being mislabelled as needing additional help. Screening batteries have not proven practical for classroom use because of unacceptably high (20-30%) false positive and false negative rates (Silver & Hagin, 1990).

The present study differed from most other predictive studies in its use of three achievement groupings (High Average, Average, and Low Average) rather than two and in its goal of predicting regular classroom achievement. Because of the use of three achievement groupings in this study, three groups of misclassifications emerged. Also, the social cost and practical import of misclassifications using regular classroom children differ from studies using children with difficulties such as reading failure.

The social cost of misclassifying High Average or Average scores as Low Average is that it may unfairly lower expectations of the children on the part of the teacher, parents, or the children themselves, and may result in poorer school performance. Also, the children may sense the changes in their teachers' or parents' attitudes and it could lower the children's sense of self esteem. Probably to a lesser degree, a High Average score mistakenly classified as Average also could lower expectations for the children and negatively affect their self esteem. Misclassifying Low Average or Average scores as High Average could result in teachers and parents developing negative impressions of the children as being unmotivated, and this factor in turn, could negatively affect the children's emotional functioning and achievement. Similarly, the misclassification of Low Average scores as Average also could result in a child possibly being seen as not working up to his potential.

The study sought to focus equally on misclassifications of High Average and Average achievement scores, as well as Low Average scores. Overall, Low Average scores had the lowest misclassification rate followed by the High Average scores, with Average scores having the highest misclassification rate (as was expected because of restricted range of the middle group). Overall, misclassifications of the Low Average and High Average group "extremes" were predominantly in the Average group, consistent with the finding of better classification for two-group comparisons, and suggesting that the predictor variables functioned fairly well across the ranges of achievement. The battery misclassified Low Average and High Average scores at nearly equivalent rates. Whether this finding was an advantage or not depends upon the purpose for which the battery was being used. There was no consistent pattern across analyses of the Average group being misclassified as either High Average or Low Average.

The misclassification rate for Language-Related achievement was higher than the rates of the Combined Academic and Arithmetic criteria. The predictors in the battery failed to do a good job of capturing the variance of Language-Related achievement. The Intelligence predictor was especially important in mitigating misclassifications for the Average group's Arithmetic and Language-Related achievement (the highest "miss rates" occurred when Intelligence was excluded). Thus Intelligence measures appears important for inclusion in prediction batteries examining mathematics and language achievement, in order to minimize Average score misclassifications. In contrast, perhaps the overall mathematical stability of the Combined Academic criterion was such that exclusion of Intelligence did not increase misclassifications. (Ethnic group difference misclassifications and age difference misclassifications are discussed below within their respective sections.)

Predicting two versus three levels of achievement. There were possible violations of the assumptions of the discriminant analysis procedure in three of six three-group (High Average versus Average versus Low Average) analyses for the full sample. Therefore, discriminant analyses also were performed using only two criterion groups (Low Average and Average-Plus) in the attempt to increase cell size and to examine comparability with other two-group studies. Three of six two-group full sample (Average-Plus versus Low Average) analyses of the assumptions of the discriminant analysis procedure also indicated possible violations. The battery achieved a two-group Combined Academic true-positive rate for the full sample (83%) that exceeded the 70-75% achieved by other batteries using two groups. The Arithmetic and Language-Related two-group classification rates for the full sample (72% and 75%, respectively) were equivalent to those of other batteries. The prediction interval used in the present study (6-months) is shorter than the 2-4 years employed by most studies, which would be expected to increase predictive accuracy. Nevertheless, these results do support the validity of the present neuropsychological battery for predicting achievement of young children and suggest that its short-term accuracy is comparable to that of previously developed neuropsychological batteries.

Teeter (1985) also made two-group comparisons, but used the 50th percentile as the cutoff between groups, in contrast to other studies' comparisons of an "impaired" group to all other children. The present study compared Low Average (below standard score of 95) to Average-Plus groups (at or above standard score of 95), so results were not fully comparable with Teeter's two-group approach. Teeter's concurrent success rate of 93% is higher than any of this study's two-group rates, but her study's one-year follow-up rate of 76% is lower than this study's (6-month follow-up) Combined Academic rate (83%).

The comparison of two-group versus three-group discriminant analyses suggests two methodological considerations. First, the number of criterion groups necessarily affects predictive accuracy and should be considered explicitly in prediction research. Secondly, the cutting point defining group membership is an important variable that needs to be formally considered. The desire to study comparisons between groups that are clinically different, in some sense, was the reason most investigators studied "impaired" versus all other children. Even though the definition of "impaired" varies and is arbitrary in many studies, such clinical comparisons appear more meaningful than artificial demarcations such as "above or below the 50th percentile" which have no relevance to potential intervention targets. Thus, comparing Low Average to Average and/or High Average achievers makes more sense than using the 50th percentile as a cutoff (dividing Average achievers in half).

Given the difficulty in increasing classification success rates of three- and two-group achievement beyond the levels achieved by prior batteries, perhaps variables other than neuropsychological ones account for variance that gives rise to the remaining 25-30% misclassification rate. As Spreen (1978) and Gaddes (1981) have contended, perhaps motivation, ability of the individual teacher or relationship between teachers and students, and home support for education are sources of the variance unaccounted for by typical predictor variables.

Issues in Test Selection and Strength of Predictors

Consistency between present and previous neuropsychological batteries. A third aim of the study was to explore the utility of using theories of processing, and/or of neuroanatomical bases of learning, in selecting tests for use with regular classroom students. These theories suggested that at least six separate neuropsychological domains contribute unique variance in the prediction of academic achievement. Results of the present study indicate that in general, Intelligence, Attention-Memory, Somesthetic, and Fine Motor were the significant unique predictors of achievement in that order, for the overall three-group and two-group analyses. Language and Visual-Spatial also became significant unique predictors of achievement when Intelligence was excluded. (These full sample results of analysis of predictor significance are the most reliable estimates of predictor significance in this study because they are based upon the largest samples.) These results are highly consistent with neuropsychological domains identified as good predictors in previous batteries (Jansky & de Hirsch, 1972; Satz & Friel, 1974; Spreen 1978b).

To the extent that there are differences among batteries as to which areas are identified as good predictors, one important source of such differences is the lack of consensus among investigators about how to operationalize neuropsychological domains. As was evident in Table 3, tests may be considered as primarily tapping one domain and secondarily tapping other domains, and confusion may exist regarding which type of measure(s), in what combination(s) is the best way to assess a particular domain. This issue emerged in developing measures for the present study. Initially it had been supposed that a "Sensation-Perception" domain would be used, consisting of a combination of somesthetic and visual-spatial tests. In fact, the tests selected for that domain were not strongly intercorrelated in the present sample even though other samples (mostly normal adult) have yielded such a factor. The Somesthetic and Visual-Spatial tests therefore were treated separately and six, not five, areas of neuropsychological functioning were identified and operationalized.

This issue highlights the importance of attention to developmental differences in patterns of neuropsychological abilities (i.e., it may be that the measures selected yield six factors for young children but five for adults) as well as underscoring how each test may reflect multiple component cognitive abilities (as opposed to one discrete ability). For example, Satz (Satz & Friel, 1974) emphasizes the predictive importance of the VMI (Benton, 1983), which he describes as a "fine motor" test. The most similar measure in the present study (i.e., the Bender Gestalt) had almost no unique predictive utility. In this study, the Bender was considered primarily a "visual-spatial" measure. The "fine motor" tests in this study (e.g., Finger Tapping) had small predictive value, and had increasing importance when Intelligence was excluded. These findings suggest that it may be the motor component, not the visual-spatial component, of performance on tests such as the VMI and Bender, which predicts individual differences in achievement. This distinction also exemplifies the importance of considering how neuropsychological "areas" are operationalized. If different researchers operationalized neuropsychological areas differently, results may vary due to subtle differences among tests.

Regular classroom utility of predictors reflecting LD subtypes. Children having learning disabilities were not studied in this project, but the importance of attending to functional learning disability subtypes to guide selection of measures used to assess regular classroom children was evident from results of the present study. The subtypes include attention, language, and visual-spatial impaired groups (Denckla, 1979), as well as a subtype of generally low intelligence children (Rutter, 1978). These functional subtypes were used to select Intelligence, Attention-Memory, Language, and Visual-Spatial predictors, with the idea that these constructs might prove useful in predicting achievement in regular classroom children. Evidence for the significant unique importance of Intelligence, Attention-Memory, Somesthetic, and Fine Motor predictors emerged, but little support was found for the unique importance of Language predictors, except in use with Younger children. Also, little evidence for the unique importance of Visual-Spatial was found. While children with specific visual-spatial deficits are rare (e.g., Mattis et al., 1975), the implication from the current study is that visual-spatial measures have little in the way of unique predictive power in use with regular classroom children. Also, Language and Visual-Spatial may have had a smaller unique role in prediction because of their correlation with Intelligence; excluding Intelligence from the battery typically resulted in Attention-Memory becoming the best predictor and occasionally allowed for the selection of Language or Visual-Spatial as a predictor.

Utility of predictors chosen to reflect discrete brain systems. As noted in the Literature Review, the tests also were selected to represent brain systems considered functionally critical for academic achievement. The best evidence of the importance of the neuropsychological approach comes from the success of the predictors, which fell into two groups. The first group consisted of Intelligence and its highest three correlates (Attention-Memory, Language, and Visual-Spatial). Attention-Memory, Language, and Visual-Spatial did not correlate with each other, suggesting that they had discriminant validity as measures of academic achievement. That is, the three areas were all related to general intelligence, but each represented unique attention-memory (or sequential-processing) based, language-based, and visual-spatial based clusters. The two best predictors in the battery were Intelligence and Attention-Memory. The second group consisted of Somesthetic and Fine Motor, which correlated with each other but not with Intelligence and its correlates. Somesthetic and Fine Motor were the third and fourth best predictors in the battery.

That the predictors from these two groups (intelligence and its correlates and sensorimotor) were selected in the stepwise function demonstrates that each of the two groups has unique predictive ability, and suggests that two independent factors underlie (or are associated with) academic achievement. That Attention-Memory, and to a lesser extent, Language, and Visual-Spatial emerged in stepwise analyses, illustrates their unique contributions to prediction. Thus, the findings are congruent with a neuropsychological approach emphasizing selection of tests that tap each of these distinct functional domains.

Further, current data also suggest that tests that reflect a variety of discrete neural areas or systems are related to academic achievement. Tests presumed to be functionally associated with global (Intelligence), bi-frontal (Attention-Memory, Fine Motor), bi-parietal (Somesthetic), left hemisphere (Language), and to a lesser extent measures of right hemisphere functioning (Visual-Spatial) occasionally proved to be important selected predictors. The results demonstrated that successful predictions of Combined Academic and Arithmetic scores generally using three to four predictors and of Language-Related achievement using one to two predictors. Hence, an approach emphasizing the importance of multisite brain functioning also is supported for predicting Combined Academic and Arithmetic achievement, and to a lesser extent Language-Related achievement.

Contributions of predictor variables. Intelligence was the best predictor for the full sample, for Ethnic Majority (four of five analyses) and Older subgroups, as well as for certain Ethnic Minority analyses (three of six). However, Intelligence was not selected for the Younger children analyses, and certain Ethnic Minority analyses (two of six). Including the Intelligence predictor substantially improved prediction for Low Average scores across all analyses except Younger children, and Intelligence improved overall three-group prediction for the full sample Arithmetic and Language-Related areas. Exclusion of Intelligence as a predictor lowered three-group Arithmetic and three- and two-group Language-Related achievement but not Combined Academic.

The emergence of the Intelligence measure as the best predictor contradicted the expectation that Intelligence would not predict achievement well. The order effects for the other variables contradicted the expectation that the other five measures would predict achievement equally well. The Attention-Memory measure was the second best predictor next to Intelligence.

One explanation is that these measures and the criteria share similar content. The Intelligence and Attention-Memory measures are comprised of different IQ subtests (WISC-R: Wechsler, 1974) and correlate strongly with each other ($r = .28 - .54$). Children who have been exposed to and have absorbed a wide range of acquired verbal knowledge (high Information, which constituted the Intelligence estimate subtest score) may have had more exposure to spelling, word identification, and arithmetic stimuli. Children having the ability to sustain concentration while remembering and processing sequential written, verbal, and nonverbal material (high Attention-Memory score) also achieved good academic skills. Academic skills require use of letters, words, and numbers. Sequential processing is required for learning to recognize and read strings of letters and words, and for carrying out arithmetic operations. One of three subtests comprising the Attention-Memory variable is a measure of verbal arithmetic computation, and thus is similar to the Arithmetic criterion in that both require mathematical skills. The two arithmetic measures differ in that the verbal measure is confounded with auditory attention, while the Arithmetic criterion is confounded with visual-perceptual processing.

Alternatively, Intelligence and Attention-Memory may be the best predictors because they appear to be the most complex and multifactorially determined scores in this battery. Any impairment, in any component skill, could impair Intelligence and Attention-Memory to some degree. Thus, compared to tests that target more specific abilities, these measures may be more sensitive to deficits or specific talents in the complex array of component skills required to perform academic achievement tests.

The Somesthetic measure and Fine Motor measures were the 3d and 4th best predictors. Somesthetic predicted Combined Academic in the overall two- and three-group analyses that included Intelligence and predicted Arithmetic if Intelligence was excluded. Somesthetic was a significant predictor for all analyses (including the full sample, Older and Younger, and Ethnic Minority) except several Ethnic Minority analyses. Perhaps Ethnic Minority children's Low Average scorers represent a more heterogenous mix of sociocultural and neurodevelopmental problems, whereas the other groups have low scorers with relatively few sociocultural problems such that the "pure" sensorimotor measure has more predictive power.

The primary importance of Fine Motor emerged in predicting Combined Academic and Arithmetic in the overall two-group analysis. These results suggest that Fine Motor differentiates Low Average scores from all other scores. One hypothesis was that Fine Motor may be a marker of neurodevelopmental immaturity or impairment that fails to distinguish Average from High Average scores because of a ceiling effect, but the follow-up tests examining the distributions do not indicate a ceiling effect. The conclusion that seems evident is that Low Average achieving children may be identified by their problems in motor performance, but that even though there is a range of fine motor performance for the Average and High Average children, average and high average motor performance is not associated with the equivalent levels of achievement.

The Somesthetic and Fine Motor predictors did not correlate with Intelligence, suggesting that they represent a domain independent of the aspects of intelligence. Somesthetic correlated with Fine Motor, presumably because both tactile perception and hand movement involve frontal-parietal systems that are closely linked (Lezak, 1983). The Somesthetic and Fine Motor variables are in theory, fairly simple, direct measures of neurodevelopmental maturity or brain integrity. These findings suggest that assessment of intactness of the sensorimotor systems, in ways not assessed by intelligence tests, is important in predicting academic achievement.

Language was the 5th best predictor, but emerged only with Intelligence excluded (in the full sample analyses). (Visual-Spatial did not emerge as significant in any of the full sample analyses.) Children who are better able to read words, letters, and numbers, and who can perform simple arithmetic computations (high Language score), were better able to perform the closely related academic achievement tasks (i.e., spelling, reading words and letters, performing computations).

While these findings are correlational, the overall pattern of results suggests that, for the full sample, children who were having some academic difficulty (performing below average), as well as those who were doing unusually well (above average), may have primarily differed in their levels of attentional control and/or working memory as well as in general intelligence. The Somesthetic and Fine Motor variables, which bear no "face-valid" relationship to the academic criteria, may have served to assess individual differences in general neurodevelopmental integrity. That is, tests of developmental maturation of sensorimotor brain systems may serve as "markers" of the overall level of maturation or efficiency of brain systems, and therefore function fairly well at predicting academic achievement.

Exploratory Issues in Prediction

Operationalized definition of cultural fairness. A fourth aim of the study was to gather data on the cultural fairness of the criterion and predictor measures in this study, by comparing scores of Ethnic Minority children to scores of Ethnic Majority children. These results are based upon small sample sizes and must be viewed with caution. The two ethnic groups of children in this study were drawn from a Midwestern, lower Middle-Class, urban environment and the two groups were equivalent with respect to age, gender, hand dominance, and grade, so it was assumed that group differences were not pre-existing, real differences in their underlying neuropsychological abilities. Therefore, group differences were ascribed to lack of some kind of cultural fairness in the measures, but other variables cannot be totally excluded as possible explanations for the differences. This operational definition of cultural fairness admittedly simplifies a complex, controversial issue. For the present purposes, however, a simple definition may be useful in the effort to identify whether measures on which two ethnic groups obtain similar mean scores will serve to predict academic achievement equally well for those two groups. Thus, the focus is on how well "biased" or "unbiased" measures (defined in a simple way) predict achievement in Ethnic Minority versus Ethnic Majority children.

Culture-fairness of criteria. There were no significant differences detected between Ethnic Minority and Ethnic Majority children on the Combined Academic, Language-Related, or Arithmetic criteria, and there was fair power to discern such differences. It was expected that the language-based Combined Academic and the Language-Related measures, in particular, might be culturally unfair, because language is one of the most important areas of ethnic differences in achievement measures (Sattler, 1992). These expectations were not supported, perhaps because the Language-Related criterion measures employed did not tap the relevant dimensions upon which differences would have emerged, such as reading comprehension ability. Alternatively, perhaps the ethnic group differences found on language measures in other studies also are associated with SES differences that were not associated with ethnic group differences in the present study.

Classification within ethnic groups. Total classification rates within the separate groups of Ethnic Minority and Ethnic Majority children were equivalent for three-group Combined Academic analyses (69-71%) and Arithmetic analyses (69-66%). Ethnic Majority children's scores were classified better than Ethnic Minority children's scores for other analyses including Language-Related (61% versus 50%), (two-group) Combined Academic (89% versus 81%), and (two-group) Arithmetic (84% versus 69%). The different rates of correct classification for ethnic groups are associated with differences in selected predictors (see below).

Culture-fairness of predictors. The Intelligence and Language measures were found to be culturally unfair (i.e., Ethnic Minority children scored significantly lower on these measures than Ethnic Majority children). While large effect sizes were found for the differences on these measures between ethnic groups the clinical significance of the differences was relatively small. The search for culturally fair intelligence and language predictors of achievement must continue, as these constructs are important, and unfortunately neither the Aphasia Screen (Reitan, 1974) or Information WISC-R subtest (Wechsler, 1974), nor the often-used Peabody Picture Vocabulary Tests (PPVT & PPVT-R) (Dunn, 1965; Dunn & Dunn, 1981), provide adequate measures. It would be interesting and worthwhile to construct culturally fair intelligence and language tests predictive of achievement by developing or selecting test items that correlate strongly with academic achievement but that do not differ among ethnic groups.

Attention-Memory, Visual-Spatial, Fine Motor, and Somesthetic variables were demonstrated to be culturally fair. These findings are consistent with results from previous studies (Gutkin & Reynolds, 1981; Hinshaw et al., 1966; Jansky & de Hirsch, 1972; Morrison & Hinshaw, 1988; Satz & Friel, 1978). These measures are less affected by social or cultural variance than most language-based or general intelligence tests.

Ethnic-group differences in predictor rankings. Contrary to expectations, the Intelligence measure was the best predictor for Ethnic Minority as well as Ethnic Majority children's Combined Academic achievement. For the two-group analyses, Intelligence predicted Combined Academic and Arithmetic scores for both ethnic groups, and was a significant predictor of Ethnic Minority children's Language-Related achievement. Intelligence was a somewhat better predictor for Ethnic Majority children, as excluding Intelligence rendered their three-group Arithmetic and Language-Related discriminant functions nonsignificant. Ethnic Minority children's two-group Language-Related discriminant function, excluding Intelligence, also was nonsignificant. These data suggest that a culturally unfair predictor can have predictive utility with Ethnic Minority children with respect to culturally fair outcome measures. The apparent conclusion seems that while the variance of the Intelligence scores is most strongly related to the achievement of Ethnic Majority children, Intelligence also is a significant predictor for Ethnic Minority children.

Interestingly, Intelligence was not selected as a predictor of Ethnic Minority Arithmetic and Language-Related achievement in the three-group analyses. Also, the exclusion of Intelligence in the Ethnic Minority analyses made little difference in overall achievement criteria, but had a differential effect for the Low Average group, compared to Average and High Average (or Average-Plus) groups. In this study, in contrast to most previous studies that emphasized the lowest group's misclassification rate, equal value was placed on Average, High Average, and Low Average misclassification rates. Exclusion of Intelligence typically lowered Low Average classification success rates, and raised Average, High Average, and Average-Plus success rates. These data suggest that Intelligence is an important predictor of Low Average scores for Ethnic Minority children, but is a poorer predictor of their Average and High Average scores. The combination of Attention-Memory, Language, and Visual-Spatial scores classified Average-Plus achievement scores for Ethnic Minority children very well (89% correct) when Intelligence was excluded. This finding suggests that Minority children who perform extremely poorly on achievement tests are likely to have low

Intelligence scores, but that Average and High Average achieving Ethnic Minority children may have Intelligence scores that are relatively lower than their level of achievement. That is, low Intelligence scores for Ethnic Minority children correlate well with low average achievement, but Intelligence scores do not correlate as well with achievement for the Average and High Average achieving children. These findings confirm the difficulties of using verbal Intelligence measures with minority children (Sattler, 1992).

In contrast to the Ethnic Minority analyses, in two of three Ethnic Majority analyses, exclusion of Intelligence dropped overall predictive success rates by a fair margin (10%). Excluding Intelligence tended to decrease accurate classification of both Low and Average-Plus groups of Ethnic Majority children. Thus, the Intelligence predictor, particularly verbal intelligence, appears most strongly linked to all levels of achievement for Ethnic Majority children.

Attention-Memory measure was the best predictor for Ethnic Minority children's Arithmetic and Language-Related achievement and similarly was a significant predictor for these Ethnic Majority children's analyses. Somesthetic significantly predicted Ethnic Majority children's Combined Academic achievement, and Arithmetic achievement in both ethnic groups with or without Intelligence included in the analyses. Somesthetic is assumed to reflect general neurodevelopmental integrity or maturation. One might conclude that the importance of Somesthetic emerged in Combined Academic only for Ethnic Majority children because of the lesser role of sociocultural or language variance in their achievement. That is, Ethnic Minority children may score lower on tests for a greater variety of sociocultural reasons than their counterparts. Therefore, low scorers are a heterogenous group in which some children score low due to neurodevelopmental problems, others due to sociocultural differences (despite intact brain systems). Thus, measures that tap simple neurodevelopmental domains are not consistently associated with variance in achievement for these children.

Language emerged as a three-group predictor (with or without Intelligence) of Combined Academic achievement for Ethnic Majority children and Arithmetic for Ethnic Minority children (fourth of four predictors in both analyses) and was selected in several other Ethnic Minority analyses when Intelligence was excluded. Similarly, the Visual-Spatial predictor only emerged as a unique predictor (Intelligence included) of Ethnic Minority children's Arithmetic achievement, and was selected when Intelligence was excluded in Ethnic Minority Combined Academic. The Language and Visual-Spatial measures had relatively little utility in unique predictability in these analyses, as it seemed that their predictive variance was largely shared with Intelligence and Attention-Memory. However, it is interesting that these predictors demonstrated more utility for Ethnic Minority children, when Intelligence was excluded, than for Ethnic Majority children. These data suggest that the shared variance between Intelligence and these other predictors are captured by the other predictors for Ethnic Minority children, but not as much for Ethnic Majority children.

The Language predictor may be more important for Ethnic Minority than Ethnic Majority children because of a ceiling effect for scores for the Ethnic Majority children. Originally, it was hoped the Language predictor would be a more culture fair verbal measure for Ethnic Minority children (and perhaps be more important for this reason), but this was demonstrated not to be the case. (In fact, Language was <u>not</u> culture fair, by the present definition.) The Language predictor may tap into or assess language and arithmetic skills very similar to those measured by the achievement variables and may correlate more with Ethnic Minority children's performance more as a measure of acculturation. The distribution of scores for Ethnic Minority children when examined was wider than that for Ethnic Majority children, and examination of the Ethnic Majority distribution suggested the presence of a ceiling effect.

That is, acculturated Ethnic Minority children knew the language and math items on the Language predictor and did better on similar items on the academic tests than Ethnic Minority children who were not so acculturated.

The importance of the Visual-Spatial test for Ethic Minority children appears more straightforward. Visual-Spatial may have more predictive utility for Ethnic Minority children than Ethnic Majority children because it is a nonverbal measure associated with intelligence, as opposed to verbal measures that may underestimate minority youngsters' abilities.

Fine Motor was important in only two ethnic group analyses and was selected last in both analyses: Ethnic Majority Arithmetic (with Intelligence) and Ethnic Minority Language-Related (without Intelligence). The implication is that Fine Motor had little unique predictive variance in these analyses, but the Somesthetic predictor demonstrated that measures of basic neurodevelopmental maturity can be significant predictors. Perhaps Somesthetic captured the unique variance that was associated with Fine Motor in these analyses, as the two predictors were correlated.

In summary, culturally fair variables that have significant predictive utility (e.g., Attention-Memory and Somesthetic) were identified in this study. The goal of developing culture-fair predictors of academic achievement is important because without them the Average-Plus Ethnic Minority children's potentials may be underestimated. On the other hand, if the primary goal of screening is accurate identification of children who will score below average on academic achievement tests, then the present results suggest that the "culturally unfair" Intelligence measure should be included among the predictors.

Age differences. A fifth aim of the study was to explore age differences related to prediction of achievement for young children. Conclusions based upon these exploratory analyses must be tempered with caution for several reasons. An important methodological concern for all developmental studies, including this study, is that the specific characteristics of the predictors and criteria may differ at different age levels, rendering comparisons across ages problematic. For example, the Intelligence predictor may be considered as comprised of items relating to everyday life (outside of school) for Younger children, but as tapping school-related learning for Older children; the predictor may not be measuring the same ability at different ages. Analogous developmental validity problems also may exist in the outcome criteria. For example, for Younger children, the Arithmetic criterion generally consists of items requiring counting and very simple addition and subtraction problems, while for Older children, the criterion generally taps higher-level, more abstract mathematical concepts such as fractions, use of carrying operations, and multiplication and division.

The underlying theoretical issue is whether the neuropsychological abilities associated with various academic skills really change with development, or whether such abilities only appear to change as artifacts of developmental changes in the nature of the predictors and criteria. The item homogeneity of content across ages for these predictors varies. The Fine Motor, Somesthetic, and Visual-Spatial measures are comprised of virtually identical items, within their respective tasks, across ages and thus their content and construct validity does not seem to change with age. In contrast, the items for the Intelligence and Language measures vary significantly across the age ranges used in this study and arguably measure different abilities at different ages. The Intelligence measure was discussed above. The Language task utilizes letter, number, and shape identification for adequate Younger children performance but requires reading and spelling of words (very different tasks) for adequate Older children performance. Two of three measures comprising the Attention-Memory predictor (i.e., Digit Span and Coding) are composed of tasks that do not differ across ages; the items in the third Attention-Memory subtest, Arithmetic, vary conceptually across the age-ranges in this study with Younger children only expected to perform the simplest counting, addition, and subtraction operations, while Older children are expected to perform more advanced mathematical calculations.

It also should be noted that developmental predictors that were not tapped in this battery could prove more useful in prediction than the selected variables. For example, listening comprehension in kindergartners might prove more predictive of later reading ability than letter identification.

All of the criterion measures in this study appear vulnerable to the methodological critique of validity problems associated with different ages of the children. The Language-Related criterion measures for the Younger children predominantly consist of copying symbols and identifying individual letters, while the measure for Older children consists of the more complex and language-related activities of spelling and verbal word calling of individual words. As noted above, similar developmental problems may characterize the Arithmetic criterion.

A related but separate methodological concern is whether social-emotional or educational age-related changes contribute to or cause changes in predictor or criterion scores across age-groups. Motivational factors and emotional functioning that differed across age-groups may partly account for the obtained results. Other possible causes for the findings in this study include classroom-to-classroom differences across age-groups in areas such as teacher interactions with their class or in test-taking attitudes. This study did not examine or control these emotional, social, and educational factors, and should be studied in future research.

In addition to the theoretical problems for developmental predictive research, there are several reasons specific to this study for viewing the conclusions cautiously. First, sample sizes were very small for the analyses. Second, these analyses frequently did not meet the conservative test of multivariate normality and homogeneity of covariance (in contrast to the other sets of analyses in which half to two-thirds may be interpreted as meeting the assumptions). Third, the Arithmetic and Language-Related Younger children's analyses were nonsignificant, in contrast to many other sets of analyses. This latter finding suggests that Younger children's Arithmetic and Language-Related abilities are poorly predicted by the battery.

Given these severe limitations, the following discussion may be viewed as a heuristic device to begin exploring the issues of age differences. Also, the recent neuropsychological academic screening literature extensively discusses age differences in prediction, so the present findings may be compared to other studies' results. The selected measures of the battery worked slightly better for Younger children's Combined Academic (74%) than for Older children's achievement (67%). Results using two achievement groups produced generally equivalent rates (85-87%) for Younger and Older children's Combined Academic achievement. The battery worked better for Older children in predicting Arithmetic (76%) and Language-Related achievement (55%).

Differences in selected predictors of Younger and Older children's achievement were strikingly apparent. Older children's achievement was associated exclusively with Intelligence (WISC-R Information), Attention-Memory (WISC-R Digit Span, Arithmetic, and Coding subtests), and Somesthetic predictors, whereas the younger children's selected predictors included sensorimotor abilities (Somesthetic, Fine Motor, and Visual-Spatial) as well as Language. This finding is consistent with data suggesting that Older children's achievement may be better predicted by intelligence tests and younger children's achievement may be better predicted by discrete component measures such as sensorimotor and visual-spatial tests

(Hinshaw et al., 1986).

The increasing role of intelligence, attention, and working memory as development proceeds makes sense from the perspective of curricula. The rudiments of reading, mathematics, and other academic skills including letter and number identification, beginning reading and mathematical operations are taught in the early grades. The coursework becomes increasingly complex and abstract over the years, and depends increasingly on the sorts of abilities presumed to underlie the Intelligence and Attention-Memory predictors.

A second possible explanation why Intelligence and Attention-Memory are poor predictors of achievement for Younger children is that the predictors suffered from floor effects in this group, but examination of the distribution of Younger children's Intelligence and Attention-Memory scores indicated that no floor effect was present.

Language predicted achievement well (when Intelligence was excluded) in most analyses except for the Older children. While it was originally expected that, consistent with Satz's theory (Satz & Friel, 1974, 1978), the Language measure (Aphasia Screening Test) should better predict achievement for older than younger children, obtained results were the opposite from those expected. The Aphasia Screening Test (AST) was the best predictor for Younger children. The AST also predicted 5-7 year old children's achievement in Townes et al. (1980) and Teeter's (1985) studies. The constellation of activities in younger children's classrooms (e.g., learning letters, basic numbers) is closely related to the AST items. Also, the items comprising the AST may parallel the Language-Related criterion measure items more closely at the Younger ages than the Older ages. As noted earlier, the AST is not an ideal language measure as it includes writing and calculation.

As noted earlier, it may be that Somesthetic and Fine Motor capture individual differences in neurodevelopmental integrity or maturation. While it was initially thought that perhaps that most Older children topped out on the Fine Motor measure, examination of the distribution of Older children's scores indicated that there was no ceiling effect present. The variance in Motor scores appears to have predictive value only for the Younger children, and while the distribution of Motor scores is not skewed for Older children, the variance no longer has predictive value. Inspection of the separate distributions of Older and Younger children's Somesthetic scores indicated that there were ceiling effects for both distributions. It would seem that Somesthetic is an important marker for the children in this study that do not perform well on the measure, irrespective of age. The Visual-Spatial measure's importance for Younger children confirms prior batteries' heavy reliance on such measures for screening the youngest (kindergarten through first grade) children (e.g., Satz & Friel, 1974).

The results of this study provide support for the hypothesis there are unique neuropsychological correlates of achievement for Older versus Younger children, in accordance with the rationale underlying Satz's theory (Satz & Friel, 1974). While the notion that different predictors are important during different developmental periods appears valid, the issue is complicated by methodological and theoretical concerns. Currently, the relationships among the functional domains do not appear adequately addressed by the existing theories. Perhaps the best neuropsychological developmental theory may combine Satz's, Jansky's (1978), and Silver and Hagin's (1990) theories in that language, sensorimotor, and visual-spatial predictors are very important early on, while more complex intellectual and mnemonic factors appear to increase in importance across time, at least through third grade.

Predicting Different Achievement Criteria

A sixth aim of the study was to explore prediction of academic abilities other than reading. As noted above, the selected measures of the battery produced better true-positive predictions of Combined Academic and Arithmetic achievement (63% for each) than

Language-Related achievement (55%) for the full-sample three-group analyses including Intelligence. In the full-sample two-group analyses (including Intelligence), however, correct classification of Arithmetic and Language-Related achievement were similar (72% and 75%); both were poorer than the Combined Academic rate (83%).

Intelligence was a key predictor of Language-Related achievement; without it, classification accuracy dropped by about 10% (and only Attention-Memory emerged as a significant predictor). With Intelligence included, the battery performed quite well (80% correct classification) at identifying average-to-above-average scores on the Language-Related criterion; apparently the deficits that produced poor performance on these Language-Related tasks (i.e., word identification and spelling-to-dictation) were not adequately identified by the present battery, however, since prediction of low scores was relatively poor.

Prediction of Language-Related achievement (Word Recognition and Spelling) generally was poorer than for the other two criteria, perhaps in part because the Language measure (AST) was an inadequate measure of language ability, as suggested above. Perhaps a lack of other language-related material in the battery, such as word attack, phonetic, or semantic measures, may have contributed to poor prediction of Language-Related achievement. Clearly, a fair proportion of Language-Related achievement must depend upon other intellectual, memory, or visual-spatial abilities not sampled in this battery (e.g., memory for words or visual-verbal learning). Further studies examining the use of word attack, phonic, and semantic predictors, as well as more intellectual and verbal memory tests are needed to shed light on these issues.

Among previous neuropsychological studies, only Teeter (1985) has studied prediction of arithmetic across time. Hinshaw and Morrison's group (Hinshaw et al., 1986; Morrison & Hinshaw, 1988) have studied the concurrent prediction of neuropsychological tests and arithmetic (and reading) ability. Thus, this study is one of the first to demonstrate predictive utility across time with Arithmetic.

Intelligence, Attention-Memory, and Fine Motor were the important predictors of Arithmetic. Emergence of Fine Motor is of interest given neuropsychological theory regarding the role of parietal and frontal functioning in Arithmetic (Weinberg & McLean, 1986; Welsh & Pennington, 1988). It is interesting to note that whereas this battery (including Intelligence) tended to overestimate Language-Related achievement, it tended to underestimate Arithmetic achievement. In other words, prediction of poor scores on Arithmetic was markedly superior to prediction of Average-Plus scores (82% versus 68%, respectively). This finding must mean that some children who performed poorly on the predictors nevertheless scored adequately on Arithmetic. Inspection of the exploratory analyses suggests this finding probably was the case for Ethnic Minority and for Younger children (discussed below).

If Intelligence was excluded, accuracy rates decreased for identifying poor Arithmetic, but actually increased somewhat (7%) for identifying Average-Plus Arithmetic. A complex array of predictors (i.e., Attention-Memory, Language, Fine Motor, Somesthetic) emerged reflecting the complexity of the Arithmetic criterion.

Limitations of the Study and Further Suggestions for Future Research

There were several limitations associated with this study. First, the sample size of this study was very small and thus the conclusions must be tempered with caution. Similarly, several of the main discriminant function analyses and certain exploratory analyses (especially the age-related analyses) are of questionable reliability, because of potential violations of the assumptions, from a strictly conservative viewpoint, underlying discriminant function analysis. Studies using larger sample sizes are needed.

A second limitation concerns the representativeness of the sample and includes the use of volunteers as participants (from schools which volunteered to participate). Combining the group of ethnic minority children into a single sample may have obscured differences among ethnic groups. Exploration of specific Ethnic Minority groups' predictive measures of achievement is needed. Also, among the Ethnic Majority sample were various ethnic groups in which differences such as ethnic subgroup variations in ability and educational background of parents may have existed. Future research studies should focus on better defined and differentiated ethnic-group samples.

The use of volunteers introduced possible selection biases into the study. For example, children having greater academic problems might have participated at a higher rate than other children. Such selection biases may have affected the results and conclusions of this study by lowering the generalizability to the participating schools, other regular classrooms, and other schools.

Further, the results from the parochial school sample probably do not generalize to urban public schools, for several reasons. Parochial schools typically require parents to have greater financial resources and parents of parochial school children may have different attitudes regarding education than public school parents. Finally, results would not be expected to generalize to suburban or rural school districts because of their vastly different community environments and ethnic group compositions compared to those used in the present study.

A better way to recruit would be to obtain commitments from communities that would allow entry into all schools in a region and commitment from the schools fostering the expectation that all children would participate (within ethical guidelines). Techniques such as random sampling or stratified random sampling could be employed to select a sample representative of the regions and populations desired to be studied.

A third limitation of the current study was the relatively short test-retest time interval (6 months) between the administration of the predictor battery and the criteria. Several studies (e.g., Jansky & de Hirsch, 1972; Satz & Friel, 1978; Spreen, 1978b) have used test-retest time intervals of at least 2 - 4 years and studies conducted over even longer periods of time are needed (Satz & Fletcher, 1988). The short predictive time interval used in the present study increased the accuracy of the battery but decreased its clinical utility. The ideal complete interval would be fourteen years (preschool through high school), with predictive comparisons made at kindergarten, 2, 4, 6, and 12 years. Obviously the longer the delay in prediction, the more intervening variables may become present and increasingly influential, but follow-up from preschool through high school would provide the optimal opportunity for understanding prediction of achievement.

A fourth group of limitations involved the selection of specific criterion measures. The use of Word Identification subtest from the WRAT-R (Jastak & Wilkinson, 1984) as one of two measures of language achievement was restricted, as the test does not involve reading comprehension. One may suggest that this error was not so egregious in this study, because the children were of young ages such that they would not be expected to be reading connected text to any great extent. However, the test might overestimate academic ability in children who are proficient "word callers" (e.g., in the extreme case, hyperlexic children). The present results cannot be considered to apply to reading comprehension, particularly for older children. Selection and use of better criterion measures of reading are needed to explore whether other components of reading ability, including word attack, phonetic, visual-verbal learning, and reading comprehension skills, for example, would be predicted by these neuropsychological measures.

Another problem in the selection of criterion measures was the use of the WRAT-R

Arithmetic subtest as a measure of math achievement. The Arithmetic subtest has a complex visual-spatial configuration and it has been demonstrated that visual-spatial problems, not arithmetic problems per se, can impair Arithmetic WRAT-R scores (Siegel & Linder, 1984). Also, the only mathematical abilities explored by the test are operational and algebraic abilities (the latter only at older ages). Future research should study more components of mathematics achievement, using measures that assess such dimensions as computation of time, counting, and geometry abilities, for example.

A further limitation of the criterion measures was that all three WRAT-R subtests include pretests (often used with the Younger children) that are arguably only remotely related to the constructs assessed by the tests themselves. For example, the Spelling pretest task consists of copying simple geometric figures; it has no language component, as does spelling. The Word Recognition pretest of letter "reading" (verbal identification) is not the same cognitive activity as single word pronunciation. There is a need for better assessment of the preschool basis of reading, spelling, and arithmetic abilities.

A fifth group of limitations concerns predictor selection in relation to neuropsychological theory. Initial reading of the theoretical literature in neuropsychology led to identification of five neuropsychological areas but the "Sensation-Perception" measures were separated into Somesthetic and Visual-Spatial because these two areas were not highly correlated. Thus, six domains were identified: Intelligence, Language, Attention-Memory, Fine Motor, Visual-Spatial, and Somesthetic. While the importance of Intelligence and Attention-Memory, and to a lesser extent, Somesthetic, Fine Motor, and Language domains were demonstrated, the six domains, considered as a group, did not function as effectively in capturing unique variance as hoped.

Identifying six rather than five domains of prediction does not challenge the basic concepts of the neuropsychological theories that guided test selection for this battery but does highlight the importance of several conceptual and methodological issues discussed previously, namely (a) consideration of developmental change in the interrelationships among neuropsychological abilities such that abilities which intercorrelate strongly in adults may not be strongly associated in young children; (b) consideration of how domains are operationalized across different studies; and (c) consideration of the relative levels of complexity of different tests, such that some tests reflect a broad array of component abilities whereas others reflect a simple circumscribed set of abilities. More work in developing neuropsychological theory (especially in the area of developmental neuropsychology) is needed as is more work in the application of neuropsychological theory to prediction and predictor selection.

With respect to limitations in selection of the predictors themselves, one measure was dropped from the analyses because of lack of correlation. The problem with the Language predictors were noted above. The AST is a poor measure of language abilities in young children as it fails to assess important areas of language (e.g., syntactic comprehension) and includes assessment of irrelevant topics (e.g., arithmetic comprehension). The Visual-Spatial measure contributed little unique prediction of achievement. Better selection and development of language predictors could be implemented in future studies by focusing on language measures that incorporate such constructs as syntactic comprehension or phonetic ability.

Finally, other sources of achievement variance should be explored in further studies. In addition to using predictors of neuropsychological functioning, formally assessing students's motivation for schooling, teachers's abilities, the relationship between teachers and their classes, home support for education, and students's emotional functioning may give rise to better predictions of achievement.

A sixth limitation was that a measure of the undergraduate examiners' interrater reliability in accuracy of test administration and scoring was lacking. However, neuropsychologists have used specially trained individuals, with equivalent levels of training, to administer neuropsychological batteries. Also, the comparability of the score distribution for five of six predictors (i.e., Intelligence, Attention-Memory, Language, Visual-Spatial, and Fine Motor) with population norms in the present sample suggested no systematic bias towards lower or higher scores for these variables. However, one Somesthetic variable had an elevated mean and the median was even more elevated (Mdn = 63). All the Somesthetic scores were log-transformed prior to using them in the analyses, and thus there may have been problems with the administration of the Somesthetic measures or problems with the norms themselves. The scoring of the tests was very straightforward and did not appear problematic. Use of direct observational measures of inter-examiner reliability and accuracy would be important for future investigations.

Conclusion

In conclusion, the theory-generated predictive screening battery developed for this study was able to classify academic achievement in ethnically diverse young urban school-age children over a 6-month period with accuracy commensurate with that of other neuropsychological batteries. Unfortunately, this study's results do not exceed previous levels of predictive accuracy. The protocol predicted Combined Academic and Arithmetic achievement better than Language-Related achievement, and as is inevitable in a correlational analysis predicted extreme groupings better (Low Average and High Average scores) better than those of the middle group (Average scores). Generally, the Intelligence and Attention-Memory measures predicted best, in that order, for the combined sample of Ethnic Minority and Ethnic Majority children (despite the cultural unfairness of Intelligence and Language measures). Somesthetic and Fine Motor predictors also were frequently selected in analyses, perhaps because they are less highly correlated with Intelligence. The latter two measures' predictive utility is of interest because they presumably more directly reflect brain integrity and are less affected by social or educational variation.

Ethnic Minority children's achievement were predicted best by Attention-Memory and Intelligence; Somesthetic, Language, Visual-Spatial, and Fine Motor variously emerged as predictors in selected analyses. Ethnic Majority children's achievement were best predicted by Intelligence, Attention-Memory, and Somesthetic. Younger children's achievement on simple preschool and early academic tasks was best predicted by Language, Somesthetic, and Visual-Spatial measures, while Older children's achievement (word recognition, spelling, and computation) was related to Intelligence, Attention-Memory, and Somesthetic measures.

The most "basic" or "simplest" tasks (Somesthetic and Fine Motor) contributed unique variance to most of the analyses, although they were ranked lower as predictors than Intelligence and Attention-Memory. Thus, these simple sensorimotor tests are useful in identifying the three levels of achievement across ethnic groups and young ages and may be markers of overall brain integrity or maturation.

From a neuropsychological perspective, these findings suggest that general intelligence, attentional and working memory factors, and sensorimotor functioning each contributes unique variance to the mental abilities that underlie academic achievement. This basic tripartite conceptualization of mental abilities (i.e., intelligence, memory, sensorimotor) is consistent with neuropsychological theories of Luria (1973) and later investigators. It may be that further division of intellectual abilities (e.g., verbal-language versus visual-spatial) is more directly relevant in cases of learning disability (where a specific ability may be impaired) than in predicting achievement among the general "regular" classroom population.

The predictive validity of the Attention-Memory, Somesthetic, Fine Motor, and Visual-

Spatial variables suggest that it may be possible to identify culturally fair predictors of academic achievement. While it still is the culturally biased measure (i.e., Intelligence) that predicts best for the full sample, Intelligence was not selected in several Ethnic Minority analyses and when excluded in certain others, did not lower the classification success rates. Despite mean differences between groups in predictor scores, children within each ethnic group who do better on the predictors will do better on the criterion. Nevertheless, it is important to identify and use culture-fair predictors because low SES and Ethnic Minority students should have the benefit of being assessed by the best available tests that do not penalize them for differences in language or sociocultural education.

Finally, these data suggest that more complex multifactorially determined predictors (Intelligence and Attention-Memory), which, theoretically, reflect functioning of the entire brain and its interconnections, are the most sensitive predictors of early achievement. This conclusion is consistent with neuropsychological data for adults showing that IQ scores are the most sensitive indicator of brain damage (Lezak, 1983). On the other hand, discrete areas of functioning (Somesthetic, Fine Motor, Language, and Visual-Spatial) can be differentiated and contribute independent variance to prediction. Therefore, the findings are congruent with a theory-guided selection of tests that assess both general intelligence and discrete abilities.

APPENDIX 1

LETTER OF PERMISSION

Stephen R. Clingerman, M.A.
5037 N. St. Louis
Chicago, Illinois 60625
(312) 267-5074

April 30, 1996

J. Clifford Kaspar, Ph.D., A.B.P.P.
5817 Howard Avenue
LaGrange Highlands 60525

Dear Dr. Kaspar:

This letter will confirm our recent telephone conversation. I am completing a doctoral dissertation at Loyola University of Chicago entitled, "Predictions of Urban, Ethnically Diverse, Early School-Age Children's Academic Achievement Using a Neuropsychological Screening Battery." I would like your permission to reprint in my dissertation excerpts from J.L. Schulman, J.C. Kaspar, and F.M. Throne (1965) Brain Damage and Behavior (Springfield, IL: Thomas). These excerpts to be reprinted are: The directions and two sample cards used in the Cards Test.

The requested permission extends to any future revisions and editions of my dissertation, including non-exclusive world wide rights in all languages, and to the prospective publication of my dissertation by University Microfilms, Inc. These rights will in no way restrict republication of the material in any other form by you or others authorized by you. Your signing of this letter also confirms that you own the copyright to the above-described material.

Sincerely,

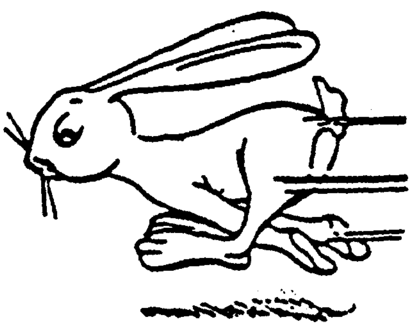
Stephen R. Clingerman

PERMISSION GRANTED FOR
USE REQUESTED ABOVE:

J. Clifford Kaspar, Ph.D., A.B.P.P.
5817 Howard Avenue
LaGrange Highlands 60525
Date May 4, 1996

APPENDIX 2

CARDS TEST



CARD SORTING

Here are some cards (Show). On some of them there is a rabbit, like this (Show), and on others there is a baby, like this (Show). They are all mixed up together. We do not want the rabbit to catch the baby. I will turn the cards over one by one and I want you to tell me every time you see a baby and I will take him out so the rabbit can not catch him. Do you understand? (If the child says `No,' repeat the instructions.)

[After the first correct response say "Good."]

Babies are Card #

2	70
6	87
12	89
16	94
29	99
33	117
40	133
43	139
54	167
59	193

REFERENCES

Ayers, A. J. (1975). Sensorimotor foundations of academic ability. In W. M. Cruikshank & D. P. Hallahan (Eds.), Perceptual and learning disabilities in children: Volume 2 Research and theory (pp. 301-358). Syracuse, NY: Syracuse University Press.

Bakker, D. J. (1983). Hemispheric specialization and specific reading retardation. In M. Rutter (Ed.), Developmental neuropsychiatry (pp. 498-506). New York: Guilford.

Bannatyne, A. (1971). Language, reading, and learning disabilities. Springfield, Il.: Charles C. Thomas.

Beery, K. E., & Buktenica, N. A. (1967). Developmental Test of Visual-Motor Integration. Chicago: Follett.

Bender, L. A. (1946). Bender Motor Gestalt Test: Cards and manual of instructions. The American Orthopsychiatric Association, Inc.

Bender, L. A. (1957). Specific reading disabilities as a maturational lag. Bulletin of the Orton Society, 7, 9-18.

Benson, D. F. & Denckla, M. B. (1969). Verbal paraphasia as a source of calculation disturbance. Archives of Neurology, 21, 96-102.

Benson, D. F. & Weir, W. F. (1972). Acalculia: Acquired anarithmetia. Cortex, 8, 465-472.

Benton, A. L. (1963). The Revised Visual Retention Test. New York: Psychological Corporation.

Benton, A. L. (1977). Reflections on the Gerstmann Syndrome. Brain & Language, 4, 45-62.

Benton, A. L. (1978). Some conclusions about dyslexia. In Benton, A. L. & Pearl, D. P. (Eds.) (1978). Dyslexia: An appraisal of current knowledge (pp. 453-476). New York: Oxford University Press.

Benton, A. L. (1979). Body schema disturbances: Finger agnosia and right-left disorientation. In K. M. Heilman & E. Valenstein (Eds.), Clinical neuropsychology (pp. 141-158). New York: Oxford University Press.

Birch, H. G. (1962). Dyslexia and maturation of visual function. In J. Money (Ed.), Reading disability: Progress and research in dyslexia. Baltimore: Johns Hopkins Press.

Blank, M., Weider, S., & Bridger, W. (1968). Verbal deficiencies in abstract thinking in early reading retardation. American Journal of Orthopsychiatry, 38, 823-834.

Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading-spelling patterns. Developmental Medicine and Child Neurology, 15, 663-687.

Brown, S. J., Rourke, B. P., & Cicchetti, D. V. (1989). Reliability of tests and measures used in the neuropsychological assessment of children. The Clinical Neuropsychologist, 3, 353-368.

Caramazza, A. & Berndt, R. S. (1978). Semantic and syntactic processes in aphasia: A review of the literature. Psychological Bulletin, 85, 898-918.

Cohen, J. (1957). The factorial structure of the WAIS between early adulthood and old age. Journal of Consulting Psychology, 21, 283-290.

Cohen, J. & Cohen, P. (1975). Applied multiple regression for the behavioral sciences. New York: Wiley.

Cohn, R. (1961). Dyscalculia. Archives of Neurology, 4, 301-307.

Critchley, M. (1966). Is developmental dyslexia the expression of minor cerebral damage? Clinical Proceedings of the Children's Hospital Washington D.C., 22, 213-222.

Cruikshank, W. M. (1968). The problems of delayed recognition and its correction. In A.H. Keeney & V. T. Keeney (Eds.), Dyslexia: Diagnosis and treatment of reading disorders (pp. 84-89). St. Louis: C.V. Mosby.

Dahmen, W., Hartje, W., Bussing, A., & Sturm, W. (1982). Disorders of calculation in aphasic patients: Spatial and verbal components. Neuropsychologia, 20, 144-153.

DeFries, J. C. & Gillis, J. J. (1991). Etiology of reading deficits in learning disabilities: Quantitative Genetic analysis. In J. E. Obrzut & G. W. Hynd (Eds.), Neuropsychological foundations of learning disabilities (pp. 29-47). New York: Academic Press.

de Hirsch, K., Jansky, J. J. & Langford, W. S. (1966). Predicting reading failure. New York: Harper & Row.

Denckla, M. B. (1979). Childhood learning disabilities. In K. M. Heilman & E. Valenstein (Eds.), Clinical neuropsychology (pp. 535-573). New York: Oxford University Press.

Denckla, M. B., LeMay, M., & Chapman, C. A. (1985). Few CT scan abnormalities found even in neurologically impaired learning disabled children. Journal of Learning Disabilities, 18, 132-135.

Denckla, M. B., Rudel, R. G., Chapman, C., & Krieger, J. (1985). Motor proficiency in dyslexic children with and without attentional disorders. Archives of Neurology, 42, 228-231.

Dennerll, R. D. (1964). Cognitive deficits and lateral brain dysfunction in temporal lobe epilepsy. Epilepsia, 5, 177-190.

Deysach, R. E. (1986). The role of neuropsychological assessment in the comprehensive evaluation of preschool-age children. School Psychological Review, 15, 233-244.

Doehring, D. G. (1978). The tangled web of behavioral research on developmental dyslexia. In A. L. Benton & D. Pearl (Eds.) Dyslexia: An appraisal of current knowledge (pp. 123-135). New York: Oxford University Press.

Downing, J. (1973). Cognitive factors in dyslexia. Child Psychiatry and Human Development, 4, 115-120.

Drake, W. (1968). Clinical and pathological findings in a child with a developmental learning disability. Journal of Learning Disabilities, 1, 486-502.

Duane, D. D. (1991). Biological foundations of learning disabilities. In J. E. Obrzut & G. W. Hynd (Eds.), Neuropsychological foundations of learning disabilities (pp. 7-27). New York: Academic Press.

Duffy, F., Denckla, M. B., Bartels, P., & Santini, G. (1980). Dyslexia: Regional differences in brain electrical activity by topographic mapping. Annals of Neurology, 7, 412-420.

Dunn, L. M. (1965). Expanded manual for the Peabody Picture Vocabulary Test. Circle Pines, MN: American Guidance Service.

Dunn, L. M. & Dunn, L. M. (1981). Peabody Picture Vocabulary Test-Revised: Manual for Forms L and M. Circle Pines, MN: American Guidance Service.

Ellison, P. H. (1983). The relationship of motor and cognitive function in infancy, pre-school, and early years. Journal of Clinical Child Psychology, 12, 81-90.

Fletcher, J. M. & Satz, P. (1982). Kindergarten prediction of reading achievement: A seven year longitudinal follow-up. Educational and Psychological Measurement, 42, 681-685.

Fletcher, J. M. & Satz, P. (1984). Test-based versus teacher-based predictions of academic achievement: A three-year longitudinal follow-up. Journal of Pediatric Psychology, 9, 193-203.

Fletcher, J. M., Taylor, H. G., Morris, R., & Satz, P. (1982). Finger recognition skills and reading achievement: A developmental neuropsychological analysis. Developmental Psychology, 18, 124-132.

Frank, J. & Levinson, H. (1973). Dysmetric dyslexia and dyspraxia. Journal of Child Psychiatry, 12, 690-701.

Frostig, M. & Maslow, P. (1973). Learning problems in the classroom: Prevention and remediation. New York: Grune & Stratton.

Gaddes, W. H. (1980). Learning disabilities and brain function: A neuropsychological approach. New York: Springer-Verlag.

Gaddes, W. H. (1981). An examination of the validity of neuropsychological knowledge in educational diagnosis and remediation. In G. W. Hynd & J. E. Obrzut (Eds.), Neuropsychological assessment of the school-age child (pp. 27-84). New York: Grune & Stratton.

Gaddes, W. H. (1983). Applied educational neuropsychology: Theories and problems. Journal of Learning Disabilities, 16, 511-514.

Gates, A. I. & MacGinitie, W. H. (1965). Gates-MacGinitie Reading Tests. New York: Teachers College Press.

Gesell, A. & Thompson, H. (1938). The psychology of early growth. New York: Macmillan.

Gonzales, R. & Roll, S. (1985). Relationship between assimilation, cognitive style, and intelligence. Journal of Cross Cultural Psychology, 16, 190-205.

Green, C., Lyles, C., & Eissenfeldt, A. (1980). Kindergarten diagnosis and training: A program report. In D. J. Sawyer (Ed.), Disabled readers: Insight, assessment, and instruction (pp. 50-56). Newark, DE: International Reading Association.

Gutkin, T. B., & Reynolds, C. R. (1981). Factor similarity of the WISC-R for white and black children from the standardization sample. Journal of Educational Psychology, 73, 227-231.

Harris, E. L. (1986). The contribution of twin research to the study of the etiology of reading disabilities. In S. D. Smith (Ed.), Genetics and learning disabilities (pp. 1-19). San Diego: College-Hill Press.

Haslam, R. H., Dalby, J. T., Johns, R. D., & Rademaker, A. W. (1981). Cerebral asymmetry in developmental dyslexia. Archives of Neurology, 38, 679-682.

Heilman, K. M. & Valenstein, E. (Eds.). (1985). Clinical neuropsychology (2nd ed.). New York: Oxford University Press.

Hermann, K. (1959). Reading disability: A medical study of word-blindness and related handicaps. Copenhagen: Munksgaard.

Hier, D. B., LeMay, M., Rosenberger, P. B., & Perlo, V. P. (1978). Developmental dyslexia: Evidence for a subgroup with a reversal of cerebral asymmetry. Archives of Neurology, 35, 90-92.

Hinshaw, S. P., Carte, E. T., & Morrison, D. C. (1986). Concurrent prediction of academic achievement in reading disabled children: The role of neuropsychological and intellectual measures at different ages. International Journal of Clinical Neuropsychology, 8, 3-8.

Hiscock, M. & Kinsbourne, M. (1987). Specialization of the cerebral hemispheres: Implications for learning. Journal of Learning Disabilities, 20, 130-143.

Hynd, G. W., & Cohen, M. (1983). Dyslexia: Neuropsychological theory, research, and clinical differentiation. New York: Grune & Stratton.

Hynd, G. W. & Semrud-Clikeman, M. (1989). Dyslexia and brain morphology. Psychological Bulletin, 106, 447-482.

Hynd, G. W., Semrud-Clikeman, M., Lorys, A. R., Novey, E. S., & Eliopulos, D. (1990). Brain morphology in Developmental Dyslexia and Attention-deficit Disorder/Hyperactivity. Archives of Neurology, 47, 919-926.

Ingram, T. T., Mason, A. W., & Blackburn, I. (1970). A retrospective study of 82 children with reading disability. Developmental Medicine and Child Neurology, 12, 271-281.

Jansky, J. J. (1978). A critical review of "Some developmental and predictive precursors of reading disabilities: A six-year follow-up." In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 379-394). New York: Oxford University Press.

Jansky, J. & de Hirsch, K. (1972). Preventing reading failure: Prediction, diagnosis, intervention. New York: Harper & Row

Jastak, S. & Wilkinson, G. S. (1984). Wide Range Achievement Test-Revised. Wilmington, DE: Jastak Associates.

Johnson, D. J. & Mykleburst, H. (1967). Learning disabilities. New York: Grune & Stratton.

Johnson, D. J. & Mykleburst, H. (1971). Learning disabilities (2nd ed.). New York: Grune & Stratton.

Joschko, M. & Rourke, B. P. (1985). Neuropsychological subtypes of learning-disabled children who exhibit the ACID pattern on the WISC. In B.P. Rourke (Ed.), Neuropsychology of learning disabilities: Essentials of subtype analysis (pp. 40-64). New York: Guilford.

Kaspar, J. C., Agresti, A., Reich, J., Wills, K., Cleland, J., Arnemann, C., & Clingerman, S. R. (1992, August). Memory trace functions and academic achievement in older children. Paper presented at the 100th American Psychological Association Meeting, Washington, D.C.

Kaspar, J. C. & Koshaba, J. E. (1974, February). Correlation of academic performance with measures of distractibility. Paper presented at the meeting of the International Neuropsychological Society, Boston, MA.

Kaspar, J. C. & Sokolec, J. (1980). Relationship between neurological dysfunction and a test of speed of motor performance. Journal of Clinical Neuropsychology, 2, 13-21.

Kaufman, A. S. (1979). Intelligence testing with the WISC-R. New York: Wiley.

Kavale, K. (1982). Meta-analysis of the relationship between visual perceptual skills and reading achievement. Journal of Learning Disabilities, 15, 42-51.

Kerasotes, D. & Walker, C. E. (1983). Hyperactive behavior in children. In C. E. Walker & M. C. Roberts (Eds.), Handbook of clinical child psychology (pp. 498-523).

Kinsbourne, M. & Warrington, E. K. (1963). The Developmental Gerstmann Syndrome. Archives of Neurology, 8, 490-501.

Klonoff, H. & Low, M. (1974). Disordered brain functions in young children and early adolescents: Neuropsychological and electroencephalographic correlates. In R. M. Reitan & L. A. Davison (Eds.), Clinical neuropsychology: Current status and applications (pp. 121-178). Washington, D.C.: Winston & Sons.

Knights, R. & Norwood, J. (1980). Revised smoothed normative data on the Neuropsychological Test Battery for Children. Privately published.

Kolb, B. & Wishaw, I. Q. (1980). Fundamentals of human neuropsychology. San Francisco: W. H. Freeman.

Koppitz, E. M. (1970). The Bender Gestalt Test for young children. New York: Grune and Stratton.

Kroll, L. G. (1984). L.D.'s: What happens when they are no longer children? Academic Therapy, 20, 133-148.

Laosa, L. M. (1984). Ethnic, socioeconomic, and home language influences upon early performance on measures of abilities. Journal of Educational Psychology, 76, 1178-1198.

LaTorre, R. A. (1985). Kindergarten screening: A cross-validation of the Florida Kindergarten Screening Battery. The Alberta Journal of Educational Research, 31, 174-190.

Levin, H. & Spiers, P. A. (1985). Acalculia. In K. M. Heilman & E. Valenstein (Eds.), Clinical neuropsychology (pp. 97-114) (2nd ed.). New York: Oxford University Press.

Lezak, M. (1983). Neuropsychological assessment (2nd ed.). New York: Oxford University Press.

Lubs, H. A., Rabin, M., Saucier-Carland, Wen, Gross-Glenn, K., Duara, R., Levin, B., & Lubs, M. L. (1991). Genetic bases of developmental dyslexia: Molecular studies. In J. E. Obrzut & G. W. Hynd (Eds.), Neuropsychological foundations of learning disabilities (pp. 49-77). New York: Academic Press.

Luria, A. R. (1966). Human brain and psychological processes. New York: Harper & Row.

Lyon, G. R., Stewart, N., & Freedman, D. (1982). Neuropsychological characteristics of empirically derived subgroups of learning disabled readers. Journal of Clinical Neuropsychology, 4, 343-365.

Mattis, S. (1978). Dyslexia syndromes: A working hypothesis that works. In A. L. Benton and D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 44-58). New York: Oxford University Press.

Mattis, S., French, J. H., & Rapin, I. (1975). Dyslexia in children and young adults: Three independent neuropsychological syndromes. Developmental Medicine and Child Neurology, 17, 150-163.

McFie, J. (1961). Recent advances in phrenology. Lancet, ii, 360-363.

Morris, R., Blashfield, R., & Satz, P. (1986). Developmental classification of reading-disabled children. Journal of Clinical and Experimental Neuropsychology, 8, 371-392.

Morrison, D. C. & Hinshaw, S. P. (1988). The relationship between neuropsychological/perceptual performance and socioeconomic status in children with learning disabilities. Journal of Learning Disabilities, 21, 124-128.

Muehl, S., & Forrell, E. R. (1973). A follow-up study of disabled readers: Variables related to high school reading performance. Reading Research Quarterly, 9, 110-123.

Naglieri, J. A., & Yazzie, C. (1983). Comparison of the WISC-R and PPVT-R with Navajo children. Journal of Clinical Psychology, 39, 598-600.

Obrzut, J. E. (1989). Dyslexia and neurodevelopmental pathology: Is the neurodiagnostic technology ahead of the psychoeducational technology. Journal of Learning Disabilities, 22, 217-218.

Obrzut, J. E. & Hynd, G. W. (Eds.). (1991). Neuropsychological foundations of learning disabilities. New York: Academic Press.

Orton, S. T. (1937). Reading, writing, and speech problems in children. New York: Norton.

Pirozollo, F. J. (1981). Language and brain: Neuropsychological aspects of developmental reading disability. School Psychology Review, 10, 350-355.

Pirozollo, F. J. & Campanella, D. J. (1981). The neuropsychology of developmental speech disorders, language disorders, and learning disabilities. In G. W. Hynd & J. E. Obrzut (Eds.), Neuropsychological assessment and the school-age child (pp. 167-191). New York: Grune & Stratton.

Rabinovitch, R. D. (1968). Reading problems in children: Definitions and classification. In A. H. Keeney & V.T. Keeney (Eds.), Dyslexia: Diagnosis and treatment of reading disorders (pp. 1-10). St. Louis: C. V. Mosby.

Raven, J. C. (1965). Guide to using the Coloured Progressive Matrices. London: H.K. Lewis.

Reitan, R. M. (1969). Manual for administration of the Halstead-Reitan Neuropsychological Test Batteries for adults and children. Privately published, Indianapolis. (Available from R. M. Reitan, 1338 E. Edison Street, Tuscon, AZ 85719).

Reitan, R. M. (1974). Psychological effects of cerebral lesions in children of early school age. In R. M. Reitan & L. A. Davison (Eds.), Clinical neuropsychology: Current status and applications (pp. 53-89). Washington, D.C.: Winston & Sons.

Reitan, R. M. (1987). Neuropsychological evaluation of children. Tuscon: Neuropsychology Press.

Reitan, R. M. & Wolfson, D. (1985). Neuroanatomy and neuropathology: A guide for clinical neuropsychologists. Tuscon, Arizona: Neuropsychology Press.

Roeltgen, D. P., Sevush, S., & Heilman, K. M. (1983). Pure Gerstmann's Syndrome from a focal lesion. Archives of Neurology, 40, 46-47.

Rourke, B. P. (1978). Reading, spelling, and arithmetic disabilities: A neuropsychological perspective. In H.R. Myklebust (Ed.), Progress in learning disabilities (Vol. 4, pp. 97-120). New York: Grune & Stratton.

Rourke, B. P. (Ed.). (1985). Neuropsychology of learning disabilities: Essentials of subtype analysis. New York: Guilford.

Rourke, B. P. (1989). Nonverbal learning disability: The syndrome and the model. New York: Guilford.

Rourke, B. P. (Ed.) (1991). Neuropsychological validation of learning disability subtypes. New York: Guilford.

Rourke, B. P. & Strang, J. D. (1978). Neuropsychological significance of variations in patterns of academic performance: Motor, psychomotor, and tactile perception abilities. Journal of Pediatric Psychology, 3, 62-66.

Rourke, B. P. & Strang, J. D. (1983). Subtypes of reading and arithmetic disabilities: A neuropsychological analysis. In M. Rutter (Ed.), Developmental Neuropsychiatry (pp. 473-488). New York: Guilford.

Rudnick, M., Sterritt, G. M., & Flax, M. (1967). Auditory and visual rhythm perception and reading ability. Child Development, 38, 581-587.

Rumsey, J. M., Dorwart, R., Vermess, M., Denckla, M. B., Kruesi, M. P., & Rapoport, J. L. (1986). Magnetic resonance imaging of brain anatomy in severe developmental dyslexia. Archives of Neurology, 43, 1045-1046.

Rutter, M. (1978). Prevalence and types of dyslexia. In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 5-28). New York: Oxford University Press.

Rutter, M. & Yule, W. (1975). The concept of specific reading retardation. Journal of Child Psychology and Psychiatry, 16, 181-197.

Sattler, J. M. (1988). Assessment of Children (3rd ed.). San Diego: Jerome Sattler, Publisher, Inc.

Sattler, J. M. (1992). Assessment of Children (Revised and updated 3rd ed.). San Diego: Jerome Sattler, Publisher, Inc.

Satz, P. & Fletcher, J. M. (1982). Manual for the Florida Kindergarten Screening Battery. Odessa, FL: Psychological Assessment Resources.

Satz, P., & Fletcher, J. M. (1988). Early identification of learning disabled children: An old problem revisited. Journal of Consulting and Clinical Psychology, 56, 824-829.

Satz, P. & Friel, J. (1974). Some predictive antecedents of specific reading disability: A preliminary two-year follow-up. Journal of Learning Disabilities, 7, 437-444.

Satz, P. & Friel, J. (1978). Predictive validity of an abbreviated screening battery. Journal of Learning Disabilities, 11, 347-351.

Satz, P. & Morris, R. (1981). Learning disability subtypes: A review. In F. J. Pirozollo & M. C. Wittrock (Eds.), Neuropsychological and cognitive processing in reading (pp. 109-141). New York: Academic Press.

Satz, P., Taylor, H. G., Friel, J., & Fletcher, J. M. (1978). Some developmental and predictive precursors of reading disabilities: A six year follow-up. In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 457-501). New York: Oxford.

Schulman, J. L., Buist, C., Kaspar, J. C., Child, D., & Fackler, E. (1969). An objective test of speed of fine motor functioning. Perceptual and Motor Skills, 29, 243-255.

Schulman, J. L., Kaspar, J. C., & Throne, F. M. (1965). Brain damage and behavior. Springfield, IL: Thomas.

Selz, M. (1981). Halstead-Reitan Neuropsychological Test Battery for Children. In G. W. Hynd & J. E. Obrzut (Eds.), Neuropsychological assessment and the school-age child (pp. 195-235). New York: Grune & Stratton.

Senf, G. M., & Freundl, P. C. (1971). Memory and attention factors in specific learning disabilities. Journal of Learning Disabilities, 4, 94-106.

Semrud-Clikeman, M., & Hynd, G. W., (1990). Right hemisphere dysfunction in nonverbal learning disabilities: Social, academic, and adaptive functioning in adults and children. Psychological Bulletin, 107, 196-209.

Shapiro, B. K., Palmer, F. B., Wachtel, R. C., & Capute, A. J. (1984). Issues in the early identification of specific learning disability. Developmental and Behavioral Pediatrics, 5, 15-20.

Shenk, B., Fitzsimmons, J., Bullard, P. C., & Satz, P. (1980). A prevention model for children at risk of reading failure. In R. M. Knights & D. J. Bakker (Eds.), Treatment of hyperactive and learning disordered children (pp. 31-48). Baltimore, MD: University Park Press.

Siegel, L. S., & Linder, A. (1984). Short term memory processes in children with reading and arithmetic disabilities. Developmental Psychology, 20, 200-207.

Silver, A. A. (1978). Prevention. In A. L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 351-376). New York: Oxford University Press.

Silver, A. A. & Hagin, R. A. (1964). Specific reading disability: Follow-up studies. American Journal of Orthopsychiatry, 34, 95-102.

Silver, A. A. & Hagin, R. A. (1990). Disorders of learning in childhood. New York: Wiley.

Silver, L. B. (1971). A proposed view on the etiology of the neurological learning disability syndrome. Journal of Learning Disabilities, 4, 123-132.

Small, N. (1969). Levels of perceptual functioning in children: A developmental study. Unpublished Master's thesis. University of Florida.

Smith, S. D. (Ed.). (1986). Genetics of learning disabilities. San Diego: College-Hill Press.

Smith, S. D., Pennington, B. F., Fain, P. R., & Ing, P. S. (1989). Genetic linkage analysis with specific dyslexia: Use of multiple markers to include and exclude possible loci. In G. T. Pavlidis's (Ed.), Dyslexia: A Neuropsychological and Learning Perspective. London: Wiley.

Spellacy, F. & Peter, B. (1978). Dyscalculia and elements of the Developmental Gerstmann Syndrome in school children. Cortex, 14, 197-206.

Spreen, O. (1978a). The dyslexias: A discussion of neurobehavioral research. In A.L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge (pp. 174-194). New York: Oxford University Press.

Spreen, O. (1978b). Prediction of school achievement from kindergarten to grade five: Review and report of a follow-up study. Victoria, B.C.: Department of Psychology, University of Victoria, Research Monograph No. 33.

Spreen, O. (1989). Learning disability, neurology, and long-term outcome: Some implications for the individual and for society. Journal of Clinical and Experimental Neuropsychology, 11, 117-140.

Spreen, O. & Haaf, R. G. (1986). Empirically derived learning disability subtypes: A replication attempt and longitudinal patterns over 15 years. Journal of Learning Disabilities, 19, 170-180.

Spreen, O. & Strauss, E. (1991). A compendium of neuropsychological tests: Administration, norms, and commentary. New York: Oxford University Press.

Strauss, A. A. & Lehtinen, L. E. (1947). Psychopathology and education of the brain-injured child. New York: Grune & Stratton.

Sutter, E., Bishop, P., & Battin, R. R. (1987). Psychometric screening for attention deficit disorder in a clinical setting. Journal of Psychoeducational Assessment, 3, 227-235.

Tallal, P., Townsend, J., Curtiss, & Wulfeck, B. (1991). Phenotypic profiles of language-impaired children based on genetic/family history. Brain and Language, 41, 81-95.

Tallal, P., Stark, R., & Mellits, D. (1985). The relationship between auditory temporal analysis and receptive language development: Evidence from studies of developmental language disorder. Neuropsychologia, 23, 527-536.

Tarter, R. E. (1972). Intellectual and adaptive functioning in epilepsy. Disease of the Nervous System, 33, 759-770.

Taylor, G. H. & Fletcher, J. M. (1983). Biological foundations of "specific developmental disorders": Methods, findings, and future directions. Journal of Clinical Child Psychology, 12, 46-65.

Teeter, P. (1985). Neurodevelopmental investigation of academic achievement: A report of years 1 and 2 of a longitudinal study. Journal of Consulting and Clinical Psychology, 53, 709-717.

Terman, L. M. & Merrill, M. A. (1937). Measuring intelligence. Boston: Houghton Mifflin.

Townes, B. D., Turpin, E. W., Martin, D. C., & Goldstein, D. (1980). Neuropsychological correlates of academic success among elementary school children. Journal of Consulting and Clinical Psychology, 48, 675-684.

Uno, H., Tarara, R., Else, J. G., Suleman, M.A., & Sapolsky, R. M. (1989). Chronic stress and hippocampal damage in primates. Journal of Neuroscience, 9, 1705-1711.

Vellutino, F. R. (1978). Toward an understanding of dyslexia. In A. L. Benton & D. Pearl (Eds.) Dyslexia: An appraisal of current knowledge (pp. 62-111). New York: Oxford University Press.

Vellutino, F. R. (1979). The validity of perceptual deficit explanations of reading disability: A reply to Fletcher and Satz. Journal of Learning Disabilities, 12, 160-167.

Vellutino, F. R. (1991). Introduction to three studies on reading acquisition: Convergent findings on theoretical foundations of code-oriented versus whole-language approaches to reading instruction. Journal of Educational Psychology, 83, 437-443.

Vellutino, F. R. & Scanlon, D. M. (1985). Free recall of concrete and abstract words in poor and normal readers. Journal of Experimental Child Psychology, 39, 363-380.

Warrington, E. K. (1982). The fractionation of arithmetical skills: A single case study. Quarterly Journal of Experimental Psychology, 34A, 31-51.

Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children-Revised. San Antonio: Psychological Corporation.

Weinberg, W. A., & McLean, A. (1986). A diagnostic approach to developmental specific learning disorders. Journal of Child Neurology, 1, 158-172.

Welsh, M. C. & Pennington, B. F. (1988). Assessing frontal lobe functioning in children: Views from developmental psychology. Developmental Neuropsychology, 4, 199-230.

Welsh, M. C., Pennington, B. F., & Groisser, D. B. (1991). A normative-developmental study of executive function: A window on prefrontal function in children. Developmental Neuropsychology, 7, 131-149.

Wepman, J. (1958). Auditory Discrimination Test. Chicago: Language Research Associates.

Whalen, C. K., Henker, B., & Hinshaw, S. P. (1985). Cognitive-behavioral therapies for hyperactive children: Premises, problems, and prospects. Journal of Abnormal Child Psychology, 13, 391-410.

VITA

Stephen R. Clingerman was born in Goshen, Indiana and is the son of Phyllis Drummond and the late Robert E. Clingerman. Mr. Clingerman is the stepson of William R. Drummond. Mr. Clingerman completed his elementary education in the public schools of Goshen, Indiana and graduated from Mount Prospect High School in Mount Prospect, Illinois. He received his Bachelor of Arts degree from the University of Illinois in May of 1978. He received the Master of Science degree from George Williams College in December of 1983 and the Master of Arts degree from Loyola University of Chicago in May of 1989. Mr. Clingerman is the co-author of eleven publications that emerged from his work as Research Assistant for Ronald T. Brown, Ph.D. at the University of Illinois-Chicago Pediatric Psychopharmacology Project. He is co-author of papers written during his tenure as Loyola University Research Assistant for Richard Bowen, Ph.D., and J. Clifford Kaspar, Ph.D.

DISSERTATION APPROVAL SHEET

The dissertation submitted by Stephen R. Clingerman has been read and approved by the following committee:

Karen Wills, Ph.D., Director
Associate Professor, Clinical Psychology
Loyola University Chicago

Alan DeWolfe, Ph.D.
Professor, Clinical Psychology
Loyola University of Chicago

Joseph Durlak, Ph.D.
Professor, Clinical Psychology
Loyola University of Chicago

Martha Ellen Wynne, Ph.D.
Assistant Professor, Educational Psychology
Loyola University of Chicago

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the committee with reference to content and form.

The dissertation is, therefore, accepted in partial fulfillment of the requirements for the degree of doctor of philosophy.

26 November 1996
Date

Karen E. Wills
Director's Signature