PROFILE ANALYSIS OF THE KAUFMAN ASSESSMENT BATTERY FOR
CHILDREN, SECOND EDITION WITH AFRICAN AMERICAN AND CAUCASIAN
PRESCHOOL CHILDREN

A DISSERTATION
SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

BY
BRITTANY ANN DALE

APPROVED BY:

_____ Committee Chairperson	_____ Date
_____ Committee Member	_____ Date
_____ Committee Member	_____ Date
_____ Committee Member	_____ Date
_____ Dean of Graduate School	_____ Date

BALL STATE UNIVERSITY
MUNCIE, INDIANA
JULY 2009

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of the many people in my life who have provided me with love and encouragement throughout my education. My journey to a completed dissertation and doctoral degree began with the love of my parents and has concluded with the support of my committee.

I would like to extend my heartfelt thanks to my committee members: Dr. David McIntosh, Dr. Paul Spengler, Dr. Holmes Finch, and Dr. Kathy Segrist. Their constructive feedback, willingness to answer questions, and support of my academic and career goals has been invaluable to this experience. I would especially like to thank Dr. McIntosh for his patience, encouragement, and availability while guiding me through this process. He has challenged me to be confident in my work and has inspired me to pursue all of my goals. I am especially grateful for Dr. McIntosh's understanding as I attempted to balance being a student, writing a dissertation, and maintaining my responsibilities to my family.

Furthermore, my family has been my strongest support. This dissertation would not have been possible without my parents. They have emphasized the value of education and always saw my full potential. I am grateful to my brother, as sibling rivalry has been extended to the doctorate level. Lastly, my husband has been my rock throughout this process. Without his encouragement, love, and patience, I am not sure I would have found the strength to be persistent and accomplish my goals.

DEDICATION

To Alan.

Love is Patient.

Love is Kind.

It bears all things, believes all things, hopes all things, endures all things.

Love Never Fails.

Your love is my strength.

TABLE OF CONTENTS

LIST OF TABLES.....	iv
LIST OF FIGURES.....	vi
ABSTRACT.....	vii
CHAPTER I: INTRODUCTION.....	1
<i>Measures of Intelligence</i>	2
<i>Theories of Intelligence</i>	4
<i>Racial Differences in Intelligence</i>	5
<i>Preschool Assessment</i>	10
<i>Purpose of the Present Study</i>	13
<i>Significance of Study</i>	14
CHAPTER II: REVIEW OF THE LITERATURE.....	17
<i>Historical Perspective of Assessment and Theory</i>	17
<i>Intelligence Theory</i>	21
<i>Contemporary Influences on Understanding Intelligence</i>	23
<i>Cattell-Horn Gf-Gc Theory</i>	23
<i>Carroll's Three Stratum Theory</i>	24
<i>Contemporary Cattell-Horn-Carroll (CHC) Theory</i>	25
<i>Empirical Support for the CHC Model</i>	28
<i>Contemporary Assessment Aligned with Theory</i>	29
<i>Racial Differences in Intelligence</i>	30
<i>The Mean IQ Score Discrepancy</i>	31
<i>Implications of an IQ Gap</i>	33
<i>Nature of the IQ Gap</i>	34
<i>Contemporary Tests of Cognitive Ability and Race</i>	40
<i>Measuring Bias in Testing</i>	41
<i>History of Preschool Assessment</i>	44
<i>Factors of Intelligence in Preschoolers</i>	45
<i>Accountability of Education for Preschool Children</i>	48
<i>Cognitive Batteries for Preschool Children</i>	51
<i>Chapter Summary</i>	54
CHAPTER III: RESEARCH METHODOLOGY.....	57
<i>Participants</i>	57
<i>Sample Size</i>	59
<i>Missing Data</i>	59
<i>Procedures</i>	60

<i>Instrumentation</i>	61
<i>Kaufman Assessment Battery for Children, Second Edition (KABC-II)</i> ...	61
<i>Chapter Summary</i>	67
CHAPTER IV: RESULTS.....	69
<i>Descriptive Statistics</i>	69
<i>Assumptions</i>	72
<i>Profile Analysis</i>	74
<i>Research Question One</i>	76
<i>Research Question Two</i>	77
<i>Follow-up Analyses</i>	79
<i>Chapter Summary</i>	81
CHAPTER V: DISCUSSION.....	84
<i>Summary of Results and Implications</i>	85
<i>Research Question One</i>	86
<i>Research Question Two</i>	89
<i>Follow-up Analyses</i>	94
<i>Limitations of the Present Study</i>	97
<i>Recommendations for Future Research</i>	98
<i>Chapter Summary</i>	99
REFERENCES.....	101
APPENDICES.....	117
<i>Appendix A: Letter Given to Parents of the Participants</i>	117
<i>Appendix B: Permission Form Completed by Parents of the Participants</i>	118
<i>Appendix C: Demographic Questionnaire</i>	119
<i>Appendix D: T-Test Mean Comparison for the KABC-II Clusters and Subtests by Ethnicity</i>	120

LIST OF TABLES

3.1	Parental Education by Race.....	59
3.2	Kaufman Assessment Battery for Children, Second Edition Subtests.....	64
4.1	Means, Standard Deviations, and Ranges of the KABC-II FCI, Cluster, and Subtest Scores for Total Sample ($n = 98$).....	71
4.2	Means, Standard Deviations, and Ranges of the KABC-II FCI, Cluster, and Subtest Scores for African American and Caucasian Groups.....	73
4.3	Follow-up ANOVA Results for the KABC-II Subtests	80
4.4	T-Test Mean Comparison for the KABC-II Clusters and Subtests by Ethnicity..	82

LIST OF FIGURES

4.1	Mean CHC Broad Ability Factor Scores for the African American and Caucasian Groups.....	77
4.2	Mean Subtest Scores for the African American and Caucasian Groups.....	79

ABSTRACT

The purpose of the present study was to determine if African American and Caucasian preschool children displayed similar patterns of performance among the Cattell-Horn-Carroll (CHC) factors measured by the Kaufman Assessment Battery for Children, Second Edition (KABC-II). Specifically, a profile analysis was conducted to determine if African Americans and Caucasians displayed the same patterns of highs and low and scored at the same level on the KABC-II composites and subtests. Forty-nine African American (mean age = 59.14 months) and 49 Caucasian (mean age = 59.39) preschool children from a Midwestern City were included in the study and were matched on age, sex, and level of parental education. Results of a profile analysis found African American and Caucasian preschool children had a similar pattern of highs and lows and performed at the same level on the CHC broad abilities as measured by the KABC-II. Comparison of the overall mean IQ indicated no significant differences between the two groups. The overall mean difference between groups was 1.47 points, the smallest gap seen in the literature. This finding was inconsistent with previous research indicating a one standard deviation difference in IQ between African Americans and Caucasians. A profile analysis of the KABC-II subtests found the African American and Caucasian groups performed at an overall similar level, but did not show the same pattern of highs and lows. Specifically, Caucasians scored significantly higher than African Americans on the Expressive Vocabulary subtest which measures the CHC narrow ability of Lexical Knowledge.

Results of this study supported the KABC-II's authors' recommendation to make interpretations at the composite level. When developing hypotheses of an individual's

strengths and weaknesses in narrow abilities, clinicians should be cautious when interpreting the Expressive Vocabulary subtest with African Americans. Overall, results of this study supported the use of the KABC-II with African American preschool children. When making assessment decisions, clinicians can be more confident in an unbiased assessment with the KABC-II.

Future research could further explore the CHC narrow abilities in ethnically diverse populations. Additionally, more research should be conducted with other measures of cognitive ability designed to adhere to the CHC theory, and the appropriateness of those tests with an African American population. Furthermore, future research with the KABC-II could determine if the results of the present study were replicated in other age groups.

CHAPTER I

Introduction

As the American population continues to become more ethnically diverse, diversity issues have become popular within psychological research. Specifically, as measures of cognitive ability are revised, test developers have made attempts to match the standardization samples to current demographic variables of the United States population (Kaufman & Kaufman, 2004; McGrew & Woodcock, 2001; Roid, 2003). The attempt to create unbiased tests has shifted the research focus to exploring whether similar factor structures are seen among different ethnic groups, especially African Americans. Historically, research discussing test bias has concentrated on the mean IQ difference between Caucasians and African Americans. Specifically, research has found a one standard deviation difference between the two groups indicating Caucasians perform approximately 15 points better than African Americans on traditional tests of cognitive ability (Jensen, 1998). Since current research has shifted away from the mean IQ difference, information regarding the utility of newly revised measures of cognitive ability with diverse populations is lacking.

Furthermore, early childhood assessment has been rooted in legislation and public programs aimed at early identification of children at-risk for learning problems and intervention have been important topics in research. Beginning with the child study

movement in the early 1900's, researchers searched for a comprehensive understanding of both normal and atypical developing children (Ford & Dahinten, 2005). With the advancement of funding for public preschool programs (e.g., Project Head Start) and legislation regarding special education placement and programming, appropriate assessment of young children became a necessity. Controversy over the appropriate use of cognitive ability measures with preschool children continues to be discussed in the research (Bracken, 1994) because of the implications involved with the use of assessment results. Results of cognitive ability measures carry important weight when determining special education placement, especially mental handicaps and specific learning disabilities (Holdnack & Weiss, 2006). Moreover, cultural bias within cognitive ability tests has threatened their appropriateness with specific populations, especially African Americans (Jensen, 1980). As tests developers publish revised versions of commonly used cognitive ability measures with preschoolers, issues of test bias must be considered (Edwards & Fuller, 2005).

Measurement of Intelligence

Although the assessment of individual abilities for educational purposes has come to the forefront during the 20th century, the measurement and models of intelligence have been around since the 19th century with the famous works of Sir Francis Galton and Alfred Binet (Wasserman & Tulskey, 2005). Considered the “father of assessment,” Galton was the first to use intelligence tests as a way to systematically study intelligence (Sattler, 2001; Wasserman & Tulskey, 2005). Galton assumed sensory and motor characteristics positively correlated with intelligence, and he developed different methods

to study individual differences. Although his assumptions were not validated, he paved the way for the study of individual differences.

The first intelligence test is credited to Binet, along with his colleague Simon, as they looked at intelligence from a different perspective than Galton (Suen & French, 2003). They viewed intelligence in terms of higher mental processes, such as memory, and developed different tasks to help them identify school-aged children with mental retardation. In 1905, the Simon-Binet Scale was developed, which measured factors such as memory, attention, concentration, and comprehension (Brody, 2000; Suen & French, 2003). Lewis Terman developed a version of Simon and Binet's test for use in the United States, the Stanford-Binet Intelligence Scales (Terman, 1916).

The Stanford-Binet highly influenced David Wechsler as he attempted to develop a test that measured both verbal and nonverbal abilities (Reynolds & Kaufman, 1985; Zhu & Weiss, 2005). Borrowing from the Army Alpha and Beta tests, Wechsler published the Wechsler-Bellvue Scale in 1939, and it soon became the most popular test of intelligence (Wasserman & Tulskey, 2005). As the popularity of the Wechsler scale increased, Wechsler developed several other versions of his test for use with different ages (i.e., Wechsler Intelligence Scale for Children [WISC; 1949], Wechsler Preschool and Primary Scale of Intelligence [WPPSI; 1967], and Wechsler Adult Intelligence Scale [WAIS; 1955]). The Wechsler Scales have been highly influential within the Field of Psychology as they are the most widely used and researched measures of cognitive ability (Sattler, 2001; Zhu & Weiss, 2005).

Theories of Intelligence

Today, the most popular tests used in the assessment of preschool children are based on the works of Carroll (1993) and Horn and Cattell (1966; Tusing & Ford, 2004). Cattell and Horn's (1966) original *Gf-Gc* Theory of intelligence proposed two abilities involved in human intelligence: fluid intelligence (*Gf*) and crystallized intelligence (*Gc*). Fluid intelligence refers to nonverbal, relatively culture-free, mental efficiency including mental processes and operations. Crystallized intelligence refers to acquired skills and knowledge that are dependent on exposure to culture (Sattler, 2001). Carroll's Three Stratum Theory of Intelligence (1993) divided intelligence into three categories: many distinct narrow abilities, eight broad abilities, and one general intelligence factor, *g*.

Carroll recognized his Three Stratum Theory had many similarities to Cattell and Horn's work but recognized the importance of the general ability factor. Carroll's (1993) Stratum I included several narrow abilities; Stratum II included eight broad abilities, fluid intelligence (*Gf*), crystallized intelligence (*Gc*), general memory and learning (*Gy*), broad visual perception (*Gv*), broad auditory perception (*Ga*), broad retrieval ability (*Gr*), broad cognitive speediness (*Gs*), and processing speed (*Gt*); and Stratum III included the general ability factor, *g*. These two theories contained many similarities and were later integrated to become the Cattell-Horn-Carroll *Gf-Gc* Theory (CHC) and is currently the dominant theory of intelligence (Alfonso, Flanagan, & Radwan, 2005). In 1999, Horn and Carroll informally agreed to the Cattell-Horn-Carroll theory terminology (McGrew, 2005). The CHC model is a three tier model with specific narrow abilities, eight broad abilities, and one general intelligence factor, *g*, resembling Carroll's Three Stratum Theory. The assessment measure used in current study, the Kaufman Assessment Battery

for Children, Second Edition (KABC-II; Kaufman & Kaufman, 2004) was developed to closely adhere to the CHC theory of intelligence. Current research involving preschoolers focuses on determining what, if any, CHC factors are involved in their cognitive ability (Tusing & Ford, 2004).

Racial Differences in Intelligence

Documented differences in mean intelligence scores between Caucasians and African Americans have been reported in research (Brown, Reynolds, & Wihitaker, 1999; Rushton & Jensen, 2005). These differences date back to the use of wide-spread mental testing during World War I (Rushton & Jensen, 2005). This large scale test administration yielded a three-year difference in mental age between Caucasians and African Americans suggesting Caucasians were intellectually superior to African Americans (Rury, 1988). More modern standardized tests of intelligence have found a one standard deviation difference in mean IQ between the two populations (Jensen, 1998). Specifically, research using the popular Wechsler Scales found the African American- Caucasian IQ gap ranged from 17 to 25 points (Arinoldo, 1981; Tuttle, 1966).

Studies utilizing the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967) indicated a smaller, but still significant, difference between African Americans and Caucasians across all composite IQ scores (Arinoldo, 1981; Tuttle, 1966). Preschoolers displayed a 10.8 point difference on the Full Scale IQ, a 10.5 difference on the Verbal IQ, and an 8.9 point difference on the Performance IQ (Tuttle, 1966). The IQ gap also was seen in other measures of cognitive ability (Vincent, 1991). Thorndike et al. (1986) found a 10-12 point advantage for Caucasian children from the standardization sample of the Stanford Binet Intelligence Scale, Fourth Edition (SB-IV).

Standardization data from the Kaufman Assessment Battery for Children (K-ABC) indicated a 7 point gap in children aged 2-12 (Kaufman & Kaufman, 1983). Given the smallest IQ gap seen in a modern test of intelligence, the K-ABC was seen as an appropriate test to use with an ethnically diverse population (Kaufman & Lichtenberger, 2002). Evidence regarding the closure of the IQ gap in recent years has been mixed with some researchers seeing progress (Vincent, 1991) and others indicating no decrease during the 20th century (McGurk, 1982; Shuey, 1966).

By age three or four, ethnic and racial differences in intellectual functioning become apparent and are stable throughout the school years (Sattler, 2001). Research using the WPPSI and the Stanford Binet Intelligence Scale, Form L-M (Terman & Merrill, 1960) has shown a gap between African American and Caucasian preschool children's mean IQs (Smith, Duncan, & Lee, 2003). Prior to controlling for demographic factors, the authors found a 19-point gap between the two ethnic groups. Evidence of an IQ gap beginning in preschool that lasts through the school-aged years and beyond has strong implications for the educational system. Naglieri and Rojahn (2001) administered the Wechsler Intelligence Scale for Children, Third Edition (WISC-III; Wechsler, 1991) and the Cognitive Assessment System (CAS; Naglieri & Das, 1997) to a group of ethnically diverse, previously identified special education students. They found the WISC-III over-identified African American children as mentally handicapped compared to the Caucasian children. Furthermore, they found the WISC-III identified more African American children as mentally retarded compared to the CAS. Other research with similar findings (Ebersole & Kapp, 2007; Skiba, Simmons, Ritter, Kohler, Henderson, &

Wu, 2006) indicated a direct link to the overrepresentation of African American children in programs for the mentally retarded and the strict cutoff of a 70 IQ.

Additionally, Edwards and Oakland (2006) identified the mean IQ gap between African Americans and Caucasians on the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJIII COG; McGrew & Woodcock, 2001). Although they acknowledged the gap existed in tests designed to measure the modern theory of intelligence (i.e., CHC theory), they concluded the test was not biased because similar factor structures were identified for both ethnic groups. Similarly, Kush and colleagues (2001) found a 15-point difference in mean IQ between the African American and Caucasian groups on the WISC-III; however, they concluded the WISC-III displayed good construct validity for both African American and Caucasian special education students because they found invariant latent intellectual traits. Support for the hierarchical structure of the Differential Ability Scales (DAS; Elliot, 1990) across African American and Caucasian groups also was established by determining an invariant factor structure across groups (Keith, Quirk, Schartzer, & Elliot, 1999). These researchers provide additional support for Shuey's (1966) and McGurk's (1982) conclusions that the IQ gap still exists and has not decreased. Additionally, this research appears to indicate current studies no longer focus on defining the gap, as there is well established evidence it exists, but exploring other methods of measuring test bias (i.e., construct validity).

Furthermore, significant differences in subtest scores between African Americans and Caucasians have been found in the literature. A two to three point difference on the Wechsler Adult Intelligence Scale-Revised (WAIS-R) was found for all age groups on nearly every subtest (Kaufman, McLean, & Reynolds, 1988). In other words, Caucasian

adults performed approximately one standard deviation above African Americans at the subtest level. The largest subtest differences were seen on Block Design and Vocabulary, with Vocabulary scores the lowest for African Americans. Ethnically diverse children also show similar subtest score discrepancies. Specifically, the Information, Comprehension, Arithmetic, Object Assembly, Picture Arrangement, and Block design subtests of the WISC and the WISC-R yielded differences between two to four points, with the largest difference displayed on Picture Arrangement and Object Assembly (Munford, Meyerowitz, & Munford, 1980). Furthermore, the research indicated an increasing gap in subtest performance between the revisions and researchers questioned the utility of the WISC-R with African American children. Most of the research on subtest discrepancies in ethnically diverse populations has been conducted using the Wechsler Scales. Current research at the subtest level of instruments designed to measure the broad and narrow abilities of the CHC model is lacking and should be investigated. Additionally, research is needed to study whether the African American-Caucasian IQ gap on recent versions of cognitive ability tests has improved to further aid in test interpretation and understand test bias.

Test bias can be measured in several different ways. As reported above, a test's fairness is questioned when one ethnic group consistently performs better than the other on a specific measure (Sattler, 2001). Additionally, differential construct validity has been used to determine if a test measures the same factors across ethnic groups. Specifically, bias is considered when "test's constructs or factors result in systematically different meanings across examinee subgroups" (Edwards & Oakland, 2006, pp. 358).

Current research focuses on the latter definition of bias, especially when studying the CHC factors, and ignores whether one ethnic group performs better.

The reason why differences between Caucasian and African American children's intelligence exist is a topic of debate. Two hypotheses have been most popular: the culture-only model and the hereditarian model (Rushton & Jensen, 2005). According to the cultural test bias hypothesis, the differences in intelligence in minority populations reflect no real differences in ability from the minority group, but rather with specific problems with the tests. Specifically, inappropriate standardization samples, inappropriate content, language bias, questions with predictive validity, and measurement constructs all have been suggested to explain the mean score difference (Brown et al, 1999). Conversely, Jensen (1998) argues cultural bias is not entirely to blame. He found that differences on intelligence tests between African American and Caucasian children were closely related to the test's *g* loading (Edwards & Fuller, 2005).

As tests of cognitive ability are revised and linked to intellectual theory, test authors attempt to eliminate any bias that may be associated with inappropriate standardization samples, inappropriate content, and language bias. The KABC-II was normed by matching the stratification of the U.S. population on gender, race, ethnicity, socioeconomic status, region, and special education status (Kaufman & Kaufman, 2004). Although the standardization sample reportedly matches the U.S. population more closely than previous versions of cognitive ability measures, it does not ensure the fairness of such test across the various ethnic groups. Additional research is needed to determine the performance of African Americans as compared to Caucasians on overall intelligence,

which is most closely related to the general ability factor, as well as on the specific subtests which claim to measure the narrow abilities described by the CHC model.

Preschool Assessment

Understanding childhood cognitive development, especially during the dramatically changing preschool years, continues to be extensively researched. Jean Piaget (1896-1980) described preschoolers as preoperational thinkers. He contended preschool children do not have the logical operations required in reasoning and lack the concepts of time, space, causality, and number (Nuttall, Romero & Kalesnik, 1992). Preschoolers have acquired symbolic functions, including the ability to search for hidden objects, perform delayed imitation, engage in symbolic play, and use language (Sattler, 2001). Piaget's theory aroused controversy and several theories of childhood cognitive development emerged. For example, Piagetian theory assumes mental growth is qualitative and significant differences in young and older children's thinking exists (Sattler, 2001). In contrast, psychometric theories (e.g. Spearman, Thurstone, Carroll, etc.) view mental growth on a curve suggesting that mental ability at an older age can be predicted from intelligence at a preceding age (Sattler, 2001).

As theories of intellectual development of preschool children were emerging, the need for specific tests for this population became necessary. Driven by legislation, the 1960's brought an interest in early childhood intervention. Specifically, publicly funded programs began to develop to foster positive growth during the preschool years. With these programs, namely Project Head Start, came the increased need for appropriate, multidisciplinary, early assessment practices (Ford & Dahinten, 2005). Project Head Start was founded to help prepare low-income children for successful entry into school. Head

Start contended that intellectual development would be fostered by adequate nutrition, early educational experiences, and intellectual stimulation. To evaluate the effectiveness of the program and continue the funding, assessment tools were needed to measure the growth these children had made by their entry into school. Further federal support for early assessment of children with special needs came from the Handicapped Children's Early Education Assistance Act of 1968. This law included all children from birth to age eight who were identified with disabilities. It also was the precursor to early childhood special education and the Education for All Handicapped Children Act (EHA; Ford & Dahinten, 2005).

As special education legislation began to develop, specific categories of eligibility for special education funding were spelled out. PL 94-142 of 1975 (EHA) was amended in 1990 and became the Individuals with Disabilities Education Act (IDEA). IDEA included several categories of eligibility (mental disability and specific learning disability) requiring the use of a cognitive ability measure as part of the multidisciplinary assessment (Jacob & Hartshorne, 2003). Although the 2004 reauthorization of IDEA (IDEIA 2004) does not require the use of cognitive ability measures, such information may still be helpful in determining specific strengths and weaknesses in the area of learning and will continue to be used in clinical practices (Holdnack & Weiss, 2006). Greater accountability for the identification of children in need of support, as described by the No Child Left Behind Act of 2001, further adds to the need of theoretically grounded, well-validated, and empirically supported measures of cognitive ability (Ford & Dahinten, 2005). Findings regarding the overrepresentation of minorities in special

education (Holdnack & Weiss, 2006) poses strong evidence for eliminating cultural bias in testing measures.

Empirically supported instruments for the use with preschool children are especially needed because of controversy over the use of cognitive ability measures in preschoolers (Bracken 1994). Previous versions of intelligence measures have been shown to be less technically adequate with preschool populations, thus requiring the establishment of sound psychometric properties in the latest versions (Bracken 1987, Bracken 1994). Research has consistently shown intelligence tests that assess a broad range of ages typically find fewer factors at younger levels than at school-aged levels (Carroll, 1993; Tusing, Maricle, & Ford, 2003). The traditional verbal-nonverbal dichotomy (verbal vs. symbolic) has long been held as the way to discriminate young children in terms of intelligence (Sattler, 2001; Tusing & Ford, 2004). However, research now shows more broad factors are involved in the cognitive abilities of young children (Hunt, 2007; Morgan, 2008). Research with the Woodcock-Johnson Tests of Cognitive Abilities-Revised and the Differential Ability Scales: Upper Preschool Level suggested a five factor model of cognitive abilities in preschoolers (Tusing & Ford, 2004). They suggested all eight broad abilities described by the CHC model can be assessed in preschoolers through popular cognitive ability measures used with preschoolers, but there is yet a single measure that assesses all eight abilities. Moreover, a cross-battery approach has been suggested as the best way to assess all eight constructs with preschoolers (Alfonso, Flanagan, & Radwin, 2005). More research involving the latest versions of cognitive ability measures is needed to confirm the broad ability factor structure in preschoolers, especially among different ethnic groups.

Purpose of the Present Study

Accurate assessment techniques for use with preschool aged children have become necessary when determining children's placement in special education (Ford & Dahinten, 2005). Even with the recent changes in special education law, assessment of cognitive abilities will continue to be an important tool in clinical practice as well as an available tool in a multidisciplinary approach within the schools (Ofiesh, 2006). However, with the documented mean score difference in intelligence between Caucasian and African American children (Arinoldo, 1981; Edwards & Oakland, 2005; Jensen, 1980; Smith, Duncan, & Lee, 2003; Thorndike, Hagen, & Sattler, 1986; Tuttle, 1966; Vincent, 1991), accurate interpretation of current cognitive ability measures is vital. Furthermore, evidence of a one standard deviation mean difference on subtest performance (Kaufman, McLean, & Reynolds, 1988; Munford, Meyerowitz, & Munford, 1980) questions the utility of narrow ability interpretation when evaluating children. Research is needed to determine how recent versions of preschool intelligence tests address bias at the broad and narrow ability level so psychologists can accurately interpret these tests with diverse populations.

The purpose of the present study was to determine if African American and Caucasian preschool children display similar patterns among the CHC factors on one recently revised cognitive ability instrument. Specifically, the patterns of performance of preschool children were studied among the CHC narrow and broad ability factors measured by the Kaufman Assessment Battery for Children, Second Edition (KABC-II). In addition, this study sought to determine whether African American and Caucasian

preschool children display different profiles across the CHC factors; therefore, a profile analysis was conducted. This study addressed two research questions:

1. When conducting a profile analysis of the KABC-II composite scores:
 - a. Do African American and Caucasian preschool children display the same patterns of highs and lows (Parallelism test) across the CHC broad factors?
 - b. Regardless of whether the profiles are parallel, does the Caucasian group, on average score higher (Levels test) across the CHC broad factors, as a set, compared to the African American group?
2. When conducting a profile analysis of the KABC-II subtest scores:
 - a. Do African American and Caucasian preschool children display the same patterns of highs and lows (Parallelism test) across the CHC narrow factors?
 - b. Regardless of whether the profiles are parallel, does the Caucasian group, on average score higher (Levels test) across the CHC narrow factors, as a set, compared to the African American group?

Significance of Study

The present study provided a clearer understanding of the use of the KABC-II with preschoolers. Specifically, performance profiles from two racial groups, African Americans and Caucasians, were analyzed to determine whether differences exist between the two groups on the specific CHC factors of the KABC-II. Profiles were generated at the subtest level as well as the composite level. A significant contribution of this study was to provide further information regarding the patterns of performance of

Caucasian and African American preschool children on the CHC factors. Previous research has shown the CHC model is invariant across racial groups (Keith, et al., 1999; Edwards & Oakland, 2006). Most research, however, does not specifically show how the CHC factors are displayed among different preschool racial groups. This present study was hoped to contribute to new research by exploring whether construct bias is shown in a recently revised test for preschool children. Psychologists can better understand the use of the KABC-II when assessing children within school or clinical environments (Holdnack & Weiss, 2006).

The research questions provided important information for clinicians when interpreting KABC-II results with diverse students. If preschool children perform differently at the subtest level, but show similar patterns of performance at the composite level, interpretation at the appropriate level can be determined. Kaufman and Kaufman (2004) recommended interpretation of the KABC-II at the composite level. Specifically, Kaufman and colleagues (2005) indicated the global score provides a norm-based overall view of the child's performance and can serve as a comparison point to assess other abilities, but it does not tell anything about strengths and weakness in ability. Furthermore, Kaufman and colleagues (2005) stated "scores on specific subtests are of little value" (p.79) because they are intended to complement each other and provide a thorough measure of the theoretical construct of the composite." Differences on the subtests that make up a broad ability can help a clinician generate hypotheses based on the narrow abilities described by the CHC model, and these hypotheses can be verified with other data and behavioral observations obtained in the assessment process. This

study was hoped to provide evidence for or against the recommendation of interpretation at the composite level.

Furthermore, this study was hoped to add to the body of research addressing the gap in IQ between African Americans and Caucasians. A literature review of articles published until 1980 found no decrease in the gap overtime (Shuey, 1966; McGurk, 1982). Little research has been conducted to study the IQ gap in the most recent versions of intelligence tests. The K-ABC was praised for significantly smaller differences than those seen on most other intelligence tests (Kaufman & Lichtenberger, 2002; Kaufman et.al, 2005) and promise was seen in the KABC-II (Kaufman & Kaufman, 2004). If the Levels test is not significant at either the composite or subtest level, then further evidence can be provided for a decreased gap in IQ scores between the two populations when using the KABC-II. Additionally, evidence of decreased test bias would be seen if the same pattern of highs and lows is displayed between the two groups. This study further explored the benefits of using the KABC-II with diverse populations. The different levels of performance between African Americans and Caucasians on the KABC-II were determined and added to the current research on the IQ gap.

CHAPTER II

Review of the Literature

This chapter provided a review of the literature relevant to the current study. It included a historical review of the development of intelligence testing and intellectual theory. Additionally, it addressed contemporary advances in intellectual theory and assessment, specifically the Cattell-Horn-Carroll Theory, and how it has been integrated into cognitive assessment. This study specifically focused on the assessment of preschool cognitive ability among African American and Caucasian populations; therefore, this chapter included the historical underpinnings of cognitive assessment with preschool aged African Americans. Furthermore, the review discussed the research comparing the difference between African Americans and Caucasians on measures of cognitive ability. The development of cognitive ability measures for use with preschool children and the adherence of these instruments to contemporary theory were discussed.

Historical Perspective of Assessment and Theory

“Intelligence is a fascinating and important topic,” (Cattell, 1971, p. 1) and vast amounts of research has been conducted to study cognitive functions over the last 150 years. The latter half of the 19th century proved to be a very important time in the development of cognitive ability assessment. Sir Francis Galton (1869) is regarded as the

“father of assessment” as he was the first to describe the concept of intelligence tests and their use to systematically study cognitive ability (Sattler, 2001; Wasserman & Tulskey, 2005). The study of intelligence was aided by his development of the statistical concepts of regression to the mean and correlation. In 1882, in London, Francis Galton opened his anthropometric laboratory “for the measurement in various ways of Human Form and Faculty” (Sattler, 2001, p.129). For a small fee, the public could get measurements of physical characteristics such as vision, hearing, and reaction time. These sensory and motor characteristics, Galton assumed, positively correlated with human intelligence, and higher sensory discrimination abilities meant higher intelligence. Although his assumption was not validated, it became apparent that individual differences in human ability should be considered. Galton’s assumptions may not have been accurate, but his use of physical differences and attributes to measure mental functioning remained constant in most early attempts to measure intelligence (Suen & French, 2003).

James McKeen Cattell was influenced by Galton’s research and left Wilhelm Wundt’s laboratory in Germany to work in Galton’s London laboratory. Unlike Cattell, Wundt believed psychology should focus on introspection rather than measurement of individual differences (Sattler, 2001). Similar to Galton, Cattell felt mental ability could be assessed through measurement of sensory characteristics. Cattell returned to the United States and established a psychological laboratory at the University of Pennsylvania to further study mental ability based on his experiences in Wundt’s and Galton’s laboratories. Consequently, he developed a testing battery that included ten tests measuring sensory discrimination, strength, motor speed, perceptual judgment, reaction time, color naming, and immediate memory (Wasserman & Tulskey, 2005). His battery

was described in his paper “Mental Tests and Measurements” (1890), which gave him credit as the first to coin the term “mental tests.”

As Americans were experimenting with mental measurement of physical and sensory attributes, Alfred Binet was measuring intelligence from a different perspective. He believed one needed to look at higher mental processes, such as memory, to be more accurate at mental measurement. He developed new tasks which required children to look at figures for a few seconds and then draw them from memory, add numbers together, and read and copy sentences (Suen & French, 2003). Along with Victor Henri and Theodore Simon, two men who shared his ideas of mental functioning, Binet developed the 1905 Binet-Simon Scale as a way to identify school-aged children who had mental retardation. This test appeared to be the first to allude to intelligence in terms of a hierarchical structure with three levels consisting of a higher order factor, lower order factors, and first order factors. These cognitive factors included memory, imagination, attention, comprehension, coordination skills and quick visual judgment. His scale measured these cognitive functions as they worked together, rather than distinguishing between separate abilities as modern day tests attempt to accomplish. Binet’s famous test was developed into an English version by American psychologist Lewis Terman (Suen & French, 2003).

Testing intelligence in America became important during World War I. The Army Alpha and Beta tests were administered to army recruits to gather a quick assessment of their intelligence (Kaufman, 2000). Robert Yerkes led the effort of understanding the data collected by these tests. Several important consequences that impacted the field of intelligence testing came from the Army testing efforts. Specifically, IQ tests like Binet’s

scale were used with children, but the Army testing showed IQ tests were also very useful for adults. Additionally, the measurement of problem solving joined verbal assessments of intelligence. The Army testing effort provided massive data samples which accounted for the validity of IQ tests to help differentiate populations (e.g., officers from recruit). Lastly, IQ tests were found to be important in the identification of giftedness, not just “feeble-mindedness” and proper interpretation of the data was essential (Kaufman, 2000).

David Wechsler is considered the link between the first attempt of mass IQ data collection during WWI and modern assessment tests and practices (Kaufman, 2000). Highly influenced by Terman’s version of the Stanford-Binet, Wechsler accepted the idea of global intelligence, but felt it could be measured in different ways, specifically verbal and nonverbal ability (Reynolds & Kaufman, 1985). He insisted that all ages should be administered the same tasks and borrowed tests from the Stanford-Binet and Army Alpha tests to develop his verbal scale and from the Army Beta and Army Performance Scale Examination for his performance (nonverbal) scale (Kaufman, 2000). He published the Wechsler-Bellevue Scale in 1939 consisting of 11 subtests that had “tryout data” on individuals with known intelligence levels. Many of these borrowed tasks can be seen on current versions of Wechsler’s test (i.e., Vocabulary, Digit Span, and Comprehension). Since most of these subtests were developed prior to the 1920s, they were developed devoid of theoretical models (Reynolds & Kaufman, 1985).

The Wechsler-Bellevue Scale soon became the most popular intelligence test, surpassing the Stanford-Binet (Wasserman & Tulskey, 2005). Wechsler’s popularity was based upon the integration of the verbal and performance scales in one test, the use of a normative sample, the use of tests that were commonly seen in practice, and the emphasis

on psychometrics. The decades following the introduction of the Wechsler-Bellevue Scale included more advances in psychological assessment for the different age ranges to make assessment practical and align clinical practice with psychometric rigor (Wasserman & Tulskey, 2005). The Wechsler Intelligence Scale for Children (WISC; Wechsler, 1949), the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) and the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967) have been developed and revised to become the most widely used tests of intelligence (Zhu & Weiss, 2005). Additionally, the Wechsler Scales are the most researched instruments and have had a great impact on the field of assessment and clinical practice (Sattler, 2001).

Intelligence Theory

Spearman was the first to use psychometric data to develop a theory of intelligence (Brody, 2000). He recognized people who tended to perform well on one of the early sensory-discrimination ability measures, also tended to perform well on another; thus, their scores were positively correlated. Spearman assumed there was a common ability factor that accounted for the positive correlation in all of the areas assessed. This general factor, g , is a fundamental characteristic of intelligence that has remained constant throughout the continued study of intelligence. The g factor “is a mathematically derived general factor, stemming from the shared variance that saturates batteries of cognitive/intelligence tests” (Wasserman & Tulskey, 2005, p. 16). His ideas lead to a dichotomous understanding of intelligence as Spearman assumed a second specific source of variance also must exist (Brody, 2000). Although Spearman’s theory was known as the two-factor theory of intelligence, he focused on the study of g .

Factor analysis, developed by Spearman, determined how much of the general intelligence factor was measured by tests (e.g., subtests) of various mental ability. The amount the test correlated with *g*, or loaded on *g*, was categorized into low or high loadings. By comparing these low and high loadings, Spearman concluded tests requiring reasoning to solve novel problems better reflected *g* than those that measured acquired knowledge (Jenson, 1998). Spearman believed the *g* factor encompassed three different mental processes that lead to the production of new knowledge from sensory or cognitive experience. These three neologisms, as he called them, included: to apprehend new experiences and stimuli; to make relations between two or more lived experiences; and to make associations between new stimuli and related experiences (Carroll, 1993; Jensen, 1998).

As Spearman focused on *g*, Louis Thurstone believed individual intelligence could not be explained as a unitary trait (Sattler, 2001). Consequently, he developed a method of multiple factor analysis that studied the independent variables accounting for the variance within the correlations among factors (Brody, 2000). Obtaining test scores from university students, Thurstone utilized his method of factor analysis to obtain what he termed *primary mental abilities*: spatial visualization, perceptual speed, numerical facility, verbal comprehension, associative memory, word fluency, and reasoning (Wasserman & Tulskey, 2005). His original study provided initial evidence of a multiple factor structure for intelligence and did not find evidence of Spearman's *g*. Stemming from Thurstone's research was the idea individual intelligence could be described as a profile of strengths and weaknesses (Brody, 2000), an idea that persists with present day intelligence test interpretation. Although Thurstone's initial denial of the existence of the

g factor was at complete odds with Spearman and fueled growing debate, Thurstone's later research with higher-order factor analysis allowed him to admit the existence of *g* (Carroll, 1993; Wasserman & Tulsy, 2005).

Contemporary Influences on Understanding Intelligence

Although Spearman helped initiate the development of intellectual theory, many began to believe intelligence was more complicated (Carroll, 1993; Cattell, 1941; Horn, 1968). In 1950, Vernon, and his colleague Burt, became the first to suggest a hierarchical structure of intelligence (Guilford, 1985). The top of the hierarchy consisted of Spearman's *g*, and immediately below *g* were two broad abilities: verbal-educational *v:ed*, and spatial-practical, *k:m*. At the bottom of Vernon's hierarchy were verbal and numerical abilities under *v:ed*, and spatial and mechanical abilities under *k:m*. Contemporary models of intelligence (i.e., CHC theory and those leading to its development) reflect a similar hierarchical structure.

Cattell-Horn Gf-Gc Theory

Stemming from the work of Thurstone, Raymond Cattell (1963) and John Horn (1976) argued that Spearman's *g* was overly simplistic and lacked the ability to adequately understand the underlying constructs of intelligence (Wasserman & Tulsy, 2005). Cattell proposed that *g* was made of two distinct factors, fluid intelligence and crystallized intelligence. Horn (1976) and Cattell (1963) defined *fluid ability* as a facility in reasoning, where crystallized abilities are of little use and where adaptation to new environments is required. *Crystallized ability*, in contrast, is the accessible store of acquired knowledge, and the ability to acquire further knowledge through learning.

The *Gf-Gc* Theory was expanded when further research (Carroll & Horn, 1981) suggested the existence of eight broad abilities (McGrew, 2005). The original factors did not appear to adequately account for visual, auditory, and basic memory abilities (Horn & Blankson, 2005). Identified through factor analysis, these second-order abilities were called acculturation knowledge (*Gc*), fluid reasoning (*Gf*), short-term memory and working memory (*Gsm*), long-term memory (*Glm*), processing speed (*Gs*), visual processing (*Gv*), auditory processing (*Ga*), and quantitative knowledge (*Gq*). Cattell and Horn's theory did not account for *g* as an underlying source of individual differences in intelligence (Davidson & Downing, 2000). Horn and Blankson (2005) further indicate human intelligence is a combination of many abilities that are interrelated, and if one central factor exists, its influence is weak.

Carroll's Three Stratum Theory

The Three Stratum Theory was developed as a result of a survey over 60-70 years of factor analytic research (Carroll, 1993; Carroll 2005). Carroll's work was considered the most comprehensive, empirically-based, factor analytic research utilizing 460 data sets (McGrew, 2005). This theory purported that the totality of human intelligence could be classified and structured within one of three stratum of abilities: the third, or highest stratum contains the general factor, *g*; the second stratum abilities are eight broad abilities including fluid intelligence, crystallized intelligence, general memory and learning, broad visual perception, broad auditory perception, broad retrieval capacity, broad cognitive speediness, and processing speed (decision speed); and the narrow, or first-factor abilities, which are more specific abilities grouped under each broad ability. Admitting that many abilities may be difficult to assign to a specific stratum, Carroll emphasized

that the stratum were not rigidly defined (Davidson & Downing, 2000). Carroll's model and the *Gf-Gc* Theory differ primarily in the existence of *g*.

Contemporary Cattell-Horn-Carroll (CHC) Theory

The late 1990s marked an important time in the development of the Cattell-Horn-Carroll Theory. Kevin McGrew (1997) recognized the similarities between Cattell and Horn's *Gf-Gc* theory and Carroll's Three Stratum Theory. According to McGrew (1997), a lack of unity in intelligence theory existed and necessitated the development of a single model for assessment to follow. The similarities between the theories included the broad and narrow abilities (McGrew & Flanagan, 1998). McGrew thought the theories needed integration for the accurate study of the broad and narrow abilities (McGrew, 2005); specifically, in order for tests to accurately measure these Stratum II and I abilities, a single taxonomy was needed. According to McGrew (2005), Horn and Carroll informally agreed to include their theories under the Cattell-Horn-Carroll (CHC) Theory terminology in 1999.

Although recognized as an integrated theory in 1999, differences between Cattell and Horn's *Gf-Gc* model and Carroll's Three Stratum model existed (McGrew & Flanagan, 1998). First, the highest stratum in the Three Stratum Theory included the general factor. Horn and Cattell's model, however, did not recognize *g* because studies conducted by Horn (1991) indicated perfect loading of *Gf* on *g* (Bickley, Keith, & Wolfle, 1995). Horn (1991) concluded the loading indicated intelligence was much more than the general factor. Secondly, some abilities considered broad abilities in the *Gf-Gc* model (e.g., quantitative knowledge) were considered narrow abilities in the Three

Stratum Theory. Lastly, memory was treated differently in the theories with short-term and working memory separated by Cattell and Horn but not by Carroll (McGrew, 2005).

CHC Broad Abilities. The following section defines the current broad abilities according to recent publications by Kevin McGrew (2005).

Fluid Intelligence (Gf). The *Gf* broad ability includes deliberate and controlled mental operations used to solve novel problems. Inductive and deductive reasoning are considered the hallmark of fluid reasoning. Analytic ability is emphasized in fluid reasoning (Cattell & Horn, 1978), and mental operation includes classification, drawing inferences, problem solving, comprehending implications, generating and testing hypotheses, concept formation, identifying relations and transforming information.

Crystallized Intelligence (Gc). Crystallized intelligence is associated with knowledge gained overtime, or a person's "achievement in that it is the accumulated knowledge of an individual" (Cattell & Horn, 1978, p. 140). *Gc* is an individual's bank of knowledge and is highly influenced by culture. It is further defined as a person's knowledge of language, and information and concepts of a specific culture including the application of that knowledge. Crystallized knowledge is acquired by using other mental processes in formal and informal educational and general life experiences.

Quantitative Knowledge (Gq). *Gq* is defined as an individual's wealth of acquired store of declarative and procedural quantitative knowledge. In other words, *Gq* represents an individual's acquired bank of mathematical knowledge. Like *Gc*, quantitative knowledge is acquired by utilizing other cognitive processes as well as educational experiences. McGrew cautions against incorrectly defining *Gq* as reasoning with the

above acquired knowledge, an ability accounted for in the narrow ability *RQ* (quantitative reasoning.)

Reading/Writing Ability (Grw). *Grw* is defined as an individual's declarative and procedural reading and writing skills. It includes basic skills such as reading and spelling single words, as well as the ability to read and write complex connected discourse.

Short-term Memory (Gsm). Short-term memory consists of an individual's ability to "apprehend and maintain awareness of elements and information in the immediate situation" (McGrew, 2005, p. 153). Time frame is limited to within the last minute. This memory system loses information quickly unless the person utilizes other mental processes to maintain it in immediate awareness.

Visual-Spatial Abilities (Gv). *Gv* includes an individual's ability to generate, store, retrieve, and transform visual images. It is the broad ability that allows one to complete visual-spatial tasks.

Auditory Processing (Ga). *Ga* depends on the input of sounds and on an individual's hearing ability. It is described as the extent to which one can control the perception of auditory information such as discriminating and analyzing sounds in a pattern or group of sounds

Long-term Storage and Retrieval (Glr). Long-term storage and retrieval is described as the broad ability to consolidate and store new information in long-term memory and to retrieve that information fluently and efficiently through association

Processing Speed (Gs). *Gs* is defined as the ability to automatically and fluently complete relatively easy cognitive tasks, especially when high attention and focused concentration are necessary.

Decision/Reaction Time or Speed (Gt). *Gt* is the ability to react and/or make decisions quickly in response to stimuli, typically measured by chronometric measures of reaction and inspection time.

Empirical Support for the CHC Model

Over the past few decades, many studies have been conducted to validate the CHC model and provide support for its structure. Confirmatory factor analyses were conducted (Bickley et al., 1995; Taub & McGrew, 2004; Reynolds, Keith, Fine, Fisher, & Low, 2007) to confirm the hierarchical structure of the CHC model across various age groups. These studies found support for the stability of the general factor as well as seven to eight broad abilities across the life span. Specifically, Bickley and colleagues (1995) studied eight different age groups ranging from 6-79 and found that the factor structure was invariant across groups. Additionally, the broad abilities described by the Three Stratum Theory had similar loadings on *g* across the life span. Furthermore, Taub and McGrew (2004) provided more evidence for the invariant factor structure of the CHC model across the life span using the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJ-III COG; McGrew & Woodcock, 2001). Tusing and Ford (2004) expanded the research on the CHC model across the lifespan by studying preschoolers ages four and five. They found five broad ability factors were reliably identified as distinct providing evidence that preschool cognitive ability can be interpreted through CHC theory. Research on the KABC-II (Reynolds, et al., 2007) indicated the instrument measures the same constructs across the ages (3 to 18), and the factor structure is closely aligned with the CHC model. Additionally, the theories included in the CHC model have

been validated across gender and ethnic groups (Carroll, 1993; Flanagan & McGrew, 1998; Horn & Noll, 1997).

Contemporary Assessment Aligned with Theory

Kaufman (1979) was one of the first to argue subtests should be organized into clusters that conformed with theory. The Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983) defined intelligence as mental processing, an idea grounded in neuropsychological theory (Kaufman, Kamphaus, and Kaufman, 1985). Although it was the first to align with a specific theory, it was not until its second revision that the Kaufman Battery Assessment for Children, Second Edition (KABC-II; Kaufman & Kaufman, 2004) utilized the CHC model as the primary method of interpretation.

Today, tests of cognitive ability have been designed to more closely reflect theories of intelligence, specifically the CHC model. As this alignment occurs, interpretation becomes more accurate (Kamphaus, Winsor, Rowe, & Kim, 2005). Prior to 1989, no individually administered test of cognitive ability reflected the findings of Horn, Carroll, and Cattell (McGrew, 2005). In an attempt to bridge theory and assessment, McGrew (1997, 2005) served as a consultant in the development of the Woodcock-Johnson-Revised (Woodcock & Johnson, 1989) along with John Horn and John Carroll (Taub & McGrew, 2004). Through data analysis conducted by Carroll on the WJ correlation matrices, Woodcock and Johnson were able to determine which existing subtests adequately measured *Gf-Gc* factors, and which factors needed subtests developed. These efforts resulted in the first individually administered and nationally normed assessment battery that adhered closely to a contemporary psychometric theory

(McGrew, 2005). Currently, the WJIII COG, the Stanford-Binet, Fifth Edition (SB-V; Roid, 2003), and the KABC-II are popular tests of intelligence used with children that are based upon CHC theory. The popular Wechsler Scales were not designed to adhere to this particular theory but recent versions have made attempts to incorporate subtests that are better measures of the CHC broad abilities (Zhu & Weiss, 2005). All of the recent tests based on the CHC model lack the assessment of all broad abilities (Alfonso et al., 2005).

Racial Differences in Intelligence Testing

Cultural differences in the measurement of IQ have dated back to the time of Alfred Binet. His 1905 scale had become a popular way to measure mental abilities without the use of physical and sensory skills, and was being used by researchers across several countries. A study done in Belgium showed significantly higher scores on the scale by Belgium children compared to those tested by Binet in France. After an examination of the results, he discovered the sample was taken from a higher social class. He recognized normal children from an enriched background performed significantly better than similar ability children from impoverished homes on his scales (Hynd & Semrud-Clikeman, 1993).

Documented differences in IQ between African Americans and Caucasians date back to the use of wide-spread mental testing during World War I (Rushton & Jensen, 2005). The Army Alpha and Beta tests were administered to soldiers to test the IQ of army recruits. These tests were administered to over 1.7 million men, and the administration was considered a “grand opportunity” to systematically study the general intelligence of American men. The project was headed by Harvard Psychologist Robert

Yerkes and used an instrument modeled after Lewis Terman's modified version of Binet's intelligence test. Controversy erupted as the data lead to ideas of racism and inferiority, and the social implications of IQ tests were raised (Kaufman, 2000). Results from this massive testing indicated the mean mental age for whites was 13.08, which Yerkes concluded was pulled down by immigrant whites. The mean mental age for native blacks was 10.41, suggesting whites were intellectually superior to blacks (Rury, 1988).

The Mean IQ Score Discrepancy

By the ages of three and four, race and social class differences in IQ emerge and remain stable throughout the lifespan (Sattler, 2001). Caucasians score 1.2 standard deviations higher, on average, than African Americans on nationally standardized tests of cognitive ability. This mean difference increases slightly from childhood to maturity (Jensen, 1998). Jensen reported most standardized tests of intelligence used in the United States have a mean of 100 and a standard deviation of 15; therefore, the mean IQ found throughout the literature for African Americans is approximately 85. Additionally, the standard deviation has been found to be approximately 12 for African Americans (Jensen, 1998). This mean difference moves the normal curve for the African American population causing implications when thresholds for IQ are set in place for certain categories, especially mental retardation. Although the majority of the curves overlap, the implications still exist. He explains that the percentages of African Americans and Caucasians with IQs below seventy are 15.9% and 2.3%, respectively.

The gap in cognitive ability between African Americans and Caucasians is displayed throughout the literature. Research using the Wechsler Scales found a gap that ranged from 17 to 25 points (Arinoldo, 1981; Tuttle, 1966). Preschool children

administered the WPPSI displayed a smaller, but significant gap in overall, verbal, and performance IQs (10.8, 10.5, and 8.9 respectively). Additionally, significant differences in subtest scores between African Americans and Caucasians have been found. Ethnically diverse children displayed subtest score discrepancies on several Wechsler subtests from the WISC and WISC-R (Munford, Meyerowitz, & Munford, 1980). Specifically, Caucasian children preformed approximately one standard deviation, or 2-4 points, higher than African American children on the Information, Comprehension, Arithmetic, Object Assembly, Picture Arrangement, and Block Design subtests.

Although research on contemporary measures of intelligence focuses on different aspects of test bias, mean differences persist. For example, Edwards and Oakland (2006) studied the factor structure of the CHC theory in ethnically diverse populations. They identified the mean IQ gap between African American and Caucasian children on the WJIII COG; however; they concluded the test was not biased because similar factor structures were identified for both ethnic groups. Similarly, Kush and colleagues (2001) found a 15-point difference in mean IQ between the African American and Caucasian groups on the WISC-III. They concluded the WISC-III displayed good construct validity for both African American and Caucasian special education students because they found invariant latent intellectual traits. These studies indicate a shift in the focus of research with ethnically diverse populations. Research on contemporary cognitive ability measures focuses on factor structure rather than the existence of mean score differences at either the composite or subtest level.

Implications of an IQ Gap

Accurate and unbiased assessment of all ethnic groups is vital as assessment is usually considered the first part of the intervention process (Nagel, 2007). Information gathered from the assessment process is examined and leads to program planning. Further assessment takes place to document the child's progress while the intervention takes place. Inaccurate evaluation of a child's progress may occur if assessment methods are found to be biased towards a particular ethnic group. Moreover, Jensen's (1998) work highlighted the importance of intelligence test interpretation as related to intervention placement. For example, special education placement would be made for a child with an IQ below 70 and special accommodations would be made. An IQ of 70 places an individual two standard deviations below the mean of 100; however, when African Americans historically show a mean around 85 for various IQ tests, an IQ of 70 would only be one standard deviation below the mean. Questions were raised as to the accuracy of a cut off of 70 for mental retardation within the African American community and other considerations should be made (Jensen, 1998). Specifically, Jensen (1998) suggested adaptive skills must be used in conjunction with a low IQ.

These questions are particularly important considering the overrepresentation of African Americans in special education across the country (Skiba et al., 2006). Research indicated a direct link to the overrepresentation of African American children in programs for the mentally retarded and the strict cutoff of a 70 IQ (Ebersole & Kapp, 2007). Jensen's (1998) research indicated African American children in a special education class displayed higher ability in social and motor adaptive skills. The Caucasian children, on the other hand, displayed under-developed adaptive skills. Jensen

questioned the accuracy of the mental retardation label for the African American children although their cognitive ability scores were similar to the Caucasian children.

Nature of the IQ Gap

The two most researched positions regarding the reason for the African American-Caucasian IQ gap are the hereditarian position and the culture-only position (Rushton & Jensen, 2005). The hereditarian position attests that individual and group differences in behavioral traits have a substantial genetic component, approximately 50%. The culture-only view states that if all environments could be equalized, the group differences in IQ would disappear. In other words, the hereditarian view is 50% genetic, 50% environmental, and the culture-only view is 0% genetic and 100% environmental. Ruston and Jensen (2005) indicated the defining difference to these views was whether there was any significant genetic component to the difference. Furthermore, the hereditarian model does not state that heritability is the final determination of intelligence because, as a phenotype, it is expressed within an environment (Sattler, 1992). Environmental factors that promote or restrict intellectual development must be accounted for when studying the nature of intelligence.

The controversial book *The Bell Curve* (Herrnstein & Murray, 1994) asserted support for the hereditarian position of the IQ gap. Its readership was not limited to the research community, but expanded to the general public causing heated discussions about race and IQ. Controversy stemmed from the book's assertion that social injustice did not solely account for economic and social status differences in America. Links between IQ and race and IQ and income were made, and the authors argue that people with high intelligence would be in higher paying jobs as those jobs required more intelligent

personnel. Furthermore, these jobs were usually those that provided higher salaries. The authors described “fundamental individual differences” rather than the economy as the reason for many societal differences. Furthermore, Herrnstein and Murray made public policy recommendations that appeared to go against government efforts of equality (e.g., scaling back affirmative action).

The concept of intelligence as genetic originated with Galton’s ideas that the laws of genetics apply to the mind and behavior instead of just the body (Loehlin, Lindzey, & Spuhler, 1975). In his book *Hereditary Genius* (1869), he ranked the worth of different races according to the number of documented geniuses. Athenian Greeks ranked above Anglo Saxons who ranked above Negros. Although he did not account for the lack of records for different races during this time, people were willing to use these records as evidence that Negros were less intelligent than other races (Loehlin, Lindzey, & Spuhler, 1975). Galton’s ideas were based on a self-devised ability criterion called eminence. Eminence was determined by such achievements that would warrant encyclopedia articles, biographies, and other published work on a particular individual. Through Galton’s review of documented geniuses, he noted that a person was more likely to obtain eminence if their relatives and ancestors obtained eminence. Moreover, he found that the probability of eminence decreased if the relationships to the previous geniuses were more remote (Jensen, 1998). These ideas introduced genetics to the field of psychology.

Arthur Jensen has had a primary role in the debate over the reason for the documented difference in intelligence. In an article published in the *Harvard Educational Review*, Jensen (1969) drew many conclusions about IQ. Specifically, he concluded IQ tests measured general ability, *g*, which had high heritability and reported educational

programs designed to raise IQ and achievement were ineffective. His report drew controversy when he concluded the mean African American-Caucasian difference in IQ probably had a genetic component. Jensen (1998) asserted that the difference lied in the *g* loading of the specific tests. He utilized twin studies, physiological tests (e.g., reaction time, nerve conduction velocity, etc.), and review of scores from inbreeding (cousin-marriages) to show that the test's *g*-loading was the best predictor of heritability coefficients. Additionally, the mean African American-Caucasian IQ gap was more pronounced on high-*g*-loaded tests than on low-*g*-loaded tests, suggesting the difference was not attributable to the unique cultural peculiarities of the test (Jensen, 1998; Rushton & Jensen, 2005).

While Spearman conducted his original research to discover the existence of *g*, he was the first to postulate that the African American-Caucasian IQ difference would be most pronounced on tests that highly loaded on *g* (Spearman, 1927). Jensen (1980) developed the method of correlated vectors to test Spearman's hypothesis. His method correlated the standardized African American-Caucasian mean difference on cognitive tests to their respective *g*-loadings (Rushton & Jensen, 2005). Using a large sample from 17 different studies, he found that the *g*-loadings consistently predicted the magnitude of the mean African American-Caucasian IQ difference (Jensen, 1998). Jensen (1998) concluded research studying environmental variables could not account for the complete variance seen within the African American-Caucasian IQ gap; therefore, the remaining variance was attributable to genetic factors.

Jensen's method of correlated vectors relied on the assumption that the *g*-factor was the same for African Americans and Caucasians. Research conducted using the

standardization sample of the Kaufman Assessment Battery for Children (K-ABC) found that the first principle component, g , accounts for the same percentage of variance (averaging 58 percent) among subtest scores for various ages groups and for African American and Caucasian children (Fan, Willson, & Reynolds, 1995). Additionally, the g congruence coefficient for African American children was .991 and .998 for Caucasian children. Similarly, research conducted on military samples concluded similar factor structures were seen in Caucasian and African American service men and woman, with the g factor accounting for the largest amount of variance in both groups (Carretta & Ree, 1995). They further found that the g -loadings differed little between the two groups.

The heritability factor of intelligence also has been controversial within the research field. Traditionally, the heritability of intelligence was only studied on Caucasian populations and has been reported as high as .75 (Herrnstein & Murry, 1994); however, most studies show a heritability coefficient of approximately .5 (Suzuki & Valencia, 1997). Research utilizing both African American and Caucasian samples is limited, but provides some evidence of similar heritability coefficients between the groups. Twin studies conducted in 1980 concluded that African Americans and Caucasians had similar heritability patterns, with 50-75% of the variance for both groups due to genetic factors (Osborne, 1980).

Environmental factors have also been associated with intelligence. Socioeconomic status has been linked to African American-Caucasian IQ differences (Brooks-Gunn, Klebanov, & Duncan, 1996). African American children are more likely than Caucasian children to be poor in the United States; specifically, three times as many black children than white children live in families below the U.S. poverty level (Bane &

Ellwood, 1989). When socioeconomic status was studied, poverty level accounted for more variance than maternal education and household structure combined. These two factors, along with family resources, may not be differentiated from poverty because they were associated with poverty. The authors (Brooks-Gunn, Klebanov, & Duncan, 1996), however, cautioned that their study should not be provided as evidence against a genetic basis of the IQ difference; maternal characteristics and home environment may be related to a combination of genetics, environment, and the interaction between the two. The amount poverty accounts for the variance in IQ test scores, however, cannot be ignored. Furthermore, paternal education has been linked to differences in IQ scores. Lower levels of parental education level corresponded with lower performance of their preschool children on the verbal, nonverbal, and overall IQs of the WPPSI-R (Sellers, Burns & Guyrke, 2002). Specifically, these researchers found parent education had the largest association with the three IQs as compared to gender, age, ethnicity, and geographic region.

Research conducted on the WPPSI and the Stanford-Binet Intelligence Scale, Form L-M examined the possible reasons for the African American-Caucasian IQ difference in preschool children (Smith, Duncan, & Lee, 2003). The authors examined how much the gap lessened as economic variables such as poverty and parent education were added. Additionally, maternal and family structure variables, and home variables such as stimulating environment were included. The 19-point gap was reduced when economic variables were added, but no further reduction was found when family structure variables were added. Home variables further reduced the gap by 4-points. Controlling for economic status and stimulations within the home environment reduced

the gap between African American and Caucasian preschool children, but did not eliminate it. This current research provided some evidence for cultural factors influencing IQ. Additionally, they found that cultural variables could not account for all the variation in scores.

Additionally, Scarr and Weinberg (1976) utilized a sample of black and biracial children adopted by white families to study the impact of environment on IQ. They found that African American and biracial children adopted into Caucasian families scored, on average, one standard deviation higher than African American children reared in their biological families. The researchers concluded IQ could be significantly shaped by the environment.

A literature review of all of the research articles published from 1966 to 1980 involving research conducted with African American preschool children concluded the Stanford Binet was the most popular instrument to measure IQ, followed by the PPVT and then the WPPSI (Shuey, 1966; McGurk, 1982). Additionally, McGurk concluded that all the research conducted in that period showed no evidence that African American and Caucasian preschool and school-aged children, as well as adults, were closing the IQ gap. Additionally, there was no support that the difference in IQ for the preschool population was due to environmental, cultural, or economic factors.

Alternative theories of intelligence have been proposed to account for racial differences. Fagan (2000) suggests intelligence should be defined by one's ability to process information, rather than the amount of knowledge a person possesses. In his theory, processing is measured by a person's performance on elementary cognitive tasks, and processing information leads to knowledge. Culture provides the information, and

therefore, what one knows depends on how one processes and on what their culture teaches them. Defining intelligence as processing, Fagan contends, eliminates the loading of culture that appears when IQ tests are used, which can be thought of as culture fair.

Based on the assumption that African Americans perform poorer on IQ tests than European Americans, test developers have attempted to create tests that were “culture fair.” All of these attempts have failed to create a measure that assessed reasoning rather than knowledge of culture (Sattler, 2001). For example, the Black Intelligence Test of Cultural Homogeneity (BITCH) was based on African American slang used throughout the United States. African American slang, however, differed depending on the region of the United States making the tool useless for universal use. Studies have shown little validity, and the BITCH has been seen as useless in the assessment of African American intelligence (Long & Anthony, 1974; Matarazzo & Wiens, 1977).

Contemporary Tests of Cognitive Ability and Race

Ethnicity differences seen in the K-ABC were significantly smaller than differences seen on most other intelligence tests (Kaufman & Lichtenberger, 2002; Kaufman et al., 2005). When the KABC-II was standardized, this smaller IQ gap between ethnic minorities and Caucasians needed to be replicated. The KABC-II manual (Kaufman & Kaufman, 2004) stated reduced ethnic differences were seen in this newer version. Unadjusted means for ages 3 to 6 were approximately 96 ($SD = 13.3$) for African Americans and 103 ($SD = 14.6$) for Caucasians; ages 7-18 yielded means of approximately 94 ($SD = 13.5$) for African Americans and 103 ($SD = 14.3$) for Caucasians.

Similar to the standardization sample of the KABC-II, the standardization sample for the WJIII COG was stratified for race according to the US Census data available (McGrew & Woodcock, 2001). During the development of the WJIII COG, McGrew and Woodcock (2001) had experts review the individual items and identify items that would possibly contain bias; items were modified in consideration of the recommendations. The WJIII COG manual reports the tests load on the same factors for both white and non-white groups (McGrew & Woodcock, 2001). The non-white group, however, was not broken down into different ethnicities making specific data for Caucasians and African Americans unavailable. The manual supported the idea that the WJIII COG measures the same factors in both white and non-white samples, showing consistency with Carroll's (1993) assertion that the model is invariant across race and age.

Similarly, the SB-V utilized a standardization sample stratified for race (Roid, 2003). During the piloting of the SB-V, differential item functioning (DIF) was studied. DIF refers to the differences in the item's functioning after groups are matched on ability or the attribute the item supposedly measures (Roid, 2003). Five items were deleted from the final revision of the SB-V, four of which differed between African Americans and Caucasians. Additionally, the SB-V manual provided evidence of construct validity between all groups except for Hispanics and Caucasians. Tests for construct validity were not significant between Caucasians and African Americans.

Measuring Bias in Testing

Besides mean differences, bias can be measured by looking at whether a test is an equally good predictor of two or more minority groups (Sattler, 2001). Research conducted on previous versions of intelligences tests has indicated equally good

prediction of school achievement for African American and Caucasian children.

Specifically, the Stanford-Binet Intelligence Scale did not differ in predicting academic achievement as measured by the Wide Range Achievement Test (Bossard, Reynolds, & Gutkin, 1980). Overall, research (Bossard, Reynolds, & Gutkin, 1980; Poteat, Wuensch, & Gregg, 1988) has supported the conclusion that tests of cognitive ability are equally good at predicting the achievement of European Americans and African Americans (Sattler, 2001).

Test bias also can be evaluated through differential construct validity. This method investigates if intelligence tests measure similar abilities in culturally different groups (Sattler, 2001). Moreover, bias is considered when “test’s constructs or factors result in systematically different meanings across examinee subgroups” (Edwards & Oakland, 2006, pp. 358). This method of measuring test bias can be done by comparing the factor structure of the different groups. Research conducted with the WISC-III found that the FSIQ and verbal and performance factors were found in both African American and Caucasian groups (Kush & Watkins, 1997). Similarly, the broad abilities described in the CHC theory of intelligence were found to be invariant across both ethnic groups (Edwards & Oakland, 2006). Findings from research indicated intelligence tests measure similar constructs among African Americans and Caucasians.

In dealing with bias in testing, people have attempted to make the case that one can judge the content of a test item and determine if the content is biased to a specific group (Sattler, 2001). This can be measured by looking at item performance statistics, measuring percent correct between two groups on a specific item, and rank order of the percent passing an item (Jensen, 1974). Investigations into the Stanford-Binet: Form L-

M (Terman & Merrill, 1972) and the WISC-R have shown no evidence that differences in performance between African Americans and Caucasians can be due to cultural bias (Jensen, 1974; Sandoval, Zimmerman, & Woo-Sam, 1983). Attempts have been made to study how accurate judges were in selecting items that may be more difficult for diverse students.

In summary, traditional standardized tests of intelligence show similar patterns of internal item consistency and predictive validity for all groups indicating the tests do not seem to be biased against African Americans (Jensen, 1980; Neisser et al., 1996). Furthermore, IQ tests are considered valid measures of racial differences (Jensen, 1980; Neisser et al., 1996).

The debate over the reason to the IQ gap remains unsolved. Brody (1992) summarized what the research showed regarding the discrepancies seen in cognitive ability between African Americans and Caucasians. He indicated the difference was not attributable to bias within the tests and the differences reflected more general, abstract knowledge. Additionally, Brody (1992) stated the IQ gap was not attributable to a distinct experience of African Americans, and there was no definitive evidence of a direct or indirect relationship to genetics. He further wondered whether research could actually discover a definitive reason for the gap and whether or not the gap could be eliminated. If it could not be eliminated, he questioned whether an environment could be designed where individual differences in intelligence were alleviated so that they did not determine, to the extent they do now, the difference in performance in school and other social contexts. Additionally, Sattler (2001) contended that one cannot make a genetic inference for the difference in tests scores when there may be relevant systematic

differences between races in their cultural and psychological environment. Since no one has been able to succeed in estimating or eliminating the influence of environment on cognitive processes, a definitive conclusion cannot be made.

History of Preschool Assessment

Although the mental testing movement can be traced back approximately 200 years ago, preschool assessment is relatively young. During the early part of the twentieth century, laws were established in Europe and America requiring school aged children to attend school. Prior to the 20th century, few citizens attended school. In 1870, a small number of secondary schools had an approximate enrollment of 80,000 children. Growth within the American school system spiked and by 1910, 900,000 children attended school, and by 1922 nearly 2 million were in attendance (Suen & French, 2003). With the rapid influx of many children into American schools, questions of the appropriate classification were raised. Specifically, “feebleminded” children and normal children in school needed to be distinguished and segregated to give appropriate instruction for their individual needs (Goodenough & Mauer, 1942). Thus, the need for appropriate tests for school aged children was introduced in the early 1900s.

Beginning with the child study movement in the early 1900’s, researchers searched for a comprehensive understanding of both normally- and atypically-developing children (Ford & Dahinten, 2005). Jean Piaget (1896-1980) described preschoolers as preoperational thinkers. He contended preschool children did not have the logical operations required in reasoning, or the concepts of time space, causality, and number (Nuttall, Romero & Kalesnik, 1992). Preschoolers have acquired symbolic functions including the ability to search for hidden objects, perform delayed imitation, engage in

symbolic play, and use language (Sattler, 2001). Piaget's theory aroused controversy because of the focus on stage differences and several theories of childhood cognitive development emerged.

Many scientists looked to intelligence tests to understand the nature of childhood mental development; however, most of the intelligence tests developed in the early 1900s were for use with school-aged children. Under the leadership of G. Stanley Hall at Clark University, scientists discussed the need for additional revisions for use with preschool children (Kelley, & Surbeck, 2007). With funding in hand, many universities established institutes of child welfare for the study of child development. The Yale Clinic for Child Development established by Gesell (1925) and his colleagues was a leader in the understanding of preschool development. Gesell operated under the principle that development was biologically predetermined, which became controversial among child development researchers (Kelley & Surbeck, 2007). Wellman (1932) found children's IQ increased when provided a stimulating environment. Another prominent child development researcher, Goodenough (1942), dismissed evidence of environmental influences on mental development and did not find any differences between nursery school and non-nursery school children. A psychometric focus on understanding intelligence at all age levels became the forefront within the research field (Sattler, 2001).

Factors of Intelligence in Preschoolers

The 1960's marked the period when psychologists began to discover that younger children's cognitive abilities were different than older children's abilities (Siegler, 1991). Prior to contemporary research, age-differentiation was taken into account when examining intelligence over the life-span. Age-differentiation suggested that, overtime,

general ability, *g*, narrows itself into more specific abilities. As an individual developed, his or her abilities became more distinct. This theory implied a quantitative difference in intelligence at different developmental periods (Tusing & Ford, 2004). In other words, the number of factors, or abilities, increased with age. Gardner and Clark (1992) suggested differentiation was the most popular way to account for development within intelligence theory.

Traditionally, clinicians have assumed young children's (i.e., preschoolers') cognitive abilities are defined by a verbal-nonverbal dichotomy (Tusing & Ford, 2004). Research conducted on past versions of intelligence tests consistently showed the two factor structure was adequate to conceptualize intelligence in preschoolers (LoBello & Gulgoz, 1991; Blaha & Wallbrown, 1991). General intelligence, *g*, also was found to account for a significant amount of variance in ability. This research was conducted on tests which were not designed to follow any theoretical structure (e.g., the Wechsler Scales) but has nonetheless influenced the conceptualization of preschool cognitive ability.

In addition, study of the standardization sample of the Differential Ability Scales (DAS; Elliot, 1990) suggested two factors, verbal and nonverbal, best defined intelligence for children from ages 3 years, 6 months to 5 years, 11 months. Using the Stanford-Binet, Fourth Edition (SB-IV; Thorndike, et al., 1986), Keith and colleagues (1988) found verbal reasoning and nonverbal/abstract reasoning was the prominent factor structure for two to six year olds. Furthermore, they found a memory factor could be differentiated in older groups (ages 7-11 and 12-23) providing evidence that more factors were apparent at older ages.

Research has consistently shown intelligence tests that assess a broad range of ages typically find fewer factors at younger levels than at school-aged levels (Carroll, 1993; Tusing et al., 2003). The traditional verbal-nonverbal dichotomy (verbal vs. symbolic) has long been held as the way to discriminate young children in terms of intelligence (Sattler, 2001; Tusing & Ford, 2004). However, research now shows more broad factors involved in the cognitive abilities of young children, such as the broad abilities defined in the CHC model.

The CHC model of intelligence has been studied in preschool children. Research with the Woodcock-Johnson Tests of Cognitive Abilities-Revised and the Differential Ability Scales: Upper Preschool Level suggested a five factor model accurately described cognitive abilities in preschoolers (Tusing & Ford, 2004). Included in this five-factor model were crystallized intelligence, long-term storage and retrieval, short-term memory, auditory processing, and nonverbal ability. The reserachers suggested more broad abilities described by the CHC model could be assessed in preschoolers through the popular cognitive ability measures used with preschoolers; however, there is yet to be a single test to incorporate all the broad abilities. Moreover, cross-battery approaches have been suggested as the best way to gather cognitive ability information as related to all broad abilities because each recent test based on the CHC model lacks assessment of all broad abilities (Alfonso et al., 2005).

Research using the same data utilized in the present study provided additional evidence of multiple factors at the preschool level. Morgan (2008) found the existence of five broad ability factors measured by the Stanford-Binet, Fifth Edition in a preschool population. These factors included *Gf*, *Gc*, *Gq*, *Gv*, and *Gsm*. Additionally, the factor

structure in preschool children most closely represented the *Gf-Gc* Theory in that the existence of *g* as a distinct factor was not found. Additionally, the KABC-II measured four broad abilities factors at the preschool level and supported a general factor. Therefore, the factor structure of the KABC-II was best described by the CHC theory. Moreover, Hunt (2007) found that multiple cognitive abilities could be assessed at the preschool level with the WJIII COG. *Gf*, however, was not found to be a distinct factor as it had a factor loading of 1.00 with *g*. Overall, this research provided more evidence for the existence of multiple broad abilities at the preschool level. More research involving the latest versions of cognitive ability measures is needed to confirm the broad ability factor structure in preschoolers, especially among different ethnic groups.

Accountability of Education for Preschool Children

Federal legislation has had the most important and influential impact on the assessment of preschool children. Education is considered a property right protected by the 14th Amendment of the United States Constitution, and is granted to every citizen under state law. Prior to two court cases in 1971 and 1972, children with disabilities were often excluded from public education. In 1971, parents of children with mental retardation won access to public education through *Pennsylvania Association for Retarded Children V. Commonwealth of Pennsylvania*. Children with mental, physical, or emotional impairments were granted a free and appropriate education from the decision of *Mills v. Board of Education of District of Columbia*. Federal legislation soon followed to ensure the education of children with disabilities.

The most important federal legislation to address the education of children with disabilities was the Education of All Handicapped Children Act of 1975, commonly

known as Public Law 94-142 (Nagel, 2007). Under this law, all school-aged children with disabilities were to receive a free and appropriate public education, as well as provide services to preschool children aged three to five with disabilities. The requirement of a free and appropriate education for preschool children with disabilities three to five was added with the passage of Public Law 99-457 in 1986. P.L. 99-457 also required that assessment of children include a multidisciplinary evaluation. As part of the mandated law, all children who qualified for services were to receive an Individualized Education Plan (IEP), which would be developed through the evaluation and diagnosis of each child's level of functioning; functioning would be assessed through various methods that ranged from behavior rating scales to standardized tests (Kelley & Surbeck, 2007). In 1990, Public Law 94-142 was renamed the Individuals with Disabilities Education Act (IDEA), and Part B specifically addressed the needs of preschool children.

Cognitive assessment instruments were critical aspects of the special education law. IDEA included several categories of eligibility requiring the use of a cognitive ability measure as part of the multidisciplinary assessment (e.g., mental handicap and specific learning disability; Jacob & Hartshorne, 2003). Although the 2004 reauthorization of IDEA (IDEIA 2004) does not require the use of cognitive ability measures, such information will still be helpful in determining specific strengths and weaknesses in the area of learning and will continue to be used in clinical practices (Holdnack & Weiss, 2006). Greater accountability for the identification of children in need of support, as described by the No Child Left Behind Act of 2001, further adds to the need of theoretically grounded, well-validated, and empirically supported measures of cognitive ability (Ford & Dahinten, 2005).

The passage of No Child Left Behind in 2001 has led to a great deal of research regarding the achievement gap crisis between diverse populations, including African Americans and Caucasians. Review of the current research showed a decrease in the achievement gap prior to the accountability legislation in the 1990s (Harris & Herrington, 2006). Higher standards of achievement appeared to coincide with a wider gap between minority and non-minority students. Research regarding achievement in preschoolers suggested that early experiences play a formative role in school readiness and explain gaps in reading and math once formal schooling begins (Chatterji, 2006). Project Head Start aims at providing these important pre-education experiences to minorities, although not all who are eligible take advantage of such programs (Chatterji, 2006). Findings regarding an increasing achievement gap (Harris & Herrington, 2006) along with the overrepresentation of minorities in special education (Holdnack & Weiss, 2006) posed strong evidence for eliminating cultural bias in testing measures.

The federal program Head Start had a significant influence on test development for preschool children (Kelley & Surbeck, 2007). Since federally funded programs such as Head Start have a performance-based evaluation component, measurement of cognitive ability and achievement is important for the program to continue to receive funding. Programs such as Head Start were designed to give the most economically disadvantaged and “those at risk for school failure” a chance to receive services in “school readiness” (Vinovskis, 2005). Head Start, along with other programs (e.g. First Steps) and federal legislation, allow for early identification of children with developmental deviations. Early intervention is essential in that it provides support for families, environmental stimulation appropriate for the child’s level of functioning, and

enhances the child's ability to function to his or her fullest capability (Kenny & Culbertson, 1993).

Cognitive Batteries for Preschool Children

The requirement of early childhood special education and the growth of school readiness programs directly contributed to the need for valid and reliable assessment of preschool and young children (Ford & Dahinten, 2005). Bracken (1987) analyzed 10 preschool instruments and their test manuals to determine if they met suggested levels of technical adequacy including internal consistency, stability, subtest floors, and validity. Included in his study were commonly used intelligence and skills tests used with preschool children including the Kaufman Assessment Battery for Children (K-ABC), Stanford Binet, 4th Edition (SB-IV), Wechsler Preschool and Primary Scale of Intelligence (WPPSI), and Peabody Picture Vocabulary Test-Revised (PPVT-R). Results indicated the K-ABC was the best psychometrically sound assessment for children under the age of four years old, especially in the areas of test-retest reliability, subtest floors, and total test reliability. Overall, these tests displayed adequate psychometric properties for children aged four to six. The author noted the K-ABC manual provided the most evidence for validity, followed by the SB-IV. When assessing preschool abilities, examiners must be aware of validity and reliability before choosing a measure for a specific age group.

The psychometric considerations just described contribute to the controversy over the use of standardized tests with preschool children. Bracken (1987) found that assessment of children under the age of four presented the most psychometric problems. Criteria recommended by the above author included: (1) median subtest internal

consistency of .80 or greater; (2) total test internal consistency of .90 or greater; (3) a total test stability coefficient of .90 or greater; (4) an average subtest floor at or below a scale score of 4; and (5) a total test floor at or below a standard score of 70 or two standard deviations below the normative total test mean score.

Several factors should be considered when choosing tests for assessment with preschoolers. Since young children and children with developmental disabilities may have difficulty maintaining focused attention, time of administration should be considered. Adequate floors are necessary especially when assessing children with developmental disabilities. A child whose chronological age is five, but is suspected to be functioning 2-years below his peers, should be tested with a measure containing an adequate floor for a 3-year old (Lichtenberger, 2005). Test developers also must provide evidence of acceptable validity (i.e., predictive, construct, and concurrent) at the preschool level so psychologists can make informed decisions as to whether the test is appropriate for the desired age group (Bracken, 1987). Additionally, Bracken (1987) indicated item gradients should be considered when selecting assessment batteries for preschool children. Item gradient is defined as “how rapidly standard scores increase as a function of a child’s success or failure on a single test item (Bracken, 1987, p. 322). In other words, a test is less effective in measuring ability when dramatic changes are seen in standard scores as minor differences in raw score occur.

The WPPSI historically has been the most widely used standardized measure of preschool intelligence (Ford & Dahinten, 2005). As more theoretically grounded tests emerge, the most recent WPPSI-III may not be as preschool friendly as other newly developed tests. Specifically, the WPPSI has been criticized for its length, emphasis on

expressive language, and strong use of fine motor abilities (Bracken, 1987). With strong technical characteristics, the Differential Ability Scale (DAS; Elliot, 1990) provides a developmentally responsive battery making it a widely used school intelligence test. The DAS includes School-Age and Preschool levels, and the Preschool level consists of an Upper and Lower sublevel.

The Stanford-Binet Intelligence Scales, Fifth Edition (SB-V; Roid, 2003) stems from the long tradition of Stanford-Binet scales first developed by Alfred Binet and Theodore Simon in 1905. Lewis Terman first revised the original scale and brought it to the United States where it has undergone several revisions. The SB-V was developed based on the CHC theory and represents a hierarchical structure of intelligence (Alfonso & Flanagan, 2007). Excellent psychometric properties were reported for the preschool population including its floors, ceilings and item gradients (Ford & Dahinten, 2005).

The K-ABC (Kaufman & Kaufman, 1983) was developed to make it appropriate to use with preschool children. Similarly, the KABC-II adequately measures cognitive ability in children between the ages of 3-years, 0-months and 18-years, 11-months. Both versions of the K-ABC are grounded in theory, and the KABC-II expanded its interpretation model to include the CHC model. While the original K-ABC utilized the Luria-Das Successive-Simultaneous Processing Dichotomy, the revision uses the CHC model as the preferred model of interpretation in most cases (Lichenberger & Kaufman, 2007). Popular among psychologists, the K-ABC attempted to reduce the effects of cultural differences making it an appropriate assessment for linguistically and culturally diverse children (Ford & Dahinten, 2005). Similarly, the authors of the KABC-II found a reduced overall IQ difference among ethnic groups (Kaufman & Kaufman, 2004). With

adequate psychometric properties, developmentally appropriate material, and a strong theoretical foundation, the KABC-II is a well-rounded instrument for use with preschool children (Lichenberger & Kaufman, 2007).

Another intelligence test grounded in the CHC theory and used with preschool children is the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJIII COG; McGrew & Woodcock, 2001). The WJIII COG can be used with children as young as 3 years old, and includes 20 tests, 14 of which measure 7 of the broad abilities defined in the CHC model; several narrow abilities are also measured. Tusing and colleagues (2003) studied the WJIII COG for use with preschool children. They found that although several broad abilities factors were distinguished in preschool children, the CHC factors changed with age. Specifically, developmental patterns of growth and decline were different for each factor. Furthermore, developmental changes impacted the age at which the CHC abilities first become evident indicating the CHC abilities may be represented differently at younger ages (Tusing et al., 2003).

Chapter Summary

This chapter outlined the historical influences on assessment and intellectual theory. Beginning with the efforts of Sir Francis Galton, the measurement of intelligence has developed from studying individual differences as measured by physical attributes to modern versions of Binet's attempts to measure higher mental processes (Sattler, 2001; Suen & French, 2003; Wasserman & Tulskey, 2005). Current tests of intelligence reflect the most widely accepted theory of intelligence, the Cattell-Horn-Carroll Theory. Kevin McGrew's proposed combination of the similar theories developed by Horn and Cattell (*Gf-Gc* Theory) and Carroll (Three Stratum Theory) has lead to the most well-validated

and empirically supported theory for intellectual assessment (Flanagan, Mascolo, & Genshaft, 2000). The KABC-II was developed to reflect the CHC model of intelligence and is considered psychometrically appropriate for use with preschool children (Kaufman & Kaufman, 2004). The relatively recent interest in preschool assessment and its roots in American legislation were reviewed.

This chapter also addressed the long standing gap in IQ between African Americans and Caucasians. Beginning with the mass testing movement that occurred during World War I, differences in intelligence between these groups has been the focus of much research (Rushton & Jensen, 2005). Since the gap has been well established over decades of research, the research now focuses on the reason the gap exists. Hereditarian and cultural models have been proposed, each having limitations within the research (Brody, 1992). Continuing research may wish to focus on interventions aimed at closing the gap rather than further studying the deadlock of a cause (Brody, 1992).

Currently, few studies using contemporary cognitive ability tests focus on the mean difference in intelligence scores between African Americans and Caucasians. The literature suggested the K-ABC showed the smallest gap among ethnically diverse children (Kaufman & Lichtenberger, 2002; Kaufman et al., 2005), and the manual of the KABC-II provided promising evidence that the effects were replicated in the revision (Kaufman & Kaufman, 2004). This study was hoped to add to the body of research addressing the IQ gap and provided important information for clinicians and psychologists when interpreting the KABC-II for ethnically diverse preschool children. Conducting a profile analysis on the KABC-II provided information on the performance of preschool children at the composite and subtest levels. Specifically, this study assisted

in understanding if significant mean differences were seen on the broad and narrow CHC abilities of preschool children. Additionally, decreased test bias would be shown if the same pattern of highs and lows is seen between the two groups.

CHAPTER III

Research Methodology

The current study examined an archival data set from the “Exploring CHC Theory’s Cognitive Constructs in Young Children” project. This chapter offers a description of the participants, procedures, and instrumentation used in the present study as well as information regarding the methods used to collect the data.

Participants

The preschoolers who participated in the “Exploring CHC Theory’s Cognitive Constructs in Young Children” project were a sample of convenience, obtained from preschools and daycares located in midsized city in the Midwest (population approximately 60,000). Two hundred children (males = 97, females = 103) ranging in age from 4 years, 0 months to 5 years, 11 months participated after parental permission was obtained (see Appendices B and C to review parent permission and demographic forms). As reported by the parents of the children, the entire sample included 124 Caucasians (62%), 49 African Americans (24.5%), 20 bi-racial students (10%), 2 Hispanics (1%), and 3 from other ethnicities (1.5%), with 2 failing to report their ethnicity (1%). The distribution of race differed from the participating area’s population, which included 85.7% Caucasian, 11% African American, 1.5% bi-racial, 1.4% Hispanic, and 1.2%

having other ethnicities (United States Census Bureau, 2001). The total research sample included an overrepresentation of African American and bi-racial children.

The preschoolers came from homes of varied education levels. Paternal education levels included 15% with four or more years of college, 22% having one to three years of college, 39% with a high school diploma, 15% having less than a high school diploma, and 9 % failing to report education. Additionally, maternal education levels included 19.5% with four or more years of college, 44.5% having one to three years of college, 24% had a high school diploma, 8% having less than a high school diploma, and 4% did not report their education level. Children with known disabilities, as reported by the facilities' teachers, were excluded from the study.

For the purpose of the present study, only 49 Caucasian and 49 African American children (males = 46, female = 52) were included in the analyses. The sample for the current study was selected to maximize the number of participants by including all of the African Americans from the total sample and matching an equal number of Caucasians. The mean age for the 98 selected participants was 59.27 months ($SD = 5.63$). Participants were matched on age, parental education, and sex to control for these variables' influence on performance. The mean age in months was 59.14 ($SD = 5.68$) for the African American children and 59.39 ($SD = 5.63$) for the Caucasians. Although the two groups were matched as closely as possible on age in months, the difference in age was no more than two months for all matched participants. Additionally, participants were matched exactly on gender and as closely as possible on parental education level. At least one parent matched on level of education, or else they were matched no more than one level difference. Specifically, 55.1% matched on paternal education, 18.4 % matched on

maternal education, 8.1 % matched maternal to paternal education, and 18.4% matched the parent within one level of education. Overall, 81.6% of the sample matched on level of education for one parent. Parental education levels by race are included in table 3.1.

Table 3.1

Parental Education by Race

Education Level	<u>Paternal Education</u>		<u>Maternal Education</u>	
	African American	Caucasian	African American	Caucasian
Some HS, no diploma	6 (12.2%)	7 (14.3%)	4 (8.2%)	3 (6.1%)
HS diploma or GED	20 (40.8%)	23 (46.9%)	16 (32.7%)	11 (22.4%)
1-3 Years college	9 (18.4%)	11 (22.4%)	21 (42.9%)	25 (51.0%)
4+ Years college	6 (12.2%)	5 (10.2%)	5 (10.2%)	8 (16.3%)
Omitted	7 (14.3%)	3 (6.1%)	3 (6.1%)	2 (4.1%)

Note. Percentages shown are percentages within race.

Sample Size

Tabacknick and Fidell (2001) indicate there should be more research units in the smallest group than there are dependent variables to use profile analysis. Commonly, sample size decides the types of analysis for those mentioned earlier. For the current study, the first analysis includes five dependent variables and 49 subjects in each group, deeming the sample size appropriate for profile analysis. The second analysis includes 9 dependent variables, making the sample size adequate in this case as well.

Missing Data

Missing data is one of the most common problems in research and data analysis (Tabacknick & Fidell, 2001). The problems of the missing data depend on the pattern, how much is missing, and why it is missing. The data missing on the KABC-II can be

explained by the age restrictions of the subtests. The preschool children aged 4 years, 0 months to 4 years 11 months, a total of 62 students, were not administered the Pattern Reasoning subtest because standardization norms were not available for this group. Since norms were not available for the 4 year old group, this subtest was excluded from the analysis.

Procedures

The present study examined an archival data set collected by graduate students in the school psychology program at a public Midwestern university of approximately 20,000 students. The “Exploring CHC Theory’s Cognitive Constructs in Young Children” study was conducted with the consent of the Institutional Review Board of Ball State University obtained in February 2004. Principals and daycare directors of local facilities were contacted to request permission to solicit parents of preschool children. Principals and directors were provided with a complete description of the CHC Theory project. The parents of preschool children from the schools that gave permission for the study to take place were sent a complete description of the study and permission forms to complete (Appendices A, B, & C). The parents who granted permission for their child to participate signed the permission form and completed the accompanying demographic form. Parents of participants were given the option of receiving the assessment results and were provided with the opportunity to ask questions prior to the data collection. Furthermore, parents were informed their child’s information would be kept confidential and names would not be released for any reason. If at any time the child no longer wanted to participate, parents were informed they could withdraw their child without penalty and their data would be destroyed upon request.

Data collection occurred from winter/spring of 2004 through the summer of 2006. The graduate students visited the preschool participants at their school or daycare. Each preschooler was individually administered the Kaufman Assessment Battery for Children, Second Edition (KABC-II). The Stanford-Binet Intelligence Scale, Fifth Edition (SB-V) and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJIII COG) were also administered to the participants, but are not included in the analysis of the present study. Each child participated in three separate testing sessions over the course of two weeks, and tests were administered in counterbalanced order. The testing occurred in a private room with only the child and examiner present. Each examiner had received graduate level training in cognitive assessment consisting of at least one full year of classroom instruction and practice on the standardized administration procedures of various cognitive ability measures. Protocols were scored by the graduate students and then reviewed by advanced graduate students and faculty for accuracy.

Instrumentation

Kaufman Assessment Battery for Children, Second Edition (KABC-II)

The Kaufman Assessment Battery for Children, Second Edition is a newly revised measure of cognitive abilities for children and adolescents aged 3 to 18 (Kaufman & Kaufman, 2004). Depending on the age of the child, the KABC-II yields scores on one to five scales, and can be interpreted using one of two theoretical models: Luria's Neuropsychological Theory or the Cattell-Horn-Carroll (CHC) model. In most cases, the CHC model is the preferred model of interpretation (Lichenberger & Kaufman, 2007) and was utilized for this study. The CHC global scale is labeled the Fluid Crystallized Index (FCI) and is composed of five CHC components. For ages four through six, the

KABC-II global index is organized into four CHC components: Short-Term Memory (*Gsm*), Visual Processing (*Gv*), Long-Term Storage and Retrieval (*Glr*), and Crystallized Ability (*Gc*) (table 3.2 provides descriptions of the KABC-II subtest for the preschool group). Fluid reasoning CHC component is not measured in this age group; therefore, all the subtests were not administered to the participants (i.e. Story Completion, Rover, and Block Counting). Story completion and Rover required an age of 6 or over, and Block Counting required an age of 5 or older.

The KABC-II standardization sample included 3,025 children aged 3 to 18. These children were chosen to match closely to the 2001 United States Census data on the variables of age, gender, geographic region, ethnicity, and parental education (Kaufman & Kaufman, 2004). The standardization sample was divided into 18 age groups consisting of 100 to 200 children in each group. Males and females were equally represented, and the children were from four regions across the United States (Northeast, North Central, South, and West).

Reliability. Reliability refers to the dependability of the test scores or the ability to reproduce the scores (Kaufman & Kaufman, 2004). Internal consistency reliability coefficients were presented in the KABC-II manual. The Fluid Crystallized Index (FCI), the global intelligence score for the KABC-II when interpreted with the CHC model, yielded a reliability coefficient of .96 for ages 3 to 6 and .97 for ages 7 to 18. The coefficients for the four broad ability scales measured at the preschool level ranged from .91 to .92, with subtests ranging from .69 to .92. The coefficients for the five broad ability scales measured for the older age group ranged from .88 to .93, with subtests ranging from .74 to .93. Overall, the subtests show very good internal consistency at all ages with

the average coefficient for the preschool group averaging .85 (Kaufman & Kaufman, 2004).

The KABC-II is considered a fairly stable measure of intelligence (Kaufman & Kaufman, 2004). During the standardization process, the instrument was administered twice to three groups of children over an average period of four weeks. These groups were divided by age. Test-retest reliability coefficients for the FCI were .90 for children ages 3 to 5, .91 for children ages 7 to 12, and .94 for adolescents ages 13-18. Adequate stability was also indicated for the broad ability scales. The test-retest coefficients for the preschool group ranged from .74 to .93, and from .76 to .95 for the school aged and adolescent groups.

Validity. Confirmatory factor analyses were conducted to establish construct validity of the KABC-II. Construct validity refers to the scales ability to measure or correlate with the theoretical construct it is proposed to measure (Kaufman et al., 2005). These analyses supported the use of different batteries for different age levels (Kaufman & Kaufman, 2004). At the preschool level, the KABC-II produces a single factor model for age 3. At age 4, confirmatory factor analyses revealed evidence of multiple factors, with Sequential/*Gsm* and Learning/*Glr* as distinct factors. Although ages 4 and 5 are assessed with a four-factor model, the Knowledge/*Gc* and Simultaneous/*Gv* factors were not distinct at the 4-year old level. The authors separated the Knowledge/*Gc* and Simultaneous/*Gv* factors on the final battery because of their distinct content (Kaufman & Kaufman, 2004). Furthermore, Planning/*Gf* is not included on the preschool battery because the analyses did not reveal it as a distinct factor until the age of 7. Fit statistics

Table 3.2

Kaufman Assessment Battery for Children, Second Edition, Subtests

Broad CHC Factors and Test	Narrow CHC Ability	Subtest Description
Short-Term Memory Sequential/Gsm		
Number Recall	Memory Span	Saying a series of numbers in the same order
Word Order	Memory Span and Working Memory	Touching silhouettes of objects after examiner names them in sequence
Long-Term Storage and Retrieval Learning/Glr		
Atlantis	Associative Memory	Learning nonsense names of pictures, pointing to correct picture when shown a series of pictures and given the name
Rebus	Associative Memory	Learning words associated with line drawings
Visual Processing Simultaneous/Gv		
Conceptual Thinking	Visualization and Induction	Looking at a group of pictures and identifying them
Face Recognition	Visual Memory	Seeing a face and identifying the same face later
Triangles	Spatial Relations and Visualization	Copying or building designs using colored triangles or plastic shapes

**Crystallized Ability
Knowledge/Gc**

Expressive Vocabulary	Lexical Knowledge	Seeing a color drawing of an object and saying its name
Riddles	Lexical Knowledge, General Reasoning, and Language Development	Retrieving the name of an object/concept with clues

Note. From *Kaufman Assessment Battery for Children, Second Edition, Manual*, by A.S. Kaufman and N.L. Kaufman, 2004

for the KABC-II all exceeded .99 supporting its CHC theory based structure (Kaufman & Kaufman, 2004).

Validity of the KABC-II is further supported by correlations with several other instruments designed to measure similar intellectual constructs (Kaufman & Kaufman, 2004). The KABC-II correlated strongly with other tests of intelligence including the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV), the Wechsler Intelligence Scale for Children, Third Edition (WISC-III), the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III), the Kaufman Adolescent and Adult Intelligence Test (KAIT), and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, (WJIII COG). The correlations among the KABC-II FCI and the global measures of these tests ranged from .72 to .91 (Kaufman & Kaufman, 2004).

Predictive validity was assessed by comparing the KABC-II with its conormed counterpart, the Kaufman Tests of Educational Achievement, Second Edition (KTEA-II) as well as other popular tests of academic achievement (Kaufman & Kaufman, 2004). Correlations between the KABC-II FCI and the global achievement score of the KTEA-II ranged from .74 to .82 for all grades. Additionally, the KABC-II FCI was compared to

the Woodcock-Johnson Tests of Achievement, Third Edition (WJIII ACH) yielding correlation coefficients of .70 (grades 2-5) and .79 (grades 6-10). Two other achievements tests, the Wechsler Individual Achievement Test, Second Edition (WIAT-II) and the Peabody Individual Achievement Test-Revised (PIAT-R) were compared to the KABC-II providing extensive evidence of predictive validity (Kaufman & Kaufman, 2004). Correlations between the KABC-II and the WIAT-II were .72 (grades 2-5) and .87 (grades 7-10), and correlations with the PIAT-R were .67 for younger grades and .73 for older grades.

Additionally, validity of the KABC-II was assessed in regards to differences among demographic groups. The tests developers attempted to establish validity by indicating whether the scores on the test have the same meaning regardless of the demographic characteristics of the individual (Kaufman & Kaufman, 2004). At the preschool ages, girls performed slightly better on the composites than boys, except for Knowledge/*Gc* in which no real difference was seen. For ages 7-18, males and females tended to perform at similar levels. Demographic information on parental education was collected during the standardization phase. Four groups were identified: less than high school graduate, high school graduate, some postsecondary schooling, and four-year college degree or more. For preschool age children, parental education level had a strong influence on performance, especially for the Knowledge/*Gc* composite. Parental education influenced the scores of school-aged children less, but still made an impact. For the present study, participants were matched on age, sex, and parental education to attempt to control for these differences.

Furthermore, ethnic differences were reported. Parental education and sex were controlled for in the analysis because parental education level differed among ethnic groups (Kaufman & Kaufman, 2004). Overall, the amount of variance in scores accounted for by ethnicity was much lower than that accounted for by parental education. On global scales, ethnicity accounts for no more than 2% of test score variance of preschoolers and 5% of school-aged children. Additionally, the mean score difference between African Americans and Caucasian Americans was reduced on the KABC-II (Kaufman & Kaufman, 2004), consistent with the previous version of the test (Kaufman & Lichtenberger, 2002; Kaufman et al., 2005). The Fluid Crystallized Index for preschool children differed 3.6 points between African Americans and Caucasian Americans when adjusted for sex and maternal education; unadjusted means yielded a 7.0 point difference. The difference in the gap was slightly higher for the school-aged children with a 7.9 point adjusted mean difference and 9.3 unadjusted mean difference. Compared to the one standard deviation (15 point) difference reported in the literature (Jensen, 1998), the KABC-II appeared to improve the assessment of cognitive ability in African Americans.

Chapter Summary

The purpose of this study was to conduct a profile analysis to determine if African American and Caucasian American preschool children display the same pattern of performance on the KABC-II. Specifically, the KABC-II measures several broad and narrow abilities as described in the Cattell-Horn-Carroll (CHC) theory of intelligence. This chapter described the participants and methods utilized in the “Exploring CHC Theory’s Cognitive Constructs in Young Children” project. A matched sample of 98 children was pulled from this archival data set of 200 preschool children. Caucasian and

African American children were matched on age, gender, and parental education level. Children were administered the KABC-II, and their scores were compiled. Furthermore, this chapter reviewed the psychometric properties, such as validity and reliability, of the KABC-II, as well as the standardization procedures. The subtests were described, and the associated CHC constructs were listed.

CHAPTER IV

Results

The focus of this study was to compare the performance profiles of African American and Caucasian preschool children on the Kaufman Assessment Battery for Children, Second Edition (KABC-II; Kaufman & Kaufman, 2004). A profile analysis was conducted to determine the performance profiles in relation to the composite and subtest scores. Moreover, this study investigated whether African Americans and Caucasians differed in performance on the broad and narrow abilities of the Cattell-Horn-Carroll (CHC) model of intelligence measured by the KABC-II. Each profile analysis measured whether the level of the profile was the same and if the same patterns of highs and lows were seen in each group. This chapter reported descriptive statistics, including means, standard deviations, and ranges of overall ability scores, broad factor scores, and individual subtest scores of the KABC-II. Furthermore, results of the profile analyses of the composites and subtests were provided.

Descriptive Statistics

Test scores were collected, and initial descriptive statistics were calculated for the KABC-II. Means, standard, deviations, and ranges were computed for the KABC-II general ability score (FCI), composite scores, and subtest scores. These analyses were

conducted using the Statistical Package for the Social Sciences (SPSS) 15.0 program. The means, standard deviations, and ranges of the KABC-II Fluid-Crystallized Index (FCI), cluster (Short-Term Memory/*Gsm*, Visual Processing/*Gv*, Long-Term Storage & Retrieval/*Glr*, and Crystallized Ability/*Gc*), and subtest scores for the sample of preschool children are presented in Table 4.1. The FCI and cluster scores have a standardized mean of 100 and standard deviation of 15. Average scores range from 85 to 115. The mean of the FCI for the matched sample of preschool children used in the present study ($n = 98$) was 96.33 with a standard deviation of 12.73. The means of the KABC-II broad factor cluster scores ranged from 94.64 to 98.84, with standard deviations ranging from 10.28 to 14.31. The means for the KABC-II FCI and factor scores fell within the average range, which was expected given the preschool children selected for the study were sampled from a general population without known disabilities. The total sample performed the lowest on the *Gsm* ($M = 94.64$, $SD = 11.36$) broad ability. Additionally, means, standard deviations and ranges were calculated for the KABC-II subtests and are presented in Table 4.1. Subtest scores have a mean of 10 and a standard deviation of 3. The KABC-II subtest mean scores for the total sample ranged from 8.36 to 10.00, with standard deviations ranging from 2.08 to 3.28. Again, the research sample scored within the average range on the KABC-II subtests. The total sample performed the lowest on the Riddles ($M = 8.36$, $SD = 2.37$) and the Word Order ($M = 8.94$, $SD = 2.12$) subtests.

Furthermore, means, standard deviations, and ranges were calculated for each ethnic group and are presented in Table 4.2. The mean of the FCI for the African American group ($n = 49$) was 95.59 ($SD = 12.81$). The African American preschool

Table 4.1

Means, Standard Deviations, and Ranges of the KABC-II FCI, Cluster, and Subtest Scores for Total Sample (n = 98)

FCI and Clusters	Mean	SD	Range
FCI	96.33	12.73	65-120
<i>Gsm</i>	94.64	11.36	68-124
<i>Gv</i>	97.68	13.59	64-124
<i>Glr</i>	98.84	14.31	67-133
<i>Gc</i>	95.48	10.28	74-123
<u>Subtests</u>			
Number Recall	9.22	2.20	3-14
Word Order	8.94	2.12	5-15
Atlantis	9.53	2.93	3-18
Rebus	10.00	3.13	4-17
Conceptual Thinking	9.71	3.28	4-17
Face Recognition	10.00	2.58	4-15
Triangles	9.02	2.49	3-15
Expressive Vocabulary	9.56	2.08	4-14
Riddles	8.36	2.37	1-14

Note. The KABC-II FCI and cluster scores have a standardized mean $M = 100$, $SD = 15$. The KABC-II subtests have a standardized $M = 10$, $SD = 3$.

children's mean scores on the KABC-II clusters ranged from 93.33 to 98.31, with standard deviations ranging from 10.08 to 14.69. The mean FCI for the Caucasian group ($n = 49$) was 97.06 ($SD = 12.74$), with the mean cluster scores ranging from 94.49 to

99.37. Standard deviations of the cluster scores for the Caucasian group ranged from 9.85 to 14.42. Subtest means, standard deviations, and ranges for each ethnic group are also shown in Table 4.2. The mean subtest scores for the African American group ranged from 8.08 to 10.18 (*SD* range = 1.77 to 3.51). The mean subtest scores for the Caucasian group ranged from 8.63 to 10.08 (*SD* range = 1.92 to 3.27). Both groups performed in the average range across all clusters and subtests.

Assumptions

Prior to running the profile analysis, several assumptions were addressed. Because the sample sizes were equal, examination of homogeneity of variance-covariance matrices was not needed (Tabachnick & Fidell, 2001). Additionally, profile analysis is robust to the violation of normality, especially when there are more cases than dependent variables and sample sizes are equal (Tabachnick & Fidell, 2001). Given that the sample sizes were equal and the number of preschoolers in each group (49) far exceeded the number of dependent variables in each analysis (5 and 9), normality was not likely to be violated. An examination of the skewness and kurtosis of the dependent variables further illustrated that the assumption of normality was not violated. Skewness refers to the symmetry of the distribution, meaning that a skewed variable is one where the mean does not lie at the center of the distribution (Tabachnick and Fidell, 2001). Kurtosis refers to the peakedness of the distribution. Positive scores reflect more scores around the median while negative scores reflect a more flat distribution. All of the skewness statistics for the composites fell between -1.0 and 1.0 and the kurtosis scores were all close to 0.0, ranging from -.455 to .061. Similarly, all of the skewness statistics for the subtests fell between -1.0 and 1.0 and the kurtosis scores were all close to 0.0, ranging from -.724 to .412.

Table 4.2

Means, Standard Deviations, and Ranges of the KABC-II FCI, Cluster, and Subtest Scores for African American and Caucasian Groups

FCI and Clusters	<u>African American (<i>n</i> = 49)</u>			<u>Caucasian (<i>n</i> = 49)</u>		
	Mean	<i>SD</i>	Range	Mean	<i>SD</i>	Range
FCI	95.59	12.81	66-120	97.06	12.74	65-120
<i>Gsm</i>	94.80	10.08	80-118	94.49	12.63	68-124
<i>Gv</i>	97.73	14.69	66-124	97.63	12.56	64-120
<i>Glr</i>	98.31	14.31	70-133	99.37	14.42	67-133
<i>Gc</i>	93.33	10.35	74-118	97.63	9.85	74-123
<u>Subtests</u>						
Number Recall	9.47	2.14	6-14	8.98	2.25	3-14
Word Order	8.76	1.77	6-12	9.12	2.43	5-15
Atlantis	9.39	2.57	4-14	9.67	3.27	3-18
Rebus	9.94	3.36	4-17	10.06	2.92	4-15
Conceptual Thinking	9.63	3.51	4-17	9.80	3.06	4-15
Face Recognition	10.18	2.75	4-15	9.82	2.43	5-15
Triangles	9.16	2.63	3-15	8.88	2.36	4-13
Expressive Vocabulary	9.04	2.12	5-14	10.08	1.92	4-14
Riddles	8.08	2.31	4-12	8.63	2.41	1-14

Note. The KABC-II FCI and cluster scores have a standardized mean $M = 100$, $SD = 15$. The KABC-II subtests have a standardized $M = 10$, $SD = 3$.

Furthermore, normal Q-Q plots in SPSS 15.0 demonstrated that the scores fell in a straight line for all composites and subtests.

Profile Analysis

A profile analysis was conducted to compare the performance of African American and Caucasian American preschool student on the KABC-II at the composite and subtest level. Profile analysis is a version of multivariate analysis of variance (MANOVA) applied when several dependent variables are measured on the same scale (Tabacknick & Fidell, 2001). In this study, the dependent variables were the composite scores (mean = 100, SD = 15) from the KABC-II for the first analysis and the subtest scores (mean = 10, SD = 3) for the second analysis. Additionally, for each analysis, the scores values have the same meaning on all the composites or subtests, a requirement for this statistical method. Profile analysis compared the means of each ethnic group and looked for interactions between race and the pattern of means across the composites and subtests. A plot of profiles was generated allowing for comparison of the two groups. Assessment of group means, standard deviations, standard errors, and confidence intervals is helpful if the performance between the groups significantly differs.

Profile analysis and repeated measures ANOVA typically are two methods to analyze data when multiple dependent variables are measured. Profile analysis is preferred, however, when the sample size is large because the assumption of sphericity is often violated in univariate analysis; sphericity is not an issue in profile analysis because of the robustness of the test statistics (i.e. Wilk's Lamda), and the other assumptions (i.e. homogeneity of variance-covariance matrices) are less likely to be violated (Tabachnick & Fidell, 2001). Additionally, there is greater power in the multivariate approach, making it more desirable (Tabachnick & Fidell, 2001).

Parallelism Test. The initial question for each profile analysis research question dealt with the parallel nature of the group profiles. The test of parallelism was conducted with adjacent segments of the profiles, and addressed the question of whether the difference between the segments from one test to another is the same for each group (Tabachnick & Fidell, 2001). Since the order of variables is often arbitrary in a profile analysis, the subtests and composites were entered into the data set in the same order as displayed on the protocol provided by the publisher of the KABC-II. For the current study, difference scores were calculated between adjacent KABC-II composite scores as well as adjacent subtest scores. A one-way MANOVA was conducted on each segment to test parallelism. Essentially, the segment represents a slope between the two adjacent scores, and a difference in the slopes on any segment between the two groups indicates the profiles are not parallel. For example, a parallel segment would mean the difference between the KABC-II Sequential/*Gsm* and Learning/*Glr* composites would be the same, or nonsignificant, between the African American and Caucasian preschool children. The profile would be parallel if no adjacent segments yielded significant differences between the groups.

Levels Test. Additionally, the levels test was conducted in regards to each research question. The levels test explored the difference in means between African American and Caucasian preschoolers combined over all composites and subtests. For the composite score analysis, the mean composite performance of African Americans was compared to the mean composite performance of Caucasians. A significant *F*-test would indicate one group performed significantly better than the other group overall. A review

of the group means determined which group performed better overall if the levels are found to be significantly different.

Research Question One

When conducting a profile analysis of the KABC-II composite scores:

- a. Do African American and Caucasian American preschool children display the same patterns of highs and lows (Parallelism test) across the CHC broad factors?
- b. Regardless of whether the profiles are parallel, does the Caucasian American group, on average score higher (Levels test) across the CHC broad factors, as a set, compared to the African American group?

An alpha level of .05 was set prior to analysis. To answer the first research question, the four CHC broad factors (*Gsm*, *Gv*, *Gl*, and *Gc*) were entered into the analysis in the order they appeared on the protocol. The test for parallelism was not significant indicating the African American and Caucasian preschool children displayed similar patterns of highs and lows (see Figure 4.1) across the CHC broad ability factors, $F(3, 94) = 2.009, p = .118$. Omega-squared ($\omega^2 = .010$) fell just at the threshold for Cohen's (1988) small range indicating ethnicity had a minimal effect on the performance of the two groups. Additionally, the levels test was not significant $F(1, 96) = .396, p = .531$, indicating that neither group performed significantly better than the other across the set of CHC broad ability factors. Omega-squared ($\omega^2 = -.006$) fell below Cohen's (1988) small range, providing further evidence that ethnicity did not affect the level of performance across the KABC-II composites.

The overall score produced by the KABC-II, the Fluid-Crystallized Index (FCI), was not included in the profile analysis as it is an overall score calculated from the composite scores. A one-way ANOVA was conducted to compare the overall performance of African American and Caucasian preschool children. No significant difference was found between the two groups on the FCI, $F(1, 96) = .324, p = .571$, indicating similar overall cognitive ability scores.

Figure 4.1. Mean CHC broad ability factor scores for the African American and Caucasian groups

Note. *Gsm* = Short-Term Memory, *Gv* = Visual Processing, *Glr* = Long-Term Storage and Retrieval, *Gc* = Crystallized ability

Research Question Two

When conducting a profile analysis of the KABC-II subtest scores:

- a. Do African American and Caucasian preschool children display the same patterns of highs and lows (Parallelism test) across the CHC narrow factors?

- b. Regardless of whether the profiles are parallel, does the Caucasian group, on average score higher (Levels test) across the CHC narrow factors, as a set, compared to the African American group?

To answer the second research question, the nine core subtests for which all participants had data were entered into the analysis in the order they appeared on the protocol. Each subtest represented a CHC narrow ability. The test for parallelism was significant, $F(8, 90) = 2.466, p = .018$, indicating African Americans and Caucasians displayed different highs and lows across the KABC-II subtests (see Figure 4.2). Through SPSS, a partial eta-squared value of .181 was obtained meaning that knowing the ethnicity of the participant accounts for 18.1% of the variance in performance across the subtests. Since eta-squared can be an overestimate of effect size (Tabachnick & Fidell, 2001), omega-squared was calculated by hand. Omega-squared ($\omega^2 = .015$) fell within Cohen's (1988) small range indicating ethnicity had a small effect on the difference in performance between African Americans and Caucasians.

Because the test for parallelism did not indicate which subtest means were significantly different, Tabachnick and Fidell (2001) suggest conducting one-way ANOVAs with the Scheffé adjustment to control for familywise Type I error. Using SPSS 15.0, the Scheffé adjusted F was calculated, $F(1, 96) = 3.94$, and a series of one-way ANOVAs was conducted (see table 4.3 for ANOVA results). Results found the Caucasian preschool children performed significantly better on the Expressive Vocabulary subtest compared to the African Americans, $F(1, 96) = 6.47, p = .013$. Caucasians performed 1.04 points better than African Americans. Effect size was calculated for the significant F value. Partial eta-squared was .063 meaning that knowing

the ethnicity of the participant accounts for 6.3% of the variance in performance on the Expressive Vocabulary subtest. Omega-squared ($\omega^2 = .053$) fell within Cohen's (1988) small range indicating ethnicity had a small effect on the difference in performance between the two groups; however, any significant differences between groups on a measure of cognitive ability must be investigated. All other subtests yielded no significant differences between groups. Additionally, the levels test was not significant $F(1, 96) = .222, p = .639$, indicating that neither group performed significantly better than the other across the set of CHC narrow ability factors. Omega-squared ($\omega^2 = -.008$) fell below Cohen's (1988) small range further indicating ethnicity had no effect on the level of performance between the two groups

Figure 4.2. Mean subtest scores for the African American and Caucasian groups

Follow-Up Analyses

Means from the research sample appeared lower than expected compared to the normative sample; therefore, follow-up *t*-tests were conducted to compare the total

sample ($n = 98$) to the KABC-II normative sample. Results are presented in table 4.4.

Because multiple t -tests were conducted, a significance value of .004 (.05 divided by the number of t -tests) was calculated as a more stringent value in order to control for alpha slippage (Type I error). The sample as a whole performed significantly weaker on *Gsm*, *Gc*, Number Recall, Word Order, Triangles and Riddles as compared to the standardization sample. The latter three subtests were the last three administered to the study sample; therefore, test fatigue and attention span may have been a factor.

Table 4.3

Follow-up ANOVA Results for the KABC-II Subtests

Subtests	<i>F</i>	Significance	Mean Difference
Number Recall	1.218	.272	-.490
Word Order	.731	.395	.360
Atlantis	.231	.632	.280
Rebus	.037	.848	.120
Conceptual Thinking	.060	.807	.170
Face Recognition	.491	.485	-.360
Triangles	.321	.571	-.280
Expressive Vocabulary	6.474	.013*	1.04
Riddles	1.335	.251	.550

Note. All scores have $df = 1, 96$.

*Significant F -test comparisons. Scheffé adjusted F of 3.94 was used to determine significance.

Additionally, the means of both the African American and Caucasian samples were compared to the normative data of the KABC-II. The manual for the KABC-II provided normative data for each ethnicity in the three to six age group. T -tests were

conducted comparing each group to the scores provided for their respective ethnicity. Results are presented in Appendix D. Because multiple *t*-tests were conducted, a significance value of .004 (.05 divided by the number of *t*-tests) was calculated as a more stringent value in order to control for alpha slippage (Type I error). Some significant differences were found between the two groups and their respective groups within the standardization sample. Specifically, both groups performed lower than the standardization sample on the *Gsm* broad ability and the Riddles and Number Recall subtests. Additionally, Caucasians in the study sample performed lower than Caucasian preschoolers in the standardization sample on the *Gc* broad ability and the Triangles subtest. African Americans performed lower than their respective group in the normative sample on the Word Order subtest.

Chapter Summary

This chapter presented the empirical results for the investigation of performance patterns of African American and Caucasian preschool children on the broad and narrow abilities of the Cattell-Horn-Carroll (CHC) theory of intelligence as measured by the KABC-II. Descriptive statistics including means, standard deviations, and ranges were reported to better describe the research sample.

Additionally, a profile analysis was conducted to compare the African American and Caucasian groups on the composites measured by the KABC-II. These composites measured the CHC broad abilities of Short-Term Memory, Visual Processing, Long-Term Storage and Retrieval, and Crystallized Ability. A one-way ANOVA was conducted to compare overall IQ (Fluid-Crystallized Index). Furthermore, another profile

analysis was run to compare the performance of African American and Caucasian children on the subtests of the KABC-II.

Table 4.4

T-Test Mean Comparison for the KABC-II Scores

FCI and Clusters	<i>t</i>	Significance	Mean Difference
FCI	-2.856	.005	-3.673
<i>Gsm</i>	-4.667	<.001*	-5.357
<i>Gv</i>	-1.686	.095	-2.316
<i>Glr</i>	-.805	.423	-1.163
<i>Gc</i>	-4.353	<.001*	-4.520
<u>Subtests</u>			
Number Recall	-3.491	.001*	-.776
Word Order	-4.946	<.001*	-1.061
Atlantis	-1.586	.116	-.469
Rebus	.000	1.000	.000
Conceptual Thinking	-.863	.390	-.286
Face Recognition	.000	1.000	.000
Triangles	-3.899	<.001*	-.980
Expressive Vocabulary	-2.087	.039	-.439
Riddles	-6.877	<.001*	-1.643

Note. A Test Value of 100 was used for the FCI and composites. A Test Value of 10 was used for all subtests. All scores have *df* = 97.

*Significant *t*-test comparisons.

Results from these analyses indicated African American and Caucasian preschool children did not differ in pattern or level of performance on the KABC-II composites. Additionally, no mean difference was found on the FCI indicating both groups displayed similar overall IQs. The Parallelism test for the subtest analysis was significant indicating African Americans and Caucasians differed in their patterns of highs and lows among the subtests. Follow-up ANVOAs corrected with the Scheffé adjustment revealed African Americans performed significantly lower than Caucasians on the Expressive Vocabulary subtest. The levels test was not significant indicating both groups performed similarly across the KABC-II subtests. Furthermore, *t*-tests were conducted to compare the research sample to the normative sample to help aid in generalizeability to a larger preschool population.

CHAPTER V

Discussion

This study investigated the interpretability of the Kaufman Assessment Battery for Children, Second Edition (KABC-II) with ethnically diverse preschool children through a profile analysis. The KABC-II was designed to reflect the contemporary Cattell-Horn-Carroll (CHC) hierarchical theory of intelligence, the most widely excepted theory of intelligence today (Flanagan et al., 2000). One objective of the present study was to determine if African American and Caucasian preschool children displayed similar patterns among the CHC factors on the KABC-II. Past research has shown the KABC-II adequately assesses multiple CHC abilities in preschool children (Hunt, 2007; Morgan, 2008; Reynolds, et al., 2007). Research also indicated the CHC factor structure is invariant across ethnic groups (Edwards & Oakland, 2006). One goal was to determine if African American and Caucasian preschool children displayed the same patterns of highs and lows on the CHC factors to aid in the interpretation of the test with these diverse populations. Additionally, this was one of the first independent studies to evaluate the use of the KABC-II with African American and Caucasian preschool students.

Investigation of the patterns of highs and lows between the groups at the composite and the subtest level adds further evidence to the interpretative nature of the KABC-II with preschool children. The publishers of the test, Kaufman and Kaufman

(2004), suggested the KABC-II be interpreted at the composite level with all age groups. The current research study evaluated the differences in patterns at both the subtest and composite levels to assess the appropriateness of the publishers' recommendation.

Another objective of this study was to determine if African Americans or Caucasians performed better than the other on the subtests and composites of the KABC-II. Research indicated Caucasians have historically performed one standard deviation better than African Americans on standardized measures of cognitive ability (Arinoldo, 1981; Rushton & Jensen, 2005; Tuttle, 1966). The discrepant scores pose problems with interpretation of the tests and question the use of cognitive ability tests with African American children. Specifically, African American children are identified as mentally retarded more often than Caucasian children when using a strict cut score of 70 (Ebersole & Kapp, 2007; Jensen, 1998; Naglieri & Rojahn, 2001; Skiba et al., 2006). The IQ gap has persisted over time, even with revisions of cognitive ability measures (McGurk, 1982; Shuey, 1966); therefore, current research focuses on whether the tests measure the same constructs between the two groups rather than evaluating if they are reducing the gap. This study aimed to determine if the KABC-II yielded similar gaps in cognitive ability as seen in the research. Controlling for the variables of age, sex, and parental education, attributes a greater confidence in the variable of race as a contributing factor to the results and corresponding interpretations of the profile analysis.

Summary of Results and Implications

Initial analysis of this sample's performance on the KABC-II confirmed an average performance indicating this sample was from the general population.

Research Question One

One objective of this study was to compare the performance of African American and Caucasian preschool children on the broad CHC factors as measured by the KABC-II composites. A profile analysis revealed African American and Caucasian preschool children display the same pattern of highs and lows across the CHC broad abilities when the variables of sex, age, and parental education were controlled. In other words, African Americans perform similarly to Caucasians on the constructs measured by the KABC-II composites. This research was consistent with previous research indicating the structure of the CHC model is invariant across different ethnic groups (Oakland & Edwards, 2006; Keith, et al., 1999; Kush et al., 2001). Furthermore, this study adds to the research by indicating similar test performance between African American and Caucasian children on the CHC broad abilities is displayed as early as the preschool years.

Additionally, the study aimed to determine if African American and Caucasian preschool children displayed similar levels of performance on the CHC broad abilities. The analysis indicated both groups obtained similar overall scores on the KABC-II composites. Caucasian preschool children do not perform significantly higher than African American preschool children on measures of the CHC broad abilities as measured by the KABC-II. This research provides further evidence that the KABC-II yielded smaller score differences between the ethnic groups. Previous research praised the K-ABC for significantly smaller differences between African Americans and Caucasian as compared to other IQ tests (Kaufman & Lichtenberger, 2002; Kaufman et.al, 2005), and the KABC-II showed similar findings.

The attention the KABC-II authors placed on ethnicity during the standardization process could be attributed to these positive findings. Like many the other tests the measure the CHC factors (e.g., Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, WJIII COG; Stanford Binet Intelligence Scales, Fifth Edition, SB-V), the standardization sample was selected to match the stratification variables of the United States. Unlike the WJIII COG, however, differential item functioning (DIF) was used during the KABC-II development to determine if different groups performed differently on the items. The WJIII COG only utilized expert review to determine bias. The WJIII COG's standardization data is also weaker because the different ethnicities are divided into white and non-white groups. The lack of differentiation causes the manual to lack validity evidence for specific ethnic groups. Follow-up studies indicated the WJIII COG yielded a 10.9 point mean difference in overall IQ between African Americans and Caucasians (Edwards & Oakland, 2006). The SB-V utilized DIF, but there is a lack of research regarding its utility with ethnically diverse populations.

Kaufman and Kaufman (2004) suggested clinicians interpret the KABC-II at the composite level. Specifically, Kaufman and colleagues (2005) indicated the global score provides a norm-based overall view of the child's performance and can serve as a comparison point to assess other abilities, but it does not tell anything about strengths and weakness in ability. Furthermore, Kaufman and colleagues (2005) stated "scores on specific subtests are of little value" (p.79) because they are intended to complement each other and provide a thorough measure of the theoretical construct of the composite. The current study provided evidence for the recommendation of interpretation at the composite level. Given that the same patterns of highs and lows were seen for both

groups along with similar mean performances on the composites, the constructs of *Gsm*, *Gv*, *Glr*, and *Gc* can be interpreted.

The use of intelligence tests with culturally diverse populations has been discussed widely in recent literature, and assessment recommendations have been made. Clinicians must be vigilant about using measures for which the background of the individual matches the normative sample of the test (Ortiz & Dynda, 2005). The authors of the KABC-II utilized a normative sample based upon stratification variables matching the United States population. This included a matching percentage of African Americans. Additionally, clinicians should choose tests that provide accurate psychometric data indicating that the constructs measured by the test are also measured in the population which the individual's background represents (Sattler, 2001). Additionally, recognition of potential bias within the test must be considered (Ortiz & Dynda, 2005). The current research study indicated interpretation at the composite level is accurate given that African American and Caucasian preschool children preformed similarly among the constructs in the area of level and pattern of scores. When making assessment decisions, clinicians can be more confident that the KABC-II will have similar interpretative quality at the broad ability level for African American and Caucasian children. This finding indicated the KABC-II may be an appropriate measure to choose when attempting to assess the cognitive ability of an African American preschooler.

Lastly, comparison of overall IQ (FCI) mean scores yielded no significant differences between African American and Caucasian preschool children. This is in contrast to previous research indicating a one standard deviation difference between the two groups on measures of cognitive ability (Arinoldo, 1981; Edwards & Oakland, 2006;

Jensen, 1998; Tuttle, 1966). Specifically, when using the WJIII COG, another modern test of intelligence that is based upon contemporary intellectual theory (CHC), Edwards and Oakland (2006) found a 10.9 point difference in the General Intellectual Ability between the African Americans and Caucasians included in the WJIII standardization sample. Preschoolers administered the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) displayed a 10.8 point difference on the Full Scale IQ., a 10.5 difference on the Verbal IQ, and an 8.9 point difference on the Performance IQ (Tuttle, 1966). Furthermore, research has found a decreased cognitive ability gap between African American and Caucasian preschool children when demographic variables of parental education and home environment were included (Smith, Duncan, & Lee, 2003), but the gap still remains between four to nine points. The current study indicated a 1.47 point, nonsignificant, mean difference between the two ethnic groups when controlling for sex, gender, and parental education. This difference represented one of the smallest mean differences found in the literature. Compared to other tests of intelligence, the KABC-II appears to be a more appropriate choice when assessing African American preschool children because no significant difference in overall IQ was found.

Research Question Two

Another objective of this study was to compare the performance of African American and Caucasian preschool children on the subtests of the KABC-II. According to test developers, the KABC-II subtests represent several of the CHC narrow abilities (refer to Table 3.2 in Chapter 3 for the narrow abilities measured by each subtest). A profile analysis indicated African American and Caucasian preschool children displayed significantly different patterns of highs and lows on the subtests of the KABC-II. Follow-

up analyses indicated the only significant difference in mean scores was seen on the Expressive Vocabulary subtest. This significant difference was considered small according to Cohen's (1988) guidelines on effect size, but any significant difference between two groups on a measure of cognitive ability warrants attention. The Expressive Vocabulary subtest asks children to say the name of a colored object presented on the stimulus page. According to the KABC-II manual (Kaufman & Kaufman, 2004), the Expressive Vocabulary subtest measures the CHC narrow ability of Lexical Knowledge. Lexical knowledge is defined as the "extent of vocabulary (nouns, verbs, or adjectives) that can be understood in terms of correct word (semantic) meaning" (McGrew, 2005, p. 151). In other words, lexical knowledge is referred to as vocabulary knowledge.

Little research has been conducted comparing African Americans and Caucasians on the CHC broad and narrow abilities with only one published study being found. Edwards and Oakland (2006) found the largest mean difference between African Americans and Caucasians across the CHC abilities was on the Verbal Comprehension subtest of the WJIII COG. According to the manual (McGrew & Woodcock, 2001), the Verbal Comprehension subtest measures two narrow abilities, Lexical Knowledge and Language Development, two highly correlated narrow abilities (McGrew, 2005). The results of the present study were consistent with this research indicating the largest mean difference in ability between African Americans and Caucasians may be related, in part, to Lexical Knowledge. Furthermore, this present study builds upon published research by indicating the difference in the CHC ability is seen as early as the preschool years.

The significant difference seen between African Americans and Caucasians on Expressive Vocabulary may be a result of that subtest's *g*-loading. Jensen (1998) found

that the mean IQ gap was greater in tests with higher *g*-loadings as compared with tests that had lower *g*-loadings. Previous research with preschool children from the same archival data as the present study indicated the *g*-loading for Expressive Vocabulary is .75 (Hunt, 2007; Morgan, 2008). According to Kaufman's (1994) classification of *g*-loadings, those higher than .70 indicate a "good" measure of *g*. Additionally, Hunt (2007) and Morgan (2008) indicated the Expressive Vocabulary subtest was a strong indicator of crystallized ability (*Gc*), and *Gc* is an excellent measure of *g* with a factor loading of .95. This hypothesis would need to be studied in future research as other subtests with high *g*-loadings (i.e., Word Order) did not show significant differences between the African American and Caucasian groups.

The present study was also consistent with past research showing African Americans perform poorer than Caucasians on tests of verbal ability and vocabulary (Bracken, Howell, & Crain, 1993; Kaufman, McLean, & Kaufman, 1995; Kaufman, McLean, & Reynolds, 1988; Sellers et al., 2002). Specifically of interest, Kaufman and colleagues (1995) studied the performance of African Americans and Caucasians on crystallized and fluid ability as defined by the Horn-Cattell *Gf-Gc* theory. They found African American children and adolescents performed 10.7 points lower than Caucasians on measures of crystallized ability, and the results were maintained after educational attainment of their parents was added as a covariate. The present study further added that the difference in narrow crystallized ability (lexical knowledge) could be seen as early as the preschool years, and existed when parental education is controlled. Additionally, the present research found a 1.04 point difference between the two groups on the Expressive Vocabulary subtest, which was lower than the historical two to four point difference

between African Americans and Caucasians on previous measures (Kaufman, McLean, & Reynolds, 1988; Munford, Meyerowitz, & Munford, 1980).

Furthermore, the study aimed to determine if African American and Caucasian preschool children displayed similar levels of performance on the CHC narrow abilities. Results indicated Caucasians and African Americans did not differ in level across the subtests. Their similar performance on the subtests as a whole was in contrast with previous research documenting two to four point differences in Caucasian and African American children (Munford, Meyerowitz, & Munford, 1980) and provided further evidence for significantly smaller differences on the KABC-II between the two ethnic groups as compared to other measures (Kaufman & Kaufman, 2004; Kaufman & Lichtenberger, 2002; Kaufman et.al, 2005).

Results from the profile analysis at the subtest level further support Kaufman and Kaufman's (2004) recommendation of interpreting the KABC-II at the composite level. Kaufman and colleagues (2005) indicated when differences on the subtests of a broad ability occurred, the clinician could formulate hypotheses and verify those hypotheses with other information. They recommend a clinician not formulate assessment decisions from the scores of individual subtests. Composite interpretation is preferred because the subtests were intended to complement each other by providing a more thorough picture of the theoretical construct measured in the composite (Kaufman et al., 2005). Subtest interpretation can be beneficial as it provides information on the student's performance on the narrow abilities and provide information on the individual's strengths and weaknesses. Without focus on the subtests, the information on an individual's performance on the CHC narrow abilities would be missed. Interpretation at the subtest

level may be problematic given the difference found on the narrow ability measured by Expressive Vocabulary. This research can further aid clinicians when interpreting the KABC-II for ethnically diverse preschool children.

In summary, when evaluating ethnically diverse preschool children, clinicians must be aware of the standardization procedures for the measure they decide to use as well as its interpretive quality. Practitioners should be aware of the cultural background of their examinees and determine how well that background matches the normative sample of the test (Ortiz & Dynda, 2005). The authors of the KABC-II appear to have taken more steps than other tests measuring the CHC factors to ensure accurate representation of African Americans and evaluate item bias. Furthermore, they included normative means on ethnically diverse groups which can aid clinicians when making assessment decisions. The current research study suggested the KABC-II is an appropriate test for use with African American preschool children. Psychologists should follow Kaufman & Kaufman's (2004) recommendation of interpretation at the composite level and can feel more confident African American preschool children do not perform significantly different than Caucasians. If psychologists want to develop hypotheses of the strengths and weakness of an examinee's narrow abilities, they should use caution when evaluating a child's performance on the Expressive Vocabulary subtest. All other subtests evaluated in this study showed similar performances between African Americans and Caucasians further aiding in generating accurate hypotheses about an individual's cognitive abilities.

Follow-Up Analyses

Comparison between the study's sample and the standardization sample indicated the current sample performed significantly lower in a few areas. Specifically, both African Americans and Caucasians performed lower than their respective group from the standardization sample on the *Gsm* broad ability and the Riddles and Number Recall subtests. Additionally, the Caucasian preschool children from the study sample performed lower than the Caucasian preschool children in the standardization sample on the Triangles subtest and the *Gc* broad ability. African Americans in the sample performed lower on the Word Order subtest as compared to their counterparts from the standardization sample.

A few hypotheses were formed to explain these results. First, these differences may be a result of a sample population whose parents were less educated than the children included in the standardization sample. Although the sample was adequately matched on parental education by race to control for the effect of education on the profile analysis, the overall sample appears to have fewer parents who received four or more years of education as compared to the standardization sample. Previous research indicated lower parental education level corresponded with lower performance of their preschool children on measures of verbal, nonverbal, and overall IQ (Sellers, Burns & Guyrke, 2002; Brooks-Gunn et al., 2003). Although some broad and narrow abilities differed from the standardization sample, it is important to remember that the sample includes children from the general education population who scored within the average range of test performance. Furthermore, the performance of both groups on the majority of subtests and composites was similar to the standardization sample.

Second, several of the differences between the sample and the normative population were in the area of memory. Research has indicated the ability to sustain attention and inhibit responses is associated with memory (Youngwirth, Harvey, Gates, Hashim, & Friedman-Weieneth, 2007), and preschool-aged children do not sustain alertness as long as older children (Morrison, 1982). The weaker performance on the memory tasks by the sample may be associated with the attention level of the participants. Upon examination of the ranges of scores on the subtests and composites, it appeared the sample had variable performances among the subtests and composites. This variability may be related to the inability of some of the participants to sustain attention throughout the assessment. Given that attention has been associated with memory, the memory subtests may have been impacted the greatest by the participants' ability to sustain attention. Additionally, children's ability to sustain attention during the preschool years increases significantly between the ages of 3 and 7 (Gathercole, 1992; Hrabok, Kerns & Muller, 2007). Hrabok and colleagues (2007) found that children ages 3-years, 5-months performed significantly worse on a computerized measure of attention as compared to children ages 4-years, 5-months. The current sample may have more children with under-developed attention skills as compared to the general population.

Furthermore, the significant difference in *Gsm* between the current sample and the normative population may be associated with the significant differences seen on *Gc* and related subtests. Research indicates memory span in preschool children predicts vocabulary knowledge (Avons, Wragg, Cupples, & Lovegrove, 1998). Additionally, higher recall of words and non-words on measures of phonological short-term memory was associated with higher vocabulary knowledge (Thorn & Gathercole, 1999). Given

the subtests that comprise *Gsm* measure the narrow ability Memory Span, this research supported the hypothesis that *Gc* may be lower as a function of *Gsm*. This hypothesis should be studied in future research.

Lastly, the overall sample, as well as the Caucasian subsample, performed significantly weaker on the Triangles subtest as compared to the standardization sample. Test fatigue and inability to sustain attention may be associated with the lower performance as this subtest is administered near the end of the test. Additionally, this difference in performance may be associated the nature of the Triangles task and the Spatial Relations and Visualization narrow abilities it measures. Previous research indicated 4-year old children were better at visual relations tasks that involve abstract stimuli rather than iconic stimuli (DeLoache, 2000; Koerber & Sodian, 2008). Iconic stimuli are realistic pictures or three-dimensional objects of real items, whereas abstract stimuli are not associated with real items. Three to five-year olds begin the Triangles subtest by manipulating colorful shapes to match a model created by the examiner. If they do not discontinue the subtest, they begin using blue and yellow foam triangles at item 11, and attempt to create linear, abstract shapes from a stimulus page. Kaufman and Kaufman (2004) state, “several of the easiest items resemble familiar objects to make the tasks less abstract” (p. 65), and created these items because the original K-ABC foam triangle items were too difficult for the youngest examiners (3-year olds). DeLoache’s (2000) and Koerber and Sodian’s (2008) research isolated 4-year olds as having strengths in different spatial relations tasks (abstract) as compared to other preschool children. The iconic nature of the earliest items of the KABC-II may have been difficult for the current

research sample's 4-year olds. Additional research should be conducted to evaluate the narrow abilities of Spatial Relations and Visualization in preschool children.

Limitations of the Present Study

The present study included some limitations regarding the composition of the sample and the ability to generalize to other populations. For instance, the sample was a sample of convenience obtained from a Midwestern city which limits its applicability to other regions of the country. Additionally, the sample included only children ages 4-years, 0-months, to 5-years, 11-months and results may not generalize to younger children or older children and adults. Results also may not be applicable to children receiving special education services as those children were excluded from the study.

Limitations were also present in attempting to match the sample to control for demographic variables. Given the limited number of African Americans, all were included in the study to maximize data, and therefore matching was limited to the fixed nature of the race variable. The most difficult variable to match was parental education. Although more Caucasians than African Americans participated in the study, and an ideal sample would have included more children for matching purposes. Ideally, all participants would have matched exactly. An adequate match was made with as 81.6% of the sample matching on parental education. Matching the participants in such a manner may have reduced the number of parents who obtained four or more years of college. The percentage of parents receiving four or more years of college was lower in the matched sample than in the same-aged group in the normative sample. Results from the present study may be limited in the ability to generalize to children whose parents have higher education.

The present research was also limited to the CHC broad and narrow abilities that are measured by the KABC-II. Specifically, only the CHC broad abilities of *Gsm*, *Gv*, *Glr*, and *Gc* are measured in preschool children. Moreover, the Pattern Reasoning subtest was excluded from the analysis because it was only administered to the 5-year olds (KABC-II administration restricts its use with 4-year olds). Pattern Reasoning is a core subtest for 5-year olds, and conclusions regarding its use with African Americans were lost through the exclusion of the subtest. Additionally, results from this study can only be generalized to the ethnic groups included in the study.

Recommendations for Future Research

Initial findings indicated the African American and Caucasian preschool children performed lower on some subtests and composites as compared to their respective group from the normative sample. A preliminary hypothesis suggests the samples' scores may be lower compared normative sample because fewer parents with four or more years of education were included in the matched sample than in the standardization sample. This means that the matched sample may not approximate U.S. population on parental education. To study these comparisons further and provide more specific reasons why the findings may exist, the sample should be compared to the data from the normative sample. For purposes of this study, only mean scores were available and used as test values in the analysis. More accurate statistical analysis would be obtained from comparing the data in the study sample to the normative data from the publisher of the KABC-II.

Additional research could focus on other cognitive ability instruments that measure CHC factors to see whether African Americans and Caucasians show similar

patterns of performance on other CHC broad and narrow abilities. Analysis of several other cognitive ability tests would be required as no one test has been developed to measure all of the abilities described by the CHC model (Alfonso et al., 2005). Furthermore, the current study only included a small age range and research should be conducted to determine if similar patterns of performance between African Americans and Caucasians in other age groups.

Moreover, the present study supports previous conclusions that the KABC-II is appropriate for use with ethnically diverse preschool children (Kaufman et al., 2005). Additional independent studies on the KABC-II with other age groups should be conducted to further support the use of the instrument with African Americans of different ages. Similar research can be conducted on other diverse cultures (e.g., Hispanics, Asian Americans, etc.) to support the use and accurate interpretation of the KABC-II with other ethnicities common in the United States.

Chapter Summary

In conclusion, the ultimate goal of the present study was to compare the performance of African American and Caucasian preschool children on the Kaufman Assessment Battery for Children, Second Edition (KABC-II). Specifically, the performance of the two groups was compared to examine group differences on the underlying constructs of the test in regards to the Cattell-Horn-Carroll (CHC) theory of intelligence. This study was unique from previous research because it was one of the first to compare the performance of African American and Caucasian preschool children on a newly revised measure of cognitive ability designed according to contemporary intellectual theory. Current research has drifted away from the historical measurement of

IQ differences between African Americans and Caucasians. Therefore, little is known about whether newer versions of cognitive ability measures specifically designed according to theory show the same gap in intelligence as previously documented.

Performance profiles at the composite (broad ability) and subtest (narrow ability) levels were generated and provided evidence supporting the test author's recommendation that the instrument be interpreted at the composite level. Results confirmed that Caucasian and African American children had similar patterns and levels of performance on the CHC broad abilities as measured by the KABC-II. Moreover, differences in performance were found at the subtest level indicating clinicians should be cautious about interpreting the Expressive Vocabulary subtest scores for African American children. Specifically, the current research supported the findings that African Americans perform significantly lower on measures of vocabulary (Kaufman, McLean, & Kaufman, 1995).

The present study contributes to the field of assessment by providing evidence of a decreased IQ gap. Specifically, overall intelligence scores were not significantly different between the groups and indicated one of the smallest gaps displayed in research. Additionally, the results confirmed that the two groups did not differ in performance of the broad abilities of *Gsm*, *Gv*, *Glr*, and *Gc*. The results suggested that the KABC-II is a valuable tool for the assessment of preschool cognitive ability.

References

- Alfonso, V.C., & Flanagan, D. P. (2007). Best practices in the use of the Stanford-Binet Intelligence Scales, Fifth Edition with preschoolers. In B. Bracken, & R. Nagle (Eds.), *Psychoeducational assessment of preschool children* (pp. 267-296). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Alfonso, V.C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 185-202). New York: Guilford Press.
- Arinoldo, C.G. (1981). Black-white differences in the general cognitive index of the McCarthy Scales and in the full scale IQs of Wechsler's scales. *Journal of Clinical Psychology*, 37(3), 630-638.
- Avons, S.E., Wraggs, C.A., Cupples, A., & Lovegrove, W.J. Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics*, 19(4), 583-601.
- Bane, M. J., & Ellwood, D. T. (1989). One fifth of the nation's children: Why are they poor? *Science*, 245, 1047-1053.
- Bickley, P.G., Keith, T.Z., & Wolfle, L.M. (1995). The three stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence*, 20, 309-328.
- Blaha, J., & Wallbrown, F. H. (1991). Hierarchical factor structure of the Wechsler Preschool and Primary Scale of Intelligence-Revised. *Psychological Assessment*, 3(3), 455-463.

- Bossard, M. D., Reynolds, C. R., & Gutin, T. B. (1980). A regression analysis of test bias on the Stanford-Binet intelligence Scale for black and white children referred for psychological services. *Journal of Clinical Child Psychology*, 2, 52-54
- Bracken, B.A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. *Journal of Psychoeducational Assessment*, 4, 313-326.
- Bracken, B.A. (1994). Advocating for effective preschool assessment practices: A comment on Bagnato and Neisworth. *School Psychology Quarterly*, 9(2), 103-108.
- Bracken, B.A., Howell, K.K., & Crain, R.M. (1993). Prediction of Caucasian and African American preschool children's fluid and crystallized intelligence: Contribution of maternal characteristics and home environment. *Journal of Clinical and Child Psychology*, 22(4), 455-464.
- Brody, N. (1992). *Intelligence*, 2nd ed. San Diego, CA: Academic Press, Inc.
- Brody, N. (2000). Theories and measurements of intelligence. In R. J. Spearman, (Ed.), *Handbook of intelligence* (pp. 16-33). New York: Cambridge University Press.
- Brooks-Gunn, J., Klebanov, P. K., & Duncan, G. J. (1996). Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development*, 67, 396-408.
- Brown, R. T., Reynolds, C.R., & Whitaker, J. S. (1999). Bias in mental testing since Bias in Mental Testing. *School Psychology Quarterly*, 14(3), 208-238.
- Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences*, 19, 149-155.

- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. Flanagan, & P. Harrison, (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 69-76). New York: Guilford Press.
- Carroll, J.B., & Horn, J.L. (1981). On the scientific basis of ability testing. *American Psychologist*, 36, 1012-1020.
- Cattell, R.B. (1941). Some theoretical issues in adult intelligence testing. *Psychological Bulletin*, 38, 592.
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1-22.
- Cattell, R.B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Cattell, R. B., & Horn, J. L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*, 15(3), 139-164.
- Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators or early reading achievement: Evidence from the early longitudinal study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology*, 98(3), 489-507.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum Associates.

- Davidson, J. E., & Downing, C. L. (2000). Contemporary models of Intelligence. In R. J. Spearman, (Ed.), *Handbook of Intelligence* (pp. 34-49). New York: Cambridge University Press.
- DeLoache, J.S. (2000). Dual representation and young children's use of scale models. *Child Development*, 71(2), 329-338.
- Ebersole, J.L., & Kapp, S.A. (2007). Stemming the tide of overrepresentation: Ensuring accurate certification of African American students in programs for the mentally retarded. *School Social Work Journal*, 31, 1-16.
- Edwards, O.E., & Oakland, T.D. (2006). Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasian Americans. *Journal of Psychoeducational Assessment*, 24, 358-366.
- Edwards, O.W., & Fuller, D.P. (2005). Implications of the Cattell-Horn-Carroll theory on ethnic differences in IQ. *Psychological Reports*, 97(3), 891-897.
- Elliot, C. D. (1990). *Differential Ability Scales*. San Antonio, TX: Psychological Corporation.
- Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law*, 6(1) 168-170.
- Fan, X., Willson, V. L., & Reynolds, C. R. (1995). Assessing the similarity of the factor structure of the K-ABC for African American and white children. *Journal of Psychoeducational Assessment*, 13, 120-131.
- Flanagan D.P., & McGrew, K.S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: Joint confirmatory factor analysis of the WJ-R and the KAIT in a non-white sample. *Journal of school psychology*, 36(2), 151-182.

- Flanagan, D.P., Mascolo, J., & Genshaft, J.L. (2000). A conceptual framework for interpreting preschool intelligence tests. In B.A. Bracken (Ed.), *The psychoeducational assessment of preschool children* (3rd ed., pp. 428-473). Boston: Allyn & Bacon.
- Ford, L., & Dahinten, S. (2005). Use of intelligence tests in the assessment of preschoolers. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 487-593). New York: Guilford Press.
- Galton, F. (1869). *Heredity genius: An inquiry into its laws and consequences*. London: Macmillan.
- Gardner, M. K., & Clark, E. (1992). The psychometric perspective on intellectual development in childhood and adolescence. In Sternberg, R. J. & Berg, C. A. (Eds.), *Intellectual development*. New York: Cambridge University Press.
- Gathercole, S.E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, 39, 3-27.
- Gesell, A. (1925). *The mental growth of the preschool child: A psychological outline of normal development from birth to the sixth year*. New York: Macmillan.
- Goodenough, F. L., & Mauer, K. M. (1942). *The mental growth of children from two to fourteen years*. Minneapolis: The University of Minnesota Press.
- Guilford, J. P. (1985). The structure-of-intellect model. In B. B. Wolman (Ed.), *Handbook of intelligence: Theories measurements and applications* (pp. 225-266). New York: John Wiley & Sons.
- Herrnstein, R. J., & Murray, C. (1994). *The bell curve*. New York: Free Press.

- Harris, D. N., & Herrington, C. D. (2006). Accountability, Standards and the growing achievement gap: Lessons from the past half-century. *American Journal of Education, 112*, 209-238.
- Holdnack, J. A., & Weiss, L. G. (2006). IDEA 2004: Anticipated implications for clinical practice—integrating assessment and intervention. *Psychology in the Schools, 43*(8), 871-882.
- Horn, J.L. (1968). Organization of abilities and the development of intelligence. *Psychological Review, 75*, 242-259.
- Horn, J. L. (1976). Human abilities: A review of research and theory in the early 1970s. *Annual Review of Psychology, 27*, 437-485.
- Horn, J.L. (1991). Measurement of intellectual capabilities: A review of theory. In K.S. McGrew, J.K. Werder, & R.W. Woodcock (Eds.), *WJ-R technical manual* (pp. 197-232). Chicago: Riverside.
- Horn, J.L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. Flanagan, & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 41-68). New York: Guilford Press.
- Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253-270.
- Horn, J.L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (pp. 53-91) New York: The Guilford Press.
- Hrabok, M., Kerns, K.A., & Muller, U. (2007). The vigilance, orienting, and executive attention networks in 4-year-old children. *Child Neuropsychology, 13*, 408-421.

- Hunt, M.S. (2007). A joint confirmatory factor analysis of the Kaufman Assessment Battery for Children, Second Edition, and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, with preschool children. Unpublished doctoral dissertation, Ball State University.
- Hynd, G. W., & Semrud-Clikeman, M. (1993). Developmental considerations in cognitive assessment of young children. In Culbertson, J. L., & Willis, D. J. (Eds.), *Testing young children: A reference guide for developmental, psychoeducational, and psychological assessments*. Austin: Pro-ed.
- Jacob, S., & Hartshorne, T. S. (2003). *Ethics and the law for school psychologists*. Hoboken, NJ: John Wiley & Sons, Inc.
- Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, 39, 1-123.
- Jensen, A. R. (1974). How biased are culture-loaded tests? *Genetic Psychology Monographs*, 90, 185-244.
- Jensen, A.R. (1980). *Bias in mental testing*. New York: The Free Press.
- Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kamphaus, R.W., Winsor, A.P., Rowe, E.W., & Kim, S. (2005). A history of intelligence test interpretation. In D.P. Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues*, 2nd edition (pp. 23-38). New York: Guilford Press.
- Kaufman, A.S., Kamphaus, R.W., & Kaufman, N.L. (1985). New directions in intelligence testing: The Kaufman Assessment Battery for Children (K-ABC). In

- B. B. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 663-698). New York: John Wiley & Sons.
- Kaufman, A. S. (1979). *Intelligence testing with the WISC-R*. New York: Wiley-Interscience.
- Kaufman, A.S. (1994). *Intelligence testing with the WISC-III*. New York: John Wiley and Sons.
- Kaufman, A. S. (2000). Tests of Intelligence. In R. J. Sternberg (Ed.) *Handbook of intelligence* (pp. 445- 476). New York: Cambridge University Press.
- Kaufman, A. S., & Kaufman, N. L. (1983). *K-ABC interpretive manual*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., & Kaufman N. L. (2004). *Manual for the Kaufman Assessment Battery for Children—Second Edition (KABC-II) Comprehensive Form*. Circle Pines, MN: American Guidance Service.
- Kaufman, A. S., & Lichenberger E. O. (2002). *Assessing adolescent and adult achievement* (2nd edition). Boston, MA: Allyn & Bacon.
- Kaufman, A.S., Lichenberger, E.O., Fletcher-Janzen, E., & Kaufman, N. (2005). *Essentials of KABC-II assessment*. Hoboken, NJ: John Wiley & Sons, Inc.
- Kaufman, A.S., McLean, J.E., & Kaufman, J.C. (1995). The fluid and crystallized abilities of white, black, and Hispanic adolescents and adults, both with and without an education covariate. *Journal of Clinical Psychology*, 51(5), 636-647.
- Kaufman, A.S., McLean, J.E., & Reynolds, C.R. (1988). Sex, race, residence region, and education differences on the 11 WAIS-R subtests. *Journal of Clinical Psychology*, 44, 231-248.

- Keith, T., Cool, V., Novak, C., White, L., & Pottebaum, S. (1988). Confirmatory factor analysis of the Stanford-Binet Fourth Edition: Testing the theory-test match. *Journal of School Psychology, 26*, 253-274.
- Keith, T., Quirk, K., Schartzer, C., & Elliot, C. (1999). Construct bias in the Differential Ability Scales? Confirmatory and hierarchical factor structure across three ethnic groups. *Journal of Psychoeducational Assessment, 17*, 249-268.
- Kelley, M. F., & Surbeck, E. (2007). History of preschool assessment. In Bracken, B., & Nagle, R. (Eds.), *Psychoeducational assessment of preschool children* (pp. 3-28). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Kenny, T. J., & Culbertson, J. L. (1993). Developmental screening for preschoolers. In Culbertson, J. L., & Willis, D. J. (Eds.), *Testing young children: A reference guide for developmental, psychoeducational, and psychological assessments*. Austin: Pro-ed.
- Koerber, S., Sodian, B. (2008). Preschool children's ability to visually represent relations. *Developmental Science, 11*(3), 390-395.
- Kush, J. C., & Watkins, M. W. (1997). Construct validity of the WISC-III verbal and performance factors for black special education students. *Assessment, 4*(3), 297-304.
- Kush, J.C., Watkins, M.W., Ward, T.J., Ward, S.B., Canivez, G.L., & Worrell, F.C. (2001). Construct validity for the WISC-III for black and white students from the WISC-III standardization sample and for black students referred for psychological evaluation. *School Psychology Review, 30*(1), 70-88.

- Lichtenberger, E. O. (2005). General measures of cognition for the preschool child. *Mental Retardation and Developmental Disabilities Research Reviews, 11*, 197-208.
- Lichtenberger, E. O., & Kaufman A. S. (2007). The assessment of preschool children with the Kaufman Assessment Battery for Children-Second Edition (KABC-II). In B. Bracken, & R. Nagle (Eds.), *Psychoeducational assessment of preschool children* (pp. 297-323). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- LoBello, S. G., & Gulgoz, S. (1991). Factor analysis of the Weschler Preschool and Primary Scale of Intelligence-Revised. *Psychological Assessment, 3*(1), 130-132.
- Loehlin, J. C., Lindzey, G., & Spuhler, J. N. (1975). *Race Differences in Intelligence*. San Francisco: W.H. Freeman and Company.
- Long, P. A., & Anthony, J. J. (1974). The measurement of mental retardation by a culture-specific test. *Psychology in the Schools, 11*, 310-312.
- Matarazzo, J. D., & Wiens, A. N. (1977) Black Intelligence Test of Cultural Homogeneity and Wechsler Adult Intelligence Scale scores of Black and White police applicants. *Journal of Applied Psychology, 62*, 57-63.
- McGrew, K. S. (1997). Analysis of major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 151-179). New York, NY: The Guilford Press.
- McGrew, K.S. (2005). The Cattell-Horn-Theory of cognitive abilities: Past, present, and future. In D. Flanagan, & P. Harrison, (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 136-181). New York: Guilford Press.

- McGrew, K S., & Flanagan, D.P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc cross-battery assessment*. Needham Heights, MA: Viacom.
- McGrew, K.S., & Woodcock, R.W. (2001). *Woodcock Johnson III Technical Manual*. Itasca, IL: Riverside Publishing.
- McGurk, F. C. (1982). Preschool children. In R. T. Osborne, & F. C. J. McGurk (Eds). *The testing of Negro intelligence*. Athens, GA: The foundation for Human Understanding.
- Morgan, K.E. (2008). The validity of intelligence tests using the Cattell-Horn-Carroll model of intelligence with a preschool population. Unpublished doctoral dissertation, Ball State University.
- Morrison, F.J. (1982). The development of alertness. *Journal of Experimental and Child Psychology*, 34, 187-199.
- Munford, P.R., Meyerowitz, B.E., & Munford, A.M. (1980). A comparison of black and white children's WISC/WISC-R differences. *Journal of Clinical Psychology*, 36(2), 471-475.
- Nagle, R. J. (2007). Issues in preschool assessment. In Bracken, B., Nagel, R. (Eds.), *Psychoeducational assessment of preschool children* (pp. 29-48). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Naglieri, J. & Das, J.P. (1997). *Cognitive Assessment System*. Itasca, IL: Riverside.
- Naglieri, J. & Rojahn, J. (2001). Intellectual classification of black and white children in special education programs using the WISC-III and the Cognitive Assessment System. *American Journal of Mental Retardation*, 106(4), 359-367.

- Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Nuttall, E. V., Romero, I., & Kalesnik, J. (1999). *Assessing and screening preschoolers: Psychological and educational dimensions*. Boston: Allyn & Bacon.
- Ofiesh, N. (2006). Response to intervention and the identification of specific learning disabilities: Why we need comprehensive evaluations as part of the process. *Psychology in the Schools*, 43(8), 883-888.
- Ortiz, S.O., & Dynda, A.M. (2005). Use of intelligence tests with culturally and linguistically diverse populations. In D.P Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 545-556). New York: The Guilford Press.
- Osborne, R. T. (1980). *Twins: Black and white*. Athens, GA: Foundation for Human Understanding.
- Poteat, G.M., Wuensch, K.L., & Gregg, N.B. (1988). An investigation of the differential prediction with the WISC-R. *Journal of School Psychology*, 26, 59-68.
- Reynolds, C. R., & Kaufman, A. S. (1985). Clinical assessment of children's intelligence with the Wechsler scales. In B. J. Wolman (Ed.), *Handbook of intelligence: Theories, measurements, and applications* (pp. 601-661). New York: John Wiley & Sons.
- Reynolds, M.R., Keith, T.Z., Fine, J.G., Fisher, M.E., & Low, J.A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children, Second Edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly*, 22, 511-539.

- Roid, G. H. (2003). *Stanford Binet Intelligence Scales: Fifth Edition: Technical Manual*. Itasca, IL: Riverside Publishing.
- Rury, J.L. (1988). Race, region, and education: An analysis of black and white scores on the 1917 Army Alpha Intelligence Test. *Journal of Negro Education*, 57, 51-65.
- Rushton, J.P., & Jensen, A.R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public, Policy and Law*, 11(2), 235-294.
- Sandoval, J., Zimmerman, I.L., & Woo-Sam, J.M. (1983). Cultural differences on WISC-R verbal items. *Journal of School Psychology*, 21, 49-55.
- Sattler, J.M. (1992). Assessment of children's intelligence. In C.E. Walker, & M.C. Roberts (Eds.), *Handbook of clinical child psychology* (pp. 85-100). New York: John Wiley & Sons, Inc.
- Sattler, J.M. (2001). *Assessment of children: Cognitive applications, fourth edition*. San Diego, CA: Jerome M. Sattler, Publisher, Inc.
- Scarr, S., & Weinberg, R.A. (1976). IQ test performance of black children adopted by white families. *American Psychologist*, 31(10), 726-739.
- Sellers, A.H., Burns, W.J., & Guyrke, J. (2002). Differences in young children's IQs on the Wechsler Preschool and Primary Scale of Intelligence-Revised as a function of stratification variables. *Applied Neuropsychology*, 9(2), 65-73.
- Shuey, A.M. (1966). *The testing of Negro intelligence*, 2nd ed. New York: Social Science Press.
- Siegler, R.S. (1991). An introduction to children's thinking. In R.S. Siegler (Ed.), *Children's thinking* (2nd ed.) (pp. 1-16). Englewood Cliffs, NJ: Prentice Hall.

- Skiba, R., Simmons, A., Ritter, S., Kohler, K., Henderson, M., & Wu, T. (2006). The context of minority disproportionality: Practitioner perspective on special education referral. *Teachers College Record*, 108(7), 1424-1459.
- Smith, J., Duncan, G. J., & Lee, K. (2003). The black-white test scores gap in young children: Contributions of test and family characteristics. *Applied Developmental Science*, 7(4), 239-252.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. New-York: Macmillan.
- Suen, H. K., & French, J. L. (2003). A history of the development of psychological and educational testing. In Reynolds & Kamphaus (Eds.), *Handbook of psychological and educational assessment of children: Intelligence, aptitude, and achievement* (pp. 3-23). New York: The Guilford Press.
- Suzuki, L.A., & Valencia, R.R., (1997). Race-ethnicity and measured intelligence: Educational implications. *American Psychologist*, 52(10), 1103-1114.
- Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics, fourth edition*. Boston: Allyn & Bacon.
- Taub, G.E., & McGrew, K.S. (2004). A confirmatory factor analysis of Cattell-Horn-Carroll theory and cross-age invariance of the Woodcock-Johnson tests of cognitive abilities III. *School Psychology Quarterly*, 19(1), 72-87.
- Terman, L.M. (1916). *The measurement of intelligence*. Boston: Houghton Mifflin.
- Terman, L.M., & Merrill, M.A. (Eds.). (1960). *Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M*. Boston: Houghton Mifflin.

- Thorn, A., & Gathercole, S.E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 52A(2), 303-324.
- Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). *The Stanford-Binet Intelligence Scale: Fourth Edition*. Itasca, IL: Riverside.
- Tusing, M.E., & Ford, L. (2004). Examining preschool cognitive abilities using a CHC framework. *International Journal of Testing*, 4(2), 91-114.
- Tusing, M.E., Maricle, D.E., & Ford, L. (2003). Assessment with the Woodcock-Johnson III and young children. In F.A. Schrank & D.P. Flanagan (Eds.), *WJIII clinical use and interpretation: Scientist-practitioner perspectives* (pp. 243-281). San Diego, CA: Academic Press.
- Tuttle, L.D. (1966). The comparative effects on intelligence test scores of Negro and white children when certain verbal and time factors are varied. In A. Shuey (Ed.), *The testing of Negro intelligence (2nd ed.)*. New York: Social Science Press.
- United States Census Bureau publications (2001). Census 2001 Summary File 1 United States. Washington, DC: Author.
- Vincent, K.R. (1991). Black/white IQ differences: Does age make the difference? *Journal of Clinical Psychology*, 47(2), 266-270.
- Vinovskis, M. (2005). *The birth of Head Start: Preschool education policies in the Kennedy and Johnson administrations*. Chicago: University of Chicago Press.
- Wasserman, J. D., & Tulskey, D. S. (2005). A history of intelligence assessment. In D. Flanagan, & P. Harrison, (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 3-22). New York: Guilford Press.

- Wechsler, D. (1939). *Wechsler-Bellevue Intelligence Scale*. New York: Psychological Corporation.
- Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.
- Wechsler, D. (1955). *Wechsler Adult Intelligence Scale*. New York: The Psychological Corporation.
- Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—Third Edition*. San Antonio, TX: Psychological Corporation.
- Wellman, B. L. (1932). The effects of preschool attendance upon the IQ. *Journal of Experimental Education*, 1, 48-49.
- Woodcock, R.W., & Johnson, M. B. (1989). *Woodcock-Johnson Tests of Cognitive Ability—Revised*. Chicago: Riverside.
- Youngwirth, S.D., Harvey, E.A., Gates, E.C., Hashim, R.L., & Friedman-Weieneth, J.L. (2007). Neuropsychological abilities of preschool-aged children who display hyperactivity and/or oppositional-defiant behavior problems. *Child Neuropsychology*, 13, 422-443.
- Zhu, J., & Weis, L. (2005). The Wechsler scales. In D. Flanagan, & P. Harrison, (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 297-324). New York: Guilford Press.

APPENDIX A

Letter Given to Parents of the Participants

Exploring CHC Theory's Cognitive Constructs in Young Children

Parent Letter

Dear Parent or Guardian:

We have recently been discussing an educational project with your daycare/preschool director. This letter is being sent to you, since your child is 4 or 5 years old and is enrolled in one of the participating daycares or preschools, to describe our intention and ask for your cooperation.

The project will focus on exploring how well three measures, the Stanford Binet, Fifth Edition, Kaufman Assessment Battery for Children, Second Edition, and Woodcock-Johnson Tests of Cognitive Abilities, Third Edition, assess similar thinking skills (such as memory, verbal and non-verbal reasoning) in young children. The questions asked will include such tasks as following directions, counting blocks, and identifying pictures. Your child's scores will remain confidential and will only be used to assist Ball State University graduate students with their dissertations.

If you approve, activities will last approximately 4 hour in total and will be completed over the course of 3 days while your child is in daycare/preschool. School psychology graduate students and professors from Ball State University will individually administer the 3 measures to your child in a one-to-one setting. The examiners have completed daycare state licensing regulations (i.e., references, criminal background check, physical examination showing the BSU student to be in good health and free of contagious diseases, and child abuse awareness training) prior to testing your child. To show our gratitude, your child's preschool class will receive a field trip to the Muncie Children's Museum upon completion of testing. The testing will be periodically supervised by the Ball State University professors.

Other 4 and 5 year olds have enjoyed participating in similar types of activities. Please sign the enclosed form, indicating whether you agree to have your child participate and return the form to the classroom teacher. These forms are necessary so that we are sure the parents and child understand the nature of the project. The parents and child are free, of course, to withdraw at any time by contacting one of us.

We will be glad to answer any questions you may have concerning the project. We can be reached at (765) 285-8500 or by email at keward@bsu.edu, demcintosh@bsu.edu, or brothlisberg@bsu.edu.

Sincerely,

Kimberly Ward, M.A., Jaime Anderson, M.A., Madeline Hunt, M.A.
Principal Investigators
Department of Educational Psychology
Muncie , Indiana 47306

David E. McIntosh, Ph.D.
Professor/Co-Principal Investigator
Department of Educational Psychology
Muncie , Indiana 47306

Barbara A. Rothlisberg, Ph.D.
Professor/Co-Principal Investigator
Department of Educational Psychology
Muncie , Indiana 47306

APPENDIX B

Permission Form Completed by Parents of the Participants

Exploring CHC Theory's Cognitive Constructs in Young Children

Permission for Child's Participation

Project Description: This form requests permission for your child to participate in a research project titled "Exploring CHC Theory's Cognitive Constructs in Young Children." The project will explore how well three cognitive measures: the Stanford-Binet, Fifth Edition, the Kaufman Assessment Battery for Children, Second Edition, and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition measure thinking skills in young children. It is important to know whether a test actually measures what it says it will measure. For example, all three of the above mentioned tests purport to measure verbal ability (e.g., vocabulary skills, expressive language, and receptive language); this study will determine if they actually assess these skills. Specifically, this study will determine if similar cognitive constructs are being assessed when using these measure with young children.

Confidentiality of Research Project: For research purposes, all information regarding your child will be kept confidential. Your child's name will not be released for any reason.

Requirements of Participation: If your child participates, he or she will participate in approximately 4 hours of individual assessment, which will be spread across three days. The testing will be conducted while your child is in the preschool/daycare class and is typical of a type of testing routinely done by schools. School psychology graduate students from Ball State University will individually administer the three tests to your child. We will notify you and your child's daycare provider the week prior to testing.

Participation is Voluntary: Your child's participation in the research project is voluntary. You and your child can withdraw from the project at any time without negative consequences and, if you request, all data pertaining to your child will be destroyed. Please feel free to ask questions of the investigators before signing the Informed Consent Form and agreeing to your child's participation in this study and any time during the study.

Questions? Please call Kim Ward, Principal Investigator, at (765) 285-8500, with any questions. If you have any questions regarding the rights of research participants, please contact Ms. Sandra Smith, Coordinator of Research Compliance, Office of Academic Research and Sponsored Programs, Ball State University, Muncie, IN 47306, (765) 285-5070.

I have read and understand the above information and give permission for _____ (insert child's name) to participate in the research project entitled, "Exploring CHC Theory's Cognitive Constructs in Young Children."

Signature of Parent or Guardian

Date

Kimberly Ward, M.A.
Dept. of Educational Psychology TC 524
(765) 285-8500
keward2@bsu.edu

Barbara Rothlisberg, Ph.D.
Dept. of Educational Psychology TC 524
(765) 285-8500
brothlisberg@bsu.edu

APPENDIX C

Demographic Questionnaire Completed by Parents of Participants

Demographic Questionnaire

If you decided to give your son or daughter permission to participate in the research project, please complete the following:

Child's Birthdate: _____

Child's Sex (circle): Male Female

Child's Race (circle): African American White (Caucasian) Other _____

Hispanic/Latino Asian

Highest grade level completed by each parent in the home (place an X where appropriate):

Father: _____ Less than 9th grade
 _____ Some high school, no diploma
 _____ High School diploma or GED
 _____ 1-3 years of college
 _____ 4 or more years of college

Mother: _____ Less than 9th grade
 _____ Some high school, no diploma
 _____ High School diploma or GED
 _____ 1-3 years of college
 _____ 4 or more years of college

Current Occupation:

Father: _____

Mother: _____

APPENDIX D

T-Test Mean Comparison for the KABC-II Clusters and Subtests by Ethnicity

FCI and Clusters	<u>African American (n = 49)</u>			<u>Caucasian (n = 49)</u>		
	<i>t</i>	Significance	Mean dif.	<i>t</i>	Significance	Mean dif.
FCI	-1.316	.195	-2.408	-2.49	.061	-4.54
<i>Gsm</i>	-4.588	<.001*	-6.604	-3.61	.001*	-6.51
<i>Gv</i>	.636	.528	1.335	-1.71	.094	-3.07
<i>Glr</i>	-.486	.629	-.994	-.161	.872	-.333
<i>Gc</i>	-1.740	.088	-2.573	-3.67	.001*	-5.17
<u>Subtests</u>						
Number Recall	-4.675	<.001*	-1.431	-3.49	.001*	-1.12
Word Order	-3.728	.001*	-.945	-2.82	.007	-.978
Atlantis	-.850	.400	-.312	-.913	.366	-.427
Rebus	-.336	.738	-.161	.866	.391	.361
Conceptual Thinking	.265	.792	.133	-.695	.490	-.304
Face Recognition	.976	.334	.384	-.241	.811	-.084
Triangles	-.098	.922	-.037	-3.33	.002*	-1.12
Expressive Vocabulary	-1.185	.242	-.359	-.795	.431	-.218
Riddles	-3.089	.003*	-1.018	-5.13	<.001*	-1.77

Note. Test Values for all t-tests were obtained from table 8.7 in the KABC-II manual (Kaufman & Kaufman, 2004). Means adjusted for sex and maternal education were used as they most accurately represented the matched variables in the current study. All scores have *df* = 48.

*Significant *t*-test comparisons.